

Dissertation
zur Erlangung des akademischen Grades des Doktors der
Naturwissenschaften (Dr. rer. nat.)

**Vergleich
sphärischer und asphärischer
Elektronendichtemodelle zur
Verfeinerung von Makromolekülen**

von
Roman Kalinowski
aus Berlin

Februar 2010

Eingereicht am Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin

Die Dissertation wurde in der Zeit von April 2006 bis Februar 2010 am Institut für Chemie und Biochemie/Kristallographie unter der Leitung von Prof. Dr. Peter Luger angefertigt.

1. Gutachter: Prof. Dr. Peter Luger

2. Gutachter: Prof. Dr. Dieter Lentz

Disputation am 25.05.2010.

Inhaltsverzeichnis

Einführung	1
Grundlagen	13
1 Sphärische und asphärische Verfeinerung	15
1.1 Sphärische Verfeinerung des Röntgenbeugungsexperiments	15
1.2 Asphärische Verfeinerung des Röntgenbeugungsexperiments	19
1.2.1 Multipolformalismus	19
1.2.2 Invariome	22
2 Proteine	25
2.1 Protein Data Bank	25
2.2 Proteinfunktion	27
2.2.1 Lysozym	27
2.2.2 Insulin	29
2.2.3 Trichotoxin	30
2.3 Proteinaufbau	31
Methoden	33
3 Programme und Auswertung	35
3.1 Statistik	35
3.1.1 Mittelwerte und Streuung	35
3.1.2 Signifikanztest	36
3.1.3 Gütefaktoren und Gewichtung	39
3.2 Strukturqualität	40
3.2.1 Procheck	40
3.2.2 Whatcheck	41
3.2.3 Shelxpro	42
3.3 Programmanpassungen	43
3.4 Datenaufbereitung und -darstellung	44

4	Verfeinerung	49
4.1	Allgemeines	49
4.1.1	Sphärische Verfeinerung	51
4.1.2	Asphärische Verfeinerung	53
4.2	Alanyl-Prolyl-Alanin	67
4.2.1	Sphärisches Modell	68
4.2.2	Asphärisches Modell	72
4.3	Triklines Lysozym	73
4.3.1	Sphärisches Modell	74
4.3.2	Asphärisches Modell	79
4.4	Tetragonales Lysozym	83
4.4.1	Sphärisches Modell	84
4.4.2	Asphärisches Modell	91
4.5	Insulin	94
4.5.1	Sphärisches Modell	95
4.5.2	Asphärisches Modell	102
4.6	Trichotoxin	105
4.6.1	Sphärisches Modell	106
4.6.2	Asphärisches Modell	109
	Ergebnisse	113
5	Alanyl-Prolyl-Alanin	115
5.1	Gütefaktoren	115
5.2	Geometrieänderungen	117
5.3	Verschiebungsparameter	119
5.4	DMSDA-Werte	121
5.5	Restdichten	123
5.6	Deformationsdichten	129
5.7	Abgeleitete Eigenschaften	130
5.7.1	Vergleich <i>ab initio</i> -ESP mit Invariom-ESP	131
5.7.2	Vergleich semiempirisches ESP mit Invariom-ESP	140
5.7.3	Punktladungs-ESP	144
6	Triklines Lysozym	147
6.1	Gütefaktoren	147
6.2	Geometrie	150
6.3	Verschiebungsparameter	156
6.4	DMSDA-Werte	163
6.5	Restdichten	173
6.6	Deformationsdichten	181

6.7	Abgeleitete Eigenschaften	183
7	Tetragonales Lysozym	187
7.1	Gütefaktoren	187
7.2	Geometrie	189
7.3	Verschiebungsparameter	194
7.4	DMSDA-Werte	200
7.5	Restdichten	204
7.6	Deformationsdichten	211
7.7	Abgeleitete Eigenschaften	213
8	Insulin	217
8.1	Gütefaktoren	217
8.2	Geometrie	219
8.3	Verschiebungsparameter	224
8.4	DMSDA-Werte	230
8.5	Restdichten	234
8.6	Deformationsdichten	241
8.7	Abgeleitete Eigenschaften	243
9	Trichotoxin	247
9.1	Gütefaktoren	247
9.2	Geometrie	248
9.3	Verschiebungsparameter	251
9.4	DMSDA-Werte	257
9.5	Restdichten	261
9.6	Deformationsdichten	265
9.7	Abgeleitete Eigenschaften	266
10	Vergleich der Strukturen und Modellverbesserungen	269
10.1	Gütefaktoren	271
10.2	Strukturgröße	272
10.3	Verschiebungsparameter	274
10.4	Restdichten	275
10.5	Geometrieänderung	281
10.6	Strukturgröße	283
10.7	Schlußfolgerungen zur erfolgreichen asphärischen Verfeinerung	284
10.8	Vergleich zu literaturbekanntem Studien	287
	Zusammenfassung	289
	Summary	291

Danksagung	293
Literaturverzeichnis	295
Anhang	317
A Alanyl-Prolyl-Alanin	319
B Triklines Lysozym	339
C Tetragonales Lysozym	355
D Insulin	371
E Trichotoxin	389
Weitere Arbeiten	401
Veröffentlichungen	403

Abkürzungs- und Formelverzeichnis

ADP	engl. <i>Anisotropic Displacement Parameter</i> : Anisotroper Verschiebungsparameter
DMSDA	engl. <i>Difference of Mean Square Displacement Amplitudes</i> : Differenz der mittleren quadratischen Verschiebungsamplituden
ESP	Elektrostatisches Potential
IAM	engl. <i>Independent Atom Model</i> : Modell unabhängiger Atome
kDa	Kilodalton, $1 \text{ Da} = 1 \text{ amu} \approx 1,661 \cdot 10^{-27} \text{ kg}$
RMS	engl. <i>Root Mean Square</i> : quadratisches Mittel
RMSD	engl. <i>Root Mean Square Deviation</i> : quadratisch gemittelte Abweichung
vdW	van der Waals
<i>mul</i>	asphärisches Verfeinerungsmodell
<i>iam</i>	sphärisches Verfeinerungsmodell
<i>iam_{rb}</i>	sphärisches Verfeinerungsmodell, verlängerte Wasserstoffabstände (<i>rb</i> = <i>reset bond</i>)
A_{vdW}	van-der-Waals-Oberfläche
F_o, F_c	beobachtete und berechnete Strukturfaktoren
<i>IQR</i>	Interquartilabstand
<i>MAD</i>	Median der absoluten Abweichungen zum Median
<i>occ</i>	engl. <i>occupancy</i> : Besetzung
ϕ	Elektrostatisches Potential

ϕ_{mul}	Elektrostatisches Potential aus Multipolelektronendichte
ϕ_{PC}	Elektrostatisches Potential aus Punktladungen
$Q1$	25 %-Quartil
$Q2$	Median oder 50 %-Quartil
$Q3$	75 %-Quartil
$\Delta\rho_{def}$	Deformationsdichte
$\Delta\rho_{res}$	Restdichte
ρ_{mul}	Multipolelektronendichte
R	engl. <i>residual factor</i> : Gütefaktor
$R1$	Gütefaktor, berechnet mit $F_o(> 4\sigma(F_o))$
R_{all}	Gütefaktor, berechnet mit F_o
$wR2$	gewichteter Gütefaktor, berechnet mit F_o^2
σ_x	Standardabweichung des arithmetischen Mittels
U_{eq}	äquivalenter isotroper Verschiebungsparameter
V_M	Matthews-Koeffizient
V_{asu}	Volumen der asymmetrischen Einheit
V_{vdW}	van-der-Waals-Volumen
\bar{x}	arithmetisches Mittel

Abbildungsverzeichnis

1.1	Multipole	20
1.2	Ableitung eines Invarioms	23
1.3	Elektronendichte eines Invarioms	24
2.1	Anzahl Strukturen der PDB	25
2.2	Gütefaktorhistogramm PDB	26
3.1	Mittelwerte und Streuungsmaße.	37
4.1.1	Überblick zur Verfeinerung der Strukturen.	49
4.2.1	Elementarzelle APA	68
4.2.2	Elektronendichte APA an Atompositionen	71
4.3.1	Strukturmodell P1LYSO	75
4.3.2	Elektronendichte P1LYSO an Atompositionen	77
4.3.3	Packung P1LYSO	78
4.3.4	Verfeinerter Strukturteil P1LYSO	81
4.4.1	Strukturmodell TETRALYSO	87
4.4.2	Elektronendichte TETRALYSO an Atompositionen	89
4.4.3	Packung TETRALYSO	90
4.4.4	Verfeinerter Strukturteil TETRALYSO	93
4.5.1	Strukturmodell INSULIN	96
4.5.2	Elektronendichte INSULIN an Atompositionen	99
4.5.3	Packung INSULIN	100
4.5.4	INSULIN-Hexamer	101
4.5.5	Verfeinerter Strukturteil INSULIN	103
4.6.1	Elektronendichte TRICHOTOXIN an Atompositionen	108
4.6.2	Packung TRICHOTOXIN	109
5.1	APA: Quotient beobachteter und berechneter Intensitäten	117
5.2	APA/APA0.83: Strukturunterschiede durch asphärische Verfeinerung	118
5.3	APA/APA0.83: Veränderung U_{eq}	120
5.4	APA/APA0.83: DMSDA-Werte und Veränderung	122
5.5	APA: Restdichten	124
5.6	APA0.83: Restdichten	125
5.7	APA: Restdichtehistogramm	127
5.8	APA0.83: Restdichtehistogramm	128

5.9	APA: Deformationsdichte	130
5.10	APA: Elektrostatisches Potential	131
5.11	APA: <i>ab initio</i> -ESP	132
5.12	APA: Unterschiede Multipol und <i>ab initio</i> -ESP	134
5.13	APA: Histogramm der <i>ab initio</i> -Multipol-Differenzen	138
5.14	APA: AM1-ESP	140
5.15	APA: Differenz AM1 – <i>mul</i>	141
5.16	APA: Histogramm der Multipol-AM1-Differenzen	141
5.17	APA: MOPAC-ESP	143
5.18	APA: Multipol-ESP und Punktladungs-ESP	145
6.1	P1LYSO: Quotient beobachteter und berechneter Intensitäten	150
6.2	P1LYSO: Strukturunterschied <i>iam</i> und <i>mul</i>	152
6.3	P1LYSO: Strukturunterschied <i>iam_{rb}</i> und <i>mul</i>	153
6.4	P1LYSO: Veränderung der U_{eq} -Werte	157
6.5	P1LYSO: Veränderung der U_{eq} -Werte, Auftragung nach U_{eq} (<i>iam</i>)	158
6.6	P1LYSO: Anisotrope Verschiebungsparamter	161
6.7	P1LYSO: U_{eq} in Abhängigkeit vom Abstand zum Massenzentrum	162
6.8	P1LYSO: DMSDA-Wertänderung	164
6.9	P1LYSO: DMSDA-Wertänderung, Auftragung nach DMSDA(<i>iam</i>)	165
6.10	P1LYSO: DMSDA-Wertänderung der CO-Bindungen	169
6.11	P1LYSO: DMSDA-Wertänderung der CC-Bindungen	170
6.12	P1LYSO: DMSDA-Wertänderung der CN-Bindungen	171
6.13	P1LYSO: DMSDA-Wertänderung der CS- und SS-Bindungen	172
6.14	P1LYSO: Restdichte <i>mul</i>	174
6.15	P1LYSO: Restdichte <i>iam</i>	175
6.16	P1LYSO: Restdichtedifferenzen <i>iam-mul</i>	176
6.17	P1LYSO: Restdichten TRP 108	177
6.18	P1LYSO: Restdichten GLU 35	178
6.19	P1LYSO: Histogramm der Restdichte	180
6.20	P1LYSO: Deformationsdichte TRP 108 und GLU 35	181
6.21	P1LYSO: Deformationsdichte	182
6.22	P1LYSO: Elektrostatisches Potential	184
6.23	P1LYSO: AM1-ESP	185
6.24	P1LYSO: Punktladungs-ESP und Multipol-ESP	186
7.1	TETRALYSO: Quotient beobachteter und berechneter Intensitäten	189
7.2	TETRALYSO: Strukturunterschied <i>iam</i> und <i>mul</i>	190
7.3	TETRALYSO: Strukturunterschied <i>iam_{rb}</i> und <i>mul</i>	191
7.4	TETRALYSO: Veränderung U_{eq}	195
7.5	TETRALYSO: Veränderung U_{eq} -Werte, Auftragung nach U_{eq} (<i>iam</i>)	196
7.6	TETRALYSO: Anisotrope Verschiebungsparamter	198

7.7	TETRALYSO: U_{eq} in Abhängigkeit vom Abstand zum Massenzentrum	199
7.8	TETRALYSO: DMSDA-Wertänderung	201
7.9	TETRALYSO: DMSDA-Wertänderung, Auftragung nach DMSDA(<i>iam</i>)	202
7.10	TETRALYSO: Histogramm der Restdichten	205
7.11	TETRALYSO: Restdichte <i>mul</i>	206
7.12	TETRALYSO: Restdichte <i>iam</i>	207
7.13	TETRALYSO: Restdichtedifferenz <i>iam-mul</i>	208
7.14	TETRALYSO: Restdichte TRP 108	209
7.15	TETRALYSO: Restdichte GLU 35	210
7.16	TETRALYSO: Deformationsdichten TRP 108 und GLU 35	211
7.17	TETRALYSO: Deformationsdichte	212
7.18	TETRALYSO: Elektrostatisches Potential	214
7.19	TETRALYSO: Punktladungs-ESP und Multipol-ESP	215
8.1	INSULIN: Quotient beobachteter und berechneter Intensitäten	219
8.2	INSULIN: Strukturunterschied <i>iam</i> und <i>mul</i>	220
8.3	INSULIN: Strukturunterschied <i>iam_{rb}</i> und <i>mul</i>	221
8.4	INSULIN: Veränderung der U_{eq} -Werte	225
8.5	INSULIN: Veränderung der U_{eq} -Werte, Auftragung nach U_{eq} (<i>iam</i>)	226
8.6	INSULIN: Anisotrope Verschiebungsparameter	228
8.7	INSULIN: U_{eq} in Abhängigkeit vom Abstand zum Massenzentrum	229
8.8	INSULIN: DMSDA-Wertänderung	231
8.9	INSULIN: DMSDA-Wertänderung, Auftragung nach DMSDA(<i>iam</i>)	232
8.10	INSULIN: Histogramm der Restdichte	235
8.11	INSULIN: Restdichten LEU 215	236
8.12	INSULIN: Restdichten TYR226	237
8.13	INSULIN: Restdichte <i>mul</i>	238
8.14	INSULIN: Restdichte <i>iam</i>	239
8.15	INSULIN: Restdichtedifferenzen <i>iam-mul</i>	240
8.16	INSULIN: Deformationsdichten LEU 215 und TYR 226	241
8.17	INSULIN: Deformationsdichte	242
8.18	INSULIN: Elektrostatisches Potential	244
8.19	INSULIN: Multipol- und Punktladungs-ESP, Orientierung 1	245
8.20	INSULIN: Multipol- und Punktladungs-ESP, Orientierung 2	246
9.1	TRICHOTOXIN: Quotient beobachteter und berechneter Intensitäten	248
9.2	TRICHOTOXIN: Strukturunterschied <i>iam/iam_{rb}</i> und <i>mul</i>	249
9.3	TRICHOTOXIN: Veränderung der U_{eq} -Werte	252
9.4	TRICHOTOXIN: Veränderung der U_{eq} -Werte, Auftragung nach U_{eq} (<i>iam</i>)	253
9.5	TRICHOTOXIN: Anisotrope Verschiebungsparameter	255
9.6	TRICHOTOXIN: U_{eq} in Abhängigkeit vom Abstand zum Massenzentrum	256
9.7	TRICHOTOXIN: DMSDA-Wertänderung	258

9.8	TRICHOTOXIN: DMSDA-Wertänderung, Auftragung nach DMSDA(<i>iam</i>)	259
9.9	TRICHOTOXIN: Histogramm der Restdichte	262
9.10	TRICHOTOXIN: Restdichte und Restdichtedifferenz	263
9.11	TRICHOTOXIN: Deformationsdichte	265
9.12	TRICHOTOXIN: Elektrostatisches Potential	266
9.13	TRICHOTOXIN: Multipol- und Punktladungs-ESP	268
10.1	<i>R</i> -Wertverbesserung, Abhängigkeit von <i>R</i> (<i>iam</i>) und Auflösung	271
10.2	<i>R</i> -Wertverbesserung, Abhängigkeit von Strukturunsicherheit	273
10.3	<i>R</i> -Wertverbesserung, Abhängigkeit vom Temperaturfaktor	274
10.4	Restdichte auf den Bindungsmitten	276
10.5	<i>R</i> -Wertverbesserung, Abhängigkeit von Restdichte	280
10.6	<i>R</i> -Wertverbesserung, Abhängigkeit von $\Delta\rho_{def}/\Delta\rho_{res}$ -Quotient	281
A.1	APA: skaliertes AM1-ESP	327
A.2	APA: PROCHECK-Ausgabe	328
A.3	APA: WHATCHECK-Ausgabe	330
A.4	APA: SHELXPRO-Ausgabe	331
A.5	APA0.83: PROCHECK-Ausgabe	333
A.6	APA0.83: WHATCHECK-Ausgabe	335
A.7	APA0.83: SHELXPRO-Ausgabe	336
B.1	P1LYSO: Anisotrope Verschiebungsparameter	342
B.2	P1LYSO: Restdichte iam_{rb}	343
B.3	P1LYSO: PROCHECK-Ausgabe	344
B.4	P1LYSO: WHATCHECK-Ausgabe	348
B.5	P1LYSO: SHELXPRO-Ausgabe	352
C.1	TETRALYSO: Anisotrope Verschiebungsparameter	356
C.2	TETRALYSO: Restdichte iam_{rb}	359
C.3	TETRALYSO: PROCHECK-Ausgabe	360
C.4	TETRALYSO: WHATCHECK-Ausgabe	364
C.5	TETRALYSO: SHELXPRO-Ausgabe	368
D.1	INSULIN: Anisotrope Verschiebungsparameter	373
D.2	INSULIN: Restdichte iam_{rb}	375
D.3	INSULIN: PROCHECK-Ausgabe	376
D.4	INSULIN: WHATCHECK-Ausgabe	380
D.5	INSULIN: SHELXPRO-Ausgabe	384
E.1	TRICHOTOXIN: Restdichte iam_{rb}	391
E.2	TRICHOTOXIN: WHATCHECK-Ausgabe	392
E.3	TRICHOTOXIN: PROCHECK-Ausgabe	395
E.4	TRICHOTOXIN: SHELXPRO-Ausgabe	398

Tabellenverzeichnis

4.1.1	Invariome und Koordinatensysteme der Aminosäuren	55
4.1.2	Wasserstoffabstände im Invariommodell	66
4.2.1	Kristallographische Daten APA.	67
4.2.2	Reflexstatistik APA	69
4.3.1	Kristallographische Daten P1LYSO	73
4.3.2	Reflexstatistik P1LYSO	76
4.4.1	Kristallographische Daten TETRALYSO.	83
4.4.2	Verfeinerungsschritte TETRALYSO.	84
4.4.3	Reflexstatistik TETRALYSO	88
4.5.1	Kristallographische Daten INSULIN.	94
4.5.2	Verfeinerungsschritte INSULIN.	95
4.5.3	Reflexstatistik INSULIN	98
4.6.1	Kristallographische Daten TRICHOTOXIN	105
4.6.2	Reflexstatistik TRICHOTOXIN	107
4.6.3	Invariome und Koordinatensysteme, Ergänzung	110
5.1	APA: Gütefaktoren	115
5.2	APA0.83: Gütefaktoren	116
5.3	APA/APA0.83: Restdichtestatistik	126
5.4	APA: Statistik der <i>ab initio</i> -Multipol-Differenzen	137
5.5	APA: Statistik der Multipol-AM1-Differenzen	142
5.6	APA: Statistik der ESP-Differenz Multipol und mopac-AM1	143
6.1	P1LYSO: Gütefaktoren	147
6.2	P1LYSO: Gütefaktoren mit beschnittenen Daten	149
6.3	P1LYSO: Peptidbindungslängen	154
6.4	P1LYSO: Differenzen der Peptidbindungslängen	155
6.5	P1LYSO: Statistik verfeinerte U_{eq}	160
6.6	P1LYSO: Statistik der DMSDA-Beträge der Peptidbindungen	166
6.7	P1LYSO: Statistik der Restdichten	179
7.1	TETRALYSO: Gütefaktoren	187
7.2	TETRALYSO: Gütefaktoren mit beschnittenen Daten	188
7.3	TETRALYSO: Peptidbindungslängen	192
7.4	TETRALYSO: Differenz der Peptidbindungslängen	193
7.5	TETRALYSO: Statistik U_{eq}	197

7.6	TETRALYSO: Statistik der DMSDA-Beträge der Peptidbindungen . . .	203
7.7	TETRALYSO: Statistik der Restdichten	204
8.1	INSULIN: Gütefaktoren	217
8.2	INSULIN: Gütefaktoren mit beschnittenen Daten	218
8.3	INSULIN: Peptidbindungslängen	222
8.4	INSULIN: Differenzen der Peptidbindungslängen	223
8.5	INSULIN: Statistik U_{eq}	227
8.6	INSULIN: Statistik der DMSDA-Beträge der Peptidbindungen	233
8.7	INSULIN: Restdichtestatistik	234
9.1	TRICHOTOXIN: Gütefaktoren	247
9.2	TRICHOTOXIN: Peptidbindungslängen	250
9.3	TRICHOTOXIN: Differenzen der Peptidbindungslängen	251
9.4	TRICHOTOXIN: Statistik U_{eq}	254
9.5	TRICHOTOXIN: Statistik der DMSDA-Beträge der Peptidbindungen .	260
9.6	TRICHOTOXIN: Restdichtestatistik	261
10.1	Vergleich ausgewählter Parameter	269
A.1	APA: Geometrieunterschiede nach <i>mul</i>	319
A.2	APA0.83: Geometrieunterschiede nach <i>mul</i>	320
A.3	APA/APA0.83: Bindungslängen	321
A.4	APA/APA0.83: Bindungslängendifferenzen	322
A.5	APA: Statistik U_{eq} und ΔU_{eq}	322
A.6	APA0.83: Statistik U_{eq} und ΔU_{eq}	323
A.7	APA: Statistik U_{eq} und ΔU_{eq} – nach Element	324
A.8	APA0.83: Statistik U_{eq} und ΔU_{eq} – nach Element	324
A.9	APA/APA0.83: DMSDA-Werte	325
A.10	APA/APA0.83: $ \Delta\text{DMSDA} $	326
B.1	P1LYSO: Bindungstypen	339
B.2	P1LYSO: Statistik U_{eq}	340
B.3	P1LYSO: Statistik verfeinerte U_{eq} – nach Element	341
C.1	TETRALYSO: Statistik U_{eq}	355
C.2	TETRALYSO: Statistik U_{eq} – nach Elementen	357
C.3	TETRALYSO: Statistik der DMSDA-Beträge	358
D.1	INSULIN: Statistik U_{eq}	371
D.2	INSULIN: Statistik U_{eq} – nach Elementen	372
D.3	INSULIN: Statistik der DMSDA-Beträge	374
E.1	TRICHOTOXIN: Statistik der U_{eq}	389
E.2	TRICHOTOXIN: Statistik U_{eq} – nach Elementen	390
E.3	TRICHOTOXIN: Statistik der DMSDA-Beträge	390

Einführung

Die Entdeckung der Beugung von Röntgenstrahlung an Kristallen im Jahr 1912 durch Max von Laue, Paul Knipping und Walter Friedrich (Friedrich u. a. 1913) und die Erkenntnis Peter Debyes drei Jahre später, daß daraus „die Elektronenanordnung im Atom experimentell festzustellen“ (Debye 1915) sei, liegt nunmehr fast ein Jahrhundert zurück. Es bedurfte etwa fünfzig Jahren, bis sich die Röntgenbeugung zu der leistungsfähigen Methode zur Beschreibung der Elektronendichteverteilung in Kristallen entwickelte, mit der seitdem eine enorme Anzahl von Strukturen aufgeklärt werden konnte. Zur Modellierung des Experiments wurde und wird aber hauptsächlich das Modell unabhängiger Atome (*independent atom model*, IAM) benutzt, welches die atomaren Streubeiträge anhand sphärischer Elektronendichteverteilungen von Grundzustandsatomen beschreibt, deren Überlagerung Dichtedeformationen und Ladungsverschiebungen durch Bindungsbildung zwischen den Atomen unberücksichtigt lassen muß.

Um der frühzeitig erkannten Limitierung dieses Ansatzes zu begegnen, wurden ihm asphärische Streumodelle, basierend auf atomzentrierten Multipolfunktionen, gegenübergestellt (Dawson 1967; Kurki-Suonio 1968; Stewart 1969; Hirshfeld 1971). Mit der Formulierung als Hansen&Coppens-Multipolformalismus (Hansen u. Coppens 1978; Coppens 1997) fand dieses Modell seitdem breite Anwendung (Koritsanzky u. Coppens 2000). Infolge der rasanten Verbesserung der technischen Möglichkeiten während der letzten 10–15 Jahre, wie der Verfügbarkeit von hochbrillianten Synchrotronquellen, Tieftemperaturmeßplätzen, Flächendetektoren sowie leistungsstarken Computern samt ebensolcher Software, wird die aufwendigere asphärische Verfeinerung von kleinen Molekülen als ausgereift betrachtet (Coppens 2005) und als zukünftige Routineanwendung gesehen (Luger 2007).

Fortschritte der experimentellen Techniken ermöglichen seit relativ kurzer Zeit atomare Auflösung auch für Proteine (Dauter u. a. 1995, 1997). Die steigende Anzahl hochaufgelöster Proteinstrukturen eröffnet wichtige strukturelle Informationen wie Lokalisierung von Wasserstoffatomen, Auflösung mehrerer Alternativkonformationen und die Verfeinerung anisotroper Temperaturfaktoren (Vrielink u. Sampson 2003). Oftmals sind die nun zugänglichen feineren Strukturdetails direkt mit der Proteinfunktion verbunden, wie beispielsweise der Protonierungszustand von katalytischen Zentren (Petrova u. Podjarny 2004).

Für sehr hoch aufgelöste und wenig fehlgeordnete Kristallstrukturen von Makromolekülen können durch Bindungseffekte verursachte Restdichteanhäufungen auf den Bindungen beobachtet werden, wie Lamzin u. a. (1999) erstmals für eine

Crambin-Struktur bei $d = 0,67 \text{ \AA}$ und Savinase bei $d = 0,9 \text{ \AA}$ beschreiben. Allerdings schränken die Autoren ein, daß individuelle Bindungen keine Abweichung vom sphärischen Modell nahelegen, erst die Mittelung über alle Peptidbindungsregionen zeigt Restdichtekarten mit signifikanten Signalen auf den Peptidbindungen. Eine Schätzung der deformierten Elektronendichtemenge ergibt etwa ein Zehntel der typischen Werte von kleinen Molekülen.

Afonine u. a. (2004) diskutieren den Einfluß der Datenauflösung und Temperaturfaktoren auf die Sichtbarkeit von Restdichteakkumulationen auf den Bindungen in einer sphärischen Verfeinerung. Die Autoren untersuchen die deformierte Dichte in Form von Restdichten eines sphärisches Modells, statt Deformationsdichten aus einem asphärischen Modell, um eine eventuelle Verzerrung der Ergebnisse durch das Multipolmodell zu vermeiden. Auf Fourierabbrucheffekte, die zusätzliche Dichte in der Bindungsregion verursachen können, wird hingewiesen, der Einfluß auf Restdichtekarten aber als gering eingestuft. Im Idealfall der Verfeinerung theoretisch berechneter Strukturfaktoren des Moleküls $\text{CH}_3\text{-NH-CO-CH}_3$ bei $d = 0,6 \text{ \AA}$ wird Bindungsrestdichte nur bis zu äquivalenten isotropen Temperaturfaktoren von $B_{\text{eq}} = 7 \text{ \AA}^2$ ($U_{\text{eq}} = 89 \cdot 10^{-3} \text{ \AA}^2$) beobachtet. Auch bei geringerer Auflösung $d = 0,9 \text{ \AA}$ bleiben die Merkmale erhalten, allerdings nur, wenn die Temperaturfaktoren aus der Verfeinerung der Daten bis $0,6 \text{ \AA}$ stammen, ansonsten wird keine Bindungsdichte beobachtet. Eine Auflösung von $d = 0,6 \text{ \AA}$ wird demnach benötigt, um eine artifizielle Vergrößerung der ADPs zu vermeiden, weil dies wiederum Bindungsrestdichten absorbiert (ein Effekt, bereits beschrieben von Coppens 1967). Im Fall realer Daten erhöhen sich die Anforderungen an die Temperaturfaktoren, respektive die Auflösung, da nun experimentelle Fehler die Beobachtbarkeit der Dichteakkumulationen erschweren.

Erfolgreiche asphärische Verfeinerungen könnten die Beschreibung von Makromolekülstrukturen verbessern, zumindest für kleine Moleküle sind die Vorteile evident. Die Beschreibung von Bindungsrestdichten im Multipolmodell führt nicht nur zu besseren Gütefaktoren, sondern sollte eine bessere Trennung von thermischer Bewegung und Verschmierung der Dichte durch Bindungseffekte bewirken, was zu kleineren und physikalisch sinnvollerem Verschiebungsparametern führt. Zudem ermöglicht die asphärische Beschreibung verbesserte Geometrien, da beispielsweise Wasserstoffatomabstände nicht verkürzt in die Verfeinerung eingehen müssen und Bindungslängen zu Sauerstoffatomen nicht mehr durch die Dichte der freien Elektronenpaare zu überhöhten Werten verfälscht werden. Zudem können aus der Multipoldichte Eigenschaften wie z. B. das elektrostatische Potential (ESP) in hoher Güte abgeleitet werden. All das wäre von großem Nutzen, insbesondere auf dem Gebiet der *life sciences* und des *drug designs*, weil molekulare Erkennungsprozesse nicht mit sterischen Wechselwirkungen allein beschreibbar sind, sondern eine einflußreiche elektrostatische Komponente haben.

Die Verfeinerung eines Multipolmodells benötigt allerdings bis zu 40 Parameter

pro Atom und stellt hohe Anforderungen an die experimentellen Daten – so sollte die Auflösung unter 0,5 Å liegen und die sphärische Verfeinerung eines vollständigen Strukturmodells bereits niedrige Gütefaktoren ($R1 < 5\%$) ergeben. Für größere Strukturen sind diese Voraussetzungen immer schwerer zu erfüllen, für Proteine nahezu ausgeschlossen. Um dennoch asphärische Verfeinerungen von Makromolekülen zu ermöglichen, können feststehende, anderweitig gewonnene Multipolparameter als Streufaktoren auf diese Strukturen übertragen werden. Wie im folgenden erläutert wird, ist dafür die Transferierbarkeit von atomaren Elektronendichten im Allgemeinen und von Pseudoatom genannten atomaren Multipolpopulationen im Besonderen die Grundlage. Dies führt direkt zu Pseudoatomdatenbanken, deren Aufbau und bisherige Verwendung ebenfalls vorgestellt wird.

Transferierbarkeit

Wie Hohenberg u. Kohn (1964) zeigten, ist die Elektronendichteverteilung eines Moleküls von fundamentaler Bedeutung, da sie alle statischen und reaktiven Eigenschaften bestimmt. Mit der „Quantentheorie der Atome in Molekülen“ („*quantum theory of atoms in molecules*“ QTAM Bader 1990) steht eine quantitative Methode zur direkten Untersuchung einer Elektronendichteverteilung zur Verfügung. Aus einer topologischen Analyse ergeben sich eindeutig bestimmte Atome, begrenzt durch Nullflußoberflächen im Gradientenvektorfeld der Elektronendichte, denen Eigenschaften wie Volumina, Ladungen und Dipolmomente zugeordnet sind. Des Weiteren resultiert die Definition von Bindungen zwischen Atomen (Bader 2009), charakterisiert durch die Eigenschaften der Dichte und ihrer Ableitungen an kritischen Punkten. Somit führt dieses Partitionierungsschema zum Konzept eines Moleküls als über Bindungen verknüpfter Atome, deren Eigenschaften additiv und zwischen verschiedenen Molekülen übertragbar sind (Bader 1991), was der chemischen Anschauung – ähnliche Atome oder funktionelle Gruppen haben eine ähnliche Elektronendichte – nahekommt (Bader 2008).

Zahlreiche Studien an Elektronendichteverteilungen aus *ab initio*-Rechnungen untermauern das Transferierbarkeitsprinzip, unter anderem anhand der Analyse von QTAM-Größen wie Volumina, Ladungen und bindungskritischer Punkte von Atomen in homologen Reihen (Stefanov u. Cioslowski 1996; Graña u. Mosquera 1999; Quinonez u. a. 2003; Alkorta u. Picazo 2005), besonders hervorzuheben ist im hiesigen Kontext die systematische Untersuchung der 20 kanonischen Aminosäuren (Matta u. Bader 2000, 2002, 2003). Darüber hinaus wurde die Torsionswinkelabhängigkeit obiger Größen betrachtet (Popelier u. Bader 1994), ebenso wie die Übertragbarkeit funktioneller Gruppen (Cortes-Guzman u. Bader 2004). Das Konzept wurde auch bezüglich weiterer Eigenschaften, wie der *source function* (Bader u. Gatti 1998; Gatti u. a. 2003; Gatti u. Bertini 2004), der atomaren Lochfunktion von Austausch-Korrelations-Dichten (Fradera u. a. 2000), der Dipol-Polarisierbarkeit

von Methyl- und Methylengruppen (Stone u. a. 1996) oder dem elektrostatischen Potential (Popelier u. a. 2004) überprüft. Zusätzliche Validierungen erfolgten anhand der Vorhersage experimentell bestimmbarer physikochemischer Eigenschaften (z. B. Kopplungskonstanten oder partielles Molvolumen) einiger biologisch relevanter Moleküle wie Aminosäuren oder Opiate (Matta 2003). Aus dem Transferierbarkeitsprinzip folgten Versuche, die Eigenschaften größerer Strukturen aus Fragmenten zu konstruieren (Chang u. Bader 1992; Matta 2001). In diesem Zusammenhang stehen die Ansätze, anhand sieben atomarer Eigenschaften die QTAM-Atomtypen in Aminosäuren zu definieren (Popelier u. Aicken 2003a,b), was zu 5 Kohlenstoffatomtypen (von insg. 213 Atomen), 7 Wasserstoffatomtypen (von 421), 6 Sauerstoffatomtypen (von 63), 2 Stickstoffatomtypen (von 57) und 6 Schwefelatomtypen (von 6) führt und dem Befund, daß die unmittelbare Koordination das ausschlaggebende Kriterium ist.

Im Prinzip ermöglicht die Multipolbeschreibung die Rekonstruktion der Elektronendichteverteilung der kristallisierten Strukturen aus Strukturparametern. Somit wurden auch für experimentell bestimmte Elektronendichten Transferierbarkeitsanalysen durchgeführt. Mit Blick auf Proteinstrukturen wurden ausführliche Studien zu atomaren und topologischen Eigenschaften von Aminosäuren (eine Zusammenfassung geben Mebs u. a. 2006), Dipeptiden (Dittrich u. a. 2000), Tripeptiden des Typs L-Ala-Xxx-L-Ala (zusammengefaßt von Grabowsky u. a. 2009) und Hexapeptiden (Dittrich u. a. 2002, 2003) durchgeführt. Weitere Vergleiche topologischer Deskriptoren mehrerer Peptide erfolgten durch Pichon-Pesme u. a. (2000) und Pichon-Pesme u. Lecomte (1998). Die Untersuchung obiger Größen ist aber nicht auf diese Substanzklassen limitiert (Scheins u. a. 2005; Madsen u. a. 1998; Tsirelson u. a. 2006).

Transferierbarkeit von Pseudoatomen

Die Frage nach einer äquivalenten Transferierbarkeit von Pseudoatomen, bzw. der sie konstituierenden Multipole wurde von einigen Autoren adressiert. Erstmals wurden Deformationsdichten, in Form von Multipolpopulationen in der Formulierung nach Hirshfeld, durch Brock u. a. (1991) von einer hochaufgelösten Perylenstruktur auf weniger gute Naphthalen- und Anthracendaten übertragen, um die Beschreibung der anisotropen Temperaturfaktoren zu verbessern.

In einem ganz ähnlichen Verfahren benutzten Pichon-Pesme u. a. (1995) die experimentell bestimmten Hansen&Coppens-Multipolpopulationen dreier hochaufgelöster Peptidstrukturen, um zwei niedrig aufgelöste Datensätze eines Tripeptids mit chemisch ähnlichen Atomen asphärisch zu verfeinern, wodurch bei diesem die Gütefaktoren verbessert, die Verschiebungsparameter und DMSDA-Werte verringert sowie sinnvolle Deformationsdichtekarten erhalten wurden. Die übertragenen Multipolpopulationen der drei Peptide zeigten für äquivalente Atome jeweils gute

Übereinstimmung, die signifikanten Parameter wurden gemittelt und, bei Beibehaltung des lokalen Koordinatensystems, auf die entsprechenden Atome des Zielmoleküls übertragen. Die Autoren befinden die Konformationsabhängigkeit als sehr gering. Final wird die Möglichkeit des Aufbaus einer Datenbank von Multipolparametern für alle organischen Atomtypen projiziert, die der asphärischen Verfeinerung von Proteinen dienen soll.

Ähnliche Untersuchungen auf theoretischem Niveau (B3LYP/6-31G(d,p)) wurden von Koritsánszky u. a. (2002) für fünf Tripeptide mit konstantem mittleren Serinrest und variierenden terminalen Resten durchgeführt, um die Möglichkeit einer Datenbank von Pseudoatomen zu eruieren, mit dem letztendlichen Ziel, damit Makromoleküle zu verfeinern. Die Populationen der Hauptkettenatome des mittleren Aminosäurerestes zeigten sich unabhängig von den umgebenden Resten. Aus Mittelwerten dieser Parameter konnte die Deformationsdichte für eines der Tripeptide rekonstruiert werden. Die QTAM-Analyse der Datenbankdichte zeigte bezüglich Bindungstopologie und elektrostatischer Momente gute Übereinstimmung zur direkt berechneten. Die Konformationsabhängigkeit der Pseudoatomparameter wurde für gering aber nicht völlig vernachlässigbar befunden. Bezüglich der Auffassung von Multipolpopulationen als eindeutige Eigenschaft eines Atoms, wird die Forderung nach enger Lokalisierung der Multipoldichte aufgestellt und das Konzept als praktikable Grundlage für weitere Datenbankansätze unterstützt.

Für ein andersartiges Schema der Multipolexpansion atomarer Ladungsverteilung aus HF/6-31G(d,p)-Rechnungen (CAMM; Sokalski u. Poirier 1983) fanden Kdzierski u. Sokalski (2001) ebenfalls geringe Konformationsabhängigkeit im Fall kleiner Rotationsbarrieren und sehr gute Transferierbarkeit der Multipole kleiner Fragmente auf größere symmetrische Moleküle aber geringere Übertragbarkeit im Fall asymmetrischer Umgebung im Zielmolekül. Die Konformationsabhängigkeit von Multipolen nach dem Formalismus von Stone (1981) (*DMA, distributed multipole analysis*) wurde allerdings bezüglich der abgeleiteten elektrostatischen Eigenschaften für signifikant befunden (Koch u. a. 1995; Koch u. Stone 1996).

Datenbanken

Mehrere Pseudoatomdatenbanken auf der Grundlage des Hansen&Coppens-Multipolformalismus wurden entwickelt: die ELMAM („*Experimental Library Multipolar Atom Model*“; Pichon-Pesme u. a. 1995, 2004; Zarychta u. a. 2007), die UBDB („*University at Buffalo Pseudoatom Databank*“; Volkov u. a. 2004c; Dominiak u. a. 2007; Volkov u. a. 2007) und die Invariom-Bibliothek („*Invariant Atoms*“; Dittrich u. a. 2004, 2006a).

Sie unterscheiden sich primär durch die Art der Erzeugung ihrer Einträge und ihre bisherigen Anwendungsgebiete. Ihre Validierung beruht zumeist auf der praktizierten Übertragung von Multipolen auf eine Struktur und deren anschließende

erfolgreiche Verfeinerung oder der Ableitung physikalischer Eigenschaften, denn, wie Volkov, Koritsánszky, Li, u. Coppens 2004b statuieren, der alleinige Vergleich von Multipolparametern ist nicht zielführend, da diese keine wohldefinierte physikalische Bedeutung tragen, im Gegensatz zu abgeleiteten physikalischen Größen.

Letztendlich stellen sie generalisierte asphärische atomare Formfaktoren dar, welche das überwiegend benutzte Modell unabhängiger Atome ersetzen können, um eine adäquatere Beschreibung eines Röntgenbeugungsexperiments zu ermöglichen und darüber hinaus elektronische Eigenschaften damit beschriebener Strukturen in feinen Details zugänglich zu machen.

ELMAM

Nur die ELMAM-Datenbank benutzt experimentell aus Verfeinerungen hochwertiger und hochaufgelöster Kristallstrukturen bestimmte Parameter, die für ähnliche Atomtypen gemittelt werden. Dieser Bezug von Multipolpopulationen birgt im Prinzip die Möglichkeit Kristalleffekte in die Datenbank zu inkludieren, obwohl diesbezüglich unterschiedliche Ansichten vertreten werden (für theoretische Studien dazu siehe bspw. Spackman u. a. 1999; de Vries u. a. 2000). Zudem besteht dadurch das Problem, für alle interessierenden Atomtypen entsprechend hochwertige Daten und Verfeinerungen zu erlangen. Aus 14 Strukturen (100 K, $d < 0,45 \text{ \AA}$, $wR(F^2) < 4\%$), die mit konsistenten lokalen Achsensystemen asphärisch gegen die experimentellen Daten verfeinert wurden, resultierten Multipolparameter, die für äquivalente Atome gemittelt wurden, wobei die Standardabweichung für die stärksten Parameter 5 % beträgt. Gering populierte ($< 0,03$) und unsichere ($< 3\sigma$) Parameter werden auf Null gesetzt. Tests an sechs kleinen, geringer aufgelösten ($0,46 \leq d \leq 0,85 \text{ \AA}$) Di- und Tripeptiden sowie dem Oktapeptid Pseudophomin A, ergaben relative R -Wertverbesserungen von 13 % bis 47 % und signifikante Verbesserungen der Verschiebungsparameter (Zarychta u. a. 2007).

Jelsch u. a. (1998) berichten die Ergebnisse der Übertragung von asphärischen Streufaktoren auf das Oktapeptid LBZ Monohydrat (1 kDa, 115 K, $d = 0,82 \text{ \AA}$, konventioneller R -Wert 6,18 %), welches recht niedrige Temperaturfaktoren B_{eq} von 1,0–3,8 \AA^2 , durchschnittlich 1,9 \AA^2 , aufweist. Es wurden geringe, zufällig verteilte Bindungsrestdichten festgestellt, die durch den Transfer von neun Pseudoatomtypen auf die 156 Atome verringert wurden. Obwohl die Gütefaktoren der asphärischen Verfeinerung wegen eines veränderten Gewichtungsschemas, Verfeinerung gegen F statt F^2 und Ausschluß von Reflexen mit $I < 3\sigma$ schlecht mit den konventionellen Werten vergleichbar sind, resultierte letztendlich eine ca. 6 %-ige relative R -Wertverbesserung. Zudem konnte eine etwa 10 %-ige Verbesserung der Verschiebungsparameter und DMSDA-Werte erreicht werden, während die Koordinaten kaum beeinflusst wurden. Die asphärische Dichte erlaubte die Ableitung eines elektrostatischen Potentials.

In der vielzitierten Studie zur asphärischen Verfeinerung des 4,7 kDa-Proteins Crambin (PDB-Eintrag 1EJG; Jelsch u. a. 2000b), welches bei 100 K bis zu der

erstaunlichen Auflösung von $d = 0,54 \text{ \AA}$ gemessen wurde, wird eine Verbesserung des sphärischen R -Wertes von 9,2 % auf 9,0 % im asphärischen Fall erreicht, was einer relativen Abnahme um 2,2 % entspricht. Die Struktur weist ausgesprochen niedrige Temperaturfaktoren B_{eq} von im Median $2,5 \text{ \AA}^2$ auf, über deren Änderung oder DMSDA-Werte wird nicht berichtet. Um ein deutliches Bild der Bindungsrestdichten nach sphärischer Verfeinerung zu gewinnen, wurden die 34 geordneten Peptidbindungen gemittelt, die maximale Restdichte wurde zu $0,12 \text{ e/\AA}^3$ bestimmt. Nach Transfer der Datenbankmultipole zeigten sich im gleichen Bild negative Restdichten (ca. $-0,8 \text{ e/\AA}^3$) auf den Bindungen, was durch eine constrainte Verfeinerung der Multipolparameter der Hauptkette behoben werden konnte, die sich um 25 % von den Datenbankwerten unterschieden. Die resultierenden Deformationsdichtekarten entsprachen den Erwartungen.

Von der asphärischen Verfeinerung eines Skorpiongift-Proteins (0,7 kDa, 300 K, $d = 0,96 \text{ \AA}$, konventioneller $R1 = 10,1 \%$ / $R_{\text{all}} = 11,8 \%$) berichten Housset u. a. (2000). Die nicht fehlgeordneten Bereiche zeigen ein durchschnittliches B_{eq} von 8 \AA^2 und geringe Bindungsrestdichten. Transfer der Datenbankeinträge und Verfeinerung gegen $F > 3\sigma$, wobei die niedrigsten Auflösungsschalen ausgeschlossen wurden, ergab bei Wiederholung der sphärischen Verfeinerung $R_{\text{all}} = 9,0 \%$, bei Benutzung der asphärischen Streufaktoren $R_{\text{all}} = 9,1 \%$ und bei Verfeinerung der atomaren Ladung $R_{\text{all}} = 9,6 \%$. Es konnte eine 5 %-ige Verbesserung der B_{eq} -Werte erreicht werden, allerdings zeigten die DMSDA-Werte keine signifikante Änderung.

Die Anwendung auf Trypsin ($d = 0,8 \text{ \AA}$, konventioneller $R_{\text{all}} = 10,7 \%$) ergab Deformationsdichte und ESP eines Histidinrestes, weitere Ergebnisse der asphärischen Verfeinerung wurden nicht berichtet (Schmidt u. a. 2003). Für die Struktur Acutohämolyisin ($d = 0,8 \text{ \AA}$, konventioneller $R_{\text{all}} = 9,5 \%$) sind die Ergebnisse unklar, es wurde ein asphärischer R -Wert von 15 % angegeben (Liu u. a. 2003).

Besonders hervorzuheben ist die Arbeit zur Verfeinerung des Proteins Aldosereduktase (Guillot u. a. 2008), ausgehend von einem Modell der bis $d = 0,66 \text{ \AA}$ aufgelösten 36 kDa-Struktur (PDB 1US0, 316 Aminosäurereste, 3042 Nichtwasserstoffatome, 500 Wassermoleküle, 1 Coenzym NADP⁺, 1 Inhibitor IDD594, 2 Citrationen), die zu $R1/R_{\text{all}} = 8,43 \%/9,38 \%$ konventionell verfeinert wurde (Howard u. a. 2004). Nach Anwendung eines elaborierteren diffusen Solvensmodells wurde eine etwa 38 % der Atome umfassende Unterstruktur mit besonders geringen Verschiebungsparametern (durchschnittlicher B_{eq} -Wert $4,23 \text{ \AA}^2$, maximal $7,98 \text{ \AA}^2$) zur freien Verfeinerung ausgewählt, der Rest blieb auf den ursprünglichen Werten. Eine Wiederholung der sphärischen Verfeinerung unter diesen Bedingungen mit MoPro (Guillot u. a. 2001; Jelsch u. a. 2005) resultierte in einem R_{all} von 9,23 %. Mit auf alle Atome der Struktur übertragenen ELMAM-Populationen und Verfeinerung der Koordinaten und Verschiebungsparameter der Substruktur konnte ein R_{all} von 8,80 % erreicht werden, entsprechend einer relativen Verbesserung von 4,7 %. Durch eine constrainte Verfeinerung der Multipole verringerte sich der Gütefaktor abermals auf $R_{\text{all}} = 8,74 \%$. Als Resultat der asphärischen Verfeinerung wurde

eine Verringerung der Restelektronendichte auf den Bindungen erzielt, wie an der graphischen Darstellung einer Peptidbindung gezeigt wurde. Der durchschnittliche B_{eq} -Wert sank um 2,4 % auf $4,13 \text{ \AA}^2$, nur 6,5 % der Atome zeigten einen Anstieg der Temperaturfaktoren. Auch der DMSDA-Wert verbesserte sich von durchschnittlich 25 auf $20 \cdot 10^{-4} \text{ \AA}$, für 20 % der Werte verschlechterte sich dieser Deskriptor. Für die Peptidbindungslängen wurde eine Veränderung von $0,001 \text{ \AA}$ festgestellt, deren Streubreite nahm um etwa den gleichen Betrag ab.

Zwischen dem Beginn der Arbeiten zu Aldosereduktase, angekündigt in Jelsch u. a. (2000b), und dieser finalen Veröffentlichung lagen acht Jahre, wiewohl vorläufige Ergebnisse (Muzet u. a. 2003; Lecomte u. a. 2004), hauptsächlich bezüglich elektrostatischer Eigenschaften, eher berichtet wurden.

UBDB

Die Herangehensweise zur Konstruktion der UBDB-Pseudoatomdatenbank beschreibt Volkov u. a. (2004c): Für ausgesuchte kleine Molekülstrukturen aus der „Cambridge Structural Database“ (CSD Allen 2002a) werden *single point*-Berechnungen auf B3LYP/6-31G(d,p)-Level durchgeführt, aus der resultierenden Wellenfunktion werden *valence only*-Strukturfaktoren bis $\sin \Theta/\lambda = 1,1 \text{ \AA}^{-1}$ für eine pseudokubische Zelle ($a = 30 \text{ \AA}$) berechnet, gegen die Multipolparameter verfeinert werden (Valenzpopulation, Hexadecapole für Nichtwasserstoffatome, bindungsgerichtete Quadrupole für Wasserstoffatome, Kontraktionsparameter κ und κ'). Die theoretischen Phasen gehen als Constraint in die Verfeinerung ein. Für zu mittelnde Pseudoatome wird ein konsistentes lokales Koordinatensystem sichergestellt. Die Parameter mehrerer Pseudoatome gleichen Typs werden gemittelt, nur Populationen $> 0,002$ und $> 1\sigma$ finden Eingang in die Datenbank. Zur Festlegung eines Pseudoatomtyps wird nach Element, nächstem Nachbarn und dessen Hybridisierungszustand, manchmal auch übernächstem Nachbarn sowie statistischer Äquivalenz bei Mittelung mehrere Pseudoatome unterschieden, so wird ein neuer Atomtyp eingeführt, wenn dessen Einbeziehung den vorherigen Mittelwert um wenigstens eine Standardabweichung verschiebt. Alle in Proteinen vorkommenden Atome werden durch 74 Pseudoatominträge abgedeckt (Dominiak u. a. 2007).

Um die Datenbank zu überprüfen, wurde damit die Dichte von drei CSD-Kristallstrukturen von Leucin, Serin und Glutamin rekonstruiert und mit mehreren *ab initio*-Modellen der jeweiligen Gesamtstruktur verglichen. Die erfolgreiche Validierung umfaßte Deformationsdichten, lokale und integrierte QTAM-Eigenschaften, elektrostatische Potentiale und Interaktionsenergien – die Unterschiede zwischen Datenbank und Rechnungen sind denen zwischen verschiedenen theoretischen Modellen gleichwertig.

Die Leistungsfähigkeit der Datenbank bezüglich elektrostatischer Interaktionsenergien wurde weiter untersucht (Volkov u. Coppens 2004) und eine neue Methode zu ihrer Berechnung (EP/MM – *Exact Potential and Multipole Method*) vorgestellt

(Volkov u. a. 2004a), die, bei Benutzung der Pseudoatomdatenbank, Resultate liefert, welche sich um nur ca. 10 kJ/mol von ADF/BLYP/TZP-Rechnungen unterscheiden. Die intermolekulare elektrostatische Interaktionsenergie, berechnet aus den Datenbankeinträgen mit der EPMM-Methode, weicht um nur 4 kJ/mol von einer direkten Berechnung auf B3LYP/6-31G(d,p)-Niveau ab. Gegenüber aufwendigeren Methoden ergeben sich höhere Abweichungen wie Volkov u. a. (2006) in einem umfassenden Vergleich theoretischer Methoden und Basissätze in der Berechnung obiger Größe für sechs zwitterionische Glycindimere zeigten.

Verfeinerungen experimenteller Daten unter Verwendung der Datenbankpseudoatome wurden für ein Tri- und Hexapeptid durchgeführt (Volkov u. a. 2007). Für diese Strukturen liegen hochaufgelöste Daten ($d \leq 0,44 \text{ \AA}$) vor, die zu Testzwecken auf $d = 0,83 \text{ \AA}$ gekürzt wurden. Für die voll aufgelösten Daten wurde damit eine relative R -Wertverbesserung von etwa 20 % erreicht, mit dem gekürzten Datensatz von ca. 42 %. Anpassung der Valenzpopulationen führte zu einer etwas stärkeren Verbesserung. Bindungsrestdichten bei hoher Auflösung werden beinahe vollständig beseitigt, die gering aufgelösten Daten zeigten im sphärischen Fall wenig Bindungsrestdichten, da diese durch vergrößerte Verschiebungsparameter überlagert wurden. Diese Verfälschung wird in der asphärischen Verfeinerung korrigiert, so daß die Verschiebungsparameter bei geringer Auflösung sich fast auf die bei voller Auflösung erzielten Werte verbesserten. Auch der DMSDA-Wert wurde signifikant verbessert, bei geringer Auflösung waren die Verbesserungen besonders ausgeprägt (−30 % bzw. −56 %),

Dominiak u. a. (2007) benutzten die Datenbank, um elektrostatische Potentiale und Interaktionsenergien von experimentell bestimmten Komplexstrukturen der Syntenin-PDZ2-Domäne und acht Peptiden zu bestimmen. Die Anwendung der Datenbank auf die Bestimmung elektrostatischer Interaktionsenergien von beeindruckenden 70 Enzym-Substrat-Komplexen zwischen Neuraminidase und Inhibitormolekülen beschrieben Dominiak u. a. (2009), wofür weitere Atomtypen bestimmt werden mußten.

Invariome

Ebenfalls auf der Ableitung von Multipolparametern aus quantenchemisch erzeugten Elektronendichten beruht der von Dittrich u. a. (2004) eingeführte Invariomformalismus, welcher nachfolgend noch verbessert und erweitert wurde (Dittrich u. a. 2006a). Allerdings werden die Pseudoatomeinträge von konstruierten, minimalen Modellverbindungen abgeleitet, statt auf experimentellen Strukturen zu basieren. Zudem wird deren Geometrie auf hohem theoretischen Niveau (B3LYP/D95++(3df,3pd)) optimiert, bevor theoretische Strukturfaktoren aus dieser Dichte abgeleitet werden, gegen die Multipolparameter unter Annahme höchstmöglicher lokaler Symmetrie verfeinert werden. Jeder Invariomeintrag der Datenbank wird aus exakt einer Modellverbindung bestimmt, eine Mittelung findet nicht statt. Die Festlegung einzigartiger übertragbarer Pseudoatome beruht auf

einem sehr geradlinigen (Über)Nächste-Nachbar-Schema, welches die Typen und die Anzahl unmittelbarer Bindungspartner berücksichtigt, bei mesomeren Systemen und Wasserstoffatomen wird die Unterscheidung inklusive der übernächsten Nachbarn getroffen, ausgedehnte delokalisierte Systeme sogar vollständig berechnet. (Für weitere Informationen siehe Abschn. 1.2.2 auf Seite 22.) Als Vorteile dieses Ansatzes werden die Abwesenheit experimenteller Fehler und die uneingeschränkten Möglichkeiten zur schnellen Generierung neuer Datenbankeinträge gesehen. Um alle in obiger Näherung vorkommenden Atomtypen der 20 kanonischen Aminosäuren abzudecken, werden 73 Invariom-Pseudoatome benötigt.

Zur Validierung des Invariomansatzes wurde zunächst eine 0,69 Å-Struktur von Trivalin asphärisch verfeinert (Dittrich u. a. 2004), deren Elektronendichte anschließend bindungstopologische Eigenschaften aufwies, die statistisch äquivalent zu Mittelwerten von sechs hochaufgelösten und multipolverfeinerten Oligopeptiden waren. Auch wurde eine signifikante Reduktion von Bindungsrestdichten und Gütefaktoren beobachtet. Trotz des fehlgeordneten Cokristallisats sank der $R1$ -Wert relativ um 13 %.

In einer ausführlichen Untersuchung zur Invariomverfeinerung von Serin (Dittrich u. a. 2005) bei drei verschiedenen Temperaturen und aufgeschlüsselt nach Auflösungsbereichen ($d_{max} < 0,5 \text{ \AA}$), waren noch deutlichere Verbesserungen zu beobachten. So sank der Gütefaktor um etwa 30 %, was mit einer Halbierung der Extrema der Restelektronendichte einherging. Der DMSDA-Test wurde in den asphärischen Verfeinerungen etwa doppelt so gut wie im sphärischen Fall erfüllt.

Eine umfassende Studie an 42 Oligopeptidstrukturen ($0,35 \leq d \leq 0,91 \text{ \AA}$) aus der CSD bestätigte diese Befunde (Dittrich u. a. 2006a) – sofern keine unerkannte Fehlordnung vorlag, wurden Verbesserungen des Gütefaktors zwischen etwa 10 % und 50 % erzielt. In diesem Zusammenhang wurde auch die nur geringe Abweichung zwischen übertragener Valenzelektronenzahl und jener der elektroneutralen Struktur quantifiziert, im Maximum betrug diese 1,1 %, was leicht durch nachträgliche Skalierung korrigiert werden kann.

Etliche weitere Anwendungsfälle belegen die Praktikabilität dieses Ansatzes, beispielsweise anhand der Verfeinerungen von Isoflavon (Kingsford-Adaboh u. a. 2006), Valinol (Dittrich u. a. 2006b), Thymidin (Hübschle u. a. 2008) oder Ornithin (Dittrich u. a. 2007). Der vorteilhafte Effekt einer asphärischen Verfeinerung auf die Temperaturfaktoren wird in einer Studie an drei Hydroxylysin-Strukturen hervorgehoben (Dittrich u. a. 2008). Zudem erhöht die Anwendung der Invariom-Streufaktoren die Sicherheit der Bestimmung des Flack-Parameters (Flack 1983), wie anhand von vier, nur C, H, N und O enthaltenden, Leichtatomstrukturen gezeigt wurde (Dittrich u. a. 2006c). Selbst bei kurzen Wellenlängen wie MoK_α ist so die absolute Aufstellung einer chiralen Struktur zu sichern. Die übertragenen Pseudoatome können auch als Startpunkt einer Multipolverfeinerung bei geringen Auflösungen dienen (Dittrich u. a. 2009a,c) oder als Referenz, um den Einfluß der kristallinen Umgebung auf die Multipoldichte herauszuarbeiten (Dittrich u.

Spackman 2007; Kalinowski u. a. 2007). Eine weitere Anwendung fand das Invariomkonzept in der Untersuchung der temperaturabhängigen Rotationsfehlordnung von Methyl-Aminoisobuttersäure (Dittrich u. a. 2009b).

Vergleich der Datenbanken

Pichon-Pesme u. a. (2004) stellen in einem Vergleich von ELMAM- und damaliger UBDB-Datenbank einige signifikante Unterschiede der verwendeten Multipolparameter und der resultierenden Deformationsdichte einzelner Atome fest. Da elektrostatische Eigenschaften als die wichtigste Datenbankanwendung gesehen werden, wurden die elektrostatischen Potentiale der Peptidgruppe verglichen. Wegen unterschiedlicher Atomladungen in den beiden Ansätzen wurden stärker negative ESPs aus der experimentellen ELMAM-Datenbank als aus der theoretischen UBDB abgeleitet. Ersteres entspricht eher den Ergebnissen aus experimentellen Multipolverfeinerungen, letzteres den Werten aus *ab initio*-Rechnungen. Volkov u. a. (2004b) lehnen die Studie wegen methodischer Unzulänglichkeiten ab und verweisen auf ausgezeichnete Übereinstimmungen elektrostatischer Interaktionsenergien, welche aus der UBDB-Datenbank beziehungsweise DFT-Rechnungen auf sehr hohem Niveau abgeleitet wurden.

In einem Vergleich von freier Multipolverfeinerung und asphärischer Verfeinerung mittels Pseudoatomen der UBDB- und Invariom-Bibliothek für die 1,2 kDa-Struktur von Cyclosporin A ($d = 0,55 \text{ \AA}$, 5 K), zeigen beide Datenbanken eine gleichermaßen gute Annäherung an die Resultate mit verfeinerten Multipolen (Johnas u. a. 2009). Während letztere z. B. den sphärischen $R1$ -Wert von 3,4 % um 27 % reduziert, ergeben beide Datenbankanwendungen eine Verbesserung um 22 %. Ganz ähnlich ist die Abnahme der mittleren Restelektronendichte mit 27 % im Fall der freien Verfeinerung etwas ausgeprägter als mit transferierten Multipolbeiträgen (21 %). Auch bindungstopologische Deskriptoren sind innerhalb der Fehler in allen drei Modellen äquivalent.

Die Möglichkeit, eine freie Multipolverfeinerung an einer partiell fehlgeordneten Paracetamolstruktur durchzuführen, falls der ungeordnete Teil mittels Datenbankpopulationen beschrieben wird, untersuchen Bak u. a. (2009). Sowohl die UBDB- als auch die Invariom-Pseudoatome fanden Anwendung – beide Varianten vermögen gleichermaßen die Verlässlichkeit der Verfeinerung zu erhöhen.

Alternativen

Anstatt übertragene Multipolparameter zur verbesserten Verfeinerung heranzuziehen, besteht auch die Möglichkeit auf den Bindungen interatomare Streuquellen (*interatomic scatterers*, IAS) zu positionieren (Afonine u. a. 2007), die Resultate der IAS-Verfeinerung von Makromolekülen werden als Multipolergebnissen vergleichbar berichtet. Bisher nur für kleine Moleküle, aber ebenfalls mit dem Ziel der Makromolekülverfeinerung, werden Kombinationen aus IAS-Verfeinerung und Kraftfeldmethoden entwickelt (Schnieders u. a. 2009).

Gegenstand und Ziel dieser Arbeit

Wie die vorherigen Ausführungen gezeigt haben, wurden zwar viele kleine Moleküle und einige mittelgroße Strukturen (≈ 1 kDa) mittels transferierter Pseudoatome sphärisch verfeinert, aber nur die beiden Proteinstrukturen Crambin (4,7 kDa; Jelsch u. a. 2000b) und Aldosereduktase (36 kDa; Guillot u. a. 2008) sind als erfolgreiche Anwendung der ELMAM zur Verfeinerung von Makromolekülen bekannt. Mittels der UBDB wurden elektrostatische Eigenschaften von Proteinkomplexen bestimmt, aber keine Verfeinerung durchgeführt.

In dieser Arbeit werden drei weitere Beispiele der sphärischen Verfeinerung von Proteinen vorgestellt: Die 14,3 kDa-Strukturen von triklinem Lysozym (Wang u. a. 2007), in dieser Arbeit mit P1LYSO bezeichnet, tetragonalem Lysozym (TETRALYSO) und ein rhomboedrisches 11,6 kDa-Dimer von Insulin (INSULIN). Ergänzt wird dies durch das 1,7 kDa-Oktadekapeptid TRICHOTOXIN mit insgesamt 3,6 kDa in der asymmetrischen Einheit (Chugh u. a. 2002), und, zu Vergleichszwecken, das Tripeptid L-Alanyl-L-Prolyl-L-Alanin (APA). Alle sphärischen Verfeinerungen greifen auf die Invariom-Datenbank zurück, was zugleich deren erste Anwendung auf Probleme dieser Größenordnung ist.

Alle Strukturen werden in einem einheitlichen, untereinander vergleichbaren Schema sphärisch verfeinert und analogen Auswertungen unterzogen. Hierbei dient APA der Einführung der verschiedenen Analysen und als Maßstab des bei hoher Strukturqualität Erreichbaren. Das in der Größe zwischen dem Tripeptid und den Proteinen liegende Oligopeptid TRICHOTOXIN komplettiert die Studie und erleichtert die Abschätzung der zu erwartenden Effekte einer sphärischen Verfeinerung von Makromolekülen.

Besonderes Augenmerk wird auf die ausführliche Analyse der Unterschiede zwischen sphärischer und asphärischer Verfeinerung gelegt. Anhand von Gütefaktoren, Restdichten, Deformationsdichten, Verschiebungsparametern, DMSDA-Werten, Strukturkoordinaten und Bindungslängen erfolgt für jede Struktur eine quantitative Auswertung der erzielten Verbesserungen in der Beschreibung des Röntgenbeugungsexperimentes durch Verwendung eines sphärischen Elektronendichtemodells. Nach Vergleich der jeweiligen Resultate wird abschließend der Versuch unternommen, den Verfeinerungserfolg zu erklären und in einen funktionalen Zusammenhang mit Parametern des sphärischen Ausgangsmodells zu bringen. Daraus folgen Abschätzungen zu Grenzwerten dieser Größen, ab denen eine sphärische Verfeinerung erfolgversprechend erscheint.

Als Beispiel einer nach Multipolbeschreibung über das sphärische Modell hinaus zugänglichen Eigenschaft wird für jedes Molekül das aus den Invariomen generierte elektrostatische Potential gezeigt und einem aus Punktladungen berechneten gegenübergestellt. Ein umfangreicherer Vergleich von *ab initio*, semiempirischen, Punktladungs- und Multipol-ESPs wird stellvertretend für APA unternommen, dessen geringe Größe es allen Methoden zugänglich macht.

Grundlagen

1 Sphärische und asphärische Verfeinerung

Elektromagnetische Strahlung wechselwirkt mit Materie, wobei der elektrische Anteil der Strahlung die Elektronenhülle der Atome der Probe in Schwingung versetzt. Die Elektronen schwingen gegenüber den relativ schwereren und daher als ruhend anzunehmenden, positiven Atomkernen mit der Frequenz der einfallenden Welle. Somit bilden Atomkern und Elektronen einen oszillierenden Dipol, welcher selbst wieder Energie in alle Raumrichtungen abstrahlt (Atkins 1996). Die Wellenlänge von Röntgenstrahlung liegt mit $0,1 \text{ \AA} - 100 \text{ \AA}$ im gleichen Bereich wie die Gitterabstände von Kristallen, so daß durch Interferenz der ausgestrahlten Kugelwellen ein Beugungsmuster entsteht. Das intensitätsgewichtete Beugungsbild stellt ein Abbild des reziproken Gitters des Einkristalls dar, jeder Reflex läßt sich einer Netzebene zuordnen. Aus diesem Beugungsbild lassen sich Periodizität (Gitterkonstanten), Symmetrie der Elementarzelle (Raumgruppe) und die Struktur der asymmetrischen Einheit (fraktionelle Koordinaten) bestimmen (Massa 2002; Müller 1996; Luger 2000).

Bei der Überlagerung der Streubeiträge vieler Atome können diese nicht einfach addiert werden, sondern es muss dabei die Phasenverschiebung der gestreuten Wellen in Abhängigkeit von der Beugungsrichtung berücksichtigt werden. Die Braggsche Gleichung 1.1 gibt an, in welcher Richtung, bei welchem Beugungswinkel Θ Maxima der Streuintensität erwartet werden können.

$$2d \sin \Theta = n\lambda \quad \text{mit } \lambda = \text{Wellenlänge,} \quad (1.1)$$

$\Theta = \text{Glanzwinkel,}$
 $n = \text{Beugungsordnung,}$
 $d = \text{Netzebenenabstand,}$

1.1 Sphärische Verfeinerung des Röntgenbeugungsexperiments

Da Röntgenstrahlung an Atomen mit mehr Elektronen stärker gebeugt wird als an Atomen mit weniger Elektronen, muß bei Kristallen, die aus mehreren Atomsorten i aufgebaut sind, diese Tatsache in Form des Atomformfaktors f_i berücksichtigt werden, da die radiale Ausdehnung δ der Elektronenhülle eine Phasenverschiebung $\delta \cdot 2\pi/d$ (von $0 - \lambda$) bewirkt. Die mit zunehmendem Beugungswinkel θ abnehmenden Streuamplituden f_i sind, auf Elektronenzahlen normiert (bei $\theta = 0$ hat die

Streuamplitude den Wert der Ordnungszahl), in Polynomform tabelliert (Maslen u. a. 1992). Diese Strukturfaktoren basieren auf Hartree-Fock-Rechnungen an den freien Atomen und beschreiben damit unbeeinflusste kugelförmige Atome.

Eine zusätzliche Schwächung erfahren die Atomformfaktoren durch die Schwingung der Atome um ihre Nullpunktslage, z.B. durch Temperaturbewegung, woraus eine zusätzliche Phasenverschiebung resultiert, die umso höher ist, je größer die mittlere Auslenkungsamplitude u ist und je kleiner d oder je größer θ sind. Mit dem isotropen Auslenkungsparameter $U = u^2$ wird dies durch den Debye-Waller-Faktor in Gleichung 1.2 berücksichtigt.

$$\begin{aligned} f' &= f \cdot \exp\left(\frac{-2\pi^2 U}{d^2}\right) \\ &= f \cdot \exp\left(-8\pi^2 U \frac{\sin^2 \theta}{\lambda^2}\right) \end{aligned} \quad (1.2)$$

Die Beachtung der radialen Ausdehnung und der Schwingungsbeiträge rechtfertigt die punktförmige Beschreibung der Atomlagen. Außerdem muss die relative Lage der verschiedenen Atomsorten in der Elementarzelle mit Hilfe einer Unterteilung des Kristallgitters in Untergitter, die aus den jeweiligen Atomsorten i aufgebaut sind, berücksichtigt werden. Wenn ein Untergitter nämlich in Reflexionsstellung ist, kommt es durch das versetzt angeordnete Untergitter zu Interferenz, da die beiden gebeugten Strahlungen phasenverschoben sind. Der Phasenunterschied ϕ ist durch folgende Gleichung 1.3 gegeben und hängt somit über die Atomparameter x_i, y_i, z_i der betrachteten Atome i von deren Lage im Kristall und über die Millerindizes von den in Reflexionsstellung befindlichen Netzebenen ab, auf denen die Atome liegen.

$$\phi_i = 2\pi(hx_i + ky_i + lz_i) \quad (1.3)$$

Für den Reflex einer Netzebene (hkl) überlagern sich die Streuwellen aller i Atome mit ihrer individuellen Phasenverschiebung zum Strukturfaktor F .

$$F(\vec{h}) = \sum_i f'_i (\cos \phi_i + i \sin \phi_i) \quad (1.4)$$

Bei einem Phasenunterschied von π (180°) zweier Atome mit gleichem Atomformfaktor interferieren die an den beiden Atomen gebeugten Wellen destruktiv. Abgesehen von der kompletten Auslöschung einiger Reflexe, führt der Phasenunterschied auch zur Verstärkung und Abschwächung einzelner Reflexe. In der Strukturamplitude $|F_{hkl}|$ werden die Auswirkungen des Phasenunterschieds ϕ_i und des Atomformfaktors f_i der einzelnen Atome berücksichtigt. Die resultierende Intensität I_{hkl} eines Reflexes ist proportional zum Quadrat der Strukturamplitude

$|F_{hkl}|$, die aus dem Strukturfaktor bei bekannter Phase erhältlich wäre.

$$I(\vec{h}) \propto |F(\vec{h})|^2 \quad (1.5)$$

$$\text{und } F(\vec{h}) = |F(\vec{h})|e^{i\phi(\vec{h})} \quad (1.6)$$

Der Strukturfaktor ist die Fouriertransformierte der Elektronendichtefunktion $F(\vec{h}) = \mathcal{F}(\rho(\vec{r}))$, umgekehrt kann $\rho(\vec{r})$ durch Fouriersummation der Strukturfaktoren gewonnen werden:

$$\rho(\vec{r}) = \frac{1}{V} \sum_{\vec{h}} F(\vec{h})e^{-i\phi(\vec{h})} \quad (1.7)$$

Aus dem Experiment ist die Phaseninformation $\phi(\vec{h}) = 2\pi(hx + ky + lz)$ nicht erhältlich – das sogenannte Phasenproblem der Strukturanalyse. Um $\rho(\vec{r})$ und damit die Lage der Atome in der asymmetrischen Einheit zu bestimmen, muß das Phasenproblem gelöst werden, z. B. durch Direkte Methoden, welche, unter Ausnutzung von statistischen Zusammenhängen zwischen Intensitäten bestimmter Reflexe und deren Phasen, eine Lösung des Phasenproblems ermöglichen.

Dadurch wird ein initialer Satz an Phasen für alle Strukturfaktoren $F_o(\vec{h})$ erhalten, womit durch Fouriersummation eine erste Elektronendichteverteilung $\rho_o(\vec{r})$ bestimmt werden kann. An Orten maximaler Dichte werden Atome platziert, dies entspricht dem Strukturmodell, welches zunächst erheblichen Näherungscharakter hat und unvollständig sein kann. Da jedem Atomtyp Streufaktoren, Koordinaten und Temperaturfaktoren zugeordnet sind, können zu diesem Modell gehörende Strukturfaktoren $F_c(\vec{h})$ berechnet werden. Zwischen beobachteten und berechneten Strukturfaktoren wird keine vollständige Übereinstimmung bestehen, da erstere mit experimentellen Fehlern behaftet sind und letztere aus einem noch unvollständigen und ungenauen Strukturmodell stammen. Ziel der Strukturverfeinerung ist die Minimierung dieser Differenz durch Anpassung der Parameter des Modells in einer Kleinste Quadrate Prozedur. Je nachdem, ob gegen $F_o(\vec{h})$ oder $F_o^2(\vec{h})$ verfeinert wird, ist die Forderung: $\sum_{hkl} w(|F_o| - |F_c|)^2 = Min$ oder $\sum_{hkl} w'(F_o^2 - F_c^2)^2 = Min$. Im isotropen Fall gibt es pro Atom vier zu variierende Parameter (x_i, y_i, z_i, U) , mit realitätsnäheren anisotropen Verschiebungsparametern neun (x_i, y_i, z_i, U_{ij}) .

Aus einer Differenz-Fouriersynthese (Gl. 1.8) wird der Unterschied zwischen beobachteter Elektronendichte und der aus dem Strukturmodell resultierenden ersichtlich. Restdichteanhäufungen können neue Atompositionen zugeordnet werden. Das so schrittweise zu komplettierende Strukturmodell wird weiter verfeinert, bis möglichst wenig Restelektronendichte verbleibt.

$$\Delta\rho_{res}(\vec{r}) = \frac{1}{V} \sum_{\vec{h}} (|F_{obs}(\vec{h})| - |F_{calc}(\vec{h})|) \exp(i\phi_{calc}) \exp(-2\pi\vec{h}\vec{r}) \quad (1.8)$$

Temperaturfaktor

Der anisotrope Temperaturfaktor wird als Tensor zweiter Ordnung in Form einer symmetrischen 3×3 -Matrix angegeben, wobei in rechtwinkligen Koordinatensystemen die Diagonalkomponenten U_{11}, U_{22}, U_{33} die Hauptachsen des Temperaturellipsoids darstellen und die gemischen Komponenten U_{12}, U_{13}, U_{23} die Orientierung relativ zum Koordinatensystem angeben. In den Verfeinerungsprogrammen `shelxl` und `XD` sind sie auf das Kristallkoordinatensystem bezogen, bei der Angabe in PDB-Dateien aber auf ein kartesisches Koordinatensystem, was im Fall orthogonaler Zellachsen identisch ist.

Ausgehend von den U_{ij} -Parametern im Kristallsystem, kann der äquivalente isotrope Temperaturfaktor U_{eq} nach Gl. 1.9 angegeben werden (Fischer u. Tillmanns 1988).

$$U_{\text{eq}} = \frac{1}{3} [U_{11}(aa^*)^2 + U_{22}(bb^*)^2 + U_{33}(cc^*)^2 + 2U_{12}aba^*b^* \cos \gamma + 2U_{13}aca^*c^* \cos \beta + 2U_{23}cb^*c^* \cos \alpha] \quad (1.9)$$

Die Zellkonstanten sind mit $a, b, c, \alpha, \beta, \gamma$, die reziproken Gitterkonstanten mit $a^* = bc \sin \alpha / V, b^* = ca \sin \beta / V, c^* = ab \sin \gamma / V$ und das Elementarzellvolumen mit $V = abc \sqrt{1 - \cos^2 \alpha - \cos^2 \beta - \cos^2 \gamma + 2 \cos \alpha \cos \beta \cos \gamma}$ bezeichnet.

Differenz der mittleren quadratischen Verschiebungsamplitude

Ob die mit der Einführung der ADPs versuchte Trennung der Gesamtelektroendichte von Effekten thermischer Verschmierung erfolgreich war, kann durch Hirshfelds *rigid bond*-Test überprüft werden (Hirshfeld 1976): Unter der Annahme, daß der Hauptteil der Vibration in einem Molekülkristall aus intermolekularen Gitterschwingungen sowie intramolekularen Torsions- und Biegeschwingungen herührt, sollten die Bindung dehrende Schwingungen vernachlässigbar klein sein, so daß die relative Verschiebung kovalent gebundener Atome zueinander verschwindet. Ausgedrückt wird dies durch die Differenz der mittleren quadratischen Verschiebungsamplituden (*difference of mean square displacement amplitudes*, DMSDA). Sind z_{AB}^2 und z_{BA}^2 die mittleren quadratischen Schwingungsamplituden der Atome A und B jeweils entlang der Bindungen von A zu B und B zu A , so sollten diese etwa gleich hoch sein. Daher sollte ihre Differenz unter 0.001 \AA^2 liegen: $|z_{AB}^2 - z_{BA}^2| < 10 \cdot 10^{-4} \text{ \AA}^2$. Niedrige DMSDA-Werte legen physikalisch sinnvoll modellierte anisotrope Verschiebungsparameter nahe, die nicht durch in die Bindungen deformierte Valenzdichte verfälscht sind. Dieser Test wird nur für Bindungen zwischen Nichtwasserstoffatomen durchgeführt, da nur diese anisotrop verfeinert werden.

1.2 Asphärische Verfeinerung des Röntgenbeugungsexperiments

1.2.1 Multipolformalismus

Da in der zuvor beschriebenen Prozedur die Modellierung der gemessenen Elektronendichte in der Annahme einer Überlagerung kugelförmiger Elektronendichteverteilungen um die Atome vorgenommen wurde, kann dieses Modell unabhängiger Atome (IAM) keine Informationen über die Deformation atomarer Ladungsdichte durch z. B. Bindungsbildung enthalten. Die Vernachlässigung der asphärischen Ladungsverteilung führt auch zu falsch bestimmten Atompositionen, wie zu kurze Wasserstoffbindungen oder zu lange Sauerstoffbindungen; des weiteren kann das IAM keinen Ladungstransfer zwischen den Atomen berücksichtigen.

Mit der Beschreibung im atomzentrierten Hansen-Coppens-Multipolformalismus (Hansen u. Coppens 1978) wird eine bessere Näherung der Elektronendichte gegeben. Hierbei unterteilt sich die atomare Elektronendichte, wie in Gleichung 1.10 angegeben, in drei Beiträge: Eine sphärische Kerndichte $\rho_{Kern}(\vec{r})$ für die inneren Elektronenschalen und eine ebenfalls sphärisch beschriebene Valenzdichte $\rho_{Val}(\kappa\vec{r})$ für die äußere Schale, deren Radialverteilungen aus Hartree-Fock-Berechnungen der freien Atome stammen, repräsentiert durch tabellierte Koeffizienten und Exponenten von Slater-Basisfunktionen (Clementi u. Roetti 1974). Mittels des Expansions-/Kontraktionsparameters κ kann die Ausdehnung der Valenzschale variiert werden. Der verbleibende dritte Term berücksichtigt nun die asphärische Deformation der Valenzdichte durch eine Linearkombination von Kugelflächenfunktionen $Y_{lm}(\theta, \phi)$ mit Radialfunktionen $R_l(\kappa'\vec{r}) = N_l r^n \exp(-\kappa'\zeta\vec{r})$, welche wieder aus einer Minimalbasis an Slaterfunktionen mit fixen, energieoptimierten *single*- ζ -Exponenten (Clementi u. Raimondi 1963) bestehen, deren Ausdehnung über einen Parameter κ' veränderbar ist. Die Laufzahl l beschreibt die Ordnung der Multipolexpansion: Monopole für $l = 0$, Dipole für $l = 1$, Quadrupole für $l = 2$, Octupole für $l = 3$ und Hexadecapole für $l = 4$. Eine Darstellung der Multipole gibt Abb. 1.1 auf der folgenden Seite.

$$\rho_{mul}(\vec{r}) = \rho_{Kern}(\vec{r}) + P_{Val}\kappa^3\rho_{Val}(\kappa\vec{r}) + \sum_{l=0}^{l_{max}} \kappa'^3 R_l(\kappa'\vec{r}) \sum_{m=0}^l P_{lm} Y_{lm\pm}(\theta, \phi) \quad (1.10)$$

In einer Kleinsten Quadrate Verfeinerung werden die Populationskoeffizienten P_{Val} und P_{lm} sowie eventuell die Ausdehnungskoeffizienten κ und κ' über die Anpassung der mittels Fouriertransformation aus $\rho_{mul}(\vec{r})$ gewonnenen Strukturfaktoren an die experimentell bestimmten Strukturfaktoren ermittelt. Ansonsten gilt das im vorherigen Abschnitt Gesagte zum Vorgehen bei der Verfeinerung.

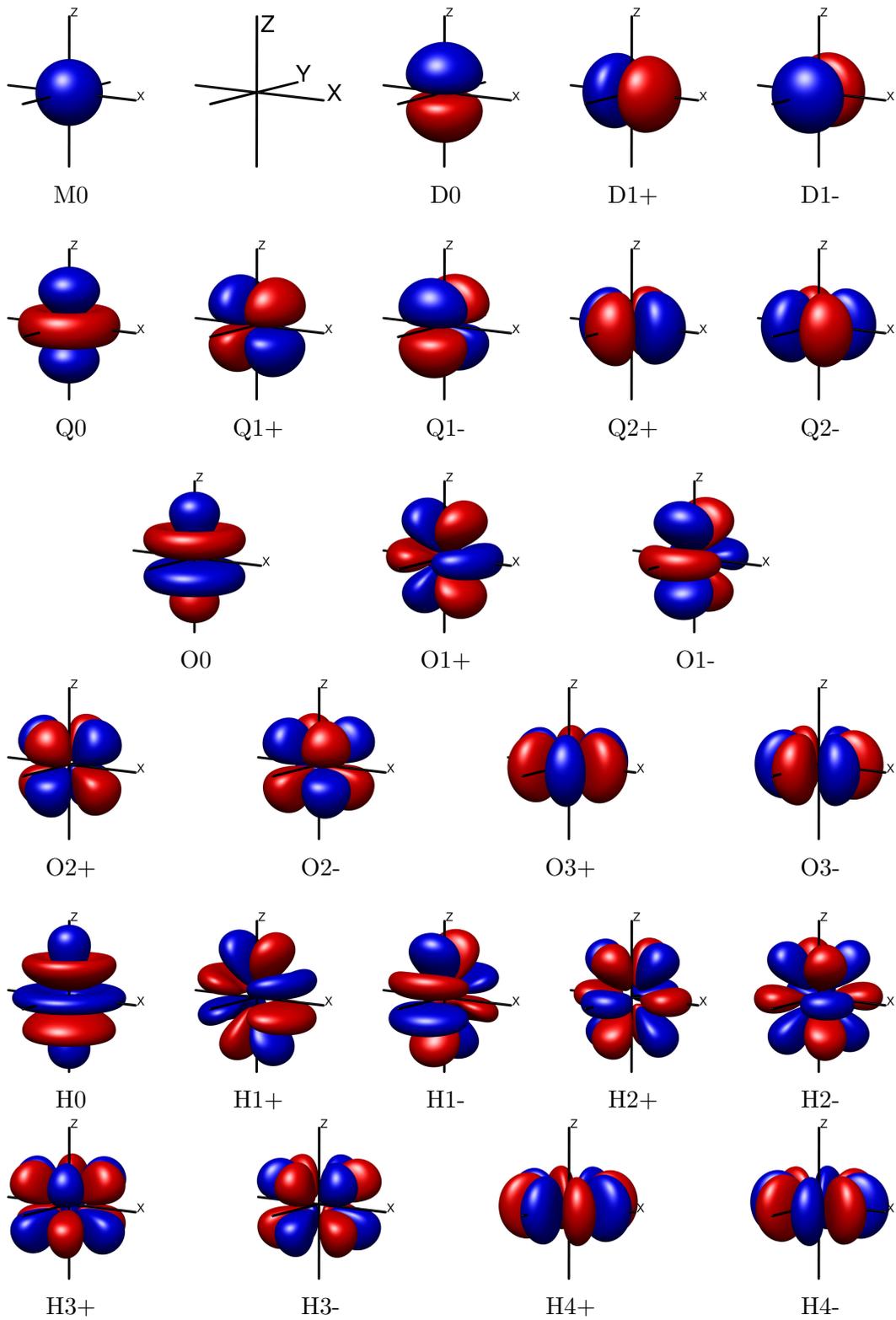


Abbildung 1.1: Multipole $R_l Y_{lm}$, $l = 0 \dots 4$, $m = -l \dots l$.

Lokale Symmetrie und Auswahlregeln

Um eine gewisse lokale Symmetrie der Elektronendichteverteilung zu erzeugen, werden nur durch Auswahlregeln (l, m, \pm) festgelegte Multipole populiert (Kurki-Suonio 1977). Dadurch kann zum einen die Parameterzahl der Verfeinerung verringert werden und zum anderen die physikalische Signifikanz des Modells erhöht werden, weil beispielsweise Parameterkorrelationen vermieden werden. Zum Beispiel lautet die Auswahlregel für Spiegelsymmetrie $(l, l - 2\mu, \pm)$ mit $\mu \in \mathbb{Z}$. Damit sind außer dem Monopol nur 14 Multipole ($D1\pm, Q0, Q2\pm, O1\pm, O3\pm, H0, H2\pm, H4\pm$) von insgesamt 24 erlaubt, weil etwa für $l = 2$ nur $m = 2 - 2 \cdot 0 = 2$ und $m = 2 - 2 \cdot 1 = 0$ möglich sind, bei anderen Werten von μ als 0 oder 1 ergeben sich für m Werte außerhalb des erlaubten Bereiches $\pm l$, so daß nur drei von fünf Quadrupole übrig bleiben.

Deformationsdichte

Zieht man von der Multipoldichte eine Superposition der kugelförmigen unabhängigen atomaren Elektronendichten (Promoleküldichte) ab, erhält man eine statische Deformationsdichte $\Delta\rho_{def}(\vec{r}) = \rho_{mul}(\vec{r}) - \rho_{pro}(\vec{r})$, welche das Ausmaß der Ladungsdichteumverteilung z. B. in Bindungsregionen beim Übergang von sphärischer zu asphärischer Modellierung anzeigt.

Abgeleitete Eigenschaften – Elektrostatisches Potential

Die im Multipolformalismus beschriebene Elektronendichteverteilung eines Moleküls aus allen Beiträgen der Pseudoatome, ist von einer Güte, die die Berechnung verschiedener abgeleiteter Eigenschaften auf *ab initio*-Niveau erlaubt. Beispielsweise können Dipol- und Quadrupolmomente, elektrische Feldgradienten, elektrostatische Potentiale und Interaktionsenergien sowie vollständige topologische Analysen nach Bader (1990) berechnet werden.

In dieser Arbeit wird nur das elektrostatische Potential $\phi_{mul}(\vec{r})$ betrachtet. Es beinhaltet eine Aussage über den Energieaufwand eine positive Einheitsladung aus dem Unendlichen an einen bestimmten Punkt im Molekül zu bringen und kann nach der Methode von Su u. Coppens (1992) aus der Multipolelektronendichte berechnet werden. In Gleichung 1.11 wird über alle Atome j des Moleküls summiert, wobei der erste Term dem Anteil der Atomkerne (Kernladungszahl Z_j , Kernposition \vec{R}_j) entspricht, von dem der Beitrag der Elektronen zum ESP abgezogen wird.

$$\phi(\vec{r}) = \sum_j \frac{Z_j}{|\vec{r} - \vec{R}_j|} - \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' \quad (1.11)$$

1.2.2 Invariome

Im Multipolformalismus stehen pro Atom, neben den neun klassischen Positions- und Verschiebungsparametern, zusätzlich 30 Parameter zur Verfügung, so daß insgesamt 39 Variablen ($x, y, z, U_{ij}, P_{val}, \kappa, P_{lm}, \kappa'_l$) verfeinert werden könnten. Nur bei ausgesprochen hochwertigen und höchst aufgelösten Daten ist eine verlässliche Anpassung aller Größen möglich. Um die Vorteile einer asphärischen Beschreibung auch für die Verfeinerung mediokrer Röntgenbeugungsdaten zugänglich zu machen, müssen die Multipol- und Kontraktionsparameter aus einer anderen Quelle als der Verfeinerung an die fraglichen Daten bezogen werden. Aufgrund der Transferierbarkeit von Hansen-Coppens-Pseudoatomen zwischen verschiedenen Verbindungen, sofern die Atome in jeweils ähnlichen chemischen Umgebungen vorliegen, können sie auch aus Verfeinerungen sehr guter experimenteller oder theoretischer Daten anderer, leichter zugänglicherer Verbindungen stammen und als fixe Beiträge in der aktuellen Verfeinerung verwendet werden.

Im Invariomformalismus wird die Übertragbarkeit anhand einer „Nächste Nachbarn“ oder „Übernächste Nachbarn“-Näherung entschieden: Sind die direkten oder auch die übernächsten Bindungspartner eines gegebenen Atoms in zwei Verbindungen jeweils identisch, so werden diese Atome als gleich angesehen. Falls für eine Bindung eine ganzzahlige Ordnung angenommen werden kann, wird nur der unmittelbare Nachbar einbezogen. Sollte die Bindung zum direkten Nachbarn gebrochenzahlig (mesomer) sein, so wird auch das übernächste Atom betrachtet, um das Ausgangsatom zu definieren. Im Fall von Wasserstoffatomen kommt ausnahmsweise immer die „Übernächste Nachbar“-Näherung zum Tragen.

Um die zugehörigen Multipolparameter zu bestimmen, wird eine Modellverbindung konstruiert, welche die chemische Umgebung in obiger Näherung nachbildet, alle offenen Valenzen werden durch Wasserstoffatome gesättigt. Diese wird auf hohem *ab initio*-Niveau (B3LYP/D95++(3df,3pd)) mit **g98** (Frisch u. a. 2002) geometrieoptimiert. Aus der resultierenden Wellenfunktion werden Strukturfaktoren für eine kubische Zelle ($a = 30 \text{ \AA}$) in der Raumgruppe $P\bar{1}$ mit **tonto** (Jayatilaka u. Grimwood 2003; Chandler u. Spackman 1978) bis $d = 0,433 \text{ \AA}$ berechnet, gegen die die Multipolparameter, unter Annahme der höchstmöglichen lokalen Symmetrie, mit **xdlsm** (Koritsanszky u. a. 2003) verfeinert werden.

Die resultierenden Pseudoatommultipolparameter werden, zusammen mit den Informationen zu verwendetem lokalem Koordinatensystem und angenommener Symmetrie, in einer Datenbank gespeichert. Sie gehen als unveränderliche Parameter in die asphärische Verfeinerung eines Atomes ein und ersetzen den sphärischen Streufaktor, der nur die Elemente unterscheidet. Die Zuordnung von Atomtypen in „Nächster“ und „Übernächster Nachbar“-Näherung führt zu einer steigenden Anzahl von Streufaktoren gegenüber dem konventionellen Formalismus. So geben die in Tab. 4.1.1 auf Seite 55 gelisteten 233 Atomtypen aller zwanzig kanonischen Aminosäuren zu 63 verschiedenen Invariomen Anlaß.

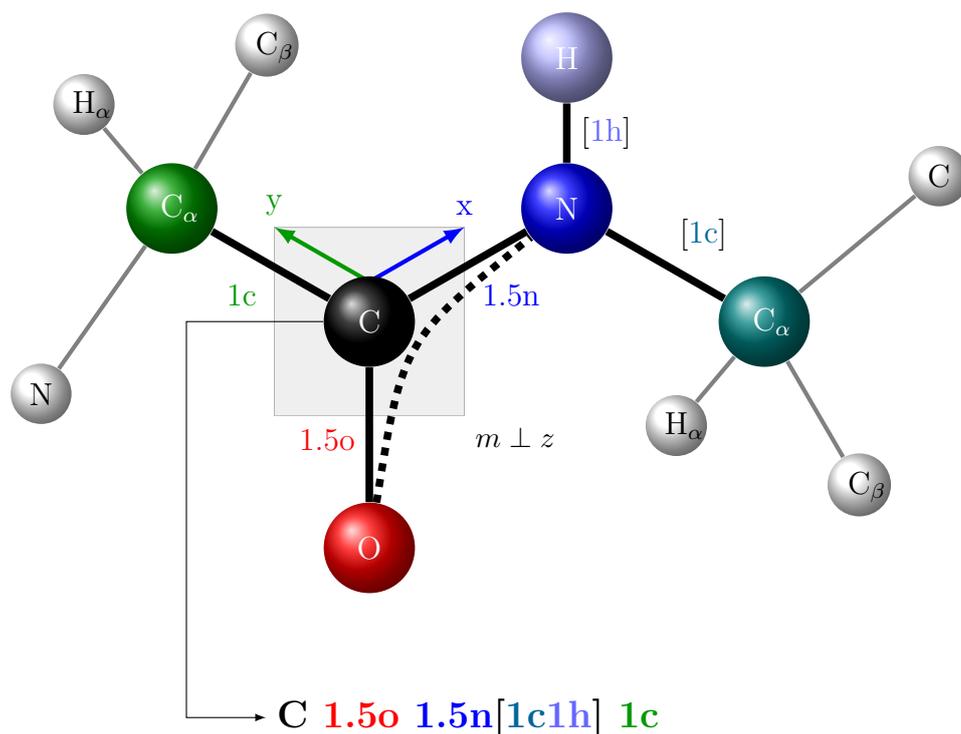


Abbildung 1.2: Ableitung des Invarioms des Carbonylkohlenstoffs der Peptidbindung.

Der Invariomname setzt sich aus dem großgeschriebenen Elementsymbol des betreffenden Atoms und kleingeschriebenen Elementsymbolen aller direkten Bindungspartner, in der Reihenfolge ihrer Ordnungszahlen, zusammen. Jeder Bindungspartner trägt dabei die modellhafte Bindungsordnung – 1 für eine Einfachbindung, 1.5 für eine mesomere Anderthalbfachbindung usw. – als Präfix. Im Falle einer nichtganzzahligen Bindungsordnung, also eines mesomer gebundenen Atoms, werden dessen Bindungspartner in eckigen Klammern diesem nachgestellt, so daß die übernächsten Nachbarn des Ausgangsatoms berücksichtigt werden. Gleiches gilt auch für Wasserstoffatome, bei ihnen werden immer die übernächsten Nachbarn berücksichtigt. Innerhalb dieser Sphäre übernächster Nachbarn können zwar wieder gebrochene Bindungsordnungen vorkommen, es werden aber keine Angaben zu der dann dritten Sphäre überübernächster Nachbarn gemacht.

Am Beispiel des Invarioms C1.5o1.5n[1c1h]1c des Carbonylkohlenstoffs der Peptidbindung soll das Konzept erläutert werden: Da es sich um das Invariom eines Kohlenstoffatoms handelt, beginnt der Name mit einem "C". Es hat als nächste Nachbarn ein "C_α", ein "O" und ein "N" mit den angenommenen Bindungsordnungen "1" und jeweils "1.5" für die Peptidbindung (siehe Abb. 1.2). Somit stellt sich der Name bisher als C1.5o1.5n1c dar. Für die mesomer gebundenen Nachbarn "1.5o" und "1.5n" müssen deren direkte Nachbarn berücksichtigt werden.

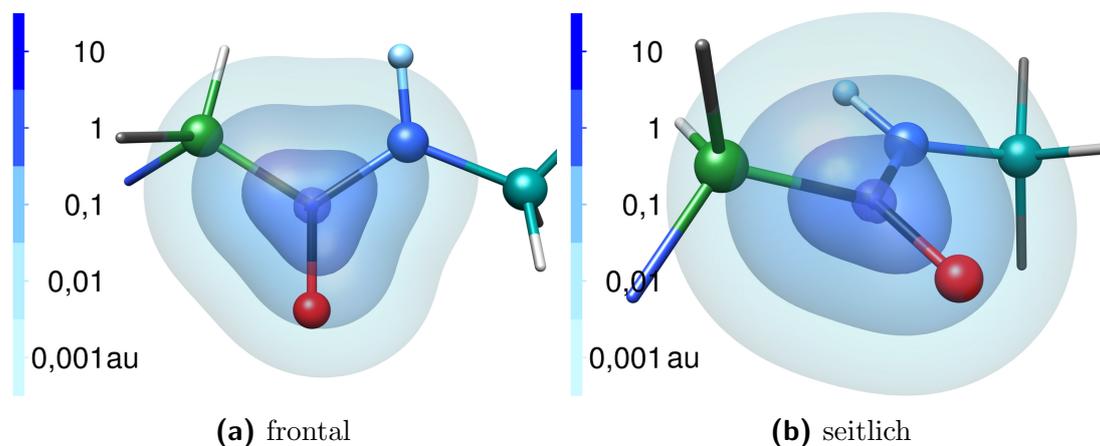


Abbildung 1.3: Elektronendichte des Invarioms C1.5o1.5n[1c1h]1c.

Der Carbonylsauerstoff ist nur an den bereits aufgeführten Carbonylkohlenstoff gebunden und fällt somit aus der weiteren Betrachtung raus. Das Stickstoffatom aber ist mit zwei weiteren Atomen mit jeweils einer Einfachbindung verknüpft: dem nächsten C_α (1c) und dem amidischen Wasserstoffatom (1h). Diese werden in eckigen Klammern dem Stickstoff nachgestellt und der vollständige Name ergibt sich als: C1.5o1.5n[1c1h]1c.

Jedem Pseudoatom ist des weiteren eine lokale Symmetrie mit einem entsprechenden lokalem Koordinatensystem zugeordnet. Für das Beispielinvariom wird bei dessen Erzeugung eine spiegelsymmetrische Elektronendichteverteilung in der C_α -C-N-Ebene angenommen. Dazu muss die Z-Achse zu dieser Spiegelebene senkrecht stehen, was vorteilhafterweise so erreicht wird, daß man die X- und Y-Achsen auf die Nachbaratome C_α und N zeigen läßt. Die Lage des Symmetrieelements und des lokalen Koordinatensystems sind ebenfalls in Abb. 1.2 angedeutet.

Aus der Abb. 1.2 leitet sich die zur Berechnung dieses Invarioms notwendige Modellverbindung dergestalt ab, daß alle grauweiß gezeichneten Atome durch Wasserstoffatome ersetzt werden, was zu N-Methylacetamid führt. Nach einer Geometrieoptimierung dieser Verbindung wird aus ihrer Elektronendichte ein Satz Strukturfaktoren erzeugt, gegen den die Struktur im Multipolformalismus asphärisch unter Beachtung der angenommenen lokalen Symmetrie verfeinert wird. Die resultierenden Multipolpopulationen werden als Datenbankeintrag gespeichert und können nun als fixe Beiträge für ein solches Carbonylkohlenstoffatom in einer anderen und größeren Struktur dienen.

Bei diesem Beispiel verfeinert sich die Monopolpopulation zu 4,0260 e, mit einem κ -Wert von 1,013766 und festem κ' von 1,0. Nach den Auswahlregeln wurden 15 der 25 Multipole populiert (siehe Abschn. 1.2.1 auf Seite 21). Deren Linearkombination mit konkreten Werten der P_{lm} ergibt eine Elektronendichteverteilung wie in Abb. 1.3 gezeigt.

2 Proteine

2.1 Protein Data Bank

In der „Protein Data Bank“ (Berman 2008; Berman u. a. 2000) sind annähernd 63000 experimentell bestimmte Strukturen hinterlegt. Den größten Anteil machen dabei Proteine mit etwa 58000 Einträgen aus, bei den restlichen handelt es sich um Nukleinsäuren oder Komplexe von Protein und Nukleinsäure. Von der Gesamtmenge an Strukturen wurde der Hauptteil (ca. 54000) durch Röntgenbeugungsexperimente bestimmt. Abb. 2.1 zeigt die Entwicklung der Anzahl von Strukturen aus Röntgenbeugungsexperimenten von den frühesten 13 Einträgen zum heutigen Stand. Jedes Jahr kommt eine steigende Zahl hinzu, waren es vor zehn Jahren noch 2237 neue Strukturen, erweiterte sich die Datenbank 2009 um 6806 Einträge. Der prozentuale Zuwachs betrug 1999 24 % und 2009 immerhin noch 15 %.

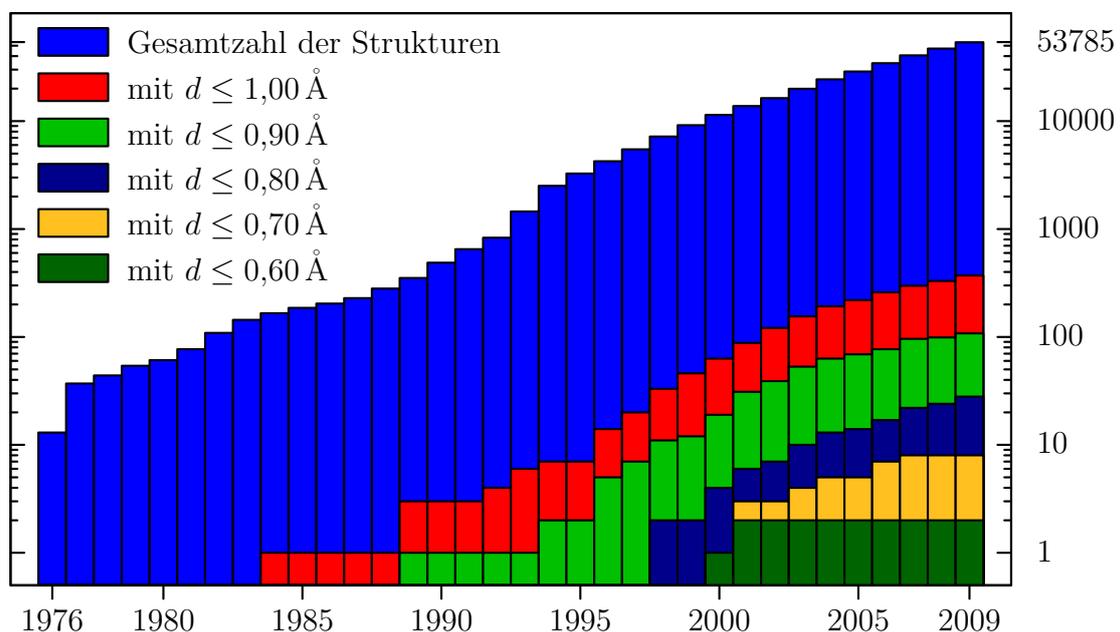


Abbildung 2.1: Jährliche Summe von Strukturen aus Röntgenbeugungsexperimenten in der PDB. Anzahl aller Einträge und für verschiedene Datenaufösungen (Stand Januar 2010).

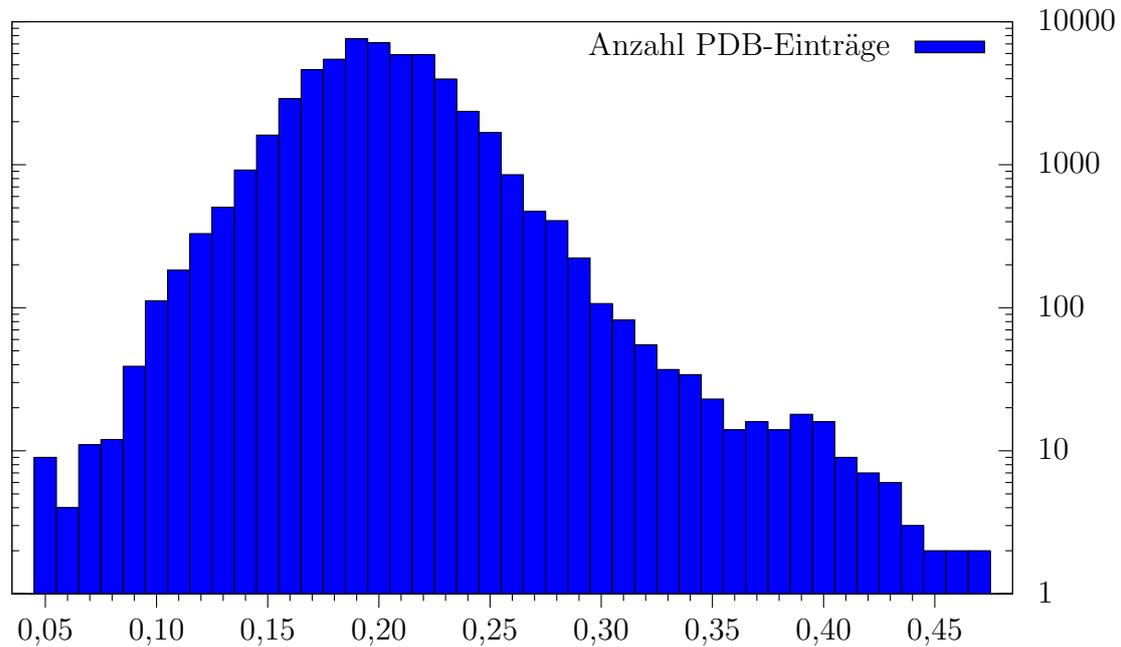


Abbildung 2.2: Anzahl von Strukturen aus der PDB pro R -Wertintervall ($R; R + 0,01$] (exklusive 4 Strukturen mit $R \geq 0,49$ und 380 ohne R -Wertangabe, Stand Januar 2010).

Die weitaus meisten Messungen haben eine Auflösung von $1,5 \leq d \leq 2,5 \text{ \AA}$. Hochaufgelöste Strukturen mit $d \leq 1,0 \text{ \AA}$ sind seit 25 Jahren in der Datenbank enthalten und summieren sich auf 371 Einträge, was 0,7% der Gesamtmenge entspricht. Der Zuwachs der letzten zehn Jahre liegt zwar über dem Durchschnitt, zuletzt aber bei nur noch ca. 12%. Besser aufgelöste Daten bis $d \leq 0,9 \text{ \AA}$ gehören zu 108 Strukturen, darunter befinden sich 28 Experimente mit einer Auflösung von $d \leq 0,8 \text{ \AA}$. Nur 8 Strukturen sind bis $d \leq 0,7 \text{ \AA}$ aufgelöst und die absolute Ausnahme sind 2 Messungen bis $d \leq 0,6 \text{ \AA}$. Mit den Jahren hat sich das Auflösungslimit zu immer geringeren Werten verschoben. Dennoch sind hoch aufgelöste Daten relativ selten und ihre Anzahl nimmt auch langsamer zu, als es für Strukturen durchschnittlicher Auflösung der Fall ist.

Nicht nur die Auflösung ist ein Qualitätskriterium, sondern auch die erreichten Gütefaktoren der Verfeinerung. Wie Abb. 2.2 zeigt, liegt das Verteilungsmaximum in einem R -Wertintervall von $(0,19; 0,20]$. Die verwendeten Werte entsprechen dem R_1 oder R_{all} , manchmal exklusive der Reflexe für die R_{free} -Validierung. Nur 75 Strukturen erreichen R -Werte unter 10%, wobei allerdings auch gering aufgelöste Daten bis zu $3,20 \text{ \AA}$ enthalten sind.

Die Anzahl von Strukturen mit sowohl $d < 1,0 \text{ \AA}$ als auch $R < 10\%$ beträgt nur 45, bei denen der Term $R \cdot d$ zwischen 4,86 und 8,91 variiert. Zehn Strukturen

zeichnen sich durch die besten Werte von $R \cdot d < 6$ aus. Nach diesem Indikator ist die kleine Proteinstruktur (4,7 kDa) Crambin (Jelsch u. a. 2000a) mit dem PDB-Bezeichner 1EJG die beste Struktur, bei einer Auflösung von $d = 0,54 \text{ \AA}$ liegt ein R_{all} -Wert von 9,2% vor. Ebenfalls ausgezeichnet ist die mit 36 kDa wesentlich größere Struktur der Aldosereduktase (Howard u. a. 2004; Guillot u. a. 2008) mit dem Kürzel 1US0, hier wurde eine Auflösung von $0,66 \text{ \AA}$ und ein R_{all} von 9,4% erreicht – für ein Protein dieser Größe absolut außergewöhnlich. Beide Strukturen wurden bereits asphärisch verfeinert. Als dritte Struktur soll eine triklin Modifikation von Lysozym unter dem Kürzel 2VB1 (Wang u. a. 2007, in dieser Arbeit P1LYSO) genannt werden, welche bis $d = 0,65 \text{ \AA}$ aufgelöst ist und zu $R_{all} = 8,4\%$ verfeinert wurde. Mit 14,3 kDa ist sie dabei wesentlich kleiner als 1US0 aber dreimal größer als 1EJG. Die restlichen sieben Strukturen haben Molekulargewichte zwischen 0,7 und 8,8 kDa.

2.2 Proteinfunktion

Proteine übernehmen in Lebewesen entscheidende Funktionen. So sind beispielsweise fast alle bisher bekannten Enzyme Proteine, deren Rolle in biologischen Systemen die eines Katalysators ist. Eine weitere Funktion von Proteinen besteht im Transport oder der Speicherung kleiner Moleküle im Körper, wie beispielsweise Hämoglobin den Sauerstofftransport im Blut ermöglicht. Auch geben sie biologischen System mechanische Festigkeit, wie z. B. das Kollagen in Haut oder Knochen. Darüber hinaus bilden sie Rezeptormoleküle des Nervensystems oder sorgen als Muskelfasern für die Bewegung von Lebewesen. Ihr Vorkommen und Zweck ist also höchst vielfältig (Stryer 1990). Dieser Abschnitt beschränkt sich auf eine kurze Behandlung der in dieser Arbeit vorkommenden Proteine Lysozym, Insulin und Trichotoxin und einen Überblick der vorhandenen PDB-Strukturen.

2.2.1 Lysozym

Lysozym ist ein in vielen Lebewesen vorkommendes Enzym. In tierischen und menschlichen Organismen wirkt es als Antibiotikum und findet sich zum Beispiel in Tränenflüssigkeit und Speichel. Auch Hühnereiweiß enthält beträchtliche Konzentrationen dieser Verbindung, aus dem es leicht zu gewinnen ist und als *hen egg-white* Lysozym (HEWL) bezeichnet wird. Dieses Protein besteht aus 129 Aminosäureresten und hat ein Molekulargewicht von 14,3 kDa. Gegenstand dieser Arbeit sind zwei kristallographische Modifikationen von HEW-Lysozym. Die erste Beschreibung von Lysozym und der antibakteriellen Wirkung stammt von Fleming (1922). In der Arbeitsgruppe um David Phillips gelang vor 45 Jahren die Aufklärung der dreidimensionalen Struktur mit Röntgenbeugungsmethoden (Blake u. a. 1965), es war somit das erste Enzym, dessen Struktur bestimmt wurde.

Als Enzym katalysiert Lysozym die Spaltung von Polysaccharidketten der Bakterienzellwand, welche dann dem osmotischen Druck des Zellinneren nicht mehr standhält, was zur Auflösung des Bakteriums führt. Angriffspunkt ist die β -1,4-glykosidische Bindung zwischen N-Acetylglucosamin (NAG) und N-Acetylmuraminsäure (NAM), zwei Zuckern, die alternierend angeordnet den Kohlenhydratanteil der Zellwand vieler Bakterien bilden. Bei der Hydrolyse dieser Bindung wird die O-glykosidische Bindung, unter Verbrauch eines Wassermoleküls, in zwei Hydroxylgruppen überführt (Stryer 1990).

Die Proteinkette ist so gefaltet, daß eine Spalte ausgebildet wird, welche sechs Zuckermomeren Platz bietet. In ihr wird das Polysaccharid durch Wasserstoffbrücken in einer Weise orientiert, welche die glykosidische Bindung zwischen dem vierten und fünften Zuckerrest in Nachbarschaft zu den Aminosäureresten Glutaminsäure 35 und Asparaginsäure 52 positioniert. Ob der unpolaren Umgebung von GLU 35, wird die Carboxylgruppe protoniert vorliegen, während die polare Nachbarschaft von ASP 52 eher eine deprotonierte Carboxylatgruppe zur Folge hat. Nach dem Phillips-Mechanismus (Phillips 1966) wird nun das Proton von GLU 35 auf den glykosidischen Sauerstoff übertragen und dadurch die Bindung zwischen den Zuckerresten gespalten, was ein Carbeniumkation zurückläßt, das durch die benachbarte Carboxylatgruppe von ASP 52 elektrostatisch stabilisiert wird. Zudem ist der betreffende Zuckerring durch die Lage in der Enzymspalte von der Sessel- in die Halbsesselkonformation verzerrt, so daß der Übergangszustand resonanzstabilisiert wird. Das Carbeniumion reagiert mit einem Wassermolekül und GLU 35 wird reprotoniert. Durch weitere Röntgenstrukturanalysen von Lysozym mit Substratmolekülen wurde dieser Mechanismus bestätigt und ergänzt (siehe z. B. Strynadka u. James 1991). Allerdings gibt es auch Hinweise auf ein kovalent an ASP 52 gebundenes Intermediat (Vocadlo u. a. 2001), statt des Oxocarbeniumions im Katalysezyklus.

Das Lysozymmolekül ist die häufigste Struktur in der PDB. Dort sind 1078 Lysozymstrukturen verschiedener Ursprungsorganismen aus Röntgenbeugungsexperimenten hinterlegt, hierin sind auch etliche hochmolekulare Komplexe mit weiteren Proteinen enthalten. Von den etwa 220 Strukturen, die hauptsächlich das HEW-Lysozymmolekül zum Gegenstand haben, ist die tetragonale Modifikation in $P4_32_12$ am stärksten vertreten, gefolgt von der triklinen $P1$ -Modifikation. Seltener sind orthorhombische ($P2_12_12_1$), monokline ($C2$, $P2_1$) und hexagonale ($P6_122$) Kristallstrukturen. Das Lysozym anderer Lebensformen kommt auch in weiteren Kristallsystemen und Raumgruppen vor.

Bezüglich Auflösung und Gütefaktoren, ist die triklinen Lysozymstruktur 2VB1 (Wang u. a. 2007) mit $d = 0,65 \text{ \AA}$ und $R_{all} = 8,4\%$ die hochwertigste, auch im Vergleich aller PDB-Strukturen sind dies außergewöhnliche Werte, wie bereits im vorherigen Abschnitt ausgeführt wurde. Auch die nächstbeste HEW-Struktur kristallisiert in $P1$, der Eintrag 3LZT (Walsh u. a. 1998) weist einen R -Wert von 9% bei einer Auflösung von $0,93 \text{ \AA}$ auf. Mit 1IEE (Sauter u. a. 2001) gibt es weiterhin

eine tetragonale Form, die bis $d = 0,94 \text{ \AA}$ gemessen und zu einem R_{all} von 12,3% verfeinert werden konnte.

In dieser Arbeit finden mit 2VB1 (P1LYSO: siehe Absch. 4.3 auf Seite 73 und Absch. 6 auf Seite 147) und dem neu gewonnenen tetragonalen Datensatz (TETRALYSO: $d = 0,80 \text{ \AA}$, $R_{all} = 13,1\%$, siehe Absch. 4.4 auf Seite 83 und Absch. 7 auf Seite 187) somit Messungen Verwendung, die zu den bestmöglichen gehören.

2.2.2 Insulin

Insulin ist ein relativ kleines Protein von 5,8 kDa. Es besteht aus zwei Ketten A und B von jeweils 21 und 30 Aminosäureresten, die über zwei Disulfidbrücken verknüpft sind. Zudem enthält Kette A noch eine dritte intramolekulare Disulfidbrücke. Die biologische Funktion ist die eines Hormones zur Stoffwechselkontrolle, es greift in den Kohlenhydrat-, Fett- und Proteinmetabolismus des Körpers ein. Insbesondere befördert es die Aufnahme von Glucose aus dem Blut in Muskel- und Fettzellen, senkt also den Blutzuckerspiegel. Daher auch die herausragende Bedeutung für die Behandlung von Diabetes mellitus, der häufigsten schweren Stoffwechselerkrankung weltweit (Stryer 1990).

Die Struktur gehört zu den ersten durch Röntgenbeugung aufgeklärten Proteinen (M.J. u. a. 1969), woran die Arbeitsgruppe um Dorothy Crowfoot Hodgkin 35 Jahre arbeitete. Zwar wurde Hodgkin bereits 1964 der Chemienobelpreis für die Aufklärung der Strukturen wichtiger biochemischer Moleküle, wie beispielsweise Penicillin oder Vitamin B12, zuteil, aber in ihrer Nobelpreisrede (Hodgkin 1972) geht sie auch auf die Arbeiten zu Insulin ein. Weitere mit Insulin verbundene Nobelpreise gingen an John James Richard Macleod und Frederick Grant Banting (geteilt mit Charles Best und James Collip) für die Entdeckung (1923, Medizin), Frederick Sanger für die Aufklärung der Sequenz (1958, Chemie) sowie Rosalyn Yalow, Roger Guillemin und Andrew V. Schally (1977, Medizin) für die Entwicklung von Radioimmunoassays zur quantitativen Bestimmung sehr kleiner Hormonmengen, welche an Insulin erstmalig praktisch erprobt wurden.

Biologisch aktiv ist die monomere Form, die Speicherung im Körper erfolgt als Hexamer in Gegenwart von Zinkionen, nach vorheriger Dimerisierung. Die hexamere Form zeigt drei allosterische Konformationen, die sich in der Ausdehnung der α -Helix am N-terminalen Ende der B-Kette unterscheiden (Kaarsholm u. a. 1989). In der T_6 -Form, welche der Struktur in dieser Arbeit entspricht, sind die ersten acht Aminosäurereste nicht Teil der nachfolgenden Helix, während sie in der R_3 -Form eine helikale Konformation einnehmen. Als Zwischenstufe liegt die $T_3R_3^f$ -Form vor, in der die ersten drei Reste nicht helikal vorliegen, die folgenden fünf aber sehr wohl. Abhängig von Kristallisationsbedingungen wie pH-Wert, Anion- und Zinkkonzentration, An- oder Abwesenheit von Phenolderivaten, als auch der Temperatur, wird die eine über die andere Form bevorzugt (Norrman u. Schluckebier 2007; Smith u. a. 2003, 2000). Hohe Chloridkonzentrationen und die

Verfügbarkeit von Phenolliganden bedingen beispielsweise die R_6 -Konformation. Ebenso zeigt Insulin eine große Vielfalt von Kristallformen. Am häufigsten ist die rhomboedrische Modifikation in der Raumgruppe R3 und die kubische Modifikation in I2₁3 des zinkfreien Insulins, wie man einer Statistik über die Strukturen in der PDB entnehmen kann. Ansonsten finden sich seltenere Insulin- oder Insulin-derivatstrukturen in triklinen, monoklinen, orthorhombischen oder tetragonalen Raumgruppen (siehe z. B. Norrman u. a. 2006; Smith u. Blessing 2003).

Die Insulindatensätze der PDB mit den bis dato besten R -Werten sind die Strukturen 2BN3 (Nanao u. a. 2005), eine zinkfreie kubische Modifikation mit einem R_{all} -Wert von 13% bei einer Auflösung von $d = 1,40 \text{ \AA}$, 3EXX (Prugovecki u. a. 2008), ein rhomboedrisches Nickelderivat mit $R_{all} = 12,6 \%$ bei $d = 1,35 \text{ \AA}$, und 1ZEG, ebenfalls eine rhomboedrische Modifikation mit einem R -Wert von 14,5% und einer Auflösung von $1,60 \text{ \AA}$. Am höchsten aufgelöst ist der 1MSO-Datensatz (Smith u. a. 2003) mit $d = 1,00 \text{ \AA}$, der R_{all} beträgt allerdings 18,3%. Insgesamt werden 338 Strukturen unter dem Schlüsselwort Insulin aufgeführt, davon sind etwa 58 Röntgenbeugungsdaten des eigentlichen Moleküls.

Somit stellt der in dieser Arbeit verwendete Datensatz (INSULIN: $d = 0,90 \text{ \AA}$, $R_{all} = 11,5 \%$, siehe Absch. 4.5 auf Seite 94 und Absch. 8 auf Seite 217), sowohl bezüglich Auflösung als auch R -Wert, die beste bisherige Messung dar.

2.2.3 Trichotoxin

Trichotoxin, konkret Trichotoxin_A50E (Chugh u. a. 2002), gehört zur Klasse der Peptaibol-Antibiotika. Dabei handelt es sich um Polypeptide mit 15–20 Aminosäureresten, sie enthalten einen hohen Anteil nichtproteinogener Aminosäuren, zumeist α -Aminoisobuttersäure (AIB) aber auch Isovalin (IVA), Hydroxyprolin (HYP) oder Ethylnorvalin (ETNOR). Sie werden von Schimmelpilzen gebildet und der N-Terminus ist im Allgemeinen acetyliert, während der C-Terminus häufig aus einem Aminoalkohol besteht (Jaworski 2000; Whitmore u. a. 2003; Chugh u. Wallace 2001). Die seltenen Aminosäuren sind wirksame Helixbildner, die resultierenden Strukturen dienen als Modelle für Ionenkanäle (Sansom 1993a,b).

Die Peptaiboldatenbank (Whitmore u. a. 2003) listet 315 Sequenzen von Molekülen dieser Substanzklasse, sechs Kristallstrukturen sind in der PDB zu finden. Neben der in dieser Arbeit verwendeten Struktur 1M24 ($d = 0,90 \text{ \AA}$, $R_{all} = 7,6 \%$, siehe Absch. 4.6 auf Seite 105 und Absch. 9 auf Seite 247) sind dies Cephaibol A, B und C (1OB4, 1OB6, 1OB7, (Bunkoczi u. a. 2003)) mit Auflösungen zwischen $0,89$ und $0,95 \text{ \AA}$ und R -Werten von 7% oder 9%, Alamethicin unter dem Code 1AMT (Fox Jr. u. Richards 1982) mit einem $R_{all} = 15,5 \%$ bei $d = 1,50 \text{ \AA}$, sowie Antiamoebin (1JOH, Snook u. a. 1998) welches bis $1,40 \text{ \AA}$ aufgelöst ist und einen R_{all} von 15,6% aufweist.

2.3 Proteinaufbau

Proteine sind lineare Ketten aus Aminosäuren, die über Amidbindungen zu einem durchgehenden Polymer kondensiert sind, manchmal quervernetzt durch Disulfidbrücken. Die Abfolge von Aminosäureresten wird Sequenz genannt und ist für jedes Protein exakt definiert, sie bezeichnet die Primaärstruktur eines Proteins. Kurze Polypeptidsequenzen werden Oligopeptid genannt, ab etwa 20 Resten kann man von einem Protein sprechen, die Einteilung ist aber willkürlich. Das Kettende mit einer Aminogruppe (N-Terminus) wird per Konvention als Beginn des Polymers betrachtet, das mit einer Carboxylgruppe (C-Terminus) als ihr Abschluß. Die natürlich vorkommenden Proteine bestehen ausschließlich aus α -Aminosäuren, bei denen Carboxyl- und Aminogruppe an das C_α gebunden sind. Diese sind, bis auf Glycin, chiral, aber nur L-Aminosäuren bilden die Bausteine von Proteinen (Stryer 1990).

Zumeist sind Proteine aus einem Satz von 20 kanonischen Aminosäuren, auch Standardamino-säuren oder proteinogene Aminosäuren genannt, aufgebaut (siehe Tab. 4.1.1 auf Seite 55), die im genetischen Code angelegt sind, während die restlichen der insgesamt 140 natürlich vorkommenden Aminosäuren durch nachträgliche Modifikation erzeugt werden (Uy u. Wold 1977). Neueren Erkenntnissen zufolge (siehe z. B. Ambrogelly u. a. 2007, und dortige Referenzen), zählen zwei weitere Aminosäuren, Selenocystein und Pyrrolysin, zum proteinogenen Satz.

Die Seitenketten der Aminosäuren unterscheiden sich in ihrem Aufbau und ihren Eigenschaften, wie Hydrophilie, Polarität, Ionisierbarkeit und der Möglichkeit als Wasserstoffbrückendonator oder -akzeptor zu wirken. Um die Energie des Moleküls zu minimieren, führt dies, zusammen mit dem Einfluß des Solvensbereichs, zu übergeordneten Faltungs- und Packungsstrukturen. Der dreidimensionale Aufbau kann in Sekundär-, Tertiär- und Quartärstruktur unterteilt werden (Stryer 1990; Richardson 1981, bzw. die aktualisierte Onlineversion: <http://kinemage.biochem.duke.edu/teaching/anatax/index.html>).

Die beiden wichtigsten Sekundärstrukturen sind α -Helix und β -Faltblatt, vorhergesagt von Pauling u. a. (1951). In der α -Helix ist die Polypeptidkette so verdreht, daß zwischen dem Carbonylsauerstoff eines Aminosäurerests n und dem Wasserstoff des Stickstoffs der Peptidbindung eines Restes $n + 4$ eine Wasserstoffbrücke geknüpft wird. Entlang der Helix ist jeder Rest um $1,5 \text{ \AA}$ gegenüber dem nächsten versetzt und um 100° verdreht – eine volle Windung umfaßt 3,6 Aminosäurereste. Da zwischen den Wasserstoffbindungspartnern 13 Atome liegen, kann dieses Strukturelement deshalb auch $3,6_{13}$ -Helix genannt werden. Im Gegensatz dazu nimmt die Polypeptidkette in β -Faltblättern (β -sheet) eine gestreckte Konformation an. Zwischen zwei gegenüberliegenden Strängen (β -strand) werden Wasserstoffbrücken aufgebaut. Wenn der zweite Strang dem ersten entgegengesetzt ausgerichtet ist (antiparalleles Faltblatt), stehen die Wasserstoffbindungen zwischen CO- und NH-Gruppen beider Ketten senkrecht auf den Strängen; ist der zweite Strang dem

ersten gleichgerichtet (paralleles Faltblatt), ergeben sich gewinkelte Wasserstoffbrückenbindungen. Wird die Wasserstoffbrücke zwischen der Carbonylfunktion eines Aminosäurerests und der NH-Gruppe einer nur drei Reste weiter befindlichen Peptidbindung gebildet, so resultiert daraus eine Haarnadel- oder β -Schleife, die zu einer Richtungsumkehr der Polypeptidkette führt. Strukturbereiche, die keiner der obigen geordneten Strukturtypen oder deren Untertypen zugeordnet werden können, werden unter dem Begriff Knäuel (*coil*) subsummiert.

Unter Erhalt der Elemente der Sekundärstruktur kann sich das Protein zu einer übergeordneten Tertiärstruktur falten. So werden in kugelförmigen globulären Proteinen sehr kompakte unsymmetrische Strukturen erreicht, in denen z. B. hydrophobe Seitenkette nach innen gekehrt sind, was sie vor der äußeren Wassersphäre verbirgt. In Faserproteinen bewirken beispielsweise Superhelices aus zwei oder drei α -Helices das stabförmige Äußere. Die spezifische Form eines Proteins, die durch die Faltung erreicht wird, ist wichtig für die jeweilige biologische Funktion.

Falls das Protein aus mehreren Untereinheiten besteht, wird deren Anordnung zueinander als Quartärstruktur bezeichnet. Ein Beispiel ist das Insulin-Hexamer.

Zu einer Proteinstruktur gehört des weiteren ein hoher Solvensanteil, insbesondere Wasser, welcher die vorliegende Faltung stabilisiert und die Konformationen von Seitenketten beeinflusst. Ein Teil des Wassers ist nahe dem Protein auf diskreten Positionen zu finden (gebundenes Wasser), während der größere Teil ungeordnet verteilt ist (*bulk solvent*).

Im Mittel liegt der Solvensanteil von Proteinen nahe 43 %, mit einem Schwankungsbereich von 27 % bis 65 % (Matthews 1968). Der Matthews-Koeffizient aus Volumen der asymmetrischen Einheit und Molgewicht (in Dalton) des Proteins

$$V_M = \frac{V_{ASU}}{M} \quad (2.1)$$

liegt üblicherweise zwischen 1,7 und 3,5 Å³/Da, der Umrechnungsfaktor in cm³/g beträgt 1/1,66. Unter der Annahme einer Proteindichte von $\rho_{prot} = 1,35 \text{ g/cm}^3$ und mit dem partiellen spezifischen Volumen $\tilde{v} = 1/\rho_{prot}$, ergeben sich Protein- und Solvensanteil am Zellvolumen zu

$$V_{prot} = \frac{1,66\tilde{v}}{V_M} \approx \frac{1,23}{V_M} \quad \text{und} \quad V_{solv} = 1 - \frac{1,66\tilde{v}}{V_M} \approx 1 - \frac{1,23}{V_M} \quad (2.2)$$

Der Wert der partiellen Proteindichte muß in separaten Dichtemessungen gewonnen werden. Die standardmäßig verwendeten 1,35 g/cm³ sind nicht unumstritten, so geben Andersson u. Hovmöller (1998) einen durchschnittlichen Wert von 1,22 g/cm³ an, woraus ca. 10 % niedrigere Solvensanteile resultieren (Andersson u. Hovmöller 2000). Umgekehrt kann diese Dichte auch zu höheren Werten berechnet werden (Quillin u. Matthews 2000), mit der Konsequenz erhöhten Solvensanteils.

Methoden

3 Programme und Auswertung

3.1 Statistik

Bis auf die R -Werte, die aus den Verfeinerungsprogrammen stammen, wurde jedwede statistische Auswertung mittels R (R Development Core Team 2009) durchgeführt. Die verwendeten Größen sind nachfolgend definiert.

3.1.1 Mittelwerte und Streuung

Als Maß für die zentrale Tendenz einer Verteilung von Daten kann man verschiedenste Mittelwerte bilden. Besonders gebräuchlich sind arithmetisches Mittel und Median. Als Maßzahl für die Streubreite einer Verteilung von Daten um einen Mittelwert werden u. a. die geschätzte Standardabweichung σ_x (engl.: *esd* - *estimated standard deviation*), der Interquartilabstand oder die Spannweite benutzt.

Das häufig nur Mittelwert genannte arithmetische Mittel \bar{x} eines Datensatzes mit n Werten und die damit assoziierte Standardabweichung σ_x sind definiert als:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.1a)$$

$$\sigma_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (3.1b)$$

Der Median, auch Zentralwert genannt, einer nach Größe geordneten Verteilung von n Werten kann wie folgt bestimmt werden:

$$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} & \text{für } n \text{ ungerade} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & \text{für } n \text{ gerade} \end{cases} \quad (3.2)$$

Es liegt also jeweils die Hälfte der Daten unter und über dem Median, der somit dem 50 %-Quartil Q_2 entspricht. Da er ein gegen Ausreißer robuster Lageparameter ist, eignet er sich somit besonders für nicht normalverteilte Daten. Analog bestimmen sich die 25 %- und 75 %-Quartile, Q_1 und Q_3 , es liegt jeweils ein Viertel der Werte darunter bzw. darüber.

Der Interquartilabstand IQR als Schätzer für die Streuung der Daten ist einfach die Differenz zwischen Q_3 und Q_1 . Es werden also vom Interquartilabstand nach

Ausschluß der kleinsten und größten 25 % der Daten die verbleibenden 50 % aller Werte rund um den Median eingeschlossen.

Des weiteren kann z. B. der Median der absoluten Abweichungen zum Median MAD als Streuungsmaß herangezogen werden. Zumeist wird er mit einem Normalisierungsfaktor 1,4826 multipliziert, um ihn auf die Standardabweichung der Normalverteilung zu skalieren:

$$MAD = 1,4826 \cdot median \sum_i |x_i - \tilde{x}| \quad (3.3)$$

Die besprochenen Parameter sind in Abb. 3.1 auf der gegenüberliegenden Seite illustriert. Gezeigt wird eine nicht normalverteilte Menge von $n = 30$ Werten mit Minimal- und Maximalwerten von $x_{min} = 1$ und $x_{max} = 24$, wobei Letzterer deutlich von den restlichen Werten abweicht (man beachte die durchbrochene Ordinate). Der nach Gleichung 3.1a berechnete Mittelwert und die Standardabweichung liegen mit $\bar{x}(\sigma_x) = 5,333(4,9782)$ über dem Median. Dieser wurde berechnet, indem die 30 Werte von 1 bis 24 ansteigend geordnet wurden; der fünfzehnte und sechzehnte Wert ist jeweils 4, der Mittelwert davon ist wieder $\tilde{x} = 4$ mit einem $MAD = 4,4478$ als Streuungsmaß. Ein Großteil der Daten hat eher kleinere Werte und der Median scheint die Verteilung besser als das arithmetische Mittel wiederzugeben.

Die Robustheit gegen den Ausreißer zeigt sich bei Berechnung der Kennziffern unter Ausschluß dieses Wertes: Der Median und der Median der absoluten Abweichungen bleiben bei $\tilde{x}(MAD) = 4(4,4478)$, während sich der Mittelwert und die Standardabweichung auf $\bar{x}(\sigma_x) = 4,69(3,5768)$ verringern. Ebenfalls verringert sich $Q3$ auf 7 und der Interquartilabstand auf 5, während $Q1$ unverändert bei 2 bleibt. Mit steigender Anzahl und Höhe von Ausreißern würde dieser Vergleich umso stärker ausfallen.

3.1.2 Signifikanztest

Beim Vergleich zweier Datensätze kann man statistische Kennwerte, wie z.B. den Mittelwert, einander gegenüberstellen und anhand dieses Vergleichs eine Aussage machen. Man kann darüber hinausgehend aber auch nach der Sicherheit einer solchen Aussage fragen und einen Signifikanztest durchführen, um z.B. herauszufinden, ob Abweichungen in den Mittelwerten durch Zufall entstanden sein können oder ob eine echte Beeinflussung der Daten vorliegt (Handl 2006).

Dazu werden eine Nullhypothese H_0 und eine Alternativhypothese H_1 aufgestellt und aus den tatsächlich beobachteten Daten eine Teststatistik errechnet. Der Teststatistik wird unter der Annahme der Richtigkeit von H_0 eine bestimmte Verteilung unterstellt, wie z.B. eine Normalverteilung oder eine t -Verteilung als Spezialfall einer Normalverteilung. Man fragt dann nach der Vereinbarkeit der unterstellten und beobachteten Verteilung. Zu einem gegebenen Signifikanzniveau gibt es kritische

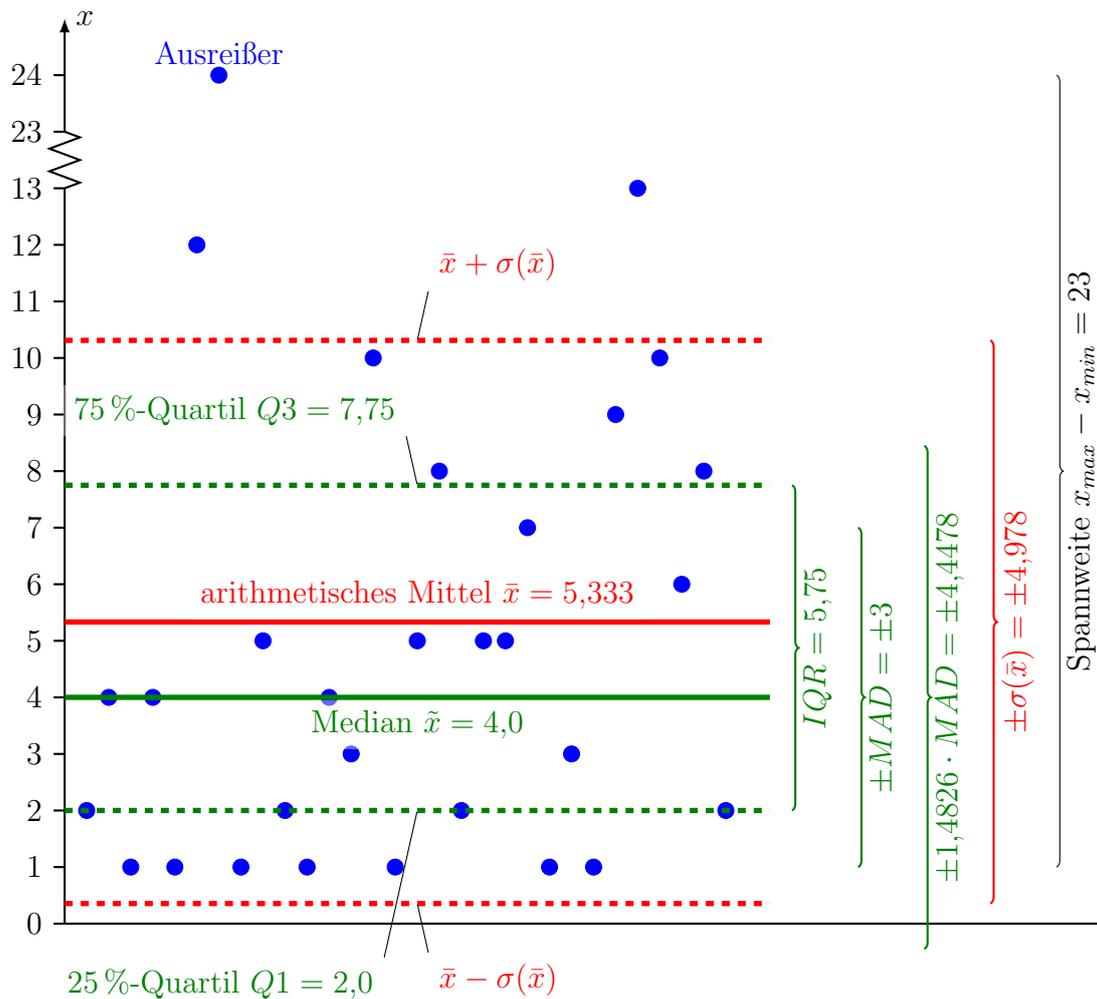


Abbildung 3.1: Mittelwerte und Streuungsmaße.

Bereiche in der Zufallsverteilung, in denen die beobachtete Verteilung nicht liegen darf, um zur Nullhypothese verträglich zu sein.

Dabei kann man sowohl einen Fehler 1. Art (Ablehnung von H_0 aufgrund der beobachteten Daten, obwohl zutreffend), als auch 2. Art (Annahme von H_0 , obwohl H_1 richtig gewesen wäre) machen. Beide Fehlerarten sind voneinander abhängig, verkleinert man z.B. den Fehler 1. Art, so begeht man häufiger einen Fehler 2. Art. Per Konvention testet man unter Minimierung des Fehlers 1. Art und bestimmt eine Übertretungswahrscheinlichkeit p , die das kleinste Signifikanzniveau angibt, zu dem man die Nullhypothese H_0 ablehnen kann. Als Gegenhypothese formuliert man das vermutete Verhalten der Daten. Der p -Wert ist also ein Maß für die Verträglichkeit der Daten mit der Nullhypothese. Je kleiner er ausfällt, desto eher

kann H_0 verworfen werden. Er wird nur zur Entscheidung über Annahme oder Ablehnung der Nullhypothese verwendet.

Es ist üblich, ab einer Irrtumswahrscheinlichkeit kleiner 5% von einem signifikanten, ab 1% von einem hochsignifikantem Ergebnis zu sprechen. Liegt die Übertretungswahrscheinlichkeit p über diesen Niveaus, so kann die Nullhypothese H_0 nicht "statistisch signifikant" zu diesem Signifikanzniveau abgelehnt werden - man riskiert einen zu großen Fehler bei Annahme der Alternativhypothese H_1 .

t-Test

Für diesen Test wird eine Normalverteilung der Daten vorausgesetzt. Die Nullhypothese geht von keinem Unterschied in den Mittelwerten zweier Datensätze A und B aus, die Gegenhypothese dazu ist eine Verschiedenheit der Mittelwerte:

$$\begin{aligned} H_0 : \quad & \mu_A = \mu_B \\ H_1 : \quad & \begin{cases} \mu_A \neq \mu_B & \text{(beidseitiger Test)} \\ \mu_A \geq \mu_B \quad \text{oder} \quad \mu_A \leq \mu_B & \text{(einseitiger Test)} \end{cases} \end{aligned} \quad (3.4)$$

Aus den m und n umfassenden Daten wird eine Teststatistik

$$t = \frac{\bar{A} - \bar{B}}{\sigma \sqrt{\frac{1}{m} + \frac{1}{n}}} \quad (3.5a)$$

$$\sigma^2 = \frac{1}{m+n-2} \left(\sum_{i=1}^m (A_i - \bar{A})^2 + \sum_{j=1}^n (B_j - \bar{B})^2 \right) \quad (3.5b)$$

berechnet und gegen eine ebensolche t -Verteilung mit $m+n-2$ Freiheitsgraden verglichen, für die $\bar{A} = \bar{B}$ gilt. H_0 wird abgelehnt, wenn $|t| > t_{m+n-2; 1-\alpha}$, dem $1-\alpha$ -Quantil dieser Zufallsverteilung, ist. Der p -Wert gibt die Wahrscheinlichkeit unter der Annahme der Richtigkeit von H_0 an, daß die zufallsverteilte Staistik t_{m+n-2} einen Wert extremer als t annimmt. Ist dann p kleiner als das gewünschte Signifikanzlevel α , kann H_0 abgelehnt werden.

Wilcoxon-Test

Dieser Signifikanztest benötigt keine normalverteilten Daten. Für die Verteilungen A und B vom Umfang m und n wird untersucht, ob es eine Lageverschiebung der Mediane gibt:

$$\begin{aligned} H_0 : \quad & M_A = M_B \\ H_1 : \quad & \begin{cases} M_A \neq M_B & \text{(beidseitiger Test)} \\ M_A \geq M_B \quad \text{oder} \quad M_A \leq M_B & \text{(einseitiger Test)} \end{cases} \end{aligned} \quad (3.6)$$

Dazu werden alle Beobachtungen A_i und B_j zusammen nach Größe geordnet (z.B. $B_{23}, B_{1275}, A_{30}, \dots$) und ihnen ein Rang R entsprechend der Position in der Liste

zugeordnet (z.B. 1,2,3, ...). Dann wird die Rangsumme von A bestimmt:

$$W = \sum_{i=1}^m R_i \quad (3.7)$$

und in eine normalverteilte Teststatistik

$$Z = \frac{W - m(m+n+1)/2}{\sqrt{mn(m+n+1)/12}} \quad (3.8)$$

eingesetzt. Wieder wird H_0 abgelehnt, wenn $|Z|$ in den zu einem bestimmten Signifikanzniveau kritischen Bereichen einer Standardnormalverteilung unter den Bedingungen von H_0 liegt, bzw., wenn die Übertretungswahrscheinlichkeit unter die gewünschte Signifikanz fällt.

3.1.3 Gütefaktoren und Gewichtung

Das `shelx1`-Gewichtungsschema umfaßt sechs mögliche Parameter (Standardwerte $a = 0, b = 0, c = 0, d = 0, e = 0, f = 1/3$). Werden nur die ersten beiden Parameter verfeinert, ergibt sich

$$w = \left(\sigma (F_o^2)^2 + \left(a \left(\frac{F_o^2}{3} + \frac{2F_c^2}{3} \right) \right)^2 + b \left(\frac{F_o^2}{3} + \frac{2F_c^2}{3} \right) \right)^{-1} \quad (3.9)$$

als Gewichtung jedes Reflexes. Die Verfeinerung wird anhand von Gütefaktoren wie dem *Goodness of fit* GoF_w

$$GoF_w = \sqrt{\frac{\sum w (F_o^2 - F_c^2)^2}{N_{ref} - N_{par}}} \quad (3.10)$$

und den R -Werten (*residual factor*)

$$wR2 = \sqrt{\frac{\sum w (F_o^2 - F_c^2)^2}{\sum w (F_o^2)^2}} \quad (3.11a)$$

$$R1 = \frac{\sum ||F_o(> 4\sigma(F_o))| - |F_c||}{\sum |F_o(> 4\sigma(F_o))|} \quad (3.11b)$$

$$R_{all} = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|} \quad (3.11c)$$

beurteilt.

3.2 Strukturqualität

Zur Überprüfung der Strukturqualität, im Sinne der Geometrie, Verschiebungsparameter, Packungsverhalten, Wasserstoffbrückenbindungen sowie Koordinatenfehlern, wurden die Diagnosen von **procheck**, **whatcheck** und **shelxpro** herangezogen. Die Ausgaben für jede der Strukturen dieser Arbeit sind auszugsweise im jeweiligen Anhang gegeben. Im Folgenden sollen die einzelnen Tests erläutert werden.

3.2.1 Procheck

Das Program **procheck** (Laskowski u. a.) versucht eine automatische Beurteilung der „stereochemischen Qualität“ einer Proteinstruktur durchzuführen, indem etliche geometrische Kriterien angezeigt und mit Referenzdaten verglichen werden, wobei ungewöhnliche Abweichungen besonders hervorgehoben werden. Im folgenden werden die einzelnen, im Anhang für jede Struktur gezeigten Tafeln erläutert.

Ramachandranauftragung

Ramachandran u. a. (1963) zeigten, daß sterische Behinderungen die Annahme beliebiger Kombination von Rotationswinkeln beidseits der Peptidbindung verhindern. Zwei Diederwinkel legen die Konformation der Peptidhauptkette fest: $\phi(C,N,C_\alpha,C)$ gibt die Rotation um die C_α -N-Bindung an, $\psi(N,C_\alpha,C,N)$ die Rotation um die C_α -C-Bindung. Die ϕ - und ψ -Torsionswinkel aller Aminosäurereste, außer der endständigen, werden gegeneinander aufgetragen. Aus der Häufigkeitsverteilung der Winkel von 121870 Aminosäureresten (exkl. Glycin und Prolin) von 463 PDB-Strukturen wurden Bereiche (*Core*, *Allowed*, *Generous*) abgeleitet (Morris u. a. 1992), die als energetisch vorteilhaft bis nachteilig betrachtet werden können und den Konformationen in rechts- und linkshändiger α -Helix sowie β -Faltblatt entsprechen. Glycin-Reste haben einen weiteren Konformationsraum, während Prolin einen eingeschränkteren ϕ -Winkel um -65° hat. Anhand der Lage der Wertepaare bezüglich der Vorzugsbereiche kann die Strukturqualität beurteilt werden, es sollten über 90 % der Reste im Kernbereich liegen.

Hauptkettenbindungslängen und -winkel

Die Werte und Häufigkeiten der vorkommenden Bindungslängen und -winkel der Hauptkette sind zusammen mit dem jeweiligen Erwartungswert und dessen Standardabweichung aufgetragen. Werte, die mehr als 2σ vom entsprechenden Mittel entfernt sind, werden hervorgehoben. Die Erwartungswerte der Bindungslängen und -winkel nach Engh u. Huber (1991) entstammen einer Auswahl der damals 80000 Strukturen der „Cambridge Structural Database“ (Allen 2002b) mit $R1 < 6\%$. Den 20 kanonischen Aminosäuren entsprechende Fragmente wurden auf die Bindungsabstände und -winkel von 31 Atomtypen untersucht und Mittelwerte und Standardabweichungen bestimmt.

Quadratisch gemittelte Abweichung der Planarität

Die Atome in aromatischen Ringen (Phe, Tyr, Trp, His) und mehreren endständigen Gruppen (Arg, Asn, Asp, Gln, Glu) sollten in einer gemeinsamen Ebene liegen. Die quadratisch gemittelte Abweichung der jeweiligen Atome eines Rings oder Gruppe von der Planarität wird angezeigt. Wenn diese höher als 0,03 Å oder 0,02 Å ausfällt, gilt dies als Ausreißer betrachtet.

Geometrieverzerrungen

Signifikante Ausreißer bzgl. der zuvor genannten geometrischen Eigenschaften.

Eigenschaften pro Aminosäurerest

Weitere geometrische und statistische Größen, aufgetragen für jeden Aminosäurerest entlang der Sequenz. Abweichungen beziehen sich auf die „Idealwerte“ nach Morris u. a. (1992).

- Absolutabweichung vom mittleren χ_1 -Torsionswinkel, der durch die Atome N, C_α , C_β , C_γ definiert ist und die Rotation um die erste Seitenkettenbindung angibt.
- Absolutabweichung vom mittleren ω -Torsionwinkel, definiert durch C_α , N, C, C_α , was der Rotation um die Peptidbindung N–C entspricht (nahe 180°).
- C_α -Chiralität, Absolutabweichung vom ζ -Torsionswinkel, definiert als „virtueller“ Diederwinkel C_α , N, C, C_β (Werte um $34^\circ \pm 4^\circ$ entsprechen einem L-Isomer).
- Zuordnung der Sekundärstruktur (Kabsch u. Sander 1983) und eine Näherung für die Zugänglichkeit des Aminosäurerestes (Nishikawa u. Ooi 1986).
- Symbole für die Lage des Restes in der Ramachandranaufrtragung.
- Farbige Repräsentation des G-Faktors für etliche Diederwinkel, Bindungslängen und -winkel sowie ein Gesamt-G-Faktor: Die Verteilung dieser Größen in einem Referenzsatzes wird zum Standard für eine normale Geometrie gemacht. Die aktuelle Verteilung wird dazu ins Verhältnis gesetzt und logarithmiert. Dieser „*log-odd score*“ wird G-Faktor genannt und als Angabe zur Normalität oder Häufigkeit der stereochemischen Eigenschaften benutzt. Je niedriger der G-Faktor, desto schlechter oder unüblicher ist die Stereochemie (unter -0,5 bedenklich, 0 Durchschnitt), eine dunklere Färbung zeigt negativere Werte an.

3.2.2 Whatcheck

Die Strukturtests des Programms `whatcheck` (Vriend 2008) basieren zumeist auf einem Vergleich mit Daten, welche aus einem regelmäßig aktualisierten Satz der 300 besten PDB-Strukturen gewonnen werden. Nur Bindungslängen und -winkel werden gegen die Referenzwerte von Engh u. Huber (1991) verglichen.

Fast alle Unterschiede zwischen Parametern der untersuchten Struktur und dem Vergleichssatz werden in Form von Z-Werten oder RMS-Z-Werten angegeben. Der Z-Wert gibt die Abweichung zwischen beobachtetem Wert x und Referenzmittelwert \bar{x} als Anzahl von Standardabweichungen des Mittelwertes $\sigma(\bar{x})$ an:

$$Z = \frac{x - \bar{x}}{\sigma(\bar{x})}$$

Bei einem Wert von Null stimmt der aktuelle Wert vollständig mit dem Erwartungswert überein. Die Tests sind so gestaltet, daß meistens eine negative Abweichung schlechter als der Durchschnitt ist, während ein positiver Wert besser als der Durchschnitt dieses Parameters im Referenzsatz ist. Außerdem wird das quadratische Mittel (RMS) mehrerer (n) Z-Werte

$$\text{RMS} - Z = \sqrt{\frac{\sum Z^2}{n}}$$

verwendet, das idealerweise bei 1,0 liegt.

Die Tests münden in drei Typen von Meldungen: Hinweise – Angabe der Werte der Statistik und Abbildungen; Warnungen – ungewöhnliche Ergebnisse; Fehler – Werte, die auf Fehler oder sehr hohe Abweichungen hinweisen. Meist ist jeder Test mit einer Einschätzung verbunden (*OK*, *unusual*, *poor*, ...). Der resultierende Bericht ist ausgesprochen lang und wird in den Anhängen stark gekürzt wiedergegeben, nur die eigentlichen Validierungen sind aufgeführt. Auch die zu jedem Test gehörenden Erläuterungen wurden entfernt. Da die einzelnen Tests mit sprechenden Titeln versehen sind und jeweils eine Beurteilung enthalten, werden sie nicht einzeln besprochen.

Neben Tests zur Stereochemie, welche recht ausführlich angelegt sind, gibt es einige weitergehende Vergleiche und Überprüfungen: Die Zellparameter werden auf konventionelle Aufstellung überprüft und anhand des Matthew-Koeffizienten die Zahl der Moleküle pro Zelle validiert. Es wird auch geprüft, ob sich nicht gebundene oder in Wasserstoffbrückenbindungen befindliche Atome zu nahe kommen und Abstände kleiner als die Summe der van-der-Waals-Radien abzüglich 0,4 Å zeigen. Anhand des Vergleichs der Umgebung von Atomgruppen in der untersuchten Struktur und der Häufigkeit einer solchen Packung im Referenzsatz, wird zudem die Packungsqualität bewertet. Ebenso, ob alle potentiellen Wasserstoffbrückenbindungsdonoren und -akzeptoren eine solche Wechselwirkung eingehen können.

3.2.3 Shelxpro

Luzzati-Auftragung

Die Entwicklung des $R1$ -wertes mit steigender Auflösung, angegeben als $1/d/\text{Å}^{-1}$, wird gezeigt. Zusätzlich sind theoretische Kurven (Luzzati 1952) für die Änderung

des R -Wertes eingezeichnet. Mit ihnen ist eine Schätzung des durchschnittlichen Fehlers der Atomkoordinaten verbunden, wobei perfekte Daten und eine unvollständig verfeinerte Struktur vorausgesetzt werden. Die Datenpunkte der überprüften Struktur sollten eine Kurve ergeben, die einer der theoretischen Verteilungen ähnelt. Dann kann an der nächst gelegenen Kurve die ungefähre durchschnittliche Unsicherheit der Atomkoordinaten abgelesen werden. Dies ist nur eine grobe Schätzung der oberen Grenze des Fehlers.

ESD-Analyse

Nur nach einer echten *least square*-Verfeinerung erhält man mit den Standardabweichungen (*esd*) der Positionsparameter eine zuverlässige Angabe zu den Koordinatenfehlern. Restraints oder Dämpfung der Verfeinerung führen allerdings zu unterschätzten Standardabweichungen. In einem Streudiagramm werden die Fehler der Atomkoordinaten und Bindungslängen gegen die korrespondierenden B_{eq} aufgetragen, getrennt nach Kohlenstoff-, Sauerstoff- und Stickstoffatomen, die nicht fehlgeordnet sind oder auf speziellen Lagen liegen.

Analyse der Verschiebungsparameter

Gezeigt werden Balkengraphiken der gemittelten B -Werte ($B = 8\pi^2U$) pro Aminosäurerest, unterteilt nach Haupt- und Seitenkette. In analoger Aufteilung wird zudem die gemittelte Anisotropie (minimaler Eigenwert des Temperaturfaktors, geteilt durch maximalen Eigenwert) jedes Rests dargestellt.

3.3 Programmanpassungen

Um Strukturen der in dieser Arbeit behandelten Größenordnung handhaben zu können, waren einige Anpassungen der verwendeten Programme notwendig, insbesondere des Programmpackets XD (Koritsanszky u. a. 2003).

Eine der wichtigsten Änderungen für die in dieser Arbeit verwendeten lokal modifizierten XD-Version betrifft die Parallelisierung des *least square*-Kerns von XD, vorgenommen von Edgar Weckert. Diese Version wurde freundlicherweise von Birger Dittrich zur Verfügung gestellt, der auch die Berechnung eines R_{free} -Faktors, des Flack-Parameters anstelle des „Overall Temperature Factors“ sowie eine Subroutine zur Ausgabe von $h_i, k_i, l_i, F_{i,o}^2, F_{i,c}^2, \phi_i$ implementierte.

Zentral verwaltet und somit leicht zu ändern, waren die Dimensionierungen der Variablen für die Anzahl der Atome (jetzt 3000), die Anzahl zu verfeinernder Variablen (9999) und die Anzahl an κ -Parametern (99). Viele andere Variablen verteilen sich dezentral über die etwa 67000 Zeilen Programmcode: Die zulässige Dummy-Anzahl wurde auf 299 erhöht, die der „reset bond“-Anweisungen auf 2000. Da das Namensschema auf 10 Stellen erweitert wurde (siehe Abschn. 4.1.2 auf Seite 53), waren entsprechende Deklarationen der Länge der zugehörigen Variablen vorzunehmen, was etliche geänderte Formatanweisungen nach sich zog

und außerdem eine Anpassung der einzulesender Stellen nach Anweisungen in der zentralen Steuerdatei `xd.mas` erforderte. Auch die Formate der Dateien `xd.res` und `xd.inp` mußten auf die längeren Atomnamen eingestellt werden, aber auch auf die vierstellige Anzahl von Atomen, wofür der Platz eines nicht verwendeten Parameters für die lokale Symmetrie und eine Nachkommastelle der Atomkoordinaten gekürzt wurde. Die Routine zum initialen Import der Atomparameter wurde so geändert, daß die „ATOM“-Liste der `shelxl-1st`-Datei direkt verwendet werden kann. Zudem wurde die Sortierung nach Elementtypen unterbunden, da die Ordnung nach Aminosäureresten praktikabler ist. Des weiteren ist die Maximalanzahl von Zeilen der `xd.mas`-Steuerdatei zu erhöhen (9999),

Es wurde eine Routine zur Ausgabe von ρ - oder ϕ -Werten in atomaren Einheiten auf einem Gitter im `gaussian-cube`-Format geschrieben. Außerdem eine Prozedur zur Ausgabe der Atomkoordinaten im `pdb`-Format, allerdings mußte das Aminosäurerestkürzel durch eine Nummerierung ersetzt werden, da diese Information nicht in `XD` übernommen wurde. Die vorhandenen Subroutinen zur Ausgabe der Strukturaktoren wurden adaptiert, um eine Datei im `shelxl-fcf`-Format auszugeben.

3.4 Datenaufbereitung und -darstellung

Restelektronendichte

Etliche Programme des CCP4-Projekts (Collaborative Computational Project 1994) wurden benutzt, um die Restdichte zu bestimmen und zur Darstellung aufzubereiten. Im Anschluß an die `XD`-Verfeinerung wurden die Informationen zu jedem Reflex ($hkl, F_o, \sigma(F_o), F_c, \phi_c$) rausgeschrieben und mittels `f2mtz` in das `mtz`-Binärformat umgewandelt, welches neben den Indices, Strukturaktoren, Fehlern und Phasen auch die Information über die Zellmetrik und die Raumgruppe enthält, sofern man diese an `f2mtz` übergibt.

Die Fouriertransformation wurde mit dem Programm `fft` über das Volumen der asymmetrischen Einheit und alle Reflexe durchgeführt. Der Typ der Summation, hier eine Differenzfouriertransformation, wird durch Angabe der passenden Koeffizienten F_o und F_c festgelegt. Mögliche Skalierungen und Gewichtungen wurden auf den Standardwerten belassen. Ebenso wurde die Standardeinstellung der Feinheit des zu berechnenden Gitters von einem Drittel der maximalen Datenauflösung beibehalten. Eine Umwandlung dieser Datei ins `ascii`-Format wurde mit `mapdump` durchgeführt. Nach weiteren Formatierungen konnten dann mit `R` die verschiedenen statistischen Größen und die Werte für die Histogramme gewonnen werden.

Das Gitter über die asymmetrische Einheit wurde mittels `extend` und entsprechenden Angaben zu fraktionellen Koordinaten auf einen Raumbereich, der die vorliegenden Koordinaten des jeweiligen Moleküls großzügig einschließt, ausgedehnt

und zur graphischen Darstellung von Isooberflächen verwendet.

Um Differenzgitter der Restelektronendichte zu gewinnen, wurden die entsprechenden Dateien in `mapman` (Kleywegt u. Jones 2008) eingelesen und zueinander punktweise addiert, nachdem der Subtrahend mit $-1,0$ multipliziert wurde. Das resultierende Gitter wurde wieder im `ccp4`-Format rausgeschrieben.

Auch die Bestimmung von Restdichtewerten an bestimmten Koordinaten des Restdichtegitters wurde mittels `mapman` durchgeführt: Einlesen eines Dichtegitters und von Koordinaten im `pdb`-Format ermöglichen die Interpolation von Werten an den Atompositionen. So wurden z. B. der Mittelpunkt M_i der Bindungen zweier Atome A und B durch die Vektoraddition

$$M_i = \left(\frac{x_{A_i} + x_{B_i}}{2}, \frac{y_{A_i} + y_{B_i}}{2}, \frac{z_{A_i} + z_{B_i}}{2} \right)$$

bestimmt und an den resultierenden Koordinaten die Werte der Restelektronendichte über eine Splineinterpolation der umgebenen 4^3 Gitterpunkte angenähert.

Sevcik-Plot

Um zu untersuchen, ob die Messung die Unterscheidung verschiedener Elemente tatsächlich zulässt oder ob nur chemisches *a priori*-Wissen zu einer Zuordnung eines bestimmten Elements zu einer Atomposition führt, trägt man die Elektronendichte an den Atompositionen über den jeweiligen Temperaturfaktor auf und sucht nach Werteballungen, die einzelnen Elementtypen zugeordnet werden können. Erwartet werden mehrere mit dem Temperaturfaktor abnehmende Kurven der Elektronendichte, die für die verschiedenen Elemente unterscheidbar sind und in ihrer Höhe proportional zur Elektronenzahl jedes Atomtypes sind (Sevcik u. a. 1996).

Die zu untersuchende Elektronendichte wird aus der Fouriertransformation der gemessenen Strukturaktoren F_o und aus dem Modell berechneter Phasen ϕ_c gewonnen. Diese wird auf einem dreidimensionalen Gitter bestimmter Maschenweite erhalten, so daß nicht notwendigerweise ein Atomort mit einem Gitterpunkt zusammenfällt, weshalb die Dichte am vorgegebenen Ort aus einer Interpolation der umgebenden Volumenelemente erhalten wird. Die praktische Umsetzung geschieht durch eine Konvertierung der von `XD` ausgegebenen Strukturaktoren und Phasen mit `f2mtz` in das `mtz`-Format, welches von `fft` gelesen wird, um die Fouriertransformation auf einem Gitter durchzuführen. Dieses kann dann in `mapman` eingelesen werden, um an den durch eine entsprechende `pdb`-Datei vorgegebenen Koordinaten die Werte der Elektronendichte über eine Splineinterpolation der umgebenen 4^3 Gitterpunkte zu ermitteln.

Die Absolutwerte der Elektronendichte sind dabei irrelevant, da sie unter anderem von der Skalierung der F_o , der Gitterfeinheit und der Interpolationsmethode abhängig sind. Entscheidend ist nur die Unterscheidbarkeit der Werteballungen untereinander und ihre Proportionalität zueinander. Beispielsweise mag ein Sauerstoffatom mit

einem kleinen Verschiebungsparameter einen Dichtewert von $20 \text{ e}\text{\AA}^{-3}$ aufweisen, ein zweites mit doppeltem Temperaturfaktor nur noch von $12 \text{ e}\text{\AA}^{-3}$. Enthält die Beispielstruktur auch ein Schwefelatom mit gleichem Temperaturfaktor wie das erste Sauerstoffatom, so sollte es einen Wert der Elektronendichte von $40 \text{ e}\text{\AA}^{-3}$ aufweisen, also dem der Ordnungszahl entsprechendem Doppelten.

Molekülvolumen

Es gibt sehr viele Möglichkeiten Moleküloberflächen und -volumina zu berechnen (Connolly 1996), in dieser Arbeit wurde die Information über das van-der-Waals-Volumen der Moleküle mit den Programmen des **3v**-Packets (Voss 2009) gewonnen. Dieses Programm berechnet solvenszugängliche (*solvent-accessible*) Oberflächen und Volumina von Molekülen nach der „*rolling probe*“-Methode von Richards (1977). Setzt man das Volumen der Testkugel auf Null, ergibt sich das Molekülvolumen bzw. die dazugehörige Oberfläche. Jedes Atom wird als harte Kugel mit bestimmtem Radius beschrieben, die verwendeten van-der-Waals-Radien sind die im **msms**-Packet (Sanner u. a. 1996) verwendeten. Aus einer **pdb**-Datei und den Standardradien wurde mit **pdb_to_xyzr** eine Datei mit Koordinaten und Radien erstellt. Die Feinheit des Gitters, auf dem die Oberfläche bestimmt wurde betrug $0,1 \text{ \AA}$, was 1000 Voxeln pro \AA^3 entspricht. Das resultierende **ezd**-Gitter wurde mit **mapman** ins **ccp4**-Format konvertiert, zwecks späterer Visualisierung.

ab initio- und semiempirische Rechnungen

Sämtliche Berechnungen mit *ab initio*-Methoden wurden mit dem **gaussian**-Programmsystem durchgeführt (Frisch u. a. 2002, 2004). Ebenso erfolgten damit semiempirische AM1-Rechnungen (Dewar u. a. 1985), welche mit **mopac** (Stewart 2008) wiederholt wurden. Ob der Atomzahl der Atome des Lysozyms, welches auch mit AM1 gerechnet wurde, mußte **formchk** an die Atomzahl angepaßt werden (Parameter **MDV** von 3000000 auf 25000000 erhöhen). Da **cubegen** entgegen der Dokumentation den Kommandozeilenparameter **memory** nicht auswertet, muß die Umgebungsvariable **\$GAUSS_MEMDEF** auf einen entsprechend hohen Wert gesetzt werden. Wählt man den gleichen Gitterursprung für die **cube**-Gitter aus **cubegen** und **xd**, können diese mit **cubman** voneinander abgezogen werden, um Differenzgitter zu erzeugen.

Molekülüberlagerung

Alle Molekülüberlagerungen und Berechnungen der quadratisch gemittelten Abweichung der Koordinaten der Atome oder entsprechender Mittelwerte für Aminosäurereste und die Gesamtstrukturen, wurden mit **ProFit** (Martin 2009) durchgeführt.

Chimera

Alle Abbildungen der Moleküle und Isooberflächen wurden mit dem Graphikprogramm Chimera (UCSF Computer Graphics Laboratory 2009) erstellt.

Gnuplot und TikZ

Alle Darstellungen von Funktionen und Daten erfolgten mittels **gnuplot** (Williams u. a. 2009) und **TikZ** (Tantau u. Wibrow 2008). Die Anpassung von Funktionen an Datenwerte erfolgte ebenfalls mit **gnuplot**.

Punktladungspotentiale

Zur Bestimmung elektrostatischer Potentiale ϕ_{PC} aus Punktladungen wurde das Programm **delphi** (Rocchia u. a. 2002) benutzt, welches die Poisson-Boltzmann-Gleichung:

$$\vec{\nabla} \left[\varepsilon(\vec{r}) \vec{\nabla} \phi(\vec{r}) \right] = -4\pi\rho(\vec{r}) + 4\pi \sum_i c_i^\infty z_i q \exp\left(\frac{-z_i q \phi(\vec{r})}{kT}\right) \lambda(\vec{r}) \quad (3.12)$$

numerisch nach der Finite-Differenzen-Methode löst (für zwei Review-Artikel siehe Fogolari u. a. 2002; Sharp u. Honig 1990). Wobei hier $\vec{\nabla}$ der Nabla-Operator ist und $\varepsilon(\vec{r})$ die ortsabhängige Dielektrizitätskonstante, $\rho(\vec{r})$ die Ladungsverteilung beschreibt, c_i^∞ die Konzentration des i -ten Ions der Ladung $z_i q$ (q ist die Protonladung) in unendlicher Entfernung vom Molekül bedeutet, $\lambda(\vec{r})$ die Zugänglichkeit eines Ortes für Ionen beschreibt und k und T die Boltzmannkonstante und die Temperatur sind.

Für kleine Potentiale $\phi(\vec{r})$ kann sie linearisiert werden:

$$\vec{\nabla} \left[\varepsilon(\vec{r}) \vec{\nabla} \phi(\vec{r}) \right] = -4\pi\rho(\vec{r}) + 4\pi \frac{\sum_i c_i^\infty z_i^2 q^2 \phi(\vec{r})}{kT} \lambda(\vec{r}) \quad (3.13)$$

$$= -4\pi\rho(\vec{r}) + \kappa^2(\vec{r}) \varepsilon(\vec{r}) \phi(\vec{r}) \lambda(\vec{r}) \quad (3.14)$$

Der zweite Term der rechten Seite berücksichtigt die Salzeffekte ionischer Lösungen ($\kappa^2(\vec{r})$ ist die inverse Debyelänge). In Abwesenheit mobiler Ionen verschwindet er und die Formel reduziert sich zur Poissongleichung, welche sich wiederum auf die Coulombgleichung zurückführen ließe, wenn die Dielektrizitätskonstante keine Funktion des Ortes wäre.

Zur numerischen Lösung der partiellen Differentialgleichung mit der Finite-Differenzen-Methode, wird ein das Molekül einschließendes Volumen in Gitterpunkte aufgeteilt, an denen das Potential eines kubischen Volumenelements 0 mit der Ladung q_0 und den Kantenlängen h nach

$$\phi_0 = \frac{\frac{4\pi q_0}{h} + \sum_{i=1}^6 \varepsilon_i \phi_i}{\sum_{i=1}^6 \varepsilon_i + \kappa^2 \varepsilon_0 \lambda_0 h^2} \quad (3.15)$$

punktweise gelöst wird. Wegen der Abhängigkeit von den umgebenden Potentialen ϕ_i , muß die Berechnung iterativ bis zur Konvergenz erfolgen. Die Ladung q jedes Voxels bestimmt sich aus der linearen, abstandsgewichteten Interpolation der an den Atomkoordinaten vorliegenden Punktladungen auf die jeweils nächsten acht Gitterpunkte.

Die Koordinaten der jeweiligen Hauptpopulation des Moleküls werden im `pdb`-Format übergeben. Zudem müssen jedem Atom Ladungen zugewiesen werden, dazu wurden die des `amber03`-Kraftfeldes (Duan u. a. 2003; Cornell u. a. 1995) verwendet, welches für jedes Atom der Standardamino­säuren differenzierte Ladungen bereitstellt. Für die Aminosäure AIB wurden entsprechende Parameter bei Dupradeau u. a. (2008) gefunden, für VOL wurden die Parameter aus denen mehrerer Standardfragmente improvisiert. Die Gesamtladung der Moleküle entsprach jeweils den formalen Ladungen durch die vorliegende Protonierung, die maximale Abweichung betrug $0,02e$. Als Atomradien wurden die in `3v` benutzten verwendet.

Die Rechnungsparameter wurden einfach gehalten, um ein isoliertes Molekül in der Gasphase zu beschreiben, was die höchste Übereinstimmung mit den Multipol-ESPs ergibt: So wurde die Dielektrizitätskonstante der Molekü­lumgebung auf den Vakuumwert gesetzt (`exdi=1.0`), auf die Möglichkeit eine Solvensumgebung zu simulieren wurde also verzichtet. Für das Molekül selbst wurde der Dielektrizitätsparameter `indi=2.0` gewählt, der einem durchschnittlichen organischen Molekül entspricht, mit höheren Werten könnte man den Effekt der Reorientierung ganzer geladener Gruppen simulieren. Als Randbedingung des Gitters wurde ein Potential von Null (`bndcon=1`) angesetzt. Die Gitterfeinheit betrug $0,3 \text{ \AA}$. Schließlich wurde das resultierende Gitter der Potentialwerte in kT/e ($1 e/a_0 = 1059,59 \text{ kT}/e$) zur Colorierung der verschiedenen Moleküloberflächen verwendet.

4 Verfeinerung

4.1 Allgemeines

Die Verfeinerung der fünf Strukturen unterteilt sich in eine klassische sphärische Verfeinerung mit dem Programm `shelx1` und die Verfeinerungen im Programmsystem `XD`. Erstere dient dem Erhalt möglichst guter Startstrukturen und unterscheidet sich in etlichen Punkten für die verschiedenen Moleküle. Letztere gestalten sich weitestgehend gleichartig für die unterschiedlichen Verbindungen, um eine maximale Vergleichbarkeit der Ergebnisse zu gewährleisten. Die später zu besprechenden Ergebnisse beziehen sich ausschließlich auf die daraus resultierenden sphärischen und asphärischen Modelle *iam*, *iam_{rb}* und *mul*. Die allgemeine Vorgehensweise soll durch folgendes Schema 4.1.1 verdeutlicht werden.

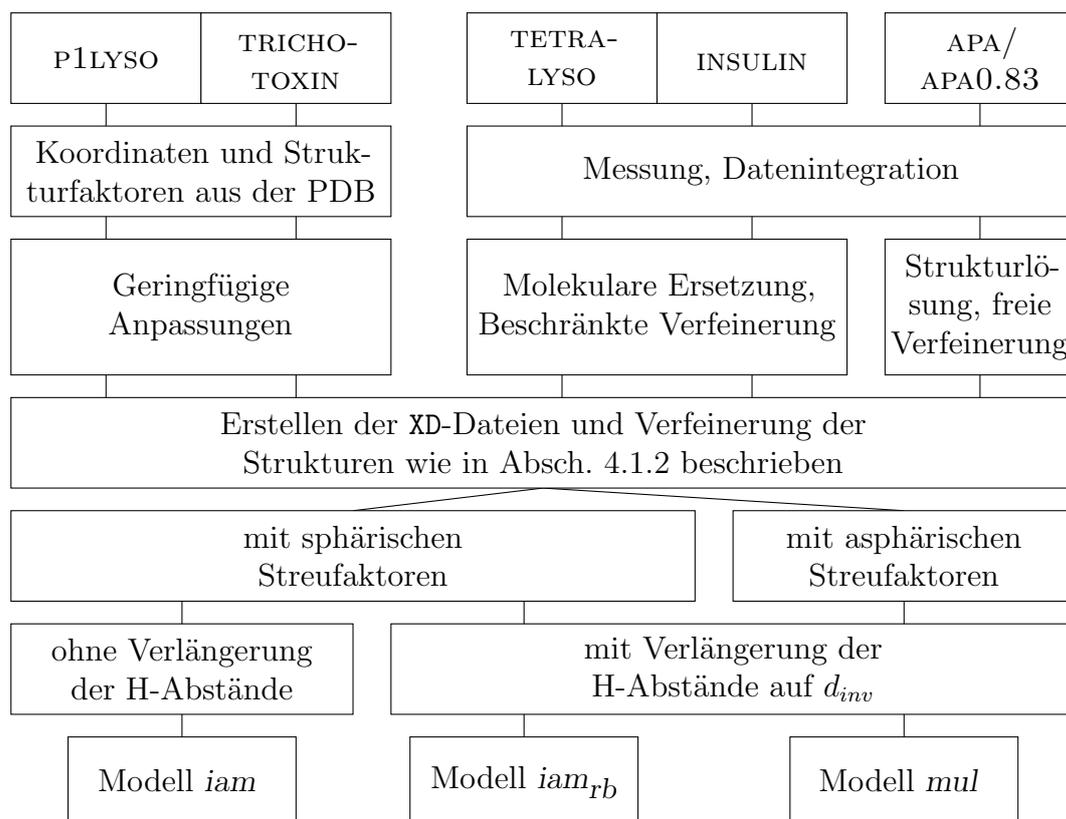


Abbildung 4.1.1: Überblick zur Verfeinerung der Strukturen.

Bezüglich des allgemeinen Vorgehens zur `shelxl`-Verfeinerung muß zwischen der Kleinmolekülstruktur APA einerseits und den Makromolekülen andererseits unterschieden werden, die sich wiederum in für diese Arbeit gewonnene und bereits literaturbekannte Strukturen unterteilen lassen.

Die Messung und Verfeinerung von P1LYSO wurde durch Wang u. a. (2007) beschrieben, die der TRICHOTOXIN-Struktur von Chugh u. a. (2002). Zugehörige Koordinaten und Strukturparameter wurden der PDB unter den Codes 2VB1 und 1M24 entnommen und weitestgehend unverändert, wie von den jeweiligen Autoren modelliert, verwendet. In den folgenden Abschnitten wird deshalb nur eine kurze Beschreibung der Strukturen und der vorgenommenen Änderungen gegeben, ergänzt um eine Wiederholung einiger wichtiger Resultate der Datenaufbereitung und -verfeinerung zu Vergleichszwecken. Anders verhält es sich mit den Strukturen TETRALYSO und INSULIN, deren Datenintegration und schrittweise sphärische Verfeinerung eingehender beschrieben wird. Die Ausführungen im folgenden Abschnitt 4.1.1 beziehen sich hauptsächlich auf diese beiden Makromoleküle, wiewohl sie sinngemäß auch für P1LYSO und TRICHOTOXIN gelten. Alle Besonderheiten und Einschränkungen der Verfeinerung von Makromolekülen sind für die Struktur APA nicht zutreffend, deren Messung und Verfeinerung in Kalinowski u. a. (2007) beschrieben ist. Insbesondere wurde die Struktur mit direkten Methoden gelöst, alle Nichtwasserstoff- und Wasserstoffatome konnten lokalisiert werden und wurden ohne Einschränkungen frei verfeinert. Ein sechstes Modell, APA0.83, resultiert schlicht aus der eingeschränkten Verwendung der Reflexe bis zu einer Auflösung von $d = 0,83 \text{ \AA}$.

Für alle Strukturen aber wurde der Übergang zum XD-Programmsystem auf die gleiche, in Absch. 4.1.2 beschriebene Weise vorgenommen. Dies betrifft die Erzeugung der Reflexdatei, die Zuordnung der Streufaktoren und Koordinatensysteme, die Übernahme des `shelxl`-Gewichtungsschemas, die Anzahl der Verfeinerungszyklen gegen F^2 , das σ -Kriterium und die Behandlung der Wasserstoffatome, die auch im Fall APA nicht frei verfeinert wurden. Nach der für jede Struktur unterschiedlichen Auswahl an freizugebenen Nichtwasserstoffatomen, wurden jeweils drei Verfeinerungen gerechnet, aus denen die später ausschließlich zu besprechenden Modelle *mul*, *iam* und *iam_rb* resultieren. Die jeweiligen Eingabedateien unterschieden sich nur durch die verwendeten Schlüsselwörter zur Auswahl der Streufaktoren und der Wasserstoffatomabstände. Mit der Wiederholung der sphärischen Verfeinerung auf diese Art (*iam*) wird sichergestellt, daß die zu diskutierenden Effekte der asphärischen Verfeinerung (*mul*) nur auf die neuen Streufaktoren und die dazu passenden Wasserstoffabstände zurückzuführen sind. Es kann also ein optimal an die sphärische Beschreibung angepaßtes Modell mit dem asphärischen Pendant verglichen werden, wobei technisch bedingte Einschränkungen beide Modelle gleichermaßen betreffen. Mit der Verfeinerung *iam_rb* schließlich, steht ein Modell zum Vergleich, daß zwar sphärisch verfeinert wurde, allerdings unter Verwendung der realistischeren längeren Wasserstoffabstände d_{inv} .

4.1.1 Sphärische Verfeinerung

Die Verfeinerung von Makromolekülen unterscheidet sich in einigen Punkten von der eines kleinen Moleküls und soll im folgenden kurz umrissen werden.

Nach Datenintegration, -skalierung und -mittelung, sowie Bestimmung der Zellkonstanten mit `xds` (Kabsch 1993), kann das Phasenproblem zur Strukturbestimmung durch das Konzept des Molekularen Ersatzes („*molecular replacement*“) gelöst werden. Dazu wird eine ähnliche Struktur auf die Proteinhauptkette reduziert und das Fragment mit `epmr` (Kissinger 2008) an die Beobachtungen angepaßt.

Ausgehend von diesem Startmodell wird die weitere Verfeinerung mit `shelxl` (Sheldrick 2008, 1997) durchgeführt. Hierbei ist es hilfreich, erst gegen die Daten geringerer Auflösung zu verfeinern und langsam zum vollen Datenumfang voranzuschreiten (STIR-Anweisung in `shelxl`). Ebenso geht man sukzessive von einer isotropen zu einer anisotropen Beschreibung der Nichtwasserstoffatome über. Die Struktur wird dabei schrittweise um fehlende Aminosäurereste, Seitenketten, Alternativkonformationen und Solvensmoleküle erweitert. Diese Modellierung wurde mit dem Programm `coot` (Emsley u. Cowtan 2004) durchgeführt und überwacht. Es ermöglicht unter anderem das Einpassen von Atomen in eine graphisch dargestellte Karte der Elektronendichte bzw. Restdichte („*real space refinement*“). Hinzugefügte und bereits vorliegende Strukturteile können darüber hinaus an Erwartungswerte angepasst, regularisiert werden. Etliche Qualitätsmerkmale, wie Abweichung von idealisierter Stereochemie, Kollisionen von Strukturteilen miteinander oder mit Solvensmolekülen, Ausmaß beobachteter Dichte an den Atompositionen, ungewöhnliche Restdichteanhäufungen oder Temperaturfaktorvarianzen, können für jeden Aminosäurerest angezeigt und korrigiert werden.

Zur Handhabung der schieren Größe und des bei Makromolekülen ausgeprägten Fehlordnungsproblems bietet `shelxl` bequeme Techniken: Jeder Aminosäurerest kann über ein Buchstaben- oder Zahlenkürzel (RESI) angesprochen werden, die Benennung von Atomen braucht nur innerhalb eines solchen Restes eindeutig zu sein. Fehlordnungen können vorteilhaft verfeinert werden, indem die Konformationen in PART-Anweisungen eingeschlossen werden, wodurch auch die korrekte Konnektivität sichergestellt wird. Die Besetzung läßt sich mittels freier Variablen (FVAR) verfeinern, deren Summen auf einen Zielwert beschränkt werden können, um z. B. für zwei Konformationen eine Gesamtbesetzung von eins sicherzustellen.

Um den Streubetrag diffus verteilter, undefinierter Solvensmoleküle zu berücksichtigen, können die Intensitäten nach $F_{c,neu}^2 = F_{c,alt}^2(1 - g \exp(-8\pi^2 U(\sin \Theta/\lambda)^2))$ korrigiert werden, wodurch der Beitrag beschriebener Strukturteile verringert wird, was dem Hinzufügen eines Streubeitrags in den leeren Regionen entspricht. Übliche Werte der Parameter g und U sind 0,7–1 und 2–5.

Eine weitere Besonderheit betrifft den Rechenaufwand für eine *full matrix least square*-Verfeinerung, welche proportional zum Produkt aus Anzahl der Reflexe und dem Quadrat der Parameteranzahl ist. Als Alternative steht mit der wesentlich

schnelleren und ressourcenschonenderen *conjugate gradient least square*-Methode eine Variante zur Verfügung, mit der die gesamte Verfeinerung durchgeführt wird.

Nicht immer erlaubt die Datenqualität oder das Reflex-zu-Parameter-Verhältnis eine freie Verfeinerung aller Parameter einer Struktur zu physikalisch sinnvollen Werten. In diesen Fällen müssen sie in der einen oder anderen Form beschränkt verfeinert werden. Bei Verfeinerung eines oder mehrerer Parameter eines Atoms können für diese Zielgrößen und erlaubte Abweichungen davon vorgegeben werden. Solcherart mit einer variablen Beschränkung verfeinerte Parameter werden „restraint“ verfeinert genannt. Dabei verringert sich die Parameterzahl nicht. Die Abweichung von der Zielgröße fließt in die zu minimierende Summe der Fehlerquadrate der Verfeinerung ein. Wird für den oder die Parameter eines Atoms ein fixer Wert, z. B. die Parameter eines anderen Atoms, vorgegeben, ist dies eine „constrainte“ Verfeinerung. Hierbei wird auch die Parameterzahl verringert.

Gerade bei Proteinen sind stereochemische Restraints zumindest für Teile der Struktur notwendig, um eine unverzerrte Geometrie zu erhalten. Die Zielgrößen für Bindungslängen und -winkel können z. B. die nach Engh u. Huber (1991) sein. Entsprechende Anweisungen (DFIX, DANG, CHIV, FLAT) sind in `shelxl` implementiert, um 1,2- und 1,3-Abstände zu einem bestimmten Wert zu bewegen oder die Planarität flacher Gruppen zu gewährleisten. Darüberhinaus stehen Restraints zur Gestaltung der Verschiebungsparameter zur Verfügung (DELU, SIMU, ISOR), die die Erfüllung des *rigid bond*-Postulats, die nur graduellen Veränderung benachbarter Temperaturfaktoren oder die Annäherung der Ellipsoide an eine Kugelform umzusetzen helfen. Eine weitere wichtige Abstandsbeschränkung ist der *anti-bumping*-Restraint, um eine zu starke Annäherung von Wassermolekülen an die Proteinatome oder untereinander zu verhindern, gleichzeitig aber typische Wasserstoffbrückenbindungsabstände zu ermöglichen.

Um das Modell um Wasserstoffatome zu ergänzen, dient der AFIX-Constraint, da ihre Lagen bei Makromolekülen entweder erst gar nicht bestimmt, zumindest aber nicht zuverlässig verfeinert werden können. Stattdessen werden einer idealisierten Geometrie relativ zum Schweratom entsprechende Koordinaten berechnet. Bei Wahl eines *riding models* folgt diese Anordnung veränderten Schweratomkoordinaten als starre Gruppe – die Idealgeometrie wird in jedem Verfeinerungszyklus wieder hergestellt. Je nach Wahl des Constraintmodells ist auch die automatische Anpassung der Torsionswinkel an eine Maximierung der potentiellen Wasserstoffbrückenbindungen oder der Elektronendichte an den Atompositionen möglich. Auch lassen sich die isotropen Wasserstoffatomtemperaturfaktoren auf z. B. das 1,2- oder 1,5-fache des äquivalenten isotropen Temperaturfaktors des sie bindenden Schweratoms constrainen.

Ein im Zusammenhang mit der Verwendung der obengenannten Verfeinerungsbeschränkungen übliches Vorgehen ist die Berechnung des sogenannten R_{free} aus einer ca. 5% aller Reflexe umfassenden zufälligen Auswahl der Beobachtungen, die zu keinem Zeitpunkt in der Verfeinerung des Strukturmodells benutzt wurden.

Durch den Vergleich mit dem $R1$, berechnet aus den verbleibenden Reflexen, an die das Modell angepaßt wurde, erhält man einen Hinweis auf die Signifikanz der eingeführten Parameter: Sofern der R_{free} gemeinsam mit dem $R1$ sinkt, handelt es sich wahrscheinlich um eine reale Verbesserung des Modells und nicht nur um „ R -Wertkosmetik“.

In einem finalen Schritt wird die Modellverfeinerung mit *full matrix least square*-Methoden für alle Parameter zugleich durchgeführt. Die übliche Vorgehensweise bei diesen Strukturgrößen wäre hingegen ein *full matrix*-Zyklus blockweiser Verfeinerung nur der Koordinaten bei gleichzeitiger Unterbindung jeglicher Parameterverschiebung (BLOC 1, DAMP 0 0), um die Standardabweichung der geometrischen Parameter zu bestimmen. Um sich aber den Bedingungen der nachfolgenden XD-Verfeinerung anzupassen, hat es sich bewährt, bereits hier *full matrix* zu verfeinern.

Letztendlich werden die beobachteten und berechneten Strukturfaktoren in eine **fcf**-Reflexdatei geschrieben (LIST 6, MERG 4), welche die skalierten und korrigierten unabhängigen Reflexe, abzüglich eventueller symmetrieverbotener, enthält. Die Korrekturen umfassen Dispersion, Extinktion und den Beitrag des diffusen Solvensmodells, falls diese verfeinert wurden. Sofern noch nicht geschehen, werden in diesem Schritt Friedelpaare gemittelt. Auch wurden die nachfolgenden Reflexstatistiken der einzelnen Verbindungen aus ebendieser Datei erstellt, womit alle genannten Korrekturen in diesen enthalten sind.

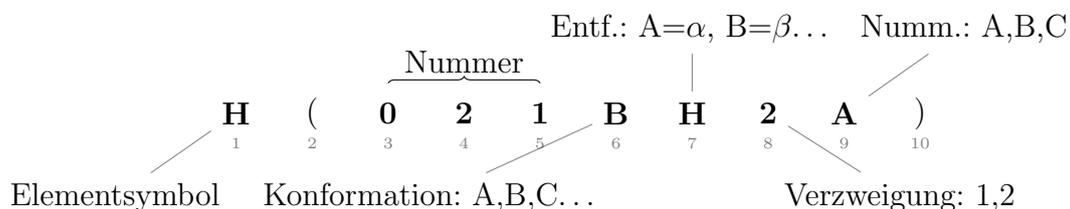
4.1.2 Asphärische Verfeinerung

Ebenso wird die XD-Reflexdatei **xd.hkl** aus der **fcf**-Datei durch entsprechendes Formatieren erhalten, so daß in **shelxl** enthaltene Korrekturen in die Verfeinerung mit XD übernommen werden können. Dies ist insbesondere für die Korrektur des Streuanteils diffuser Solvenssphäre bedeutsam, da diese nicht als Parameter in XD vorgesehen ist. Auch enthalten die Strukturfaktoren nun bereits Korrekturen für anomale Dispersion und Extinktion, so daß diese Parameter in XD nicht mehr verfeinert werden brauchen.

Die Eingabedatei **xd_ini.inp** enthält Informationen zur Zelle, Gittertyp, Wellenlänge, Symmetrioperationen und den Arbeitstitel, sowie die Atomnamen mit den zugehörigen Koordinaten, Verschiebungsparametern und Besetzungszahlen. Mittels **xdini** werden daraus die initialen Steuerungsdateien **xd.inp** und **xd.mas** erstellt. Hierfür und die folgenden Verfeinerungen fanden lokal modifizierte Programme Verwendung (siehe Absch. 3.3).

Zur eindeutigen Identifizierung jedes Atoms wurde eine neue, am PDB-Format orientierte, maximal zehnstellige Bezeichnung eingeführt. Sie besteht aus dem Elementsymbol, dem in runden Klammern zuerst eine dreistellige Zahl zur Identifizierung des Aminosäurerestes folgt, dann ein Buchstabe zur Kenntlichmachung der Konformation, jeweils ein Entfernungs- und Verzweigungsindikator, sowie eine

letzte Stelle für eventuelle Nummerierungen. Bei folgendem Beispiel handelt es sich also um das erste (A) Wasserstoffatom (H) am zweiten (2), sechs Bindungen ($H=\eta$) vom C_α entfernten Atom des minderpopulierten (B) Teils des 21. Aminosäurerestes (021):



Aus einer zuvor festgelegten Zuordnung zwischen jedem Atom einer jeden Aminosäure und dem dazugehörigen Invariomnamen, wird dann eine Datei `Invariome.in` erstellt. Diese wird zusammen mit den ursprünglichen `inp`- und `mas`-Dateien von einer angepaßten Version des Programms `invariomtool` (Hübschle u. a. 2007) gelesen, um zu jedem Atom den korrespondierenden Datenbankeintrag, multipliziert mit der jeweiligen Besetzung, in eine neue XD-Eingabedatei `xd.inp` zu schreiben. Die verwendete Zuordnung Atomtyp–Invariom ist in Tab. 4.1.1 auf der gegenüberliegenden Seite aufgeführt, ebenso der in XD verwendete Atomname.

Da die Summe der Invariomonopole nicht zwingend zu einer elektroneutralen Struktur führt, werden die Datenbankmultipole automatisch um einen entsprechenden Betrag korrigiert. Hier ist meist eine manuelle zweite Verbesserung angezeigt, um tatsächlich Elektroneutralität zu erreichen.

In Tab. 4.1.1 ist das zu jedem Invariom benötigte Koordinatensystem gelistet, welches zur Symmetrie und der Achsenorientierung bei Berechnung des Invarioms passen muß. Die verallgemeinerten Koordinatensysteme wurden im Rahmen dieser Arbeit bestimmt, da eine automatische Zuweisung momentan nicht möglich ist. Anzugeben sind zwei benachbarte Atome und die Bezeichnung der zu ihnen weisen Achsen ($AX1 \rightarrow AT1$ und $AX2 \rightarrow AT2$), die dritte Achse steht senkrecht auf den ersten beiden in die durch die Händigkeit (H) festgelegte Richtung. Manche Invariome verlangen eine Achsenlegung zu Platzhalteratomen (Dummy), um $mm2$ - oder m -Symmetrie zu erreichen, deren Koordinaten liegen auf der Winkelhalbierenden zu den benachbarten Nichtwassertoffatomen. Die verallgemeinerten Koordinatensysteme pro Aminosäurerest werden in die XD-Steuerdatei `xd.mas` übertragen und die Platzhalter „nnn“ und „X“ in die aktuelle Sequenznummerierung und das korrekte Konformationssymbol umgeschrieben. Besondere Aufmerksamkeit wurde der richtigen Konnektivität der fehlgeordneten Atome gewidmet.

Als Gewichtungsschema wird das in `shelxl` verfeinerte gewählt, ebenso wie das Beobachtungskriterium $F_o^2 > 2\sigma(F_o^2)$ übernommen wird. Falls notwendig, müssen in XD, im Unterschied zum Automatismus in `shelxl`, Constraints für Koordinaten und Verschiebungsparameter aufgrund spezieller Lagen und zur Nullpunktsfestlegung in polaren Raumgruppen explizit angegeben werden.

Tabelle 4.1.1: Invariome, Symmetrie und Koordinatensystem der 20 kanonischen Aminosäuren.

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
N-Terminus						
N	N1c1h1h1h+	3	N(nnnX)	Z→C(nnnXA)	Y→H(nnnX0A)	R
H0A	H1n[1c1h1h]+	6	H(nnnX0A)	Z→N(nnnX)	Y→C(nnnXA)	R
H0B	H1n[1c1h1h]+	6	H(nnnX0B)	Z→N(nnnX)	Y→C(nnnXA)	R
H0C	H1n[1c1h1h]+	6	H(nnnX0C)	Z→N(nnnX)	Y→C(nnnXA)	R
C-Terminus						
C	C1.5o1.5o1c-	m	C(nnnX)	X→O(nnnXT1)	Y→O(nnnXT2)	R
OT1	O1.5c[1.5o1c]	m	O(nnnXT1)	X→C(nnnX)	Y→O(nnnXT2)	R
OT2	O1.5c[1.5o1c]	m	O(nnnXT2)	X→C(nnnX)	Y→O(nnnXT1)	R
Hauptkette						
N	N1.5c[1.5o1c]1c1h	m	N(nnnX)	X→C(n-1X)	Y→C(nnnXA)	R
H0	H1n[1.5c1c]	6	H(nnnX0)	Z→N(nnnX)	Y→C(n-1X)	R
CA	C1n1c1c1h	m	C(nnnXA)	X→H(nnnXA)	Y→N(nnnX)	R
HA	H1c[1n1c1c]	6	H(nnnXA)	Z→C(nnnXA)	Y→N(nnnX)	R
C	C1.5o1.5n[1c1h]1c	m	C(nnnX)	X→N(n+1X)	Y→C(nnnXA)	R
O	O1.5c[1.5n1c]	mm2	O(nnnX)	Z→C(nnnX)	Y→N(n+1X)	R
<i>Hauptkette vor Prolin</i>						
C	C1.5o1.5n[1c1c]1c	m	C(nnnX)	X→N(n+1X)	Y→C(nnnXA)	R
<i>Prolin-Hauptkette</i>						
N	N1.5c[1.5o1c]1c1c	m	N(nnnX)	X→C(n-1X)	Y→C(nnnXA)	R
CA	C1n1c1c1h	m	C(nnnXA)	X→H(nnnXA)	Y→N(nnnX)	R
HA	H1c[1n1c1c]	6	H(nnnXA)	Z→C(nnnXA)	Y→N(nnnX)	R
C	C1.5o1.5n[1c1h]1c	m	C(nnnX)	X→N(n+1X)	Y→C(nnnXA)	R
O	O1.5c[1.5n1c]	mm2	O(nnnX)	Z→C(nnnX)	Y→N(n+1X)	R

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
Glycin						
N	N1.5c[1.5o1c]1c1h	m	N(nnnX)	X→C(n-1X)	Y→C(nnnXA)	R
H0	H1n[1.5c1c]	6	H(nnnX0)	Z→N(nnnX)	Y→C(n-1X)	R
CA	C1n1c1h1h	1	C(nnnXA)	Z→N(nnnX)	Y→C(nnnX)	R
HA1	H1c[1n1c1h]	6	H(nnnXA1)	Z→C(nnnXA)	Y→N(nnnX)	R
HA2	H1c[1n1c1h]	6	H(nnnXA2)	Z→C(nnnXA)	Y→N(nnnX)	R
C	C1.5o1.5n[1c1h]1c	m	C(nnnX)	X→N(n+1X)	Y→C(nnnXA)	R
O	O1.5c[1.5n1c]	mm2	O(nnnX)	Z→C(nnnX)	Y→N(n+1X)	R
Alanin						
CB	C1c1h1h1h	3	C(nnnXB)	Z→C(nnnXA)	Y→H(nnnXB1)	R
HB1	H1c[1c1h1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1h1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB3	H1c[1c1h1h]	6	H(nnnXB3)	Z→C(nnnXB)	Y→C(nnnXA)	R
Arginin						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG	C1c1c1h1h	mm2	C(nnnXG)	Z→DUM	Y→C(nnnXB)	R
HG1	H1c[1c1c1h]	6	H(nnnXG1)	Z→C(nnnXG)	Y→C(nnnXB)	R
HG2	H1c[1c1c1h]	6	H(nnnXG2)	Z→C(nnnXG)	Y→C(nnnXB)	R
CD	C1n1c1h1h	1	C(nnnXD)	Z→N(nnnXE)	Y→C(nnnXG)	R
HD1	H1c[1n1c1h]	6	H(nnnXD1)	Z→C(nnnXD)	Y→N(nnnXE)	R
HD2	H1c[1n1c1h]	6	H(nnnXD2)	Z→C(nnnXD)	Y→N(nnnXE)	R
NE	N1c1c1h	1	N(nnnXE)	X→C(nnnXD)	Y→C(nnnXZ)	R

Invariome, Symmetrie und Koordinatensystem (Forts.).

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
HE	H1n[1c1c]	6	H(nnnXE)	Z→N(nnnXE)	Y→C(nnnXD)	R
CZ	C1.5n[1c1h]1.5n[1h1h]1.5n[1h1h]+	mm2	C(nnnXZ)	Z→N(nnnXE)	Y→N(nnnXH1)	R
NH1	N1.5c[1.5n1.5n]1h1h+	mm2	N(nnnXH1)	Z→C(nnnXZ)	Y→H(nnnXH1A)	R
HH1A	H1n[1.5c1h]+	6	H(nnnXH1A)	Z→N(nnnXH1)	Y→C(nnnXZ)	R
HH1B	H1n[1.5c1h]+	6	H(nnnXH1B)	Z→N(nnnXH1)	Y→C(nnnXZ)	R
NH2	N1.5c[1.5n1.5n]1h1h+	mm2	N(nnnXH2)	Z→C(nnnXZ)	Y→H(nnnXH2A)	R
HH2A	H1n[1.5c1h]+	6	H(nnnXH2A)	Z→N(nnnXH2)	Y→C(nnnXZ)	R
HH2B	H1n[1.5c1h]+	6	H(nnnXH2B)	Z→N(nnnXH2)	Y→C(nnnXZ)	R
Asparagin						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG	C1.5o1.5n[1h1h]1c	m	C(nnnXG)	X→O(nnnXD1)	Y→N(nnnXD2)	R
OD1	O1.5c[1.5n1c]	mm2	O(nnnXD1)	Z→C(nnnXG)	Y→N(nnnXD2)	R
ND2	N1.5c[1.5o1c]1h1h	2	N(nnnXD2)	Z→C(nnnXG)	Y→H(nnnXD2A)	R
HD2A	H1n[1.5c1h]	6	H(nnnXD2A)	Z→N(nnnXD2)	Y→C(nnnXG)	R
HD2B	H1n[1.5c1h]	6	H(nnnXD2B)	Z→N(nnnXD2)	Y→C(nnnXG)	R
Aspartat						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG	C1.5o1.5o1c-	m	C(nnnXG)	X→O(nnnXD1)	Y→O(nnnXD2)	R
OD1	O1.5c[1.5o1c]	m	O(nnnXD1)	X→C(nnnXG)	Y→O(nnnXD2)	R
OD2	O1.5c[1.5o1c]	m	O(nnnXD2)	X→C(nnnXG)	Y→O(nnnXD1)	R

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
Cystin						
CB	C1s[1h]1c1h1h	m	C(nnnXB)	X→S(nnnXG)	Y→C(nnnXA)	R
HB1	H1c[1s1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→S(nnnXG)	R
HB2	H1c[1s1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→S(nnnXG)	R
SG	S1s1c	m	S(nnnXG)	X→S(n+nXG)	Y→C(nnnXB)	R
Glutamin						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG	C1c1c1h1h	mm2	C(nnnXG)	Z→DUM	Y→C(nnnXB)	R
HG1	H1c[1c1c1h]	6	H(nnnXG1)	Z→C(nnnXG)	Y→C(nnnXB)	R
HG2	H1c[1c1c1h]	6	H(nnnXG2)	Z→C(nnnXG)	Y→C(nnnXB)	R
CD	C1.5o1.5n[1h1h]1c	m	C(nnnXD)	X→O(nnnXE1)	Y→N(nnnXE2)	R
OE1	O1.5c[1.5n1c]	mm2	O(nnnXE1)	Z→C(nnnXD)	Y→N(nnnXE2)	R
NE2	N1.5c[1.5o1c]1h1h	2	N(nnnXE2)	Z→C(nnnXD)	Y→H(nnnXE2A)	R
HE2A	H1n[1.5c1h]	6	H(nnnXE2A)	Z→N(nnnXE2)	Y→C(nnnXD)	R
HE2B	H1n[1.5c1h]	6	H(nnnXE2B)	Z→N(nnnXE2)	Y→C(nnnXD)	R
Glutamat						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG	C1c1c1h1h	mm2	C(nnnXG)	Z→DUM	Y→C(nnnXB)	R
HG1	H1c[1c1c1h]	6	H(nnnXG1)	Z→C(nnnXG)	Y→C(nnnXB)	R
HG2	H1c[1c1c1h]	6	H(nnnXG2)	Z→C(nnnXG)	Y→C(nnnXB)	R

Invariome, Symmetrie und Koordinatensystem (Forts.).

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
CD	C1.5o1.5o1c-	m	C(nnnXD)	X→O(nnnXE1)	Y→O(nnnXE2)	R
OE1	O1.5c[1.5o1c]	m	O(nnnXE1)	X→C(nnnXD)	Y→O(nnnXE2)	R
OE2	O1.5c[1.5o1c]	m	O(nnnXE2)	X→C(nnnXD)	Y→O(nnnXE1)	R
<i>Glutaminsäure protoniert an O_{e2}</i>						
CD	C2o1o1c	m	C(nnnXD)	X→O(nnnXE1)	Y→O(nnnXE2)	R
OE1	O2c	m	O(nnnXE1)	Z→DUM	Y→C(nnnXD)	R
OE2	O1c1h	m	O(nnnXE2)	X→C(nnnXD)	Y→H(nnnXE2)	R
HE2	H1o[1c]	6	H(nnnXE2)	Z→O(nnnXE2)	Y→C(nnnXD)	R
Histidin, N_{δ1} protoniert						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG	C1.5n[1.5c1h]1.5c[1.5n1h]1c	m	C(nnnXG)	X→C(nnnXB)	Y→N(nnnXD1)	R
ND1	N1.5c[1.5n1h]1.5c[1.5c1c]1h	m	N(nnnXD1)	X→C(nnnXG)	Y→C(nnnXE1)	R
HD1	H1n[1.5c1.5c]	6	H(nnnXD1)	Z→N(nnnXD1)	Y→C(nnnXG)	R
CD2	C1.5n[1.5c]1.5c[1.5n1c]1h	m	C(nnnXD2)	X→N(nnnXE2)	Y→C(nnnXG)	R
HD2	H1c[1.5n1.5c]	6	H(nnnXD2)	Z→C(nnnXD2)	Y→N(nnnXE2)	R
CE1	C1.5n[1.5c1h]1.5n[1.5c]1h	m	C(nnnXE1)	X→H(nnnXE1)	Y→N(nnnXD1)	R
HE1	H1c[1.5n1.5n]	6	H(nnnXE1)	Z→C(nnnXE1)	Y→N(nnnXD1)	R
NE2	N1.5c[1.5n1h]1.5c[1.5c1h]	m	N(nnnXE2)	X→C(nnnXD2)	Y→C(nnnXE1)	R
Isoleucin						
CB	C1c1c1c1h	3m	C(nnnXB)	Z→H(nnnXB)	X→C(nnnXA)	R
HB	H1c[1c1c1c]	6	H(nnnXB)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG1	C1c1c1h1h	mm2	C(nnnXG1)	Z→DUM	Y→C(nnnXB)	R

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
HG1A	H1c[1c1c1h]	6	H(nnnXG1A)	Z→C(nnnXG1)	Y→C(nnnXB)	R
HG1B	H1c[1c1c1h]	6	H(nnnXG1B)	Z→C(nnnXG1)	Y→C(nnnXB)	R
CG2	C1c1h1h1h	3	C(nnnXG2)	Z→C(nnnXB)	Y→H(nnnXG2A)	R
HG2A	H1c[1c1h1h]	6	H(nnnXG2A)	Z→C(nnnXG2)	Y→C(nnnXB)	R
HG2B	H1c[1c1h1h]	6	H(nnnXG2B)	Z→C(nnnXG2)	Y→C(nnnXB)	R
HG2C	H1c[1c1h1h]	6	H(nnnXG2C)	Z→C(nnnXG2)	Y→C(nnnXB)	R
CD1	C1c1h1h1h	3	C(nnnXD1)	Z→C(nnnXG1)	Y→H(nnnXD1A)	R
HD1A	H1c[1c1h1h]	6	H(nnnXD1A)	Z→C(nnnXD1)	Y→C(nnnXG1)	R
HD1B	H1c[1c1h1h]	6	H(nnnXD1B)	Z→C(nnnXD1)	Y→C(nnnXG1)	R
HD1C	H1c[1c1h1h]	6	H(nnnXD1C)	Z→C(nnnXD1)	Y→C(nnnXG1)	R
Leucin						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG	C1c1c1c1h	3m	C(nnnXG)	Z→H(nnnXG)	X→C(nnnXB)	R
HG	H1c[1c1c1c]	6	H(nnnXG)	Z→C(nnnXG)	Y→C(nnnXB)	R
CD1	C1c1h1h1h	3	C(nnnXD1)	Z→C(nnnXG)	Y→H(nnnXD1A)	R
HD1A	H1c[1c1h1h]	6	H(nnnXD1A)	Z→C(nnnXD1)	Y→C(nnnXG)	R
HD1B	H1c[1c1h1h]	6	H(nnnXD1B)	Z→C(nnnXD1)	Y→C(nnnXG)	R
HD1C	H1c[1c1h1h]	6	H(nnnXD1C)	Z→C(nnnXD1)	Y→C(nnnXG)	R
CD2	C1c1h1h1h	3	C(nnnXD2)	Z→C(nnnXG)	Y→H(nnnXD2A)	R
HD2A	H1c[1c1h1h]	6	H(nnnXD2A)	Z→C(nnnXD2)	Y→C(nnnXG)	R
HD2B	H1c[1c1h1h]	6	H(nnnXD2B)	Z→C(nnnXD2)	Y→C(nnnXG)	R
HD2C	H1c[1c1h1h]	6	H(nnnXD2C)	Z→C(nnnXD2)	Y→C(nnnXG)	R

Invariome, Symmetrie und Koordinatensystem (Forts.).

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
Lysin						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG	C1c1c1h1h	mm2	C(nnnXG)	Z→DUM	Y→C(nnnXB)	R
HG1	H1c[1c1c1h]	6	H(nnnXG1)	Z→C(nnnXG)	Y→C(nnnXB)	R
HG2	H1c[1c1c1h]	6	H(nnnXG2)	Z→C(nnnXG)	Y→C(nnnXB)	R
CD	C1c1c1h1h	mm2	C(nnnXD)	Z→DUM	Y→C(nnnXG)	R
HD1	H1c[1c1c1h]	6	H(nnnXD1)	Z→C(nnnXD)	Y→C(nnnXG)	R
HD2	H1c[1c1c1h]	6	H(nnnXD2)	Z→C(nnnXD)	Y→C(nnnXG)	R
CE	C1n1c1h1h	1	C(nnnXE)	Z→N(nnnXZ)	Y→C(nnnXD)	R
HE1	H1c[1n1c1h]	6	H(nnnXE1)	Z→C(nnnXE)	Y→N(nnnXZ)	R
HE2	H1c[1n1c1h]	6	H(nnnXE2)	Z→C(nnnXE)	Y→N(nnnXZ)	R
NZ	N1c1h1h1h+	3	N(nnnXZ)	Z→C(nnnXE)	Y→H(nnnXZ1)	R
HZ1	H1n[1c1h1h]+	6	H(nnnXZ1)	Z→N(nnnXZ)	Y→C(nnnXE)	R
HZ2	H1n[1c1h1h]+	6	H(nnnXZ2)	Z→N(nnnXZ)	Y→C(nnnXE)	R
HZ3	H1n[1c1h1h]+	6	H(nnnXZ3)	Z→N(nnnXZ)	Y→C(nnnXE)	R
Methionin						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG	C1s[1h]1c1h1h	m	C(nnnXG)	X→S(nnnXD)	Y→C(nnnXB)	R
HG1	H1c[1s1c1h]	6	H(nnnXG1)	Z→C(nnnXG)	Y→S(nnnXD)	R
HG2	H1c[1s1c1h]	6	H(nnnXG2)	Z→C(nnnXG)	Y→S(nnnXD)	R

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
SD	S1c1c	mm2	S(nnnXD)	Z→DUM	X→C(nnnXG)	R
CE	C1s[1h]1h1h1h	3	C(nnnXE)	Z→S(nnnXD)	Y→H(nnnXE1)	R
HE1	H1c[1s1h1h]	6	H(nnnXE1)	Z→C(nnnXE)	Y→S(nnnXD)	R
HE2	H1c[1s1h1h]	6	H(nnnXE2)	Z→C(nnnXE)	Y→S(nnnXD)	R
HE3	H1c[1s1h1h]	6	H(nnnXE3)	Z→C(nnnXE)	Y→S(nnnXD)	R
Phenylalanin						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG	C1.5c[1.5c1h]1.5c[1.5c1h]1c	mm2	C(nnnXG)	Z→C(nnnXB)	Y→C(nnnXD1)	R
CD1	C1.5c[1.5c1c]1.5c[1.5c1h]1h	mm2	C(nnnXD1)	Z→DUM	Y→C(nnnXG)	R
HD1	H1c[1.5c1.5c]	6	H(nnnXD1)	Z→C(nnnXD1)	Y→C(nnnXG)	R
CD2	C1.5c[1.5c1c]1.5c[1.5c1h]1h	mm2	C(nnnXD2)	Z→DUM	Y→C(nnnXG)	R
HD2	H1c[1.5c1.5c]	6	H(nnnXD2)	Z→C(nnnXD2)	Y→C(nnnXG)	R
CE1	C1.5c[1.5c1h]1.5c[1.5c1h]1h	mm2	C(nnnXE1)	Z→DUM	Y→C(nnnXD1)	R
HE1	H1c[1.5c1.5c]	6	H(nnnXE1)	Z→C(nnnXE1)	Y→C(nnnXD1)	R
CE2	C1.5c[1.5c1h]1.5c[1.5c1h]1h	mm2	C(nnnXE2)	Z→DUM	Y→C(nnnXD2)	R
HE2	H1c[1.5c1.5c]	6	H(nnnXE2)	Z→C(nnnXE2)	Y→C(nnnXD2)	R
CZ	C1.5c[1.5c1h]1.5c[1.5c1h]1h	mm2	C(nnnXZ)	Z→DUM	Y→C(nnnXE1)	R
HZ	H1c[1.5c1.5c]	6	H(nnnXZ)	Z→C(nnnXZ)	Y→C(nnnXE1)	R
Prolin						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R

Invariome, Symmetrie und Koordinatensystem (Forts.).

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
CG	C1c1c1h1h	mm2	C(nnnXG)	Z→DUM	Y→C(nnnXB)	R
HG1	H1c[1c1c1h]	6	H(nnnXG1)	Z→C(nnnXG)	Y→C(nnnXB)	R
HG2	H1c[1c1c1h]	6	H(nnnXG2)	Z→C(nnnXG)	Y→C(nnnXB)	R
CD	C1n1c1h1h	1	C(nnnXD)	Z→N(nnnX)	Y→C(nnnXG)	R
HD1	H1c[1n1c1h]	6	H(nnnXD1)	Z→C(nnnXD)	Y→N(nnnX)	R
HD2	H1c[1n1c1h]	6	H(nnnXD2)	Z→C(nnnXD)	Y→N(nnnX)	R
Serin						
CB	C1o1c1h1h	m	C(nnnXB)	X→O(nnnXG)	Y→C(nnnXA)	R
HB1	H1c[1o1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→O(nnnXG)	R
HB2	H1c[1o1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→O(nnnXG)	R
OG	O1c1h	m	O(nnnXG)	X→C(nnnXB)	Y→H(nnnXG)	R
HG	H1o[1c]	6	H(nnnXG)	Z→O(nnnXG)	Y→C(nnnXB)	R
Threonin						
CB	C1o1c1c1h	m	C(nnnXB)	X→O(nnnXG1)	Y→C(nnnXA)	R
HB	H1c[1o1c1c]	6	H(nnnXB)	Z→C(nnnXB)	Y→O(nnnXG1)	R
OG1	O1c1h	m	O(nnnXG1)	X→C(nnnXB)	Y→H(nnnXG1)	R
HG1	H1o[1c]	6	H(nnnXG1)	Z→O(nnnXG1)	Y→C(nnnXB)	R
CG2	C1c1h1h1h	3	C(nnnXG2)	Z→C(nnnXB)	Y→H(nnnXG2A)	R
HG2A	H1c[1c1h1h]	6	H(nnnXG2A)	Z→C(nnnXG2)	Y→C(nnnXB)	R
HG2B	H1c[1c1h1h]	6	H(nnnXG2B)	Z→C(nnnXG2)	Y→C(nnnXB)	R
HG2C	H1c[1c1h1h]	6	H(nnnXG2C)	Z→C(nnnXG2)	Y→C(nnnXB)	R
Tryptophan						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG	C1.5c[1.5c1.5c]1.5c[1n1h]1c	m	C(nnnXG)	X→C(nnnXB)	Y→C(nnnXD1)	R
CD1	C1.5n[1.5c1h]1.5c[1.5c1c]1h	m	C(nnnXD1)	X→C(nnnXG)	Y→N(nnnXE1)	R
HD1	H1c[1.5n1.5c]	6	H(nnnXD1)	Z→C(nnnXD1)	Y→N(nnnXE1)	R
CD2	C1.5c[1.5n1.5c]1.5c[1.5c1c]1.5c[1.5c1h]	m	C(nnnXD2)	X→C(nnnXG)	Y→C(nnnXE2)	R
NE1	N1.5c[1.5c1.5c]1.5c[1.5c1h]1h	m	N(nnnXE1)	X→C(nnnXD1)	Y→C(nnnXE2)	R
HE1	H1n[1.5c1.5c]	6	H(nnnXE1)	Z→N(nnnXE1)	Y→C(nnnXD1)	R
CE2	C1.5n[1.5c1h]1.5c[1.5c1.5c]1.5c[1.5c1h]	m	C(nnnXE2)	X→N(nnnXE1)	Y→C(nnnXD2)	R
CE3	C1.5c[1.5c1.5c]1.5c[1.5c1h]1h	m	C(nnnXE3)	X→C(nnnXD2)	Y→C(nnnXZ3)	R
HE3	H1c[1.5c1.5c]	6	H(nnnXE3)	Z→C(nnnXE3)	Y→C(nnnXD2)	R
CZ2	C1.5c[1.5n1.5c]1.5c[1.5c1h]1h	m	C(nnnXZ2)	X→C(nnnXE2)	Y→C(nnnXH2)	R
HZ2	H1c[1.5c1.5c]	6	H(nnnXZ2)	Z→C(nnnXZ2)	Y→C(nnnXE2)	R
CZ3	C1.5c[1.5c1h]1.5c[1.5c1h]1h	mm2	C(nnnXZ3)	Z→DUM	Y→C(nnnXE3)	R
HZ3	H1c[1.5c1.5c]	6	H(nnnXZ3)	Z→C(nnnXZ3)	Y→C(nnnXE3)	R
CH2	C1.5c[1.5c1h]1.5c[1.5c1h]1h	mm2	C(nnnXH2)	Z→DUM	Y→C(nnnXZ2)	R
HH2	H1c[1.5c1.5c]	6	H(nnnXH2)	Z→C(nnnXH2)	Y→C(nnnXZ2)	R
Tyrosin						
CB	C1c1c1h1h	mm2	C(nnnXB)	Z→DUM	Y→C(nnnXA)	R
HB1	H1c[1c1c1h]	6	H(nnnXB1)	Z→C(nnnXB)	Y→C(nnnXA)	R
HB2	H1c[1c1c1h]	6	H(nnnXB2)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG	C1.5c[1.5c1h]1.5c[1.5c1h]1c	mm2	C(nnnXG)	Z→C(nnnXB)	Y→C(nnnXD1)	R
CD1	C1.5c[1.5c1c]1.5c[1.5c1h]1h	mm2	C(nnnXD1)	Z→DUM	Y→C(nnnXG)	R
HD1	H1c[1.5c1.5c]	6	H(nnnXD1)	Z→C(nnnXD1)	Y→C(nnnXG)	R
CD2	C1.5c[1.5c1c]1.5c[1.5c1h]1h	mm2	C(nnnXD2)	Z→DUM	Y→C(nnnXG)	R
HD2	H1c[1.5c1.5c]	6	H(nnnXD2)	Z→C(nnnXD2)	Y→C(nnnXG)	R

Invariome, Symmetrie und Koordinatensystem (Forts.).

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
CE1	C1.5c[1.5c1o]1.5c[1.5c1h]1h	mm2	C(nnnXE1)	Z→DUM	Y→C(nnnXD1)	R
HE1	H1c[1.5c1.5c]	6	H(nnnXE1)	Z→C(nnnXE1)	Y→C(nnnXD1)	R
CE2	C1.5c[1.5c1o]1.5c[1.5c1h]1h	mm2	C(nnnXE2)	Z→DUM	Y→C(nnnXD2)	R
HE2	H1c[1.5c1.5c]	6	H(nnnXE2)	Z→C(nnnXE2)	Y→C(nnnXD2)	R
CZ	C1.5c[1.5c1h]1.5c[1.5c1h]1o	mm2	C(nnnXZ)	Z→O(nnnXH)	Y→C(nnnXE1)	R
OH	O1c1h	m	O(nnnXH)	X→C(nnnXZ)	Y→H(nnnXH)	R
HH	H1o[1c]	6	H(nnnXH)	Z→O(nnnXH)	Y→C(nnnXZ)	R
Valin						
CB	C1c1c1c1h	3m	C(nnnXB)	Z→H(nnnXB)	X→C(nnnXA)	R
HB	H1c[1c1c1c]	6	H(nnnXB)	Z→C(nnnXB)	Y→C(nnnXA)	R
CG1	C1c1h1h1h	3	C(nnnXG1)	Z→C(nnnXB)	Y→H(nnnXG1A)	R
HG1A	H1c[1c1h1h]	6	H(nnnXG1A)	Z→C(nnnXG1)	Y→C(nnnXB)	R
HG1B	H1c[1c1h1h]	6	H(nnnXG1B)	Z→C(nnnXG1)	Y→C(nnnXB)	R
HG1C	H1c[1c1h1h]	6	H(nnnXG1C)	Z→C(nnnXG1)	Y→C(nnnXB)	R
CG2	C1c1h1h1h	3	C(nnnXG2)	Z→C(nnnXB)	Y→H(nnnXG2A)	R
HG2A	H1c[1c1h1h]	6	H(nnnXG2A)	Z→C(nnnXG2)	Y→C(nnnXB)	R
HG2B	H1c[1c1h1h]	6	H(nnnXG2B)	Z→C(nnnXG2)	Y→C(nnnXB)	R
HG2C	H1c[1c1h1h]	6	H(nnnXG2C)	Z→C(nnnXG2)	Y→C(nnnXB)	R

Tabelle 4.1.2: Differenzierte Wasserstoffabstände im Invariommodell im Vergleich zu AFIX-Abständen und gemittelten Neutronenabständen.

Invariom	Bindungstypen	$d_{inv}/\text{\AA}$	$d_{sph}/\text{\AA}$	$d_{neutron}/\text{\AA}$
H1o[1c]	$O_\epsilon-H_\epsilon, O_\gamma-H_\gamma, O_\eta-H_\eta$	0,960708	0,82	0,967
H1n[1.5c1.5c]	$N_\delta-H_\delta, N_\epsilon-H_\epsilon$	1,005134	0,86	1,009
H1n[1.5c1h]	$N_\delta-H_\delta, N_\epsilon-H_\epsilon$	1,006230	0,86	1,009
H1n[1.5c1h]+	$N_\eta-H_\eta$	1,006320	0,86	1,009
H1n[1.5c1c]	$N'-H'$	1,007910	0,86	1,009
H1n[1c1c]	$N_\epsilon-H_\epsilon$	1,011719	0,86	1,009
H1n[1c1h1h]+	$N_{term}-H_{term}, N_\zeta-H_\zeta$	1,022880	0,89	1,033
H1c[1.5n1.5n]	$C_\epsilon-H_\epsilon$	1,077606	0,93	1,094
H1c[1.5c1.5c]	$C_\delta-H_\delta, C_\epsilon-H_\epsilon, C_\eta-H_\eta, C_\zeta-H_\zeta$	1,082246	0,93	1,083
H1c[1.5n1.5c]	$C_\delta-H_\delta$	1,084434	0,93	1,094
H1c[1s1h1h]	$C_\epsilon-H_\epsilon$	1,087362	0,96	1,066
H1c[1s1c1h]	$C_\beta-H_\beta, C_\gamma-H_\gamma$	1,089368	0,97	1,092
H1c[1c1h1h]	$C_\beta-H_\beta, C_\delta-H_\delta, C_\gamma-H_\gamma$	1,091400	0,96	1,059
H1c[1c1c1h]	$C_\beta-H_\beta, C_\delta-H_\delta, C_\gamma-H_\gamma, C_\gamma-H_\gamma$	1,093723	0,97	1,092
H1c[1c1c1c]	$C_\beta-H_\beta, C_\gamma-H_\gamma$	1,095990	0,98	1,099
H1c[1o1c1h]	$C_\beta-H_\beta$	1,096252	0,97	1,092
H1c[1o1c1c]	$C_\beta-H_\beta$	1,098210	0,98	1,099
H1c[1n1c1h]	$C_\alpha-H_\alpha, C_\delta-H_\delta, C_\epsilon-H_\epsilon$	1,099321	0,97	1,092
H1c[1n1c1c]	$C_\alpha-H_\alpha$	1,101063	0,98	1,099

Nach einem ähnlichen Prinzip werden die Wasserstoffatomabstände blockweise pro Aminosäurerest festgelegt. Die aus der Invariomberechnung stammenden Bindungslängen sind wesentlich realistischer als die in einer sphärischen Verfeinerung verwendeten, welche in ihrer Verkürzung an ein Modell ohne explizite Bindungsbeschreibung angepaßt sind. Mit den zur Multipolbeschreibung passenden längeren Wasserstoffabständen werden Werte nahe gemittelter Abstände aus Neutronenbeugungsexperimenten (Allen u. a. 1992) erreicht. Tab. 4.1.2 stellt die Bindungslängen zu in „Nächster Nachbar Näherung“ verschiedenen Typen von Wasserstoffatomen nach den drei Varianten einander gegenüber: Die Abstände d_{inv} werden in den Modellen *mul* und *iam_{rb}* mittels der `reset bond`-Anweisung in XD verwendet. In der klassischen sphärischen Verfeinerung *iam* finden die Bindungslängen wie in Spalte d_{sph} Verwendung. Der Vergleich mit dem letzten Eintrag $d_{neutron}$ zeigt nun die Verbesserung der Beschreibung des Röntgenbeugungsexperimentes durch die Invariom-Abstände.

4.2 Alanyl-Prolyl-Alanin

Die Datenkollektion fand am Hamburger Synchrotronstrahlungslabor HASYLAB (DESY, Hamburg) statt. Wichtige Merkmale der bereits publizierten Struktur (Kalinowski u. a. 2007) sind in Tab. 4.2.1 aufgelistet: Die Messung erfolgte bei einer Temperatur von $T = 100$ K und einer Wellenlänge von $\lambda = 0,56$ Å. Insgesamt 186017 Reflexe bis zu einer Auflösung von $d = 0,37$ Å wurden mit *xds* (Kabsch 1993) zu einem durchschnittlich 12-fach redundanten Datensatz in der orthorhombischen Raumgruppe $P2_12_12_1$ gemittelt. Von den unabhängigen Reflexen erfüllen 87 % das Beobachtungskriterium $I > 2\sigma(I)$. Die Vollständigkeit der Daten liegt im Mittel bei nur 92 %, da diese in den höchsten Auflösungsschalen stark abnimmt und zwischen $0,38$ – $0,37$ Å nur noch 73 % beträgt. In gleicher Weise ist eine Verschlechterung des an sich sehr guten Signal-zu-Rausch-Verhältnisses und der Redundanz zu beobachten, die aber auch bei höchsten Beugungswinkeln noch akzeptable Werte erreichen. Mit der Messung zu maximaler Auflösung geht ein R_{mrgd} von 5 % einher. Lag dieser bei $d = 0,8$ Å noch bei ausgezeichneten 1,3 %, erreicht er bei $0,4$ Å bereits 15 %, um letztlich auf 38 % zu steigen.

Tabelle 4.2.1: Kristallographische Daten APA.

Summenformel	$C_{11}N_3O_4H_{19} \cdot H_2O$		
Formelmasse	275,303 g/mol	Meßtemperatur	100 K
Kristallsystem	Orthorhombisch	Wellenlänge λ	0,56 Å
Raumgruppe	$P2_12_12_1$ (Nr. 19)	max. 2Θ	99,64°
Z	4	Auflösungsbereich	5–0,37 Å
Gitterkonstanten:		Grenzreflexe	$0 \leq h \leq 15$
a	6,825(1) Å		$0 \leq k \leq 24$
b	9,042(2) Å		$0 \leq l \leq 58$
c	21,728(4) Å	unabh. Reflexe	13992
$\alpha = \beta = \gamma$	90,000°	$F_o^2 > 2\sigma(F_o^2)$	12185
V	1340,9(5) Å ³	Abdeckungsgrad [†]	91,8 (72,6) %
Dichte	1,3638(5) g/cm ³	Redundanz [†]	12,01 (2,62)
F(000)	592,0	$I/\sigma(I)$ [†]	30,08 (3,77)
Absorptionskoeff.	0,066 mm ⁻¹	R_{mrgd} [†]	4,9 (38,2) %

[†] Angabe in Klammern für Auflösungsschale $0,38$ – $0,37$ Å.

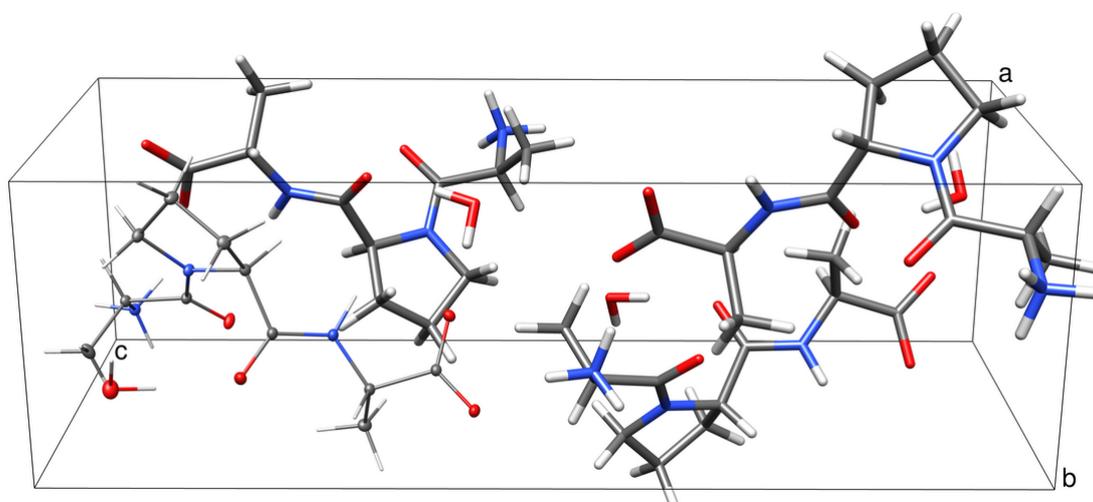


Abbildung 4.2.1: Elementarzelle APA. Eine der asymmetrischen Einheiten ist mit Temperaturellipsoiden (50 %) dargestellt.

4.2.1 Sphärisches Modell

Die Strukturlösung mit direkten Methoden in `shelxs` ergab alle Nichtwasserstoffatompositionen. Deren Verfeinerung und eine anschließende Differenzfourieranalyse ergab die Koordinaten aller Wasserstoffatome. Sowohl die Atome des Tripeptids als auch die des kokristallisierten Wassermoleküls konnten ohne jede Beschränkung frei verfeinert werden. Dabei wurden die Verschiebungsparameter der Schweratome anisotrop, die der Wasserstoffe isotrop verfeinert. Das Tripeptid besteht aus 37 vollbesetzten Atomlagen, davon 18 Nichtwasserstoffatomen. Auch das einzelne Wassermolekül ist fehlordnungsfrei. Eine Darstellung der Struktur inklusive der Temperaturellipsoide und der Zellpackung bietet Abb. 4.2.1.

Mit insgesamt 256 Parametern konvergierte die Rechnung nach acht Zyklen bei Gütefaktoren von $R1(F_o > 4\sigma(F_o)) = 0,0352$, $R_{all} = 0,0423$ und $wR2 = 0,0917$ mit einem *goodness of fit* von 1,039. Die mittleren und maximalen Parameteränderungen lagen praktisch bei Null. Es verblieb eine durchschnittliche Restelektronendichte nahe Null mit einer quadratisch gemittelten Abweichung von $0,06 \text{ e}\text{\AA}^{-3}$ und Extremwerten bei $-0,21$ und $0,51 \text{ e}\text{\AA}^{-3}$. Zur Beurteilung der Verfeinerung kann auch die Reflexstatistik in Tab. 4.2.2 auf der nächsten Seite herangezogen werden.

Durch Beschränkung der Daten auf $d = 0,83 \text{ \AA}$ wurde für spätere Vergleichszwecke eine gering aufgelöste Messung simuliert. Die gewählte Auflösung entspricht mit $\sin \Theta/\lambda = 0,6 \text{ \AA}^{-1}$ den Minimalanforderungen an Kleinmolekülstrukturen. Hierbei erreichte die Verfeinerung von APA0.83 gegen die verbliebenen 1436 Reflexe, von denen 99 % beobachtet sind, Gütefaktoren von $R1(F_o > 4\sigma(F_o)) = 0,0242$, $R_{all} = 0,0243$ und $wR2 = 0,0764$ sowie einen *goodness of fit* von 1,425. Die mitt-

Tabelle 4.2.2: Reflexstatistik APA nach `shelxl` inklusive aller Korrekturen.

Auflösung/Å	$N_{unabh.}$	$\overline{I/\sigma(I)}$	$(I > 2\sigma(I))/\%$	Vollständigkeit/%	$\sum(F_o)/\sum(F_c)$	R1(all)	R1($F_o > 4\sigma(F_o)$)
$\infty-4,000$	8	197,46	47,06	47,06	0,998	0,013	0,013
4,000–3,000	22	194,87	95,65	95,65	1,028	0,035	0,035
3,000–2,000	86	186,33	98,84	100,00	1,047	0,047	0,047
2,000–1,800	37	182,72	100,00	100,00	1,060	0,057	0,057
1,800–1,600	62	166,89	100,00	100,00	1,050	0,052	0,052
1,600–1,400	103	153,53	100,00	100,00	1,037	0,041	0,041
1,400–1,200	182	138,96	99,45	100,00	1,017	0,030	0,030
1,200–1,000	340	111,06	99,12	100,00	0,992	0,026	0,026
1,000–0,900	300	79,72	98,67	100,00	0,993	0,025	0,024
0,900–0,800	459	64,85	98,26	99,78	1,002	0,024	0,023
0,800–0,700	739	57,56	98,65	99,46	1,003	0,018	0,018
0,700–0,600	1293	44,04	95,23	97,81	0,995	0,020	0,019
0,600–0,550	1038	33,02	93,38	96,74	0,990	0,029	0,027
0,550–0,500	1426	23,48	91,02	95,58	0,990	0,036	0,033
0,500–0,450	2092	16,31	84,77	93,73	1,006	0,045	0,037
0,450–0,400	3109	9,26	74,23	89,83	1,039	0,081	0,061
0,400–0,366	2696	5,01	51,13	77,05	1,057	0,154	0,101
$\infty-0,366$	13992	29,67	78,96	90,67	1,011	0,042	0,035

lere Verschiebung-pro-Standardabweichung lag nach acht Zyklen bei 0,002. Eine Differenzfouriersynthese ergibt Maximalwerte von -0,15 und 0,18 eÅ⁻³ der Restdichte, deren Mittel Null beträgt bei einer quadratisch gemittelten Abweichung von 0,03 eÅ⁻³.

Zusätzlich zu den bereits genannten Parametern der Verfeinerung von APA und APA0.83, können die Erhebungen zur Strukturqualität A.2 bis A.7 ab Seite 328 zur Einschätzung dieser Verbindungen herangezogen werden. Diese Auswertungen sind primär für Proteine gedacht und sind für die aktuelle Struktur nur von beschränktem Nutzen. Zu Vergleichszwecken mit den in dieser Arbeit behandelten Makromolekülen sind sie dennoch angegeben.

Die Auftragung von Standardabweichungen der Atompositionen und Bindungslängen gegen den äquivalenten isotropen Temperaturfaktor zeigt für alle Atome von APA inklusive des Wassersauerstoffs Werte um 0,0006 Å. Bei APA0.83 liegt

diese Unsicherheit mit ca. $0,0022 \text{ \AA}$ zwar um das Dreifache höher als bei voller Auflösung, ist aber absolut gesehen immer noch ein sehr niedriger Wert. Der über die Aminosäurereste gemittelte B -Wert von APA liegt bei ungefähr $1-1,5 \text{ \AA}^2$, in den Seitenketten marginal höher als in der Hauptkette. Ebenso sind kaum Unterschiede zwischen den den drei Resten zu erkennen. Die Verschiebungsparameter des gekürzten Datensatzes APA0.83 sind nur unwesentlich höher als bei voller Auflösung. An den Luzzati-Auftragungen (in Abb. A.4, A.7) wird der große Unterschied zu den Makromolekülen nochmal deutlich: Über den gesamten Auflösungsbereich liegt die Kurve der R -Werte weit unter den vorgesehenen theoretischen Kurven zur Abschätzung des durchschnittlichen Koordinatenfehlers. Zumindest im Fall APA0.83 ergeben die R -Werte eine Parallele zur Abszisse, bei APA zeigt sich ein Anstieg bei höchster Auflösung.

Bei Begutachtung der *procheck*- und *whatcheck*-Ausgaben sind nun einige Besonderheiten zu beachten: Da die Struktur so klein ist, gibt es beispielsweise keine Ramachandran-Auftragung, welche nur für nicht terminale Reste, exklusive Prolin und Glycin erstellt wird, womit aber schon alle in APA vorkommenden Aminosäurereste erschöpft sind. Alle packungsbezogenen Tests liefern Warnungen, da das Tripeptid, ob der fehlenden Solvenssphäre, viel enger an symmetrieverwandten Molekülen liegt, als es für Proteine normal wäre. Die Struktur gibt auch Anlaß zu mediokren G -Faktoren, welche die Normalität der stereochemischen Eigenschaften widerspiegeln sollen. Hierbei wird auf einen statistischen Vergleich mit guten Proteinstrukturen zurückgegriffen, der hier aber unangebracht ist.

Darüberhinaus zeigen APA und APA0.83 beim direkten Vergleich mit Engh&Huber-Bindungslängen etliche mehr als 2σ vom Mittelwert abweichende Atomabstände, alle aktuellen Bindungen sind länger als der Engh&Huber-Mittelwert. Werte außerhalb zweier Standardabweichungen um den Mittelwert gibt es auch bei den Bindungswinkeln, jedoch ist keine geometrische Abweichung hoch genug, um als besonders verzerrt reportiert zu werden. Allerdings ist die Richtigkeit der Geometrie von APA angesichts der niedrigen R -Werte, geringen Restdichten und sehr niedrigen Temperaturfaktoren sowie der verschwundenen Koordinatenunsicherheit kaum in Zweifel zu ziehen. Vielmehr könnten die 1991 bestimmten Mittelwerte aus Strukturen geringerer Güte als der vorliegenden bestimmt worden sein. So lassen sich die prinzipiell größeren Abstände z. B. zwanglos mit einer verringerten scheinbaren Bindungsverkürzung bei kleinen Verschiebungsparametern erklären. Zudem sind die am stärksten abweichenden Bindungen an terminalen Resten lokalisiert, darunter auch die Bindungen zu den Carboxylatsauerstoffen, die in mehrere Wasserstoffbrückenbindungen involviert sind. Analoge Erwägungen gelten auch für weitere Faktoren, wie eine außerhalb des für Proteine Üblichen liegende Auflösung oder die niedrige Meßtemperatur.

Diese Darlegungen sollten bei der Beurteilung der entsprechenden Strukturtests der nachfolgenden Verbindungen insofern Beachtung finden, als auch diese in einigen Punkten, wie z. B. der Auflösung, ungewöhnlich sind.

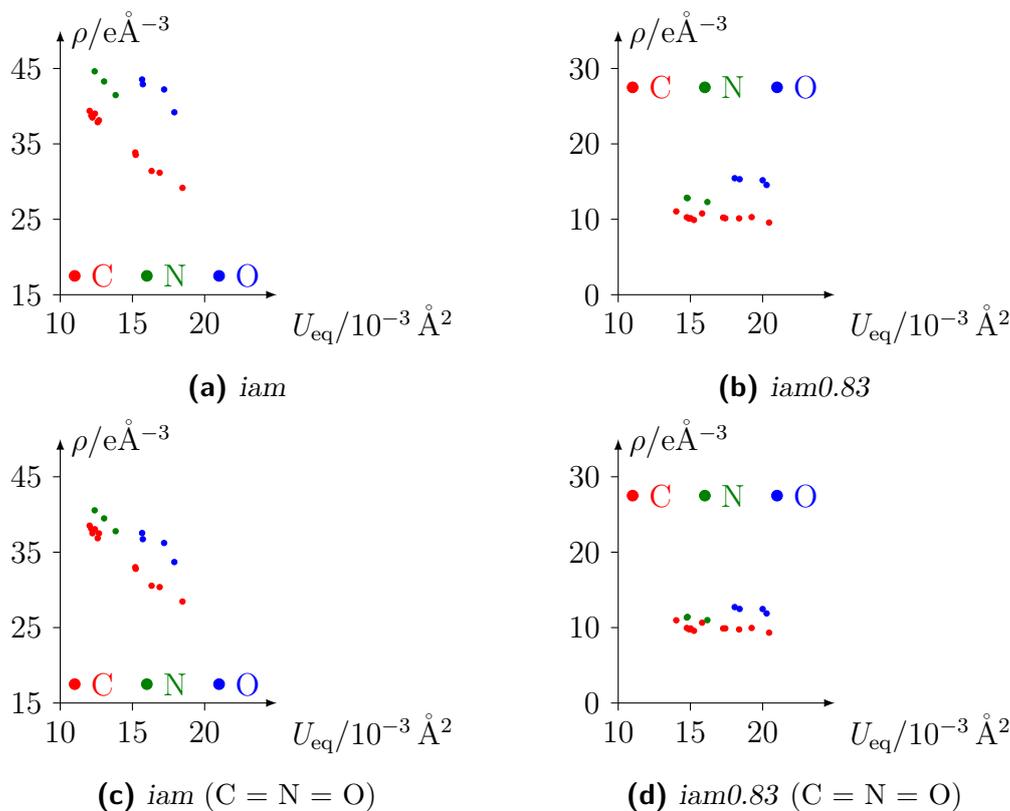


Abbildung 4.2.2: Elektronendichte des sphärischen Modells an den Atompositionen.

Inwieweit die gemessenen Intensitäten die Information über die verschiedenen Elemente beinhalten, kann anhand der Auftragung der aus einer $F_o\phi_c$ -Fouriersynthese gewonnenen Elektronendichte an den Atompositionen des jeweiligen Modells über den Temperaturfaktor abgeschätzt werden (siehe hierzu auch Abschn. 3.4 auf Seite 45). Diese sogenannten Sevcik-Diagramme zeigt Abb. 4.2.2, sowohl für den hochaufgelösten Datensatz (a) als auch für den gekürzten (b), mit den Phasen aus *iam* bzw. *iam0.83*. Bekanntermaßen tragen nicht nur die Intensitäten, sondern auch die Phasen zur Elektronendichteinformation bei. Um den Effekt abschätzen zu können, sind die aufgetragenen Werte in den beiden unteren Abbildungen 4.2.2(c) und (d) aus einer Verfeinerung mit gleichen Streufaktoren für C, N und O entstanden. Damit soll eine Beeinflussung des Ergebnisses durch *a priori*-Information über die Atomsorte vermieden werden. In den beiden linken Abbildungen (a) und (c) zum vollständigen Datensatz, ist eine eindeutige Trennung nach Elementen in drei zusammenhängende Bereiche zu erkennen, wobei die Proportionalität der Dichtewerte zueinander ungefähr den Ordnungszahlen entspricht. Die Unterscheidung bleibt auch bei Verwendung von Kohlenstoff-Streufaktoren für die Stickstoff- und Sauerstoffatome in der unteren Abbildung erhalten, wenn auch in verrin-

gertem Ausmaß. Das Diagramm 4.2.2(b) legt nahe, daß die Information über den Elementtyp auch bei geringer aufgelösten Daten erhältlich bleibt, allerdings ist ein „Zusammenrutschen“ der drei Verteilungen im Vergleich mit den vollen Daten zu verzeichnen und es gibt einen geringeren Abfall in Abhängigkeit von den Verschiebungsparametern. Nimmt man den Phasen die *a priori*-Information über den Elementtyp, erhält man die Auftragung 4.2.2(d), welche zumindest eine Unterscheidung C oder N nicht mehr erlauben würde.

In Abb. 4.2.1 auf Seite 68 wurde die Packung der vier Tripeptid- und Wassermoleküle in der Zelle gezeigt, welche durch drei parallel zu *a*, *b* und *c* liegende 2_1 -Schraubenachsen erzeugt wurden. Verglichen mit Proteinen sind die kleinen Moleküle recht eng gepackt, was sich in einem Matthews-Koeffizienten von nur $1,304 \text{ \AA}^3/\text{Da}$ äußert. Mit den Näherungen in Gl. 2.2 auf Seite 32 berechnen sich die prozentualen Anteile von „Protein“ und Solvens am Zellvolumen auf 94,3 % und 5,7 %. Auch das Verhältnis von Volumen der asymmetrischen Einheit und van-der-Waals-Volumen des Tripeptids $V_{asu}/V_{vdW} = 335,218/248,371 = 1,350 \text{ \AA}^3$ fällt, wie sich später zeigen wird, klein aus. Nach dieser Methode nimmt das Tripeptid 74 % des Volumens der asymmetrischen Einheit ein. Eine Erklärung der großen Unterschiede zwischen den beiden Berechnungen könnte in der Ungültigkeit der in Gl. 2.2 angenommenen Proteindichte für kleine Moleküle liegen. Davon ist der Weg über das vdW-Volumen unabhängig.

4.2.2 Asphärisches Modell

Der Übergang zum XD-Programmsystem vollzog sich wie in Absch. 4.1.2 beschrieben. Nach Invariomtransfer und Elektroneutralitätsausgleich verblieb eine Überschlußladung von $-0,0012e$. Es wurden für alle drei Modelle acht *full matrix least square*-Zyklen gegen $F_o^2 > 2\sigma(F_o^2)$ gerechnet, in denen die Ort- und Verschiebungsparameter aller Atome frei verfeinert werden konnten. Hierbei waren die verbleibenden Parameterverschiebungen mit einem Maximum von $2 \cdot 10^{-7}$ im Fall der Wiederholung der sphärischen Verfeinerung *iam* am geringsten. Beim zweiten sphärischen Modell *iam_{rb}* führte die Verlängerung der Wasserstoffabstände auf die in Tab. 4.1.2 genannten Invariomwerte zu höheren Restverschiebungen von $2 \cdot 10^{-2}$, gleiches gilt für die eigentliche asphärische Verfeinerung *mul* mit $6 \cdot 10^{-3}$, wobei jeweils die Koordinaten von Wasserstoffatomen betroffen sind. Die erreichten Gütefaktoren sind in Tab. 5.1 auf Seite 115 und Tab. 5.2 auf Seite 116 aufgelistet.

4.3 Triklines Lysozym

Messung, Strukturlösung und -verfeinerung von P1LYSO sind in der Publikation von Wang u. a. (2007) beschrieben, Koordinaten und Strukturaktoren als PDB-Eintrag 2VB1 veröffentlicht.

Einige Eckdaten dieses Experimentes wiederholt Tab. 4.3.1: Die Datenkollektion fand an der APS-Synchrotronbeamline bei 100 K bis zu einer Auflösung von $d = 0,65 \text{ \AA}$ statt. Hierbei wurde für die triklin Struktur ein 187165 symmetrieunabhängige Reflexe umfassender, durchschnittlich siebenfach redundanter Datensatz erhalten. Beinahe 94 % der Reflexe sind beobachtet, erfüllen also das $I > 2\sigma(I)$ -Kriterium. Dazu paßt das ausgezeichnete Signal-zu-Untergrund-Verhältnis von 36, welches auch in der höchsten Auflösungsschale nicht unter 4 sinkt. Ebenfalls sehr gut ist der interne R -Wert von 4,5 %, insbesondere bleibt die Reflexqualität auch bei höheren Beugungswinkeln erhalten, wie der moderate Anstieg auf nur 18 % bei $d < 0,67 \text{ \AA}$ belegt. Die Daten sind mit 97 % fast vollständig, daß in der letzten Schale nur noch 67 % erreicht werden, mindert den Eindruck einer hervorragenden Messung nicht, bedenkt man die für Proteine außerordentlich hohe Auflösung.

Tabelle 4.3.1: Kristallographische Daten P1LYSO (Wang u. a. 2007).

Summenformel	$C_{613}N_{193}O_{185}S_{10}H_{960} \cdot 3 C_2O_2H_6, 1 C_2O_2H_3^-, 9 NO_3^-, 170 O$	Absorptionskoeff.	$0,13 \text{ mm}^{-1}$
Formelmasse	16886,2 g/mol	Meßtemperatur	100 K
Kristallsystem	Triklin	Wellenlänge λ	$0,6526 \text{ \AA}$
Raumgruppe	P1 (Nr. 1)	max. 2Θ	$60,37^\circ$
Z	1	Auflösungsbereich	$30-0,65 \text{ \AA}$
Gitterkonstanten:		Grenzreflexe	$0 \leq h \leq 40$
a	$27,07(3) \text{ \AA}$		$-48 \leq k \leq 43$
b	$31,25(3) \text{ \AA}$		$-50 \leq l \leq 48$
c	$33,76(3) \text{ \AA}$	unabh. Reflexe	187165
α	$87,98(5)^\circ$	$F_o^2 > 2\sigma(F_o^2)$	175025
β	$108,00(5)^\circ$	Abdeckungsgrad [†]	97,6 (67,3) %
γ	$112,11(5)^\circ$	Redundanz [†]	7,12 (2,7)
V	$25057,00 \text{ \AA}^3$	$I/\sigma(I)$ [†]	36,20 (4,20)
Dichte	$1,120 \text{ g/cm}^3$	R_{mrgd} [†]	4,5 (18,4) %
F(000)	8931,9		

[†] Angabe in Klammern für Auflösungsschale $0,67-0,65 \text{ \AA}$.

4.3.1 Sphärisches Modell

Wang u. a. verfeinerten nach *ab initio*-Lösung der Struktur ein aus 2900 Atomlagen bestehendes Modell. Das Lysozymmolekül aus 129 Aminosäureresten umfaßt 1960 Atome, von denen 697 oder 35,5 % in zwei oder drei Konformationen modelliert wurden. Die Solvenssphäre besteht aus 170 Wassermolekülen mit frei verfeinerter Besetzung, neun Nitrat-Anionen, einem Acetat-Anion sowie drei Ethylenglykollmolekülen. Im Unterschied zum ursprünglichen Modell von Wang u. a. wurde die Glutaminsäure 35 neutral mit einer Carboxylfunktion modelliert, um Konsistenz mit dem Modell des tetragonalen Lysozyms (siehe Absch. 4.4 auf Seite 83) herzustellen. Die 959 Wasserstoffpositionen, davon 375 fehlgeordnet, stammen aus einem idealisierten Geometriemodell mit steter Anpassung an die Position der sie tragenden Nichtwasserstoffatome. Inklusiv des diffusen Solvensmodells ($g = 0,786, U = 3,392$) beträgt die Anzahl verfeinerter Parameter 14111 bei 20150 Restraints. Die nur in einer Konformation vorliegenden Teile der Struktur wurden ohne Restraints verfeinert, die für 45 Seitenketten und 8 Teile der Hauptkette beibehalten wurden, was in oben genannter Arbeit ausführlich begründet wird. Eine Illustration der Struktur gibt Abb. 4.3.1 auf der gegenüberliegenden Seite, dabei sind fehlgeordnete Bereiche dünner gezeichnet als die Hauptkomponente. Anhand der Färbung von blau nach rot des transparenten Bandes wird der Verlauf der Hauptkette über 129 Aminosäurereste vom N- zum C-Terminus nachgezeichnet.

Nach zehn *full matrix least squares*-Zyklen mit allen Parametern konvergiert das Modell zu einem sphärischen R -Wert von 8,12 % für $F_o > 4\sigma(F_o)$; R_{all} , $wR2$ und *goodness of fit* betragen jeweils 8,40 %, 22,97 % und 1,06. Dabei verblieben Parameterverschiebungen von maximal 0,034 Å, die Koordinaten von Wassermolekülen betreffend. Die gemittelte Verschiebung pro Standardabweichung liegt bei 0,032. Eine Differenzfouriersynthese ergibt eine mittlere Restdichte von $0,00 \text{ e}\text{\AA}^{-3}$ mit $0,10 \text{ e}\text{\AA}^{-3}$ quadratisch gemittelter Abweichung vom Mittelwert und Extremwerten bei 1,42 und $-0,57 \text{ e}\text{\AA}^{-3}$. Die auflösungsabhängigen Abweichungen von Beobachtung und Modell können anhand der Reflexstatistik in Tab. 4.3.2 auf Seite 76 (Unterschiede zu Tab. 4.3.1 wegen anderer Schaleneinteilung und der inkludierten Korrekturen) und zusätzlich in einem F_o^2/F_c^2 -Diagramm in Absch. 6.1 auf Seite 148 untersucht werden.

Die hohe Qualität dieser Proteinstruktur wird von den *procheck*-, *whatcheck*- und *shelxpro*-Strukturtests der Abb. B.3, B.4 und B.5 ab Seite 344 unterstrichen. Der Ramachandran-Plot weist 88,5 % der Aminosäurereste in der energetisch vorteilhaftesten Region und die restlichen 11,5 % in den erlaubten Konformationen aus. Die quadratisch gemittelte Abweichung von Idealwerten der Bindungslängen beträgt 0,019 Å, die der Bindungswinkel $2,154^\circ$. Es werden kaum Geometrieauffälligkeiten von den Tests aufgezeigt, nur in 14 Resten werden signifikante Abweichungen gefunden. Die Standardabweichung der Koordinaten ist für fast alle Atome des Hauptmoleküls kleiner als 0,02 Å, die der Bindungslängen liegt unter

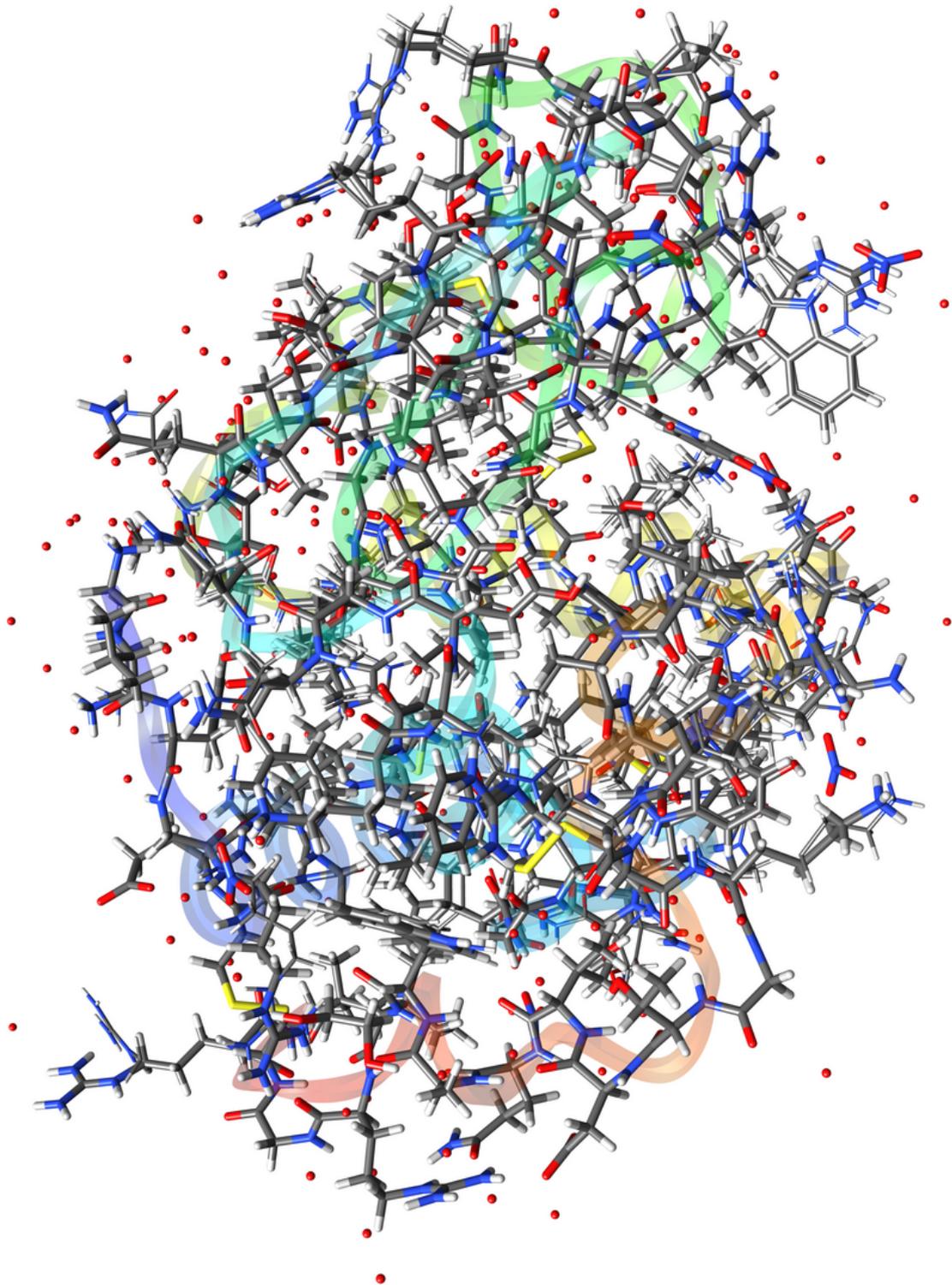


Abbildung 4.3.1: Strukturmodell P1LYSO: Farbband vom N-Terminus (blau) zum C-Terminus (rot), fehlgeordnete Bereiche schmal gezeichnet.

Tabelle 4.3.2: Reflexstatistik PLYSO nach shelx1 inklusive aller Korrekturen.

Auflösung/Å	$N_{unabh.}$	$\overline{I/\sigma(I)}$	$(I > 2\sigma(I))/\%$	Vollständigkeit/%	$\sum(F_o)/\sum(F_c)$	R1(all)	R1($F_o > 4\sigma(F_o)$)
$\infty-10,00$	48	31,41	88,68	90,57	1,054	0,297	0,266
10,000–5,000	369	36,91	98,92	99,19	1,045	0,234	0,232
5,000–4,000	389	34,97	98,73	98,98	1,027	0,132	0,125
4,000–3,000	1107	37,79	99,19	99,28	1,079	0,129	0,128
3,000–2,000	4609	42,95	99,52	99,52	1,035	0,102	0,102
2,000–1,800	2397	45,84	99,63	99,63	1,020	0,095	0,095
1,800–1,600	3830	49,31	99,95	99,95	1,013	0,085	0,085
1,600–1,400	6321	48,51	99,95	99,95	1,011	0,078	0,078
1,400–1,200	11261	44,20	99,94	99,98	0,991	0,071	0,071
1,200–1,000	22053	35,01	99,58	99,87	0,980	0,063	0,063
1,000–0,900	19460	31,97	99,14	99,89	0,998	0,062	0,061
0,900–0,850	13524	29,67	98,17	100,00	1,020	0,065	0,064
0,850–0,800	16982	26,73	96,87	99,99	1,029	0,069	0,067
0,800–0,750	21946	22,73	95,38	99,99	1,025	0,073	0,069
0,750–0,700	28605	16,41	91,76	99,99	1,015	0,079	0,072
0,700–0,649	34264	8,36	67,85	87,87	0,974	0,120	0,088
$\infty-0,649$	187165	26,05	91,14	97,48	1,010	0,084	0,081

0,01 Å, beide steigen erwartungsgemäß mit dem B -Faktor, wobei die Restraints den Anstieg des Bindungslängenfehlers dämpfen. Auch der Luzzati-Plot legt einen unterdurchschnittlich kleinen Fehler nahe.

Der mittlere B -Wert aller Atome der Struktur ohne Solvenssphäre liegt mit 7 \AA^2 sehr niedrig und erreicht in den wohlgeordneten Bereichen des Proteins fast Werte um 4 \AA^2 , was auch für kleine Moleküle akzeptabel wäre. Auch sind die pro Aminosäurerest gemittelten B -Werte recht gleichmäßig über die Sequenz verteilt und in den Seitenketten nur geringfügig höher als in der Hauptkette. Ähnliches gilt für die mittlere Anisotropie. Nur die wesentlich schlechteren Werte des weniger gut bestimmten C-terminalen Endes sind auffällig. Zwischen dem von `whatcheck` berichteten Packungsqualitätsmaß über die Sequenz und den korrespondierenden B -Faktoren kann nur eine wache Korrelation festgestellt werden, insofern in Bereichen mit abnehmendem Qualitätsindex häufiger die durchschnittlichen Verschiebungsparameter zunehmen. Allerdings sind Sequenzabschnitte mit helikaler

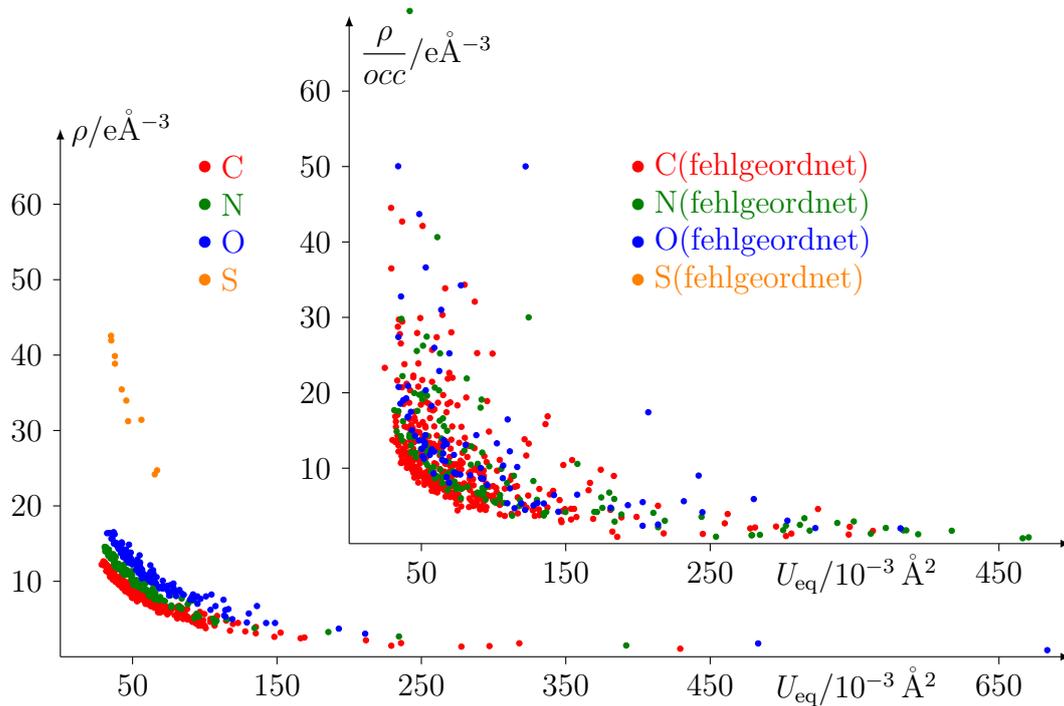


Abbildung 4.3.2: $F_o\phi_c$ -Elektronendichte des sphärischen Modells an den Positionen voll besetzter und fehlgeordneter Atome.

Sekundärstruktur und Faltblattbereiche tendenziell besser geordnet als Schleifen oder Knäulbereiche (siehe auch Abb. 4.3.4 auf Seite 81).

Die Güte der beobachteten Elektronendichteverteilung spiegelt sich in den zwei Sevcik-Diagrammen (siehe auch Abschn. 3.4) der Abb. 4.3.2 wieder. Hier wurde die Elektronendichte an den Atompositionen, sowohl der Atome mit ganzer Besetzung als auch der fehlgeordneten, gegen den jeweiligen Temperaturfaktor aufgetragen, wodurch zweierlei verdeutlicht wird: Die beobachteten Intensitäten lassen die Unterscheidung verschiedener Elemente voneinander zu, denn im unteren Diagramm erkennt man vier voneinander verschiedene Verteilungen, die den jeweiligen Atomsorten zuzuordnen sind, sofern die Auslenkungsparameter nicht zu groß werden. Dies gilt aber nur, wenn es sich um die geordnet beschriebenen Atome handelt. Bei mehreren Konformationen verschwimmen die Grenzen zwischen den Elementen wegen der diffuser verteilten Elektronendichte und der Unsicherheit in den verfeinerten Besetzungszahlen. Wenn bereits der Elementtyp nicht mehr sicher aus den beobachteten Daten bestimmt werden kann, wird die verfeinerte Unterscheidung in „Nächster Nachbar Näherung“ des asphärischen Modells umso fraglicher. Zu bemerken bleibt noch, daß die zur Fouriertransformation notwendigen Phasen ϕ_c , die ja aus dem Strukturmodell bestimmt wurden, eher eine Präferenz der so erhaltenen Dichte bezüglich eben dieses Modells bedingen, allerdings in moderatem

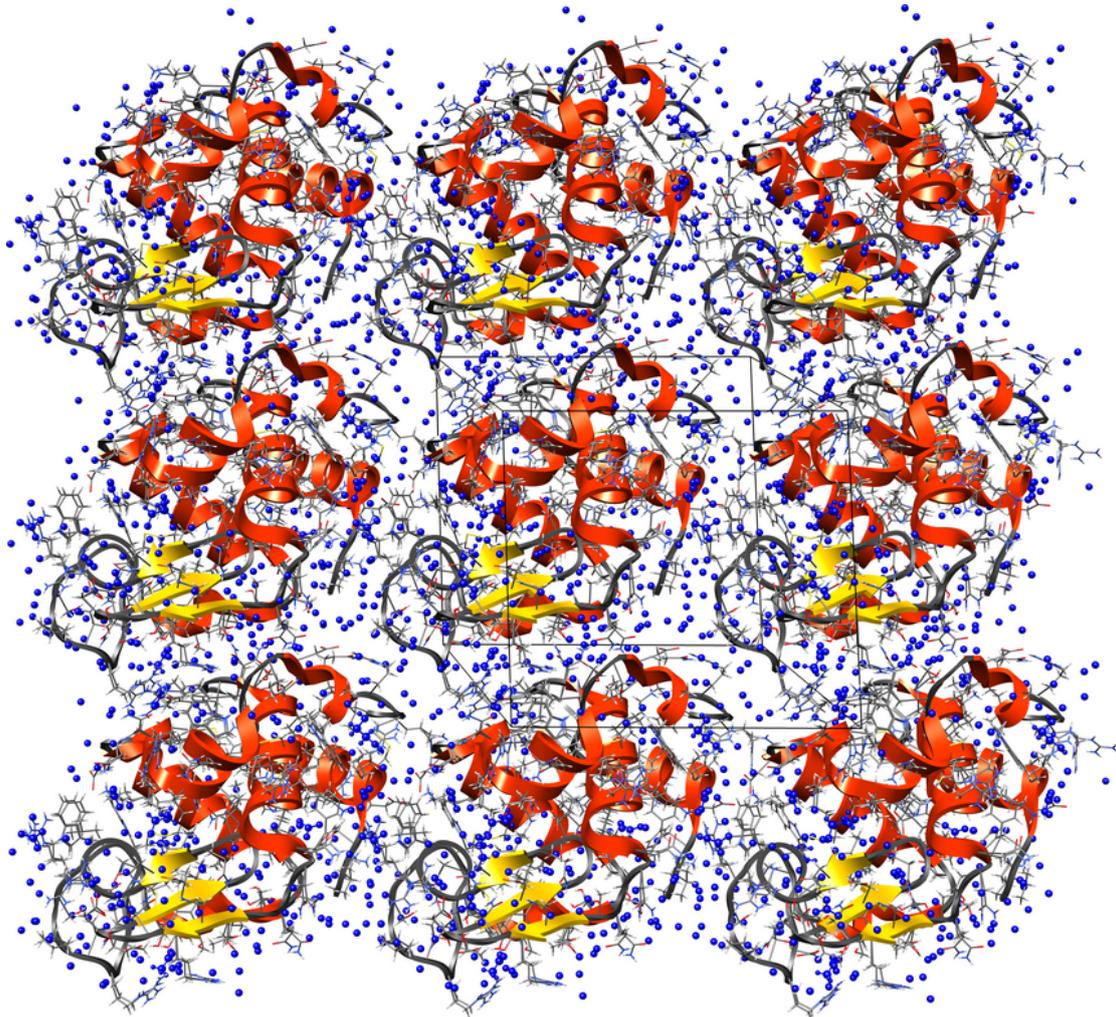


Abbildung 4.3.3: Ausschnitt der Packung von P1LYSO, neun Elementarzellen in der *cb*-Ebene. Helices (rot) und β -Stränge (gelb) in Cartoondarstellung, Seitenketten in Stabdarstellung, Solvenssphäre (blau) als Kugeln.

Umfang, wie in Absch. 4.2.1 auf Seite 71 ausgeführt wurde.

Abb. 4.3.3 vermittelt einen Eindruck der Packung, für deren Dichte der in Abb. B.4 auf Seite 348 vermerkte Matthews-Koeffizient ein Maß ist. Mit $1,751 \text{ \AA}^3/\text{Da}$ entspricht er einer eng gepackten Proteinstruktur. Legt man die Näherungsformel 2.2 zugrunde, ergibt sich daraus ein Solvensanteil von 30 %, entsprechend einem Volumenanteil des Proteins an der asymmetrischen Einheit von 70 %. Alternativ kann die Bestimmung auch über ein molekülbezogenes Volumen erfolgen: Der Quotienten aus Volumen der asymmetrischen Einheit und dem 14493 \AA^3 betragenden van-der-Waals-Volumen des Lysozymmoleküls berechnet sich zu $V_{asu}/V_{vdW} = 25057/14493 = 1,729$, woraus ein Proteinanteil von 58 % folgt.

4.3.2 Asphärisches Modell

Die Vorbereitung der asphärischen Verfeinerung wurde wie in Absch. 4.1.2 beschrieben vorgenommen. Nach Invariomtransfer und Ladungsausgleich verblieb ein Überschuß von 0,0225 e. Da in XD keine Restraints implementiert sind, können nicht alle Parameter verfeinert werden, sondern müssen auf den `shelxl`-Werten fixiert bleiben. Bereiche, die in der vorherigen Verfeinerung keine Restraints benötigten, wurden auch in der XD-Verfeinerung freigegeben. Traten dabei zu hohe Verschiebungen in den Koordinaten oder Verschiebungsparametern auf, wurden die betreffenden Atome wieder von der Verfeinerung ausgeschlossen. Umgekehrt wurden versuchsweise gut definierte Atome in die Verfeinerung aufgenommen und bei unauffälligem Verhalten beibehalten. Mittels dieses „Versuch und Irrtum“-Verfahrens wurde angestrebt, die Anzahl der verfeinerten Atome zu maximieren ohne die Struktur zu verzerren.

Alle zur Solvenssphäre zählenden Moleküle wurden fixiert. Von den 1001 Nichtwasserstoffatomen des Lysozymmoleküls gingen 639 oder 64 % in die Verfeinerung ein. Damit wurden fast alle 679 Atome im wohlgeordneten Strukturteil verfeinert, bis auf wenige Ausnahmen in nicht fehlgeordnet beschriebenen aber gleichwohl schlecht definierten Bereichen (wenig Elektronendichte, hohe und volatile Verschiebungsparameter). Haupt- oder Seitenketten, die teilweise oder vollständig in alternativen Positionen gefunden wurden, blieben auf den restrainten Atompositionen. Wasserstoffatome wurden nicht verfeinert, sondern nur, ausgehend von ihren idealisierten Positionen, entlang der Bindungsachse auf theoretische Abstände verschoben.

Die zur Verfeinerung ausgewählten Bereiche der Struktur und die Verteilung der Fehlordnung sind in Abb. 4.3.4 als farbcodierte, nach Hauptkette (HK) und Seitenkette (SK) getrennte, Pictogramme für jeden Aminosäurerest dargestellt. Darüber ist außerdem die Sekundärstruktur symbolisiert, welche von `whatcheck` nach der DSSP-Methode (Kabsch u. Sander 1983) bestimmt wurde und in Abb. 4.3.3 dargestellt ist. Unter jedem Kürzel eines Aminosäurerests ist zunächst vermerkt, ob der betreffende Strukturteil geordnet oder fehlgeordnet ist. Jeweils darunter, nach wie vor in Haupt- und Seitenkette unterteilt, befindet sich die Angabe zu Verfeinerung oder Fixierung. Hierbei steht „blau“ sowohl für „geordnet“ als auch für „verfeinert“, letzteres wird außerdem durch ein „✓“ dargestellt. Umgekehrt bedeutet „rot“, daß dieser Aminosäureteil „fehlgeordnet“ ist bzw. „nicht verfeinert“ wurde, was durch ein „X“ nochmal verdeutlicht wird. Strukturteile, die nur teilweise fehlgeordnet oder verfeinerbar sind, werden in „orange“ dargestellt und zeigen außerdem die noch freigegebenen Atome. So liest sich bspw. der Eintrag für die N-terminale Aminosäure LYS001 wie folgt: Die Atome der Hauptkette sind nicht fehlgeordnet und es wurden N', C_α, C' und O' verfeinert, während aus der Seitenkette nur C_β und C_γ freigegeben wurden, C_δ aber, obwohl nicht fehlgeordnet, C_ε und N_ζ, beide fehlgeordnet, wurden festgehalten.

4 Verfeinerung

LYS 001	VAL 002	PHE 003	GLY 004	ARG 005	CYS 006	GLU 007	LEU 008
HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK
✓ C γ	✓ ✓	✓ ✓	✓ ✓	X X	✓ ✓	✓ ✓	✓ ✓

ALA 009	ALA 010	ALA 011	MET 012	LYS 013	ARG 014	HIS 015	GLY 016
HK SK	HK SK	HK SK	HK SK				
✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ C δ	✓ X	✓ X	✓ ✓

LEU 017	ASP 018	ASN 019	TYR 020	ARG 021	GLY 022	TYR 023	SER 024
HK SK							
N X	X X	X X	X X	X X	X X	X ✓	X X

LEU 025	GLY 026	ASN 027	TRP 028	VAL 029	CYS 030	ALA 031	ALA 032
HK SK							
X X	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓

LYS 033	PHE 034	GLU 035	SER 036	ASN 037	PHE 038	ASN 039	THR 040
HK SK							
✓ ✓	✓ ✓	✓ ✓	✓ ✓	CO X	✓ ✓	CO X	✓ ✓

GLN 041	ALA 042	THR 043	ASN 044	ARG 045	ASN 046	THR 047	ASP 048
HK SK							
CO X	-N ✓	X X	✓ ✓	✓ X	✓ ✓	✓ ✓	✓ X

GLY 049	SER 050	THR 051	ASP 052	TYR 053	GLY 054	ILE 055	LEU 056
HK SK							
✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	CO X	✓ ✓

GLN 057	ILE 058	ASN 059	SER 060	ARG 061	TRP 062	TRP 063	CYS 064
HK SK							
✓ ✓	✓ ✓	✓ ✓	✓ ✓	CO X	✓ X	✓ ✓	✓ ✓

ASN 065	ASP 066	GLY 067	ARG 068	THR 069	PRO 070	GLY 071	SER 072
HK SK							
✓ ✓	✓ ✓	✓ ✓	CO X	✓ ✓	X X	✓ ✓	✓ ✓

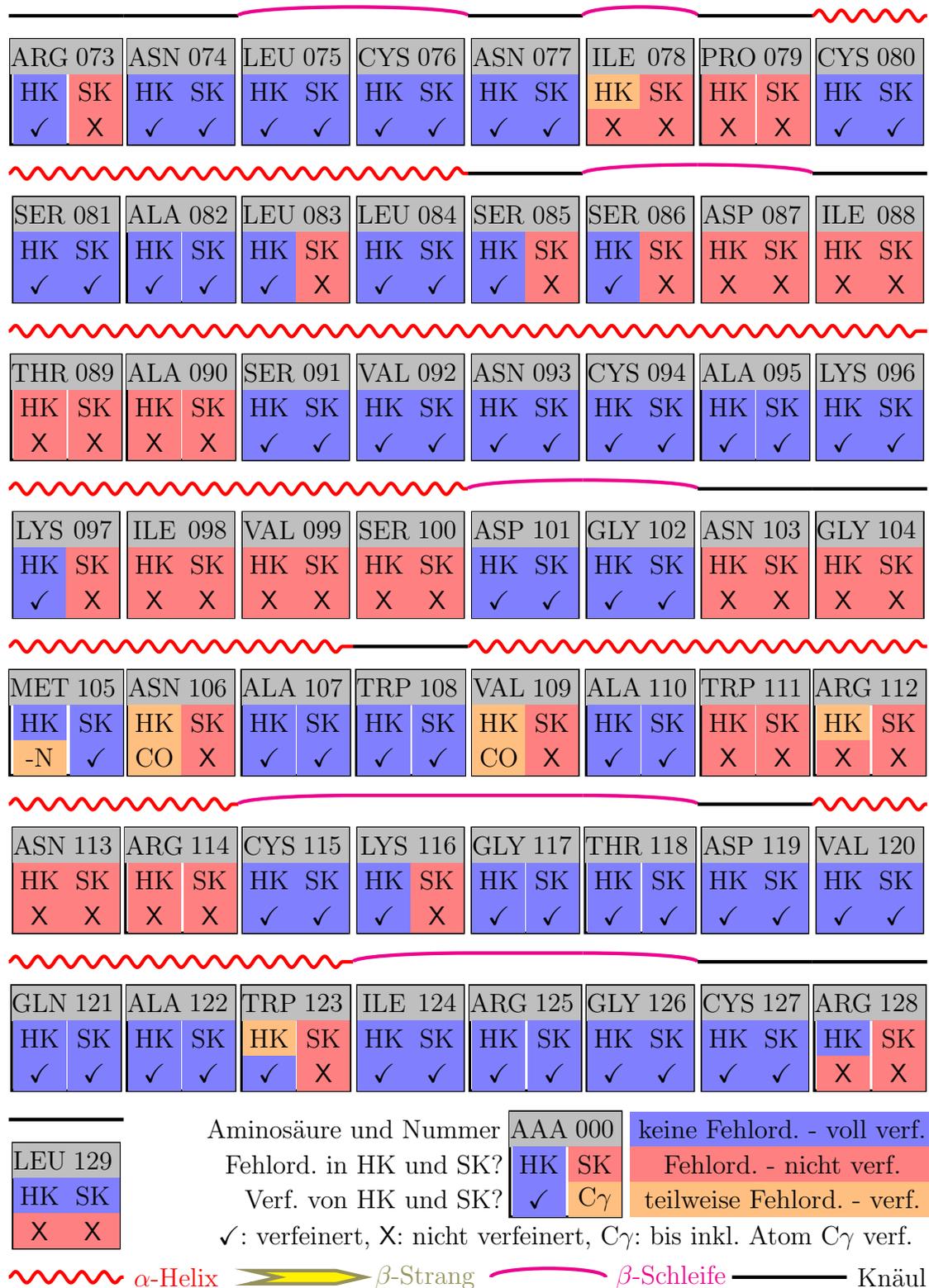


Abbildung 4.3.4: DSSP-Zuordnung der Sekundärstruktur von p1LYSO, Verteilung fehlgeordneter Bereiche und verfeinerter Strukturteil.

Das Nichtverfeinern eines Gutteils der Struktur führt zu mehr als einer Halbierung der Parameteranzahl von den obengenannten 14111 bei 20150 Restriktionen auf 5752 unrestringierte Parameter in den hier diskutierten Modellen. Die maximalen Parameteränderungen in der asphärischen Verfeinerung betreffen die anisotropen Verschiebungsparameter. Nach zwölf *full matrix least square*-Zyklen beträgt die maximal auftretende Veränderung noch 10^{-5} , nach anfänglichen $2 \cdot 10^{-2}$. Weitere Kennziffern der Verfeinerung sind in Tab. 6.1 auf Seite 147 gegeben. Der Versuch, den fehlgeordneten Strukturteil nur sphärisch zu behandeln, blieb ohne nennenswerte Änderung.

Um alle Einflüsse außer den asphärischen Streufaktoren auszuschließen, wurde mit den identischen Eingabedateien des Modells *mul*, unter Austausch der verwendeten Streufaktoren gegen sphärische, die Verfeinerung wiederholt. Dies ist das Modell *iam_{rb}*. Wird die Verlängerung der Wassertoffabstände nicht durchgeführt, erhält man das Modell *iam*, das die *shelxl*-Ergebnisse reproduzieren sollte (siehe dazu auch Abschn. 4.1.2 auf Seite 53), allerdings teilweise höhere Gütefaktoren ergibt. In den beiden sphärischen Modellen verblieben maximale Parameterverschiebungen von nur 10^{-6} , nach ebenfalls zwölf Verfeinerungszyklen.

4.4 Tetragonales Lysozym

Der TETRALYSO-Datensatz wurde von Dr. Alke Meents an der „Swiss Light Source“ des Paul Scherrer-Instituts (Villingen, Schweiz) gemessen. Die bei $T=100\text{ K}$ und einer Wellenlänge von $\lambda = 0,6888\text{ \AA}$ erreichte maximale Auflösung betrug $d = 0,65\text{ \AA}$. Von den dabei gewonnenen 8 709 706 Reflexen wurden allerdings nur 4 339 785 Beobachtungen bis zur hier verwendeten Auflösung von $d = 0,80\text{ \AA}$ verwendet, weil die zu größerem Winkel gemessenen nicht von ausreichender Qualität waren. Wie in Tab. 4.4.1 aufgeführt, wurden diese in der tetragonalen Raumgruppe $P4_32_12$ zu einem durchschnittlich 35-fach redundanten Datensatz von 122388 Reflexen gemittelt, der zu 88 % dem Beobachtungskriterium $I > 2\sigma(I)$ genügt. Datenintegration und -mittelung wurden mit `xds` (Kabsch 1993) durchgeführt. Einen Eindruck der mit steigender Auflösung abnehmenden Qualität geben die sowohl für den Gesamtbereich als auch für die höchste Auflösungsschale angegebenen Kennziffern: Der Datensatz ist über den gesamten Bereich vollständig, auch bleibt die Redundanz auf hohem Niveau. Das Signal-zu-Rausch-Verhältnis nimmt um mehr als den Faktor zehn ab, bleibt aber auf angemessenen Werten. Hingegen ist die Zunahme des internen R -Wertes auf das fünfzehnfache kein Ausweis hoher Datenqualität.

Tabelle 4.4.1: Kristallographische Daten TETRALYSO.

Summenformel	$C_{613}N_{193}O_{185}S_{10}H_{960} \cdot 4 C_2O_2H_6, 3 C_2O_2H_3^-, 6 Cl^-, 210 O$		
Formelmasse	17486,1 g/mol	Meßtemperatur	100 K
Kristallsystem	Tetragonal	Wellenlänge λ	0,6888 \AA
Raumgruppe	$P4_32_12$ (Nr. 96)	max. 2Θ	51,00°
Z	8	Auflösungsbereich	10–0,80 \AA
Gitterkonstanten:		Grenzreflexe	$0 \leq h \leq 98$
a	78,981(8) \AA		$0 \leq k \leq 69$
b	78,981(8) \AA		$0 \leq l \leq 46$
c	36,844(4) \AA	unabh. Reflexe	122388
$\alpha = \beta = \gamma$	90,000°	$F_o^2 > 2\sigma(F_o^2)$	107618
V	229832,84 \AA^3	Abdeckungsgrad [†]	100 (99,9) %
Dichte	1,011 g/cm ³	Redundanz [†]	35,5 (13,20)
F(000)	73831	$I/\sigma(I)$ [†]	47,52 (3,69)
Absorptionskoeff.	0,06 mm ⁻¹	R_{mrgd} [†]	3,4 (50,6) %

[†] Angabe in Klammern für Auflösungsschale 0,82–0,80 \AA .

Tabelle 4.4.2: Verfeinerungsschritte TETRALYSO.

Verfeinerungsstufe	$d/\text{\AA}$	N_{prot}	N_{solv}	N_{par}	N_{restr}	$R1/R_{free}/\%$
epmr: initiales Modell	4,0	517	0	—	—	71,8
Isotrop	1,5	517	0	2069	2107	47,01/50,32
Modellierung	1,5	956	0	3825	4016	30,78/33,87
Solvens	1,5	979	98	4307	4113	22,14/24,33
Auflösungserhöhung	0,80	1029	91	4484	4405	22,75/23,73
Anisotrop	0,80	1056	110	10334	13484	16,73/18,35
Solvens	0,80	1083	165	11238	14646	14,42/16,19
H-fix	0,80	2116	188	11366	14803	13,19/15,02
Solvens-Besetzung	0,80	2104	289	11944	23225	12,41/15,04
Restraints lockern	0,80	2104	289	11930	20544	12,37/15,07
<i>full matrix least square</i>	0,80	2104	289	11930	20689	12,55

4.4.1 Sphärisches Modell

Als Startmodell diente eine bei 0,94 Å verfeinerte Lysozymstruktur mit ähnlichen Zellparametern (PDB-Eintrag 1IEE, Sauter u. a. (2001)). Nach Entfernen von Solvens, Fehlordnung und Seitenketten wurden die verbleibenden Koordinaten als Suchmodell für die molekulare Ersetzung (*molecular replacement*) mit **epmr** übergeben. Aus den resultierenden Koordinaten wurde mittels **shelxpro** eine **shelxl**-Eingabedatei inklusive Restraints erstellt. In mehreren Schritten wurde das Modell aufgebaut, siehe dazu auch Tab. 4.4.2. Hierzu wurde von den *real space*-Modellierungsmöglichkeiten des Programmes **coot** (siehe Absch. 4.1.1 auf Seite 51) Gebrauch gemacht, welches auch verschiedene Qualitätstests ermöglicht, die die Modellierung steuern helfen. In allen Verfeinerungsstufen, außer der letzten, wurden die Parameter im *conjugate gradient least square*-Verfahren an die Beobachtungen angepaßt, immer wurde gegen F^2 verfeinert. Jede der genannten Verfeinerungsstufen umfaßt mehrere Verfeinerungen nach Änderungen am Modell, so daß letztendlich ungefähr 80 Verfeinerungen mit jeweils 50–100 Zyklen gerechnet wurden. Hierbei meint Modelländerung das Einfügen oder Entfernen von Atomen, die Verbesserung von Bindungs- und Torsionswinkeln anhand der beobachteten Dichte und der in **coot** implementierten Restraints, sowie das Aufspüren von Alternativkonformationen.

Beginnend mit der Verfeinerung von Koordinaten und isotropen Temperaturfaktoren wurde, bei zunächst geringer Auflösung, ein Gutteil der Seitenketten aufgebaut. Sodann wurde mit **shelxwat** eine automatische Suche nach Wasserpositionen durchgeführt. Die Auflösung der zur Verfeinerung verwendeten Daten wurde schrittweise an den kompletten Datenumfang herangeführt um dann die

Verschiebungsparameter der Nichtwasserstoffatome anisotrop zu verfeinern. Eine weitere Modellverbesserung wurde mit einer weiteren `shelxwat`-Runde erreicht. Die Struktur wurde nun mit Wasserstoffatomen komplettiert, die an idealisierten Positionen relativ zum sie tragenden Schweratom constraint wurden (AFIX). Bei basischen oder sauren Resten wurde das Protonierungsschema anhand der Elektronendichte- bzw. Restdichtekarte und durch den Vergleich von Bindungslängen mit tabellierten Werten (Wilson 1992) gestaltet. So wurde z. B. einer der zwei Glutaminsäurereste deprotoniert modelliert, da die beiden $C_\delta-O_\epsilon$ -Bindungen mit 1,23 Å und 1,25 Å ungefähr gleich lang waren und etwa einer 1,5-fachen Bindung in einer Carboxylatgruppe entsprechen, während der zweite Glutaminsäurerest GLU 35 unterschiedliche Bindungslängen von 1,23 Å für $C_\delta-O_{\epsilon 1}$ und 1,30 Å für $C_\delta-O_{\epsilon 2}$ zeigte. Da diese Abstände ungefähr einer CO-Doppelbindung und einer C–OH-Einfachbindung entsprechen, wurde $O_{\epsilon 2}$ protoniert. Ein weiterer Schritt bestand aus der Verfeinerung auch der Besetzungszahlen der Solvensmoleküle, wobei versucht wurde, die Gesamtstruktur formal elektroneutral zu halten. Anschließend wurde ein möglichst hoher Strukturanteil frei verfeinert. Hierzu wurden gut definierte Aminosäurereste ausgewählt, für die die bisherigen Restraints aufgehoben wurden. Die resultierende Struktur wurde zusätzlich mit `procheck` und `whatcheck` (siehe Abschn. 3.2 auf Seite 40) auf eventuelle Verzerrungen der Stereochemie überprüft. Für auffällige Strukturteile wurden die Beschränkungen wieder eingeführt, andere Atome wiederum freigelassen. In einem iterativen Verfahren wurde so versucht, möglichst viele Atome ohne Restraints, aber auch ohne zusätzliche Geometrieverzerrungen zu verfeinern. Letztlich genügten Restraints für folgende 37 Aminosäurereste, deren Verfeinerung mehr oder weniger stereochemischen Beschränkungen, zumeist der Seitenketten, unterlag: LYS 1, ARG 5, LYS 13–HIS 15, ASP 18–ASN 19, ARG 21, ASN 37, ALA 42–SER 50, ILE 55, ASN 59, ARG 61–TRP 63, ARG 68, PRO 70, ASN 77, ASP 87, ASN 93, LYS 97, ASP 101, GLN 121, TRP 123–GLY 126, ARG 128–LEU 129. Darüber hinaus wurde der Verschiebungsparameter des Stickstoffatoms von ASN 46 auf isotrope Form restraint und alle folgenden ADPs von ASN 46–GLY 49 darauf constraint. Ebenfalls nur eingeschränkt verfeinert wurden alle Solvensmoleküle: Ihre ADPs waren auf annähernd kugelförmige Gestalt restraint, Ethylenglykol und Acetat waren außerdem Abstands- und Winkelrestraints unterworfen. Zuletzt wurde eine *full matrix least square*-Verfeinerung der Struktur durchgeführt.

Parallel zur bereits ausgeführten Strategie wurde die Verfeinerung bis zur anisotropen Stufe zusätzlich gegen einen Datensatz durchgeführt, der im Schritt der Datenmittelung auf Strahlenschäden korrigiert worden war (*zero dose extrapolation*; Diederichs u. a. 2003). Da die Ergebnisse aber weder bezüglich Gütefaktoren noch Restdichten oder Temperaturfaktoren besser waren, wurde dieser Zweig der Verfeinerung nicht weiter verfolgt.

Das finale Modell umfaßt 2393 Atomlagen, wovon 2104 auf das Hauptmolekül und 289 auf die Solvensmoleküle entfallen. Es wurden nur 143 Atome (69 Nichtwas-

serstoff und 74 Wasserstoffatome) in einer weiteren Konformation modelliert. Zwar legten mehr Strukturteile Fehlordnung nahe, im Sinne von Restdichten, Geometrieverzerrungen und elongierten Verschiebungsparametern, allerdings führte die Modellierung eines entsprechenden minderpopulierten Teils nicht zu einer Verbesserung von Gütefaktoren, Geometrie oder Temperaturfaktoren, so daß es bei der Beschreibung einer einzigen Konformation belassen wurde. Die Hauptkonformation des Polypeptids besteht aus 1961 Lagen, die sich in 1001 Nichtwasserstoffatome und 960 Wasserstoffatome unterteilen.

Die Solvenssphäre setzt sich aus Wasser, Chlorid, Ethylenglycol und Acetat zusammen, wobei die beiden Letzteren inklusive Wasserstoffatomen beschrieben sind. Zwar ist auch die Anwesenheit von Natriumionen wahrscheinlich, die aber nicht lokalisiert werden konnten, da sie die gleiche Elektronenzahl wie Wasser haben und von diesem nur schwer anhand der jeweiligen Koordinationssphäre und evtl. dem Temperaturfaktor unterschieden werden können. In diesem Zusammenhang sei nochmal darauf hingewiesen, daß die Wassermoleküle nur durch Sauerstoffatome repräsentiert werden, weil es unmöglich ist, entsprechende Wasserstoffatompositionen zu berechnen, geschweige denn in der Dichte zu lokalisieren. Die zwischen 0,2 und 1,0 variierenden Besetzungszahlen der 210 Wassermoleküle summieren sich auf 165,84. Die sechs Chloridionen sind voll besetzt, wohingegen die zweite anionische Spezies der 3 Acetationen zum einen zu 89 % und zum anderen zu 76 % populiert ist, während das dritte Acetat in zwei Konformationen mit Besetzungen von 61 % und 39 % beschrieben ist. Der formalen Ladung des Lysozymmoleküls von +9 stehen also 8,65 negative Ladungen im Solvensbereich gegenüber. Auch von den vier Ethylenglykoleinheiten ist eines fehlgeordnet modelliert und in der Gesamtpopulation auf eins (0,7 + 0,3) constraint. Die verbleibenden drei Moleküle haben Besetzungen von 0,56, 0,62 und 0,88.

In Abb. 4.4.1 auf der gegenüberliegenden Seite ist das Gesamtmodell abgebildet. Die Atome sind nach Elementtyp koloriert und die Struktur ist mit einem semitransparenten farbigen Band überlagert, das den Verlauf der Lysozymhauptkette nachzeichnet und mit dem Farbverlauf von blau nach rot die Sequenz vom N-Terminus zum C-Terminus symbolisiert. Atome in Alternativkonformation sind als dünnere Stäbe dargestellt. Die 7 % Fehlordnung der Struktur verteilen sich auf neun Aminosäurereste: ARG 21, THR 43–ASN 44, ARG 61–TRP 62 sind vollständig fehlgeordnet, LYS 1, LYS 97, ASP 101, GLN 121 sind mit alternativen Seitenketten modelliert. Bezüglich der Fehlordnung sei noch auf Abb. 4.4.4 auf Seite 93 verwiesen, die unter anderem die fehlgeordneten Teile der Peptidsequenz parallel zu einer Darstellung der Sekundärstruktur dieses Moleküls zeigt.

In der finalen *full matrix*-Verfeinerung wurden Gütefaktoren von $R1(F_o > 4\sigma(F_o)) = 0,1255$, $R_{all} = 0,1314$ und $wR2 = 0,3083$ bei einem *goodness of fit* von 2,7 erreicht. Die Restelektronendichte beträgt im Mittel 0,00 mit einer quadratisch gemittelten Abweichung von $0,15 \text{ e}\text{\AA}^{-3}$, die Extremwerte liegen bei $-0,91$ und $1,44 \text{ e}\text{\AA}^{-3}$. Zur weiteren Beurteilung der Verfeinerung dient die nach Auflösungs-

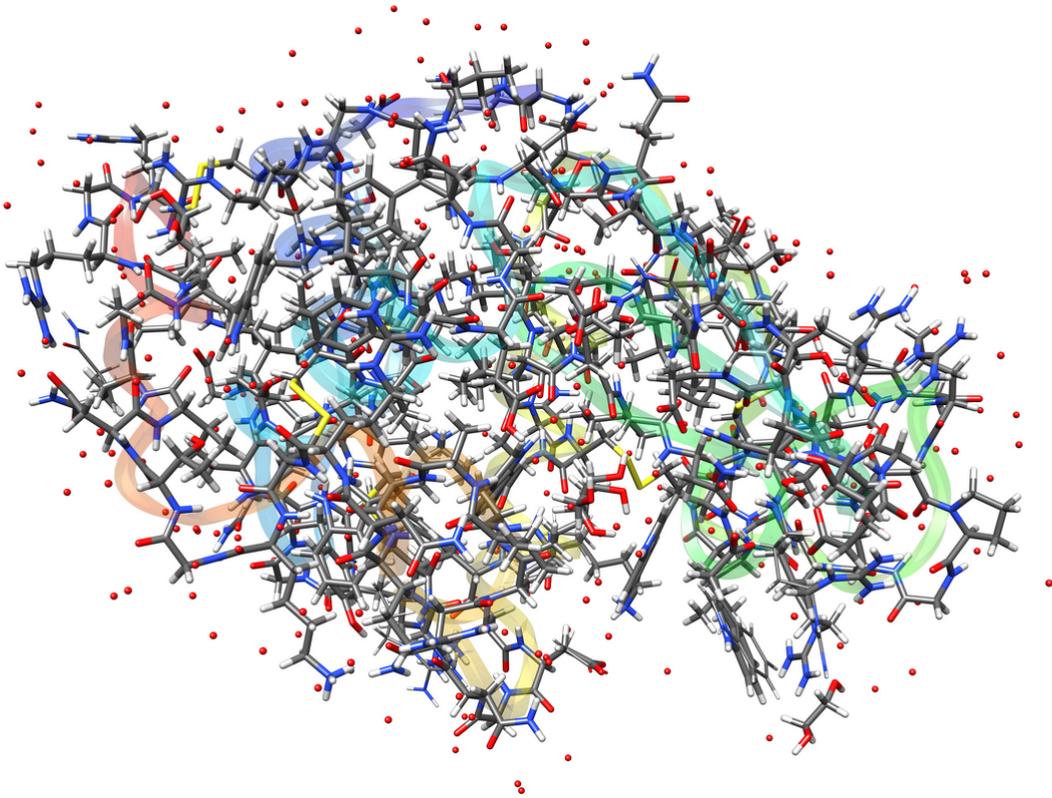


Abbildung 4.4.1: Strukturmodell TETRALYSO: Farbband vom N-Terminus (blau) zum C-Terminus (rot), fehlgeordnete Bereiche schmal gezeichnet.

schalen diskriminierte Reflexstatistik in Tab. 4.4.3 auf der folgenden Seite. Wieder ist die mit steigender Auflösung schnell abnehmende Reflexstärke am mittleren Verhältnis der Intensität zum Rauschen und dem Anteil beobachteter Reflexe zu erkennen. Interessant ist, daß sowohl bei hohen als auch bei niedrigen Beugungswinkeln die R -Werte deutlich schlechter werden. Während erstere Zunahme der allgemein geringeren Qualität der Intensitäten, bedingt durch ein schlechteres Signal-zu-Untergrund-Verhältnis, erklärt werden kann, ließe sich für letztere der mangelhaft beschriebene Einfluß der diffus verteilten Solvenssphäre geltend machen. Der gleiche Sachverhalt macht sich auch im Luzzati-Plot in Abb. C.5 auf Seite 368 bemerkbar.

Dort ist ebenfalls eine nach Aminosäureresten geordnete Darstellung der mittleren B -Werte und der mittleren Anisotropie pro Rest hinterlegt: Die Hauptkette hat meist B -Werte unter 10 \AA^2 , die Seitenketten etwas darüber. Auffallend höhere Temperaturfaktoren zeigen sich an den Resten nahe des C-Terminus und den Resten 46–49, sie erreichen einen mittleren B -Wert von 20. Die Anisotropie variiert zwischen 0,2 und 0,7, wobei die stärkere Abweichung von der Kugelform meist mit hohen B -Werten zusammenfällt. Ebenfalls in Abb. C.5 sind Diagramme des

Tabelle 4.4.3: Reflexstatistik TETRALYSO nach `shelxl` inklusive aller Korrekturen.

Auflösung/Å	$N_{unabh.}$	$\overline{I/\sigma(I)}$	$(I > 2\sigma(I))/\%$	Vollständigkeit/%	$\sum(F_o)/\sum(F_c)$	R1(all)	R1($F_o > 4\sigma(F_o)$)
$\infty-5,000$	598	148,69	98,68	99,01	0,767	0,484	0,484
5,000-4,000	526	158,51	99,62	100,00	1,026	0,170	0,170
4,000-3,000	1444	151,71	99,72	100,00	1,104	0,160	0,159
3,000-2,000	5726	159,16	99,79	99,93	1,058	0,129	0,129
2,000-1,800	2986	117,12	99,43	99,90	1,045	0,113	0,113
1,800-1,600	4623	109,96	99,29	99,91	1,029	0,103	0,103
1,600-1,400	7615	95,68	98,78	99,97	1,012	0,095	0,094
1,400-1,200	13470	75,45	98,00	99,96	0,988	0,091	0,089
1,200-1,000	26294	46,89	95,70	99,98	0,986	0,093	0,088
1,000-0,950	10365	23,48	91,33	99,98	1,002	0,110	0,098
0,950-0,900	12763	15,61	86,77	100,00	1,015	0,123	0,104
0,900-0,880	5908	10,48	81,33	100,00	1,028	0,148	0,118
0,880-0,860	6509	8,80	78,61	100,00	1,029	0,155	0,121
0,860-0,840	7121	6,67	72,48	100,00	1,042	0,177	0,132
0,840-0,820	7840	4,85	65,75	99,96	1,029	0,215	0,143
0,820-0,800	8599	3,77	59,15	99,88	1,052	0,260	0,169
$\infty-0,800$	122388	47,54	87,90	99,96	1,016	0,131	0,125

Zusammenhangs zwischen B -Wert und Unsicherheit von Atomlage und Bindungslänge gezeigt: Die Standardabweichung der Atomkoordinaten reicht von ca. 0,003 Å bis 0,075 Å und steigt mit zunehmendem Temperaturfaktor an. Die meisten Atome haben Fehler um 0,01 Å. Die Unsicherheit der Bindungslängen steigt steiler mit den B -Werten, und liegt zwischen 0,005 und 0,022 Å.

Zur weiteren Beurteilung der Qualität liegen die Ausgaben von `procheck` und `whatcheck` in Abb. C.3 und C.4 ab der Seite 360 vor. Der Ramachandran-Plot zeigt 89,4% der Aminosäurereste in der energetisch vorteilhaftesten Region und 10,6% in der zweitbesten. Bei 21 Resten werden auffällig von Engh&Huber-Werten abweichende Bindungslängen und -winkel der Hauptkette angezeigt. Wie schon bei P1LYSO werden auch hier einige Abweichungen von der Planarität der Peptidbindung gefunden: 7 Reste haben einen um zwei Standardabweichungen vom Mittel der Teststatistik abweichenden ω -Torsionswinkel. Die `whatcheck`-Ausgabe enthüllt noch einige Geometrieverzerrungen der Seitenketten, ist aber insgesamt

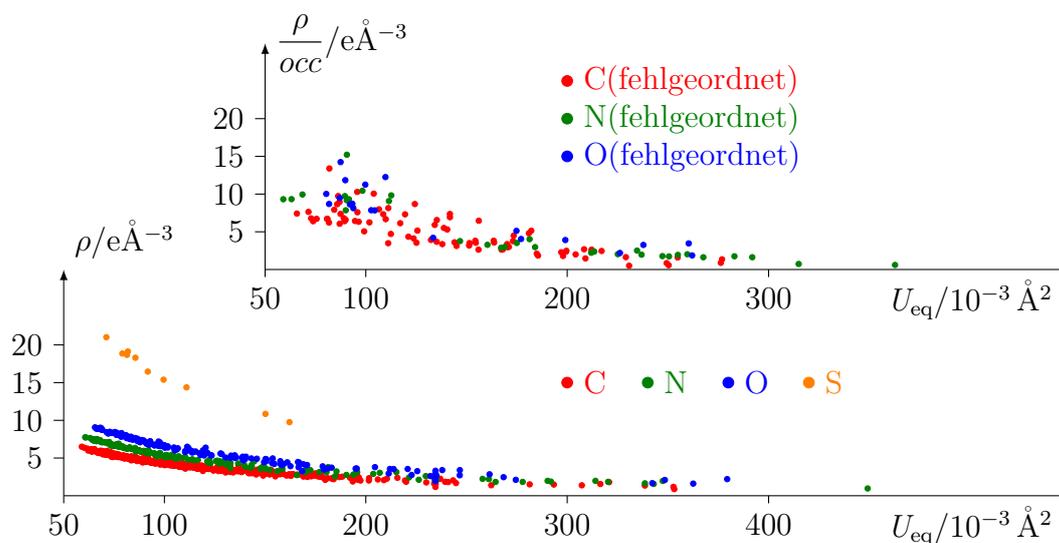


Abbildung 4.4.2: $F_o\phi_c$ -Elektronendichte des sphärischen Modells an den Positionen voll besetzter und fehlgeordneter Atome.

auch zufriedenstellend. Die problematischen Strukturteile sind zumeist solche, die oben als restraint verfeinert genannt wurden, insbesondere wurden keine zusätzlichen Verzerrungen durch die Lockerung der stereochemischen Beschränkungen verursacht.

Als letzte Eigenschaft zur Beurteilung der Struktur soll die Sevcik-Darstellung in Abb. 4.4.2 dienen (siehe dazu auch die Erläuterung in Abschn. 3.4 auf Seite 45 und die Ergänzungen in Abschn. 4.2.1 auf Seite 71). Sie zeigt die Abnahme der beobachteten Elektronendichte an den Atompositionen mit zunehmendem Temperaturfaktor und unterschieden nach Elementtyp. Die Proportionalität der Elektronendichte entsprechend der Ordnungszahl der Elemente ist zumindest bei den Atomen mit geringerem U_{eq} gut erkennbar. Bis zu einem Temperaturfaktor von ca. $0,180 \text{ \AA}^2$ lassen sich die Elementtypen der geordneten Atome anhand der beobachteten Dichte unterscheiden, darüberhinaus ist dies nicht mehr möglich. An den fehlgeordneten Atompositionen hingegen, ist keine Aussage zum Elementtyp möglich.

Einen Packungsausschnitt der ab -Ebene zeigt Abb. 4.4.3 auf der nächsten Seite. Dabei sind fraktionelle Koordinaten in a -, b - und c -Richtung von $-0,5$ bis $1,5$, $-0,1$ bis $1,1$ und 0 bis 1 dargestellt. Die Kolorierung der acht Proteineinheiten der Elementarzelle ist dabei willkürlich gewählt. Jede Zelle enthält acht Lysozymmoleküle, sie werden durch die acht Symmetrieelemente in $P4_32_12$ erzeugt: In dieser Darstellung verlaufen die 4_3 - und 4_1 -Schraubenachsen senkrecht zur Papierebene bei $0, \frac{1}{2}, z$ und $\frac{1}{2}, 0, z$. In gleicher Richtung bei $0, 0, z$ liegt außerdem eine 2_1 -Schraubenachse. Weitere 2_1 -Schraubenachsen liegen senkrecht dazu, parallel der Papierebene, entlang der a - und b -Achsen bei jeweils $x, \frac{1}{4}, \frac{1}{8}$ und $\frac{1}{4}, y, \frac{3}{8}$.

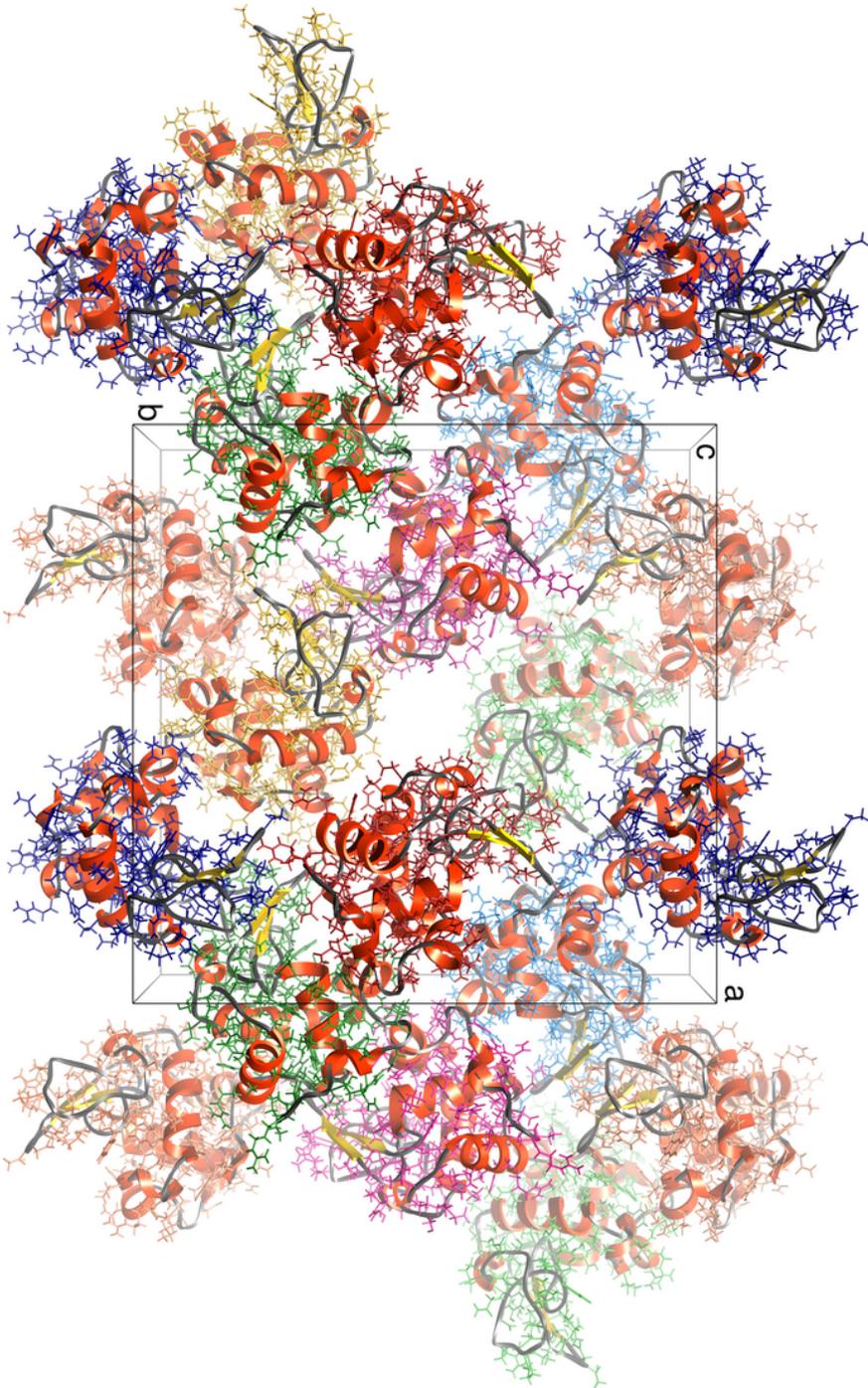


Abbildung 4.4.3: Ausschnitt der Packung von TETRAISO in der *ab*-Ebene. Helices (rot) und β -Stränge (gelb) in Cartoondarstellung, Seitenketten in Stabdarstellung.

Es verbleiben zwei diagonal orientierte 2-Drehachsen entlang $x, x, 0$ und $x, -x, 1/4$, sowie die Identität.

Wie der gekürzten `whatcheck`-Ausgabe C.3 auf Seite 360 entnommen werden kann, berechnet sich die Packungsdichte in Form des Matthews-Koeffizienten zu $1,980 \text{ \AA}^3/\text{Da}$. Nach der Näherungsformel Gl. 2.2 auf Seite 32 entspricht dies einem Solvensanteil von 38 %, mit anderen Worten, nur 62 % des Zellvolumens werden vom Protein beansprucht. Eine analoge Betrachtung mittels des Quotienten des Volumens der asymmetrischen Einheit $V_{asu} = 28729 \text{ \AA}^3$ und des van-der-Waals-Volumens $V_{vdW} = 13927 \text{ \AA}^3$ des Lysozymmoleküls, ergibt mit $V_{asu}/V_{vdW} = 2,063$ einen prozentualen Anteil des Proteins von 49 % am Kristallvolumen.

4.4.2 Asphärisches Modell

Das asphärische Streumodell wurde wie in Abschn. 4.1.2 auf Seite 53 beschrieben erstellt. Nach Invariomtransfer und Elektroneutralitätsausgleich trägt die Struktur eine Gesamtladung von 0,0785 e. Zur Auswahl des zu verfeinernden Strukturteils wurde sich an der sphärischen `shelxl`-Verfeinerung orientiert: Da nur Constraints in XD zur Verfügung stehen, wurden alle Lysozymatome festgehalten, die zuvor in Restraints involviert waren, ebenso wie sämtliche Solvensmoleküle. Keines der fehlgeordneten Atome wurde zur Verfeinerung freigegeben. Zeigten sich dann zu hohe Verschiebungen der Lageparameter oder Temperaturfaktoren, wie z. B. für ARG 112, so wurden auch diese Atome auf den `shelxl`-Werten fixiert. Die Auswahl ist in Abb. 4.4.4 auf Seite 93 dargestellt, Symbole und Farbschema wurden bereits in Abschn. 4.3.2 auf Seite 79 zur analogen Graphik 4.3.4 erläutert. Eine Entsprechung der gezeigten Sekundärstruktur findet sich in Abb. 4.4.3, bei gleicher Kolorierung der Elemente Helix, Faltblatt und Knäuel. Offensichtlich enthalten helikale Bereiche der Struktur weniger Aminosäurereste in Alternativkonformationen als Schleifen- und Knäuelbereiche. Letztendlich wurden 751 der insgesamt 2393 Atomlagen verfeinert, das entspricht 75 % aller Nichtwasserstoffatome der Hauptkonformation des Lysozymmoleküls. Hiermit ist eine Halbierung der Parameterzahl von 11930 bei 20689 Restraints im `shelxl`-Fall auf 6760 Parameter und null Restraints in der XD-Verfeinerung verbunden.

Zur Verfeinerung der Modelle `mul`, `iam` und `iamrb` wurden jeweils zwölf Verfeinerungszyklen gegen $F_o^2 > 2\sigma(F_o^2)$ mit dem `shelxl`-Gewichtungsschema gerechnet. Die dabei verbleibenden maximalen Parameterverschiebungen liegen in den zwei Modellen mit verlängerten Wasserstoffabständen in der Größenordnung von 10^{-4} , aber bei nur 10^{-5} im Fall der Wiederholung der klassischen sphärischen Verfeinerung. Dabei sind jeweils die anisotropen Verschiebungsparameter von Seitenkettenatomen betroffen. Mittlere Verschiebungen/Standardabweichung sind zusammen mit den erreichten Gütefaktoren in Tab. 7.1 auf Seite 187 gegeben.

4 Verfeinerung

LYS 001	VAL 002	PHE 003	GLY 004	ARG 005	CYS 006	GLU 007	LEU 008
HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK
✓ C δ	✓ ✓	✓ ✓	✓ ✓	✓ X	✓ ✓	✓ ✓	✓ ✓
ALA 009	ALA 010	ALA 011	MET 012	LYS 013	ARG 014	HIS 015	GLY 016
HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK
✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ X	✓ X	✓ X	✓ ✓
LEU 017	ASP 018	ASN 019	TYR 020	ARG 021	GLY 022	TYR 023	SER 024
HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK
✓ ✓	✓ C β	✓ X	✓ ✓	X X	✓ ✓	✓ ✓	✓ ✓
LEU 025	GLY 026	ASN 027	TRP 028	VAL 029	CYS 030	ALA 031	ALA 032
HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK
✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓
LYS 033	PHE 034	GLU 035	SER 036	ASN 037	PHE 038	ASN 039	THR 040
HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK
✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ X	✓ ✓	✓ ✓	✓ ✓
GLN 041	ALA 042	THR 043	ASN 044	ARG 045	ASN 046	THR 047	ASP 048
HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK
✓ ✓	X X	X X	X X	X X	X X	X X	X X
GLY 049	SER 050	THR 051	ASP 052	TYR 053	GLY 054	ILE 055	LEU 056
HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK
X X	-N ✓	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ X	✓ ✓
GLN 057	ILE 058	ASN 059	SER 060	ARG 061	TRP 062	TRP 063	CYS 064
HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK
✓ ✓	✓ ✓	✓ X	✓ ✓	X X	X X	X X	✓ ✓
ASN 065	ASP 066	GLY 067	ARG 068	THR 069	PRO 070	GLY 071	SER 072
HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK	HK SK
✓ ✓	✓ ✓	✓ ✓	✓ X	✓ ✓	✓ X	✓ ✓	✓ ✓

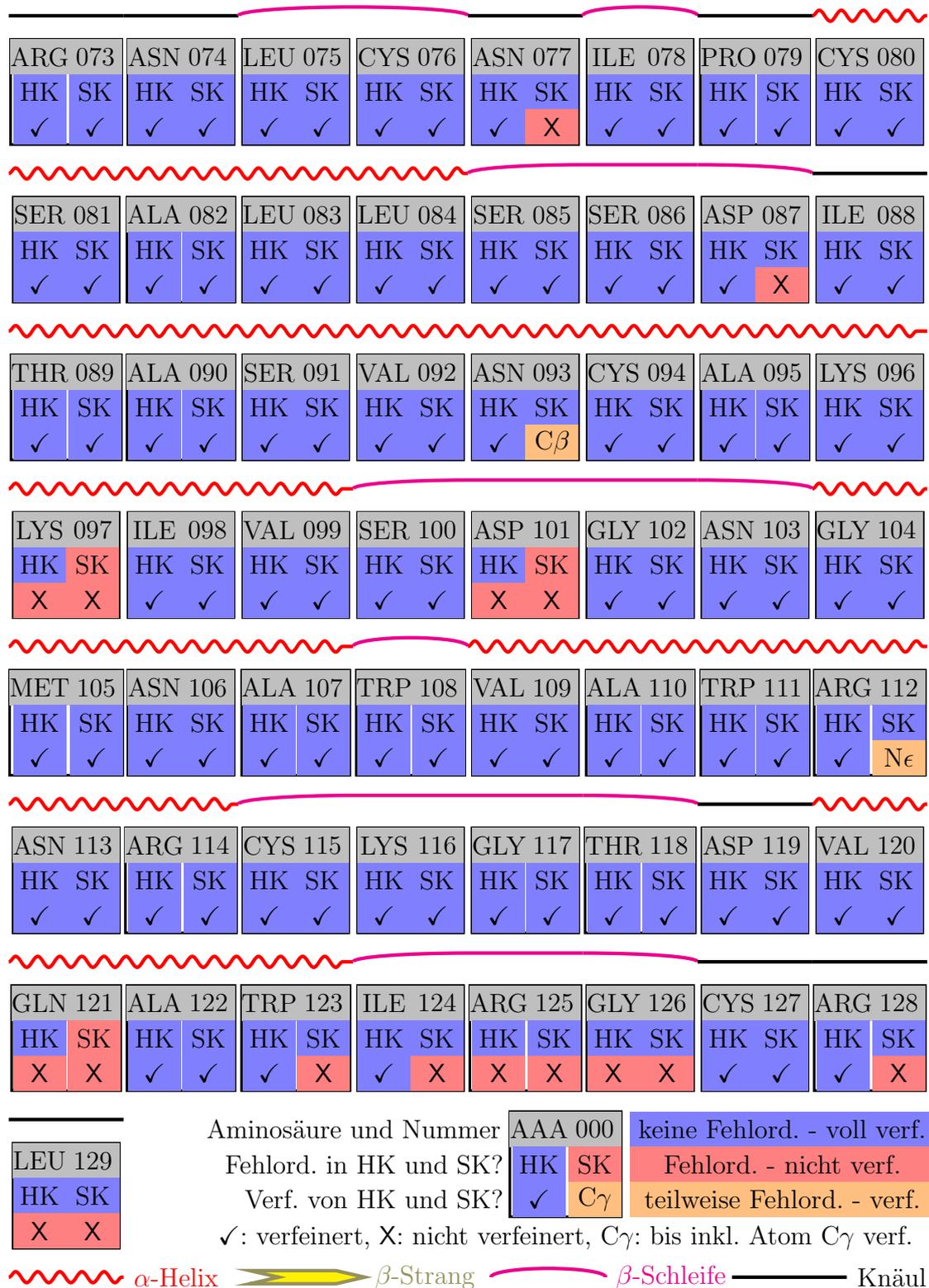


Abbildung 4.4.4: DSSP-Zuordnung der Sekundärstruktur von TETRALYSO, Verteilung fehlgeordneter Bereiche und verfeinerter Strukturteil.

4.5 Insulin

Der hier verwendete INSULIN-Datensatz wurde als bester von sechs Versuchen in Kooperation mit Dr. Armin Wagner an der Diamond Light Source (Großbritannien) gemessen. Mit einem ADSC Q315-Flächendetektor wurden im Stickstoffkaltgastrom bei einer Wellenlänge von $\lambda = 0,7293 \text{ \AA}$ insgesamt 593124 Reflexe bis zu einer Auflösung $d = 0,90 \text{ \AA}$ aufgenommen. Datenintegration, Mittelung und Skalierung in der rhomboedrischen Raumgruppe R3 mit `xds` (Kabsch 1993) ergab einen durchschnittlich 9,5-fach redundanten Satz von 62154 symmetrieunabhängigen Beobachtungen, der über den gesamten Auflösungsbereich zu 100 % vollständig ist. Hierbei erfüllen 87 % der Daten das Beobachtungskriterium von $I > 2\sigma(I)$. Der Gütefaktor der gemittelten Daten R_{mrgd} liegt bei guten 4 %, allerdings läßt die Reflexqualität ab 1 \AA Auflösung schnell nach und der R_{mrgd} steigt auf zuletzt 68 %. Kongruent dazu ist das Verhalten des Signal-zu-Untergrund-Verhältnisses, das zwar im Mittel einen guten Wert von 26 erreicht, aber bei höherer Auflösung schnell auf nur noch 2,5 sinkt. Dagegen ist die Abnahme der Redundanz auf ein Drittel des Mittelwertes moderat. Somit ist das letzte Viertel des Datensatzes von wesentlich geringerer Güte als die bis 1 \AA reichenden Daten. Allerdings wird eine hohe Auflösung für unabdingbar zur erfolgreichen Durchführung einer asphärischen Verfeinerung gehalten, so daß sie dennoch inkludiert wurden.

Tabelle 4.5.1: Kristallographische Daten INSULIN.

Summenformel	$\text{C}_{514}\text{N}_{130}\text{O}_{154}\text{S}_{12}\text{H}_{762} \cdot 2 \text{Zn}^{2+}, 1 \text{C}_2\text{O}_2\text{H}_3^-, 1 \text{Na}^+, 173 \text{O}$		
Formelmasse	13551,8 g/mol	Absorptionskoeff.	$0,08 \text{ mm}^{-1}$
Kristallsystem	Rhomboedrisch	Meßtemperatur	100 K
Raumgruppe	R3 (Nr. 146)	Wellenlänge λ	$0,7293 \text{ \AA}$
Z	9	max. 2Θ	$47,80^\circ$
Gitterkonstanten:		Auflösungsbereich	$10\text{--}0,90 \text{ \AA}$
a	$81,619(82) \text{ \AA}$	Grenzreflexe	$-90 \leq h \leq 90$
b	$81,619(82) \text{ \AA}$		$-78 \leq k \leq 78$
c	$33,786(34) \text{ \AA}$		$-37 \leq l \leq 35$
α	$90,00^\circ$	unabh. Reflexe	62154
β	$90,00^\circ$	$F_o^2 > 2\sigma(F_o^2)$	54157
γ	$120,00^\circ$	Abdeckungsgrad [†]	100 (99,9) %
V	$194917,05 \text{ \AA}^3$	Redundanz [†]	9,5 (3,5)
Dichte	$1,039 \text{ g/cm}^3$	$I/\sigma(I)$ [†]	26,27 (2,58)
F(000)	64407,1	R_{mrgd} [†]	4,1 (68,0) %

[†] Angabe in Klammern für Auflösungsschale $0,91\text{--}0,90 \text{ \AA}$.

Tabelle 4.5.2: Verfeinerungsschritte INSULIN.

Verfeinerungsstufe	$d/\text{\AA}$	N_{prot}	N_{solv}	N_{par}	N_{restr}	$R1/R_{free}/\%$
epmr: initiales Modell	4,0	412	2	—	—	65,2
Isotrop	1,5	412	2	1653	1608	43,51/50,26
Modellierung	1,5	744	8	3009	3103	22,45/28,84
Solvens	1,5	765	52	3265	3227	19,83/25,20
Auflösungserhöhung + ADP	0,90	845	48	8040	10627	15,26/17,93
Solvens + Besetzung	0,90	998	147	10392	15263	12,71/16,20
H-fix	0,90	2023	158	10788	16420	11,14/14,68
Restraints lockern	0,90	2023	158	10788	14970	11,13/14,65
<i>full matrix least square</i>	0,90	2055	183	11188	15501	10,95

4.5.1 Sphärisches Modell

Um mittels molekularer Ersetzung ein geeignetes Startmodell zu erhalten, wurden sechs bekannte Insulinstrukturen mit ähnlichen Zellkonstanten getestet. Letztlich wurden die Hauptkettenkoordinaten der bis 1 Å aufgelösten Struktur 1MSO (Smith u. a. 2003) verwandt, da sie die beste Anpassung mit epmr (Kissinger 2008) an den Datensatz und die aktuellen Gitterkonstanten ergaben. Aus den so verfeinerten Koordinaten des Molekülfragments wurde mit shelxpro die Eingabedatei für shelxl erstellt. Davon ausgehend, wurde die Struktur in insgesamt 46 Verfeinerungsrunden schrittweise aufgebaut, hierbei wurden von Anfang an Alternativkonformationen und Solvensmoleküle in das Modell übernommen, sobald sie in der Dichte- bzw. Restdichtekarte erkannt werden konnten. Bezüglich technischer Details, gilt das einleitend in Absch. 4.1.1 auf Seite 51 gesagte, sowie ergänzend die Ausführungen zur Verfeinerung von TETRALYSO in Absch. 4.4 auf Seite 83.

Tab. 4.5.2 faßt wichtige Zwischenstufen zusammen: Nach Verfeinerung der Koordinaten und Temperaturfaktoren des Proteinrückgrats mit den Daten bis 1,5 Å, wurden ungefähr 90 % des Hauptmoleküls modelliert. Daran schloß sich eine automatische Suche nach diskreten Wasserpositionen mit shelxwat an, der Beitrag diffuser Solvensverteilung wurde durch die SWAT-Parameter berücksichtigt. Mit Verfeinerung anisotroper Verschiebungsparameter bei sukzessiver Hinzunahme aller Reflexe, wurden mehr Positionen des Insulinmoleküls entdeckt. Erst durch die Erweiterung der Solvenssphäre und Verfeinerung der Besetzung aller Wassermoleküle wurde das Modell so weit verbessert, daß alle Nichtwasserstoffatome des Proteins bestimmt werden konnten. Im Anschluß wurde die Struktur um Wasserstoffatompositionen ergänzt. Sofern ohne zusätzliche Geometrieverzerrungen möglich, wurden nunmehr Restraints für Atome des Hauptmoleküls gelockert. Nach Übergang von der bisherigen *conjugate gradient least square*-Verfeinerung zu *full*

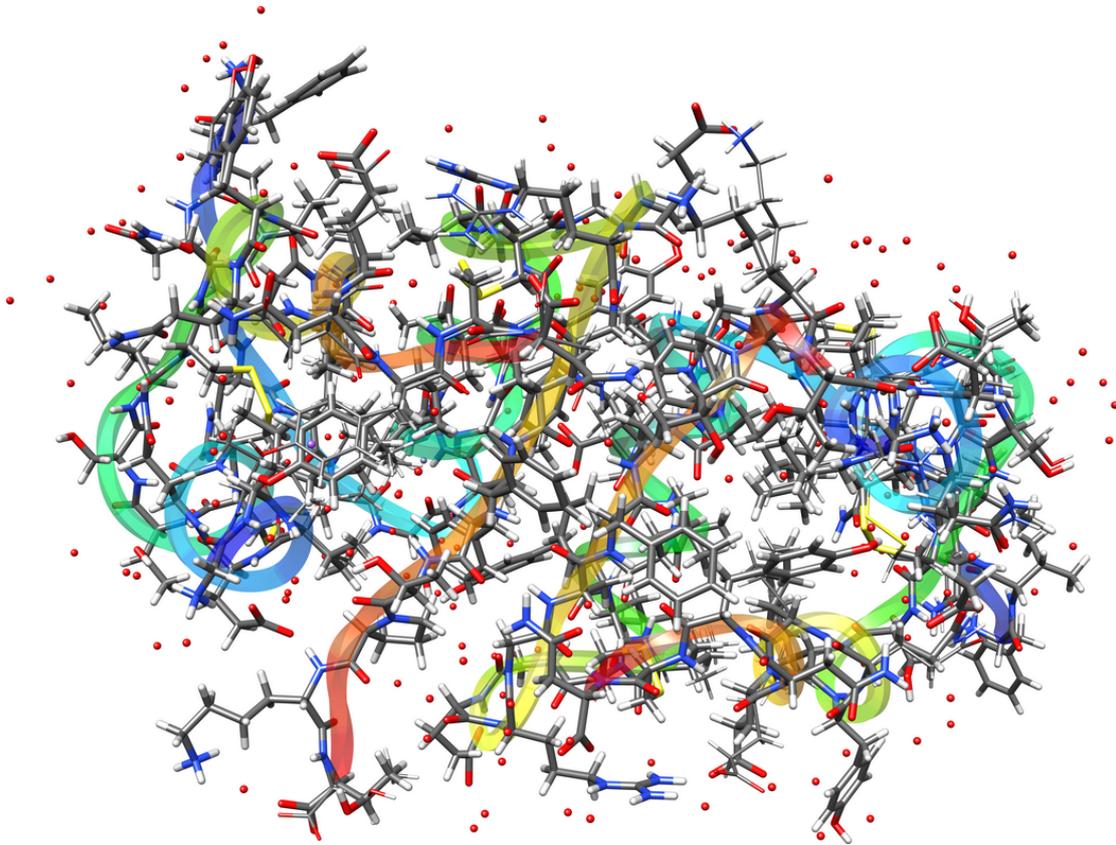


Abbildung 4.5.1: Strukturmodell INSULIN: Atome nach Element gefärbt, fehlgeordnete Bereiche schmal, Farbbänder jeweils vom N-Terminus (blau) zum C-Terminus (rot).

matrix least square-Methoden, erfolgten noch einige Anpassungen der Fehlordnungsmodellierung, des Solvensbereichs und der Auswahl an Restraints, um final ein 2238 Atomlagen umfassendes Strukturmodell zu ergeben, welches in Abb. 4.5.1 dargestellt ist.

Im Gegensatz zu allen anderen behandelten Strukturen liegt im Insulin keine durchgehende Sequenz vor, sondern vier Proteinketten A–D. Je zwei Ketten (A, B und C, D) sind über Disulfidbrücken verknüpft. Die Sequenz der Ketten A und C beinhaltet jeweils identische 21 Aminosäurereste, die der Ketten B und D jeweils 30. Allerdings gibt es Konformationsunterschiede zwischen den Ketten A,B und C,D, so daß die asymmetrische Einheit aus einem Dimer mit insgesamt 102 Resten besteht.

Auf das Hauptmolekül entfallen 2055 Positionen, wovon 1572 zur Hauptpopulation und 483 zu Alternativkonformationen zählen. Hierbei handelt es sich um 810 Nichtwasserstoffatome, von denen 229 in mehr als einer Konformation vorliegen. Inklusive Wasserstofflagen liegt der Fehlordnungsanteil also bei 30,7% und nur auf

die Schweratome bezogen bei 28,3%. Dieser verteilt sich auf 33 der insgesamt 102 Aminosäurereste: ILE 102, VAL 103, TYR 114, GLU 117, ASN 118, TYR 119, CYS 120, ASN 121 in der Kette A; LEU 206, LEU 211, GLU 213, ARG 222, THR 230 in der Kette B; GLY 301, ILE 302, VAL 303, GLU 304, GLN 305, CYS 306, CYS 307, THR 308, SER 309, CYS 311, LEU 313, GLN 315, LEU 316, GLU 317 in der Kette C und ASN 403, GLN 404, SER 409, VAL 412, VAL 418, LYS 429 in der Kette D. Offensichtlich ist die dritte Kette C doppelt so stark von Fehlordnung betroffen wie die drei anderen Ketten, siehe hierzu auch Abb. 4.5.5 auf Seite 103. Von den 102 Resten konnten nur 41 vollständig ohne Restraints verfeinert werden, ohne daß zusätzliche Verzerrungen der Stereochemie auftraten. Weitere 27 Aminosäurereste wurden teilweise von den Beschränkungen befreit und die verbleibenden 34 mußten vollständig restraint bleiben. Hierin spiegelt sich ungefähr die oben beschriebene Fehlordnung wieder, als für die Ketten B und D besonders viele Reste frei verfeinert wurden und besonders wenige in der Kette C.

An die Struktur sind zwei Zinkionen koordiniert, die, da sie auf einer dreizähligen Achse liegen, zu jeweils einem Drittel besetzt sind. Hinzu kommen ein zu je 82% populiertes Acetat- und Natriumion. Komplettiert wird die Solvenssphäre durch 173 Wasserlagen, deren Besetzungszahlen zwischen 0,15 und 1 schwanken und sich auf 115,7 summieren. Somit wird die Formalladung -4 der Insulinketten durch entgegengesetzte $+4$ des Solvensbereiches neutralisiert. Alle Solvensmoleküle wurden mit Restraints verfeinert: Die Verschiebungsparameter auf annähernde Kugelform, Bindungslängen und -winkel des Acetatanions sowie Abstände der Wassermoleküle untereinander.

Die Anpassung von 11188 Parametern bei 15501 Restraints an die Daten mündete in Gütefaktoren von $R1(F_o > 4\sigma(F_o)) = 0,1095$, $R_{all} = 0,1150$ und $wR2 = 0,2782$. Der *goodness of fit* lag bei 2,419, die SWAT-Parameter bei $g = 0,834$, $U = 2,514$. Im Mittel liegt die Restelektronendichte bei $0,00 \text{ e}\text{\AA}^{-3}$ mit einer quadratisch gemittelten Abweichung von $0,08 \text{ e}\text{\AA}^{-3}$. Das Minimum und Maximum der Differenzfouriersynthese liegen bei $-0,68$ und $0,49 \text{ e}\text{\AA}^{-3}$. Eine weitere Sicht auf die Verfeinerung ermöglicht Tab. 4.5.3 auf der folgenden Seite mit einer nach Auflösungsbereichen aufgeschlüsselten Reflexstatistik: Zu kleinsten und insbesondere zu höchsten Beugungswinkeln ist die Übereinstimmung zwischen berechneten und beobachteten Strukturparametern besonders schlecht. Letzteres korrespondiert mit der eingangs erfolgten Kritik zur Datenqualität, ebenso wie die hier wiedergegebenen Werte zur Stärke der Reflexe. Da die Daten geringer Auflösung insbesondere von diffuser Elektronendichte beeinflusst werden, ist zu vermuten, daß eine ungenügende Beschreibung des Solvensbereiches zu den schlechten R -Werten in diesen Schalen ursächlich beiträgt.

Neben den Gütefaktoren dienen die in den Abb. D.3, D.4 und D.5 ab Seite 376 aufgeführten stereochemischen Auswertungen zur Beurteilung der Struktur: Die Ramachandran-Auftragung zeigt 94% der Aminosäurereste in energetisch vorteilhaftesten Konformationen und 6% in erlaubten Regionen. Allerdings werden bei ungefähr einem Drittel der Reste einige Bindungslängen und -winkel

Tabelle 4.5.3: Reflexstatistik INSULIN nach `shelxl` inklusive aller Korrekturen.

Auflösung/Å	$N_{unabh.}$	$\overline{I/\sigma(I)}$	$(I > 2\sigma(I))/\%$	Vollständigkeit/%	$\sum(F_o)/\sum(F_c)$	R1(all)	R1($F_o > 4\sigma(F_o)$)
$\infty-5,000$	311	29,13	85,44	85,44	1,070	0,239	0,239
5,000-4,000	343	31,89	100,00	100,00	1,044	0,140	0,140
4,000-3,000	971	36,81	100,00	100,00	1,050	0,127	0,127
3,000-2,000	3981	49,94	100,00	100,00	1,018	0,109	0,109
2,000-1,800	2116	48,28	100,00	100,00	1,001	0,097	0,097
1,800-1,600	3306	43,41	99,91	100,00	0,995	0,092	0,092
1,600-1,400	5467	39,34	99,65	100,00	0,999	0,085	0,085
1,400-1,200	9720	31,43	98,23	100,00	1,006	0,091	0,090
1,200-1,000	19111	18,60	92,05	100,00	1,014	0,108	0,099
1,000-0,980	2839	7,04	76,86	100,00	1,006	0,158	0,121
0,980-0,960	3067	5,85	71,41	100,00	1,020	0,174	0,129
0,960-0,940	3315	4,77	66,00	100,00	1,016	0,204	0,139
0,940-0,930	1753	3,89	59,25	99,77	1,004	0,232	0,145
0,930-0,920	1889	3,27	55,13	99,84	1,016	0,253	0,158
0,920-0,910	1932	2,71	49,90	100,00	1,011	0,293	0,178
0,910-0,900	2033	2,46	46,04	99,90	1,021	0,302	0,182
$\infty-0,900$	62154	23,38	87,05	99,90	1,016	0,115	0,110

als signifikant von Idealwerten abweichend gefunden. Dazu kommen Warnungen ob der Nichtplanarität etlicher Peptidbindungen und Seitenketten. Konsequenterweise ergibt sich ein nicht so guter G-Faktor von -0,16 bei Beurteilung der Strukturnormalität. Auch fällt die Einschätzung der Packungsqualität mittelmäßig aus.

Die mit dem Temperaturfaktor zunehmende Standardabweichung der Atomkoordinaten liegt meist zwischen 0,01 und 0,05 Å, die der Bindungslängen zwischen 0,006 und 0,03 Å, allerdings teilweise durch die verwendeten Restraints gedämpft. Das paßt zur Luzzati-Auftragung, anhand der die durchschnittliche Positionsunsicherheit zu 0,05 Å abgeschätzt wird. Über jeden Aminosäurerest gemittelt, liegen die B -Werte der Ketten B und D bei ungefähr 10 Å², in den Ketten A und C etwas darüber. Zu Beginn und Ende einer Kette steigen sie aber auf bis zu 30 an, sind ansonsten aber recht gleichförmig verteilt. In den Seitenketten liegen die Verschiebungsparameter jeweils ungefähr um die Hälfte höher, auch treten stärkere

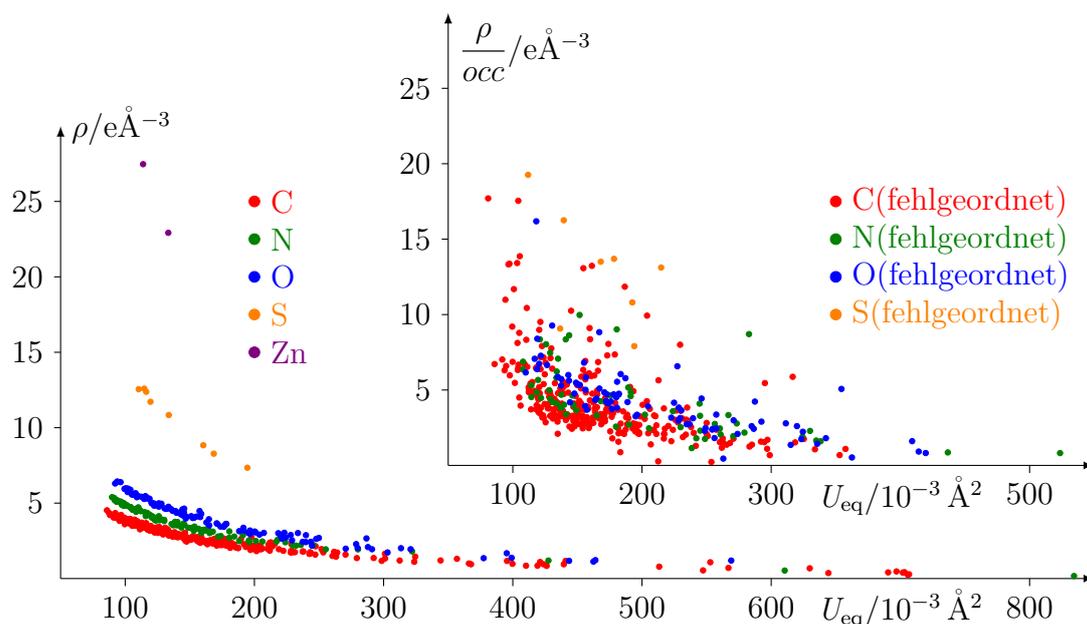


Abbildung 4.5.2: $F_o\phi_c$ -Elektronendichte des sphärischen Modells an den Positionen voll besetzter und fehlgeordneter Atome.

Schwankungen auf. In der gleichen Mittelung variieren die Anisotropiewerte zwischen 0,2 und 0,8, in den Seitenketten sind die Werte etwas niedriger, entsprechend weniger kugelförmiger Auslenkungsellipsoide.

Eine in Absch. 3.4 auf Seite 45 beschriebene Auftragung der Elektronendichte an den Atompositionen gegen den entsprechenden Temperaturfaktor findet sich in Abb. 4.5.2. Den fünf Elementen entsprechend, sind fünf Punktwolken zu erkennen, deren Werte zu gegebenem U_{eq} eine Proportionalität analog den Ordnungszahlen erkennen lassen. Des weiteren sind die Cluster bis zu einem Temperaturfaktor von ca. $200 \cdot 10^{-3} \text{ \AA}^2$ voneinander gut zu unterscheiden. Bei zu hohen U_{eq} oder bei Fehlordnung (rechtes oberes Diagramm) aber, ist in einer solchen Auftragung keine Entscheidung über den Elementtyp aus den Daten allein mehr möglich. In diesem Fall ist eine verfeinerte Betrachtung dieser Atome im asphärischen Invariomformalismus wahrscheinlich insignifikant.

Einen graphischen Eindruck der Packung soll Abb. 4.5.3 auf der nächsten Seite vermitteln, dargestellt ist die säulenartige Stapelung von Insulin-Hexameren entlang der dreizähligen Achsen senkrecht zur Papierebene. Dabei ist jedes der asymmetrischen Einheit entsprechende Dimer individuell gefärbt. Die Zinkionen und die Solvensmoleküle verteilen sich innerhalb und außerhalb der Röhren. Neben der parallel zu z liegenden dreizähligen Achse gibt es in gleicher Richtung 3_1 - und 3_2 -Schraubenachsen, so daß die drei Hexamere je Zelle, und damit die abgebildeten Stapel, entlang z um je $1/3$ und $2/3$ gegeneinander verschoben sind.

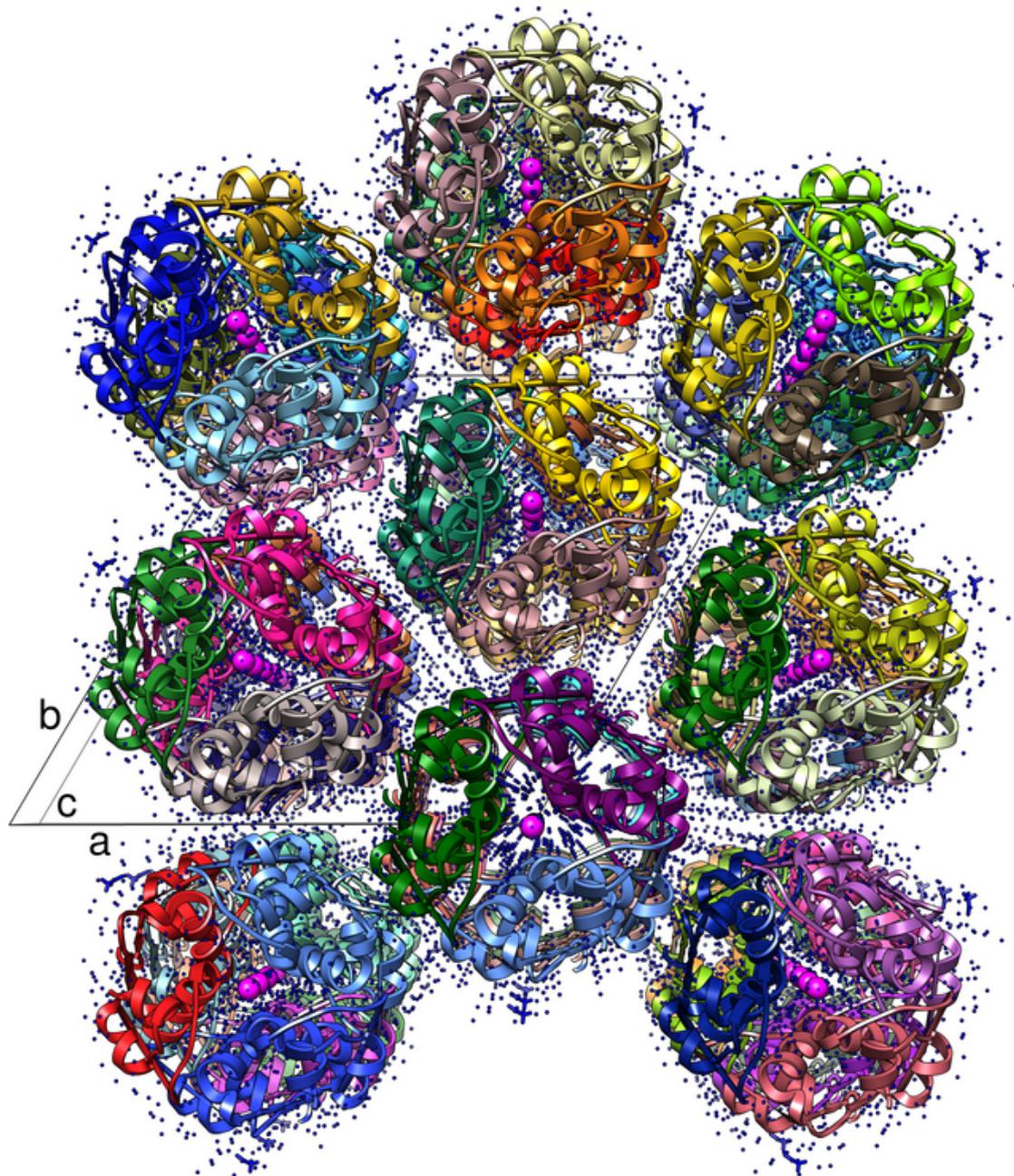


Abbildung 4.5.3: Ausschnitt der INSULIN-Packung. Blick auf ab -Ebene, drei Translations-einheiten in c -Richtung. Willkürliche Färbung der asymmetrischen Einheiten, Zink magenta, Solvessphäre blau.

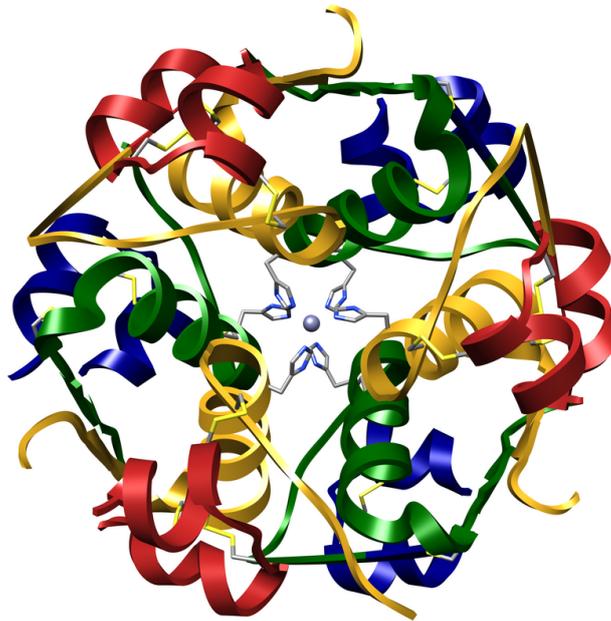


Abbildung 4.5.4: INSULIN-Hexamer: Blick entlang der dreizähligen Achse. Darstellung der Sekundärstruktur, der Disulfidbrücken und der Koordination des Zinks durch Histidin. Ketten A und C - blau und rot, Ketten B und D - grün und gelb.

In Abb. 4.5.4 ist die hexamere Quartärstruktur nochmal separat gezeigt und jede der vier Ketten ist individuell koloriert. Das Dimer aus den beiden Insulineinheiten A/B (blau/grün) und C/D (rot/gelb) wird zum Hexamer verdreifacht. Damit wird jedes der zwei Zinkionen von je drei Histidinresten umgeben, nichtabgebildete Wassermoleküle komplettieren die tetraedrische Koordination.

Der sich aus dem Quotienten des Volumens der asymmetrischen Einheit und dem Molekulargewicht der Proteinstruktur berechnende Matthews-Koeffizient beträgt laut Abb. D.4 auf Seite 380 $V_M = 1,843 \text{ \AA}^3/\text{Da}$. Mittels der Näherung Gl. 2.2 auf Seite 32 ergibt sich damit der prozentuale Volumenanteil von Protein und Solvens am Zellvolumen zu 67 % und 33 %.

Die Bestimmung dieser Anteile über das van-der-Waals-Volumen anhand des Quotienten V_{asu}/V_{vdW} ist abhängig von der Berücksichtigung, bzw. Vernachlässigung der Fehlordnung: Inklusiv der minderpopulierten Strukturanteile beträgt $V_{vdW} = 11668 \text{ \AA}^3$, das vdW-Volumen nur der Hauptpopulation ist $V_{vdW} = 10953 \text{ \AA}^3$. Teilung des Volumens der asymmetrischen Einheit V_{asu} durch diese beiden Volumina ergibt Quotienten von 1,856 und 1,977. Somit beläuft sich der prozentuale Volumenanteil des Proteins auf 54 %, wenn die fehlgeordneten Atome dem Proteinvolumen zugeschlagen werden, und auf etwas geringere 51 %, wenn ausschließlich das Volumen der Hauptpopulation verwendet wird.

4.5.2 Asphärisches Modell

Anhand der Ergebnisse der `shelxl`-Verfeinerung wurde eine Auswahl der zu constrainenden und freizugebenden Atome getroffen: Mußte der betreffende Aminosäurerest mit Restraints verfeinert werden oder war seine Geometrie verzerrt, wurden die Koordinaten fixiert, was zumeist für fehlgeordnete Reste zutrifft. Ebenso bei zu hohen B -Faktoren und einem zu geringen Ausmaß an sie umgebender Elektrodichte, sofern nicht schon über das Restrain-Kriterium abgedeckt. Ansonsten erfolgte der Übergang zum XD-Programmsystem wie in Absch. 4.1.2 auf Seite 53 ausgeführt. Nach dem Transfer der Invariome für das asphärische Modell verblieb nach Elektroneutralitätsausgleich eine Überschußladung von $-0,055e$.

Eine symbolische Darstellung der Fehlordnung und Auswahl frei zu verfeinernder und festgehaltener Bereiche der Struktur, analog den entsprechenden Abbildungen für `P1LYSO` und `TETRALYSO`, ist in Abb. 4.5.5 auf der gegenüberliegenden Seite zu finden. Ergänzend sei erwähnt, daß die hier als teilweise fehlgeordnet gekennzeichneten Hauptketten nur am C_α in Mehr- und Minderpopulation geteilt sind. Falls ebendiese Fragmente als verfeinert markiert sind, so wurde nur der der Hauptpopulation entsprechende Teil in der Verfeinerung freigegeben.

Obwohl der Anteil stabilisierender α -Helices in den Ketten A und C höher als in den beiden längeren Ketten B und D ist, sind die fehlgeordneten Aminosäurereste in ersteren zahlreicher. Es sind diese kurzen Sequenzen A und C, die im Hexamer eher nach außen gerichtet sind, während die beiden längeren Ketten nach innen orientiert sind. Bei den Ketten B und D sind wegen des höheren Anteils an ausreichend definierten Resten ohne Zweitkonformation auch wesentlich mehr Atome zur Verfeinerung freigegeben.

Insgesamt wurden die Koordinaten und Verschiebungsparameter von 403 oder 50 % der 810 Nichtwasserstoffatome der Insulinketten zur Verfeinerung zugelassen, während die andere Hälfte auf den bisherigen Werten fixiert blieb, ebenso wie der gesamte Solvensbereich. Die Wassertoffatome wurden gegebenenfalls entlang der Bindungsachse verschoben, ansonsten aber nicht verfeinert. Auf diese Weise ist auch der Ursprung in der polaren Raumgruppe festgelegt. Somit reduziert sich die Parameterzahl im XD-Fall von 11188 um ungefähr zwei Drittel auf nunmehr 3628 Parameter. Mit diesem Modell wurden sowohl asphärische (*mul*) als auch eine Wiederholung der sphärischen Verfeinerung (*iam*) gerechnet, wobei letztere nicht ganz die ursprünglichen Gütefaktoren erreicht, da wesentlich weniger Parameter zur Verfügung stehen. Dazu kommt noch die Variante *iam_{rb}* mit sphärischen Streufaktoren, aber den Wasserstoffabständen des Invariommodells für direkte Vergleichszwecke. Die dabei erreichten Gütefaktoren sind in Tab. 8.1 auf Seite 217 zusammengefaßt und liegen ob der fehlenden Parameter etwas über den bisher erreichten. Es verblieben nach jeweils zwölf Verfeinerungszyklen gegen $F_o^2 > 2\sigma(F_o^2)$ im Modell *mul* maximale Parameterverschiebungen von $1 \cdot 10^{-4}$ und etwas geringere von ca. $7 \cdot 10^{-5}$ in den beiden sphärischen Modellen.

GLY 101	ILE 102	VAL 103	GLU 104	GLN 105	CYS 106	CYS 107	THR 108
HK SK							
X X	X X	X X	✓ X	X X	✓ ✓	✓ ✓	X X

SER 109	ILE 110	CYS 111	SER 112	LEU 113	TYR 114	GLN 115	LEU 116
HK SK							
✓ X	✓ ✓	✓ ✓	✓ ✓	✓ ✓	✓ X	✓ ✓	✓ ✓

GLU 117	ASN 118	TYR 119	CYS 120	ASN 121
HK SK				
X X	X X	X X	X X	X X

PHE 201	VAL 202	ASN 203	GLN 204	HIS 205	LEU 206	CYS 207	GLY 208
HK SK							
X X	✓ ✓	✓ X	✓ X	✓ X	X X	✓ ✓	✓ ✓

SER 209	HIS 210	LEU 211	VAL 212	GLU 213	ALA 214	LEU 215	TYR 216
HK SK							
✓ ✓	✓ X	✓ X	✓ X	✓ X	✓ ✓	✓ ✓	✓ ✓

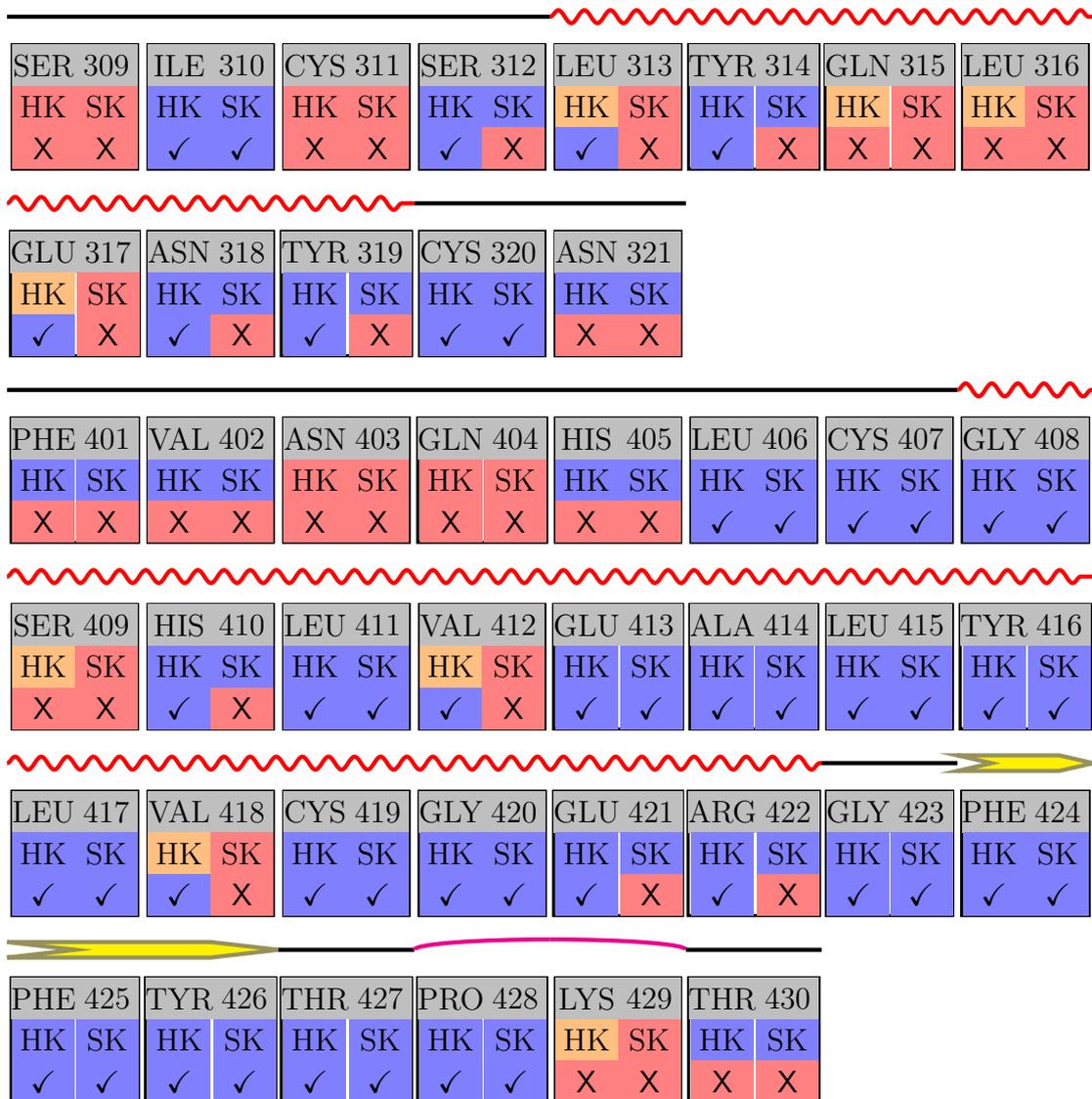
LEU 217	VAL 218	CYS 219	GLY 220	GLU 221	ARG 222	GLY 223	PHE 224
HK SK							
✓ ✓	✓ ✓	✓ ✓	✓ ✓	X X	X X	✓ ✓	✓ ✓

PHE 225	TYR 226	THR 227	PRO 228	LYS 229	THR 230
HK SK					
✓ X	✓ ✓	✓ ✓	✓ ✓	X X	X X

GLY 301	ILE 302	VAL 303	GLU 304	GLN 305	CYS 306	CYS 307	THR 308
HK SK							
X X	X X	X X	X X	X X	X X	X X	X X

Abbildung 4.5.5: Sekundärstruktur, fehlgeordnete und verfeinerte Bereiche von INSULIN.

4 Verfeinerung



Aminosäure und Nummer AAA 000 keine Fehlord. - voll verf.
 Fehlord. in HK und SK? HK SK Fehlord. - nicht verf.
 Verf. von HK und SK? ✓ C γ teilweise Fehlord. - verf.
 ✓: verfeinert, X: nicht verfeinert, C γ : bis inkl. Atom C γ verf.

α -Helix
 β -Strang
 β -Schleife
 Knäuel

Abbildung 4.5.5: Sekundärstruktur, fehlgeordnete und verfeinerte Bereiche von INSULIN (Forts.).

4.6 Trichotoxin

Koordinaten und Strukturfaktoren von TRICHOTOXIN wurden der PDB unter dem Eintrag 1M24 entnommen. Die Struktur wurde von Chugh u. a. (2002) publiziert.

Die Messung der triklinen Struktur fand an der Daresbury Synchrotron Radiation Source (Warrington, Großbritannien) bei einer Temperatur von 150 K und einer Wellenlänge von 0,6883 Å statt. Wie in Tab. 4.6.1 zusammengefaßt, enthält der Datensatz in einem Auflösungsbereich von 10–0,9 Å 14153 symmetrieunabhängige Reflexe, deren Kollektion durchschnittlich nur 2,8-fach redundant war, zu den höchsten Beugungswinkel noch etwas geringer. Die Datenvollständigkeit aber ist hoch und liegt durchgängig bei fast 100 %. Von der Gesamtzahl an Beobachtungen erfüllen 75 % das Kriterium $I > 2\sigma(I)$, es sind also recht viele schwache Reflexe enthalten. Dies deckt sich mit dem geringen Signal-zu-Untergrund-Verhältnis, welches im Mittel 15, bei hoher Auflösung nur noch 2,4 beträgt. Mit dem Intensitätsverlust geht ein steigender R_{mrgd} einher, beträgt dieser im Durchschnitt 5 %, steigt er in der letzten Auflösungsschale auf 26 %.

Tabelle 4.6.1: Kristallographische Daten TRICHOTOXIN (Chugh u. a. 2002).

Summenformel	$2\text{C}_{78}\text{N}_{20}\text{O}_{21}\text{H}_{136} \cdot 2\text{C}_2\text{NH}_3, 7\text{O}$	Absorptionskoeff.	$0,09\text{ mm}^{-1}$
Formelmasse	3475,88 g/mol	Meßtemperatur	150 K
Kristallsystem	Triklin	Wellenlänge λ	0,6883 Å
Raumgruppe	P1 (Nr. 1)	max. 2Θ	$45,06^\circ$
Z	1	Auflösungsbereich	10–0,90 Å
Gitterkonstanten:		Grenzreflexe	$-9 \leq h \leq 10$
a	9,49(1) Å		$-18 \leq k \leq 18$
b	16,85(2) Å		$-35 \leq l \leq 35$
c	31,68(3) Å	unabh. Reflexe	14153
α	$95,77(5)^\circ$	$F_o^2 > 2\sigma(F_o^2)$	10659
β	$98,07(5)^\circ$	Abdeckungsgrad [†]	99,7 (99,9) %
γ	$99,45(5)^\circ$	Redundanz [†]	2,79 (2,18)
V	$4907,83\text{ Å}^3$	$I/\sigma(I)$ [†]	15,03 (2,36)
Dichte	$1,176\text{ g/cm}^3$	R_{mrgd} [†]	5,1(26,1) %
F(000)	1875,44		

[†] Angabe in Klammern für Auflösungsschale 0,95–0,90 Å.

4.6.1 Sphärisches Modell

Die Struktur liegt in Form zweier symmetrieunabhängiger Peptidketten geringfügig unterschiedlicher Geometrie vor. Die quadratisch gemittelte Abweichung der Koordinaten beider Oktadekapeptidketten beträgt 0,5 Å. Jedes der beiden Moleküle wird durch die gleiche Sequenz von 18 Einheiten beschrieben, zudem ist der N-Terminus mit einer Acetylgruppe (ACE) geschützt: ACE–AIB–GLY–AIB–LEU–AIB–GLN–(AIB)₃–(ALA)₂–AIB–PRO–LEU–(AIB)₂–GLN–VOL. Davon sind nur fünf Einheiten kanonische Aminosäuren, verteilt auf acht Positionen. Der C-Terminus besteht aus dem Aminoalkohol L-Valinol (VOL). Ansonsten enthält die Sequenz neun Einheiten α -Aminoisobuttersäure (AIB), auch α -Methylalanin genannt. Diese achiralen Aminosäurereste wirken als Helixbildner, so daß beide Moleküle hauptsächlich als α -Helices vorliegen: In Kette A vom Rest 2 bis 16, in Kette B von 2 bis 17.

Jedes Oktadekapeptid besteht aus 119 Nichtwasserstoffatomen und 136 Wasserstoffatomen. Die insgesamt 510 Atomlagen sind voll populiert – obwohl VOL 218, der C-Terminus von Kette B, anhand der Restdichte Fehlordnung nahelegt. Diese ist schlecht definiert und von nur geringer Besetzung, so daß sie von Chugh u. a. unmodelliert blieb, was hier beibehalten wurde. Es sind auch diese endständigen Valinolreste, die neben LEU 114/214 die größten Konformationsunterschiede zwischen den beiden Ketten zeigen.

Die Solvenssphäre wird in der Originalstruktur durch fünf vollbesetzte Wasser und zwei Acetonitrilmoleküle gebildet. Eines dieser Acetonitrilmolekül zeigt eine stark verzerrte nichtlineare Geometrie. Neben den stereochemischen Restraints der Peptidkette und der Acetonitrilmoleküle, wurden die Solvensmoleküle unter Verwendung von *anti bumping*-Restraints verfeinert. In diesem Modell wurden Gütefaktoren von $R1(F_o > 4\sigma(F_o)) = 0,056$ und $R_{all} = 0,075$ erzielt.

Für diese Arbeit wurden einige Änderungen im Lösungsmittelbereich umgesetzt. Diskrete Koordinaten waren ausreichend zur Beschreibung der Solvenssphäre, wie eine testweise Verfeinerung von Korrekturparametern für diffuse Solvensbeiträge auf verschwindende Beträge ($g = 0,1, U = 0,05$) zeigte. Auch waren die Restdichten zwischen den Peptidketten recht gering. Letztendlich wurden sieben Wassermoleküle unterschiedlicher Besetzung (0,31–0,87) eingeführt, sowie nur ein Acetonitrilmolekül auf zwei Lagen, mit Populationen von 0,48 und 0,52.

Die Stereochemie des Acetonitrils und die Abstände der Wassermoleküle blieben restraint, das Hauptmolekül aber konnte ohne Einschränkungen verfeinert werden, ohne daß Änderungen der Geometrie auftraten. Aus dieser Modellierung, die im weiteren verwendet wird, resultierten ein $R1(F_o > 4\sigma(F_o)) = 0,0477$, $R_{all} = 0,0675$ und ein $wR2$ von 0,1155 bei einem *goodness of fit* von 0,918. Es wurden 2324 Parameter mit 113 Restraints *full matrix least square* verfeinert. Der Restelektrorendichtemittelwert beträgt 0,00 mit einer quadratisch gemittelten Abweichung von $0,04 \text{ eÅ}^{-3}$ und Extrema von $-0,34$ und $0,57 \text{ eÅ}^{-3}$. Ergänzend zeigt Tab. 4.6.2 die

Tabelle 4.6.2: Reflexstatistik TRICHOTOXIN nach `shelx1` inklusive aller Korrekturen.

Auflösung/Å	$N_{unabh.}$	$\overline{I/\sigma(I)}$	$(I > 2\sigma(I))/\%$	Vollständigkeit/%	$\sum(F_o)/\sum(F_c)$	R1(all)	R1($F_o > 4\sigma(F_o)$)
∞ –5,000	74	57,74	91,36	91,36	0,962	0,063	0,063
5,000–4,000	83	65,74	98,80	100,00	0,977	0,039	0,039
4,000–3,000	213	47,92	98,12	100,00	0,985	0,045	0,044
3,000–2,000	911	42,59	98,46	100,00	0,998	0,035	0,035
2,000–1,600	1209	23,27	95,29	100,00	1,014	0,042	0,039
1,600–1,400	1253	15,08	92,18	100,00	1,012	0,045	0,041
1,400–1,200	2187	10,26	86,10	100,00	1,012	0,053	0,043
1,200–1,000	4348	6,44	76,69	99,95	1,007	0,075	0,054
1,000–0,980	636	3,82	62,11	100,00	0,999	0,119	0,072
0,980–0,960	672	3,33	56,99	100,00	1,005	0,129	0,072
0,960–0,940	760	2,85	51,32	100,00	0,995	0,145	0,080
0,940–0,920	887	2,39	42,73	100,00	1,006	0,182	0,084
0,920–0,898	920	1,91	33,71	95,73	1,008	0,215	0,092
∞ –0,898	14153	11,79	75,05	99,65	1,004	0,067	0,048

Entwicklung der R -Werte und Reflexstärke mit der Auflösung. Offensichtlich sind die Daten zwischen 1 und 0,9 Å deutlich schlechter als die restlichen Meßwerte, auch Chugh u. a. berichten von der nachlassenden Datenqualität.

Gegenüber dem ursprünglichen Modell konnten sowohl einige Geometrieverzerrungen behoben als auch die durchschnittlichen B -Faktoren gesenkt werden. Diese betragen nun für die Kette A im Durchschnitt 3,93 Å² und für die Kette B 4,29 Å². Höhere Verschiebungsparameter bis zu 25 Å² treten an den Kettenanfängen und -enden auf, was auch für das N_{e2} des Restes GLN 106 gilt.

Die `whatcheck`-, `procheck`- und `shelxpro`-Tests (Abb. E.2–E.4 ab Seite 392) geben fast keine Warnungen bezüglich Bindungslängen und -winkel aus. Einzig für die CN-Bindung zwischen ACE 200 und AIB 201 findet sich eine Abweichung von Zielwerten, die aber insofern fragwürdig ist, als es sich bei der Acetylgruppe ja nicht um eine Aminosäure handelt. Folgerichtig werden alle Reste in der Ramchandran-auftragung in der energetisch vorteilhaftesten ϕ - ψ -Konformation gefunden. Die ungewöhnliche α -helikale Struktur mit der damit verbundenen dichten Anordnung im Kristall, bewirkt einige zu ignorierende Hinweise auf zu enge und unübliche Packung.

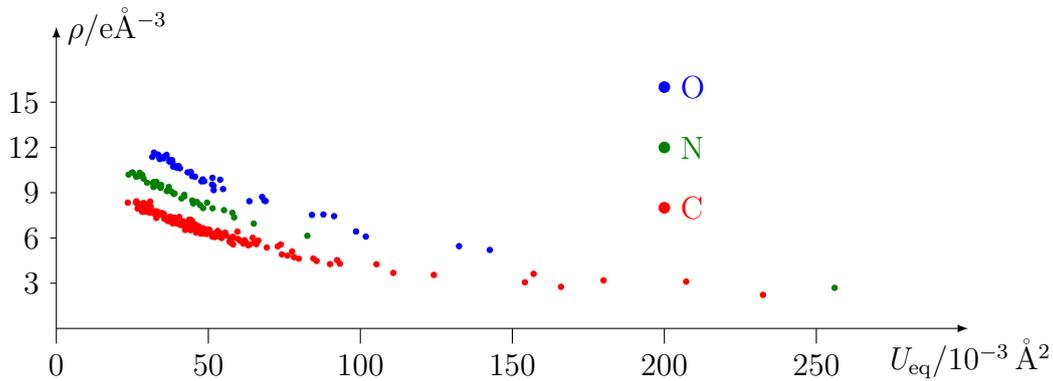


Abbildung 4.6.1: $F_o\phi_c$ -Elektronendichte des sphärischen Modells an den Atompositionen.

Das Luzzati-Diagramm legt anhand der flachen, fast parallel zur Abszisse liegenden Auftragung des R -Wertes mit der Auflösung, einen geringen Koordinatenfehler nahe. Mit der *full matrix*-Verfeinerung stehen aber auch die Standardabweichungen der verfeinerten Koordinaten zur Verfügung: Sie variieren zwischen geringen 0,005 und einigen Ausreißern bis zu 0,04 Å, der Mittelwert beträgt 0,012 Å. Noch geringer ist die Standardabweichung der Verschiebungsparameter, sie beläuft sich im Mittel auf 0,002 Å².

Anhand der Sevcik-Auftragung in Abb. 4.6.1 wird die deutliche Unterscheidbarkeit der Elemente aus der beobachteten Elektronendichte allein demonstriert. Der Einfluß der Phasen des Strukturmodells ist zwar nicht unerheblich, beeinträchtigt aber nicht die Aussage, wie die Ausführungen in Absch. 4.2.1 auf Seite 71 zeigten. Zwar nimmt auch hier die Differenzierung in drei den Elementen entsprechende Punktwolken mit zunehmendem Temperaturfaktor ab, da aber fast alle Atome kleine Temperaturfaktoren unter 0,1 Å² (B -Faktor 8 Å²) aufweisen, bleibt die Unterscheidbarkeit gewahrt.

Abb. 4.6.2 auf der nächsten Seite zeigt einen Ausschnitt der Packung. Die jeweils ungefähr 30 Å langen Helices sind etwa 10° gebogen und zueinander antiparallel orientiert. Um eine möglichst enge Packung zu gewährleisten, sind die Ketten wie ein Reißverschluß in der Längsausdehnung gegeneinander verschoben. Für diese Struktur liegt der Matthews-Koeffizient bei $V_M = 1,418 \text{ Å}^3/\text{Da}$. Mit Gl. 2.2 auf Seite 32 ergibt sich als Näherung für Protein- und Solvensanteil am Volumen der Zelle eine Angabe von 87 % bzw. 13 %. Der Quotient aus Volumen der asymmetrischen Einheit und van-der-Waals-Volumen der TRICHOTOXIN-Moleküle beträgt mit $V_{asu} = 4908$ und $V_{vdW} = 3434 \text{ Å}^3$ $V_{asu}/V_{vdW} = 1,429$. Damit errechnet sich die Raumauffüllung durch das Oktadekapeptid zu 70 %.

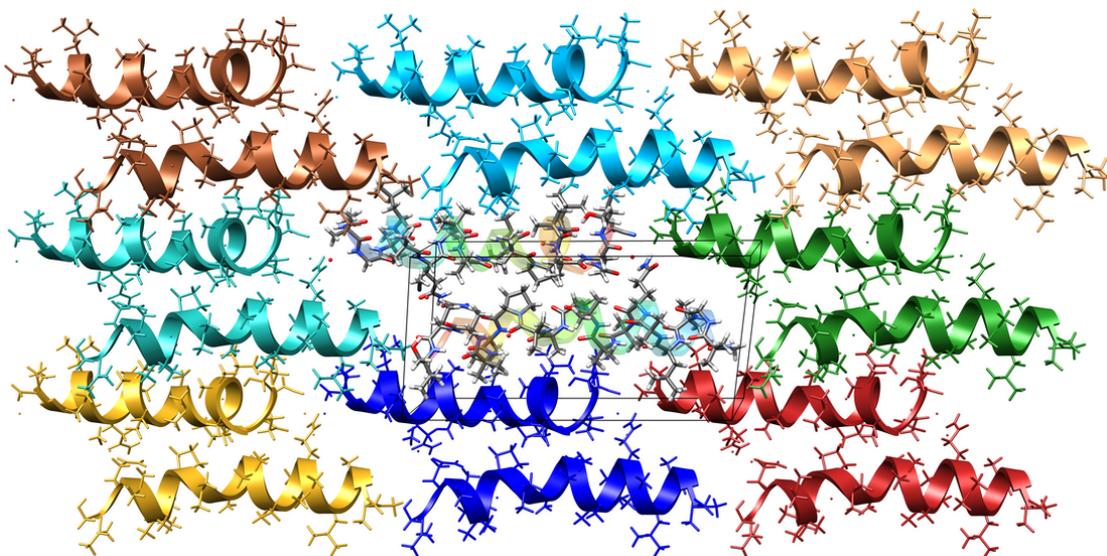


Abbildung 4.6.2: Ausschnitt der Packung von TRICHOTOXIN, neun Elementarzellen in der *cb*-Ebene. Cartoondarstellung der α -helikalen Sekundärstruktur.

4.6.2 Asphärisches Modell

Die Vorbereitung der asphärischen Verfeinerung erfolgte weitestgehend wie in Abschn. 4.1.2 auf Seite 53 beschrieben. Hinzu kommen weitere Invariomzuordnungen, um die Reste ACE, AIB und VOL zu beschreiben. Tab. 4.6.3 auf der folgenden Seite enthält die Ergänzungen zur bereits bekannten Aufstellung der Tabelle 4.1.1. Während alle 510 Atome der Peptidketten asphärische Streufaktoren zugewiesen bekamen, blieb es für alle Solvensmoleküle bei einer sphärischen Beschreibung. Nach Übertragung der Multipole und Elektroneutralitätsausgleich betrug die Ladung der Struktur 0,0087 e.

Da der Lösungsbereich in Restraints involviert war, wurden die zugehörigen Parameter auf den `shelxl`-Werten constraint. Die Nichtwasserstoffatome der Peptidketten konnten zu 98 % zur freien Verfeinerung freigegeben werden. Nur die Seitenkette von VOL 218 wurde festgehalten, was fünf Schweratome mit besonders hohem Temperaturfaktor betrifft. Die Koordinaten von Wasserstoffatomen blieben entweder auf den ursprünglichen kurzen Werten, um im Modell *iam* die sphärische Verfeinerung unter diesen Rahmenbedingungen zu wiederholen, oder wurden auf die Invariomabstände der Tab. 4.1.2 verlängert, was bei sphärischen Streufaktoren zum Modell *iam_{rb}*, bei Verwendung der Invariome zum Modell *mul* führt.

Jedes der drei Modelle wurde mit 2098 Parametern in zwölf Zyklen gegen $F_o^2 > 2\sigma(F_o^2)$ verfeinert. Die drei Rechnungen konvergierten mit maximal verbleibenden Parameterverschiebung von $3-5 \cdot 10^{-6}$, wobei in allen Verfeinerungen der bereits als problematisch bekannte Temperaturfaktor von N(106AE2) die Ursache war. In Tab. 9.1 auf Seite 247 sind die resultierenden Gütefaktoren aufgeführt.

Tabelle 4.6.3: Zusätzliche Invariome, Symmetrie und Koordinatensystem für TRICHOTOXIN.

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
Acetat						
C	C1.5o1.5n[1c1h]1c	m	C(nnnx)	X→N(n+1x)	Y→C(nnnxH3)	R
O	O1.5c[1.5n1c]	mm2	O(nnnx)	Z→C(nnnx)	Y→N(n+1x)	R
CH3	C1c1h1h1h	3	C(nnnxH3)	Z→C(nnnx)	Y→H(nnnxH3A)	R
HH3A	H1c[1c1h1h]	6	H(nnnxH3A)	Z→C(nnnxH3)	Y→C(nnnx)	R
HH3B	H1c[1c1h1h]	6	H(nnnxH3B)	Z→C(nnnxH3)	Y→C(nnnx)	R
HH3C	H1c[1c1h1h]	6	H(nnnxH3C)	Z→C(nnnxH3)	Y→C(nnnx)	R
α-Aminoisobuttersäure						
N	N1.5c[1.5o1c]1c1h	m	N(nnnx)	X→C(n-1x)	Y→C(nnnxA)	R
H0	H1n[1.5c1c]	6	H(nnnx0)	Z→N(nnnx)	Y→C(n-1x)	R
CA	C1n1c1c1c	3m	C(nnnxA)	Z→N(nnnx)	X→C(nnnx)	R
C	C1.5o1.5n[1c1h]1c	m	C(nnnx)	X→N(n+1x)	Y→C(nnnxA)	R
O	O1.5c[1.5n1c]	mm2	O(nnnx)	Z→C(nnnx)	Y→N(n+1x)	R
CB1	C1c1h1h1h	3	C(nnnxB1)	Z→C(nnnxA)	Y→H(nnnxB1A)	R
HB1A	H1c[1c1h1h]	6	H(nnnxB1A)	Z→C(nnnxB1)	Y→C(nnnxA)	R
HB1B	H1c[1c1h1h]	6	H(nnnxB1B)	Z→C(nnnxB1)	Y→C(nnnxA)	R
HB1C	H1c[1c1h1h]	6	H(nnnxB1C)	Z→C(nnnxB1)	Y→C(nnnxA)	R
CB2	C1c1h1h1h	3	C(nnnxB2)	Z→C(nnnxA)	Y→H(nnnxB2A)	R
HB2A	H1c[1c1h1h]	6	H(nnnxB2A)	Z→C(nnnxB2)	Y→C(nnnxA)	R
HB2B	H1c[1c1h1h]	6	H(nnnxB2B)	Z→C(nnnxB2)	Y→C(nnnxA)	R
HB2C	H1c[1c1h1h]	6	H(nnnxB2C)	Z→C(nnnxB2)	Y→C(nnnxA)	R

Zusätzliche Invariome, Symmetrie und Koordinatensystem (Forts.).

Atomtyp	Invariom	Symmetrie	Atomname	AX1→AT1	AX2→AT2	H
Valinol						
N	N1.5c[1.5o1c]1c1h	m	N(nnnx)	X→C(n-1x)	Y→C(nnnxA)	R
H0	H1n[1.5c1c]	6	H(nnnx0)	Z→N(nnnx)	Y→C(n-1x)	R
CA	C1n1c1c1h	m	C(nnnxA)	X→H(nnnxA)	Y→N(nnnx)	R
HA	H1c[1n1c1c]	6	H(nnnxA)	Z→C(nnnxA)	Y→N(nnnx)	R
CB	C1c1c1c1h	3m	C(nnnxB)	Z→H(nnnxB)	X→C(nnnxA)	R
HB	H1c[1c1c1c]	6	H(nnnxB)	Z→C(nnnxB)	Y→C(nnnxA)	R
CG1	C1c1h1h1h	6	C(nnnxG1)	Z→C(nnnxB)	Y→H(nnnxG1A)	R
HG1A	H1c[1c1h1h]	3	H(nnnxG1A)	Z→C(nnnxG1)	Y→C(nnnxB)	R
HG1B	H1c[1c1h1h]	3	H(nnnxG1B)	Z→C(nnnxG1)	Y→C(nnnxB)	R
HG1C	H1c[1c1h1h]	3	H(nnnxG1C)	Z→C(nnnxG1)	Y→C(nnnxB)	R
CG2	C1c1h1h1h	6	C(nnnxG2)	Z→C(nnnxB)	Y→H(nnnxG2A)	R
HG2A	H1c[1c1h1h]	3	H(nnnxG2A)	Z→C(nnnxG2)	Y→C(nnnxB)	R
HG2B	H1c[1c1h1h]	3	H(nnnxG2B)	Z→C(nnnxG2)	Y→C(nnnxB)	R
HG2C	H1c[1c1h1h]	3	H(nnnxG2C)	Z→C(nnnxG2)	Y→C(nnnxB)	R
CH2	C1o1c1h1h	m	C(nnnxH2)	X→O(nnnxH)	Y→C(nnnxA)	R
HH2A	H1c[1o1c1h]	6	H(nnnxH2A)	Z→C(nnnxH2)	Y→O(nnnxH)	R
HH2B	H1c[1o1c1h]	6	H(nnnxH2B)	Z→C(nnnxH2)	Y→O(nnnxH)	R
OH	O1c1h	m	O(nnnxH)	X→C(nnnxH2)	Y→H(nnnxH)	R
HH	H1o[1c]	6	H(nnnxH)	Z→O(nnnxH)	Y→C(nnnxH2)	R

Ergebnisse

5 Alanyl-Prolyl-Alanin

5.1 Gütefaktoren

Aus den in Absch. 4.2 auf Seite 67 beschriebenen Modellen *iam*, *iam_{rb}* und *mul* der Verfeinerung von APA resultieren die in Tab. 5.1 aufgeführten Gütefaktoren. Analoge Ergebnisse der Modelle *iam0.83*, *iam0.83_{rb}* und *mul0.83* für APA0.83 faßt Tab. 5.2 auf der folgenden Seite zusammen.

Die mit der Verwendung der Datenbankmultipole einhergehende Verbesserung der Gütefaktoren von APA ist unmittelbar ersichtlich: Trotz gleichbleibender Parameterzahl sinken die *R*-Werte in der asphärischen Verfeinerung *mul* um 16–23 % gegenüber dem sphärischen Modell *iam*. Im Fall des beschränkten Datensatzes APA0.83, ist die Abnahme, ausgehend von ohnehin niedrigeren *R*-Werten, mit ca. 50 % noch ausgeprägter. Vergleicht man die Ergebnisse der sich ausschließlich im verwendeten Streufaktor unterscheidenden Modelle *iam_{rb}* und *mul*, werden die erzielten Fortschritte nochmals höher. Im Fall APA beträgt der Gewinn nun 21–28 %, bei APA0.83 sogar 64 %. Die physikalisch sinnvollere Modellierung mit verlängerten Wasserstoffatomabständen (siehe Vergleich mit Neutronenabständen in Tab. 4.1.2 auf Seite 66) ist offensichtlich nicht mit der sphärischen Beschreibung vereinbar, weshalb bei *iam_{rb}* die Gütefaktoren auch über denen des Modells *iam* liegen. Somit muß die Verbesserung gegenüber *iam_{rb}* durch ein daran angepaßtes Modell *mul* stärker ausfallen, als die Verbesserung gegenüber einem sphärischen Modell *iam*, mit wiederum dazu passenden Wasserstoffabständen.

Tabelle 5.1: Gütefaktoren der sphärischen und asphärischen Verfeinerungen von APA.

	<i>iam</i>	<i>iam_{rb}</i>	<i>mul</i>
Reflexe $F_o^2 > 2\sigma(F_o^2)$	12185	12185	12185
Parameter	256	256	256
N_{ref}/N_{par}	47,598	47,598	47,598
$R(F)$	0,0352	0,0377	0,0271
$R_{all}(F)$	0,0423	0,0447	0,0343
$R_w(F^2)$	0,0882	0,0942	0,0740
GoF_w	1,0707	1,1436	0,8982
Mittl.(Versch./esd)	$1,19 \cdot 10^{-5}$	$8,97 \cdot 10^{-1}$	$1,92 \cdot 10^{-1}$

Tabelle 5.2: Gütefaktoren der sphärischen und asphärischen Verfeinerungen von APA0.83.

	<i>iam0.83</i>	<i>iam0.83_{rb}</i>	<i>mul0.83</i>
Reflexe $F_o^2 > 2\sigma(F_o^2)$	1423	1423	1423
Parameter	256	256	256
N_{ref}/N_{par}	5,559	5,559	5,559
$R(F)$	0,0243	0,0298	0,0109
$R_{all}(F)$	0,0245	0,0300	0,0110
$R_w(F^2)$	0,0770	0,1079	0,0387
GoF_w	1,4448	2,0253	0,7240
Mittl.(Versch./esd)	$2,02 \cdot 10^{-3}$	$6,82 \cdot 10^{-1}$	$2,10 \cdot 10^{-1}$

Der Vollständigkeit halber sei noch erwähnt, daß bei Verwendung einer $1/\sigma^2$ -Gewichtung und einem Ausschlußkriterium von $F_o^2 > 3\sigma(F_o^2)$ im Invariommodell Gütefaktoren von $R(F) = 0,0246$, $R_{all}(F) = 0,0332$ und $R_w(F^2) = 0,0349$ für die Daten bis $d = 0,37 \text{ \AA}$ erreicht werden können. Ebenso sinken die R -Werte der Verfeinerung der Daten bis $d = 0,83 \text{ \AA}$ auf $R(F) = 0,0097$, $R_{all}(F) = 0,0100$ und $R_w(F^2) = 0,0260$. Dieses Vorgehen entspräche der üblichen Praxis der Verfeinerung kleiner Moleküle, hier wurde aber Konsistenz mit der Herangehensweise für die Makromoleküle angestrebt.

Über den Vergleich der R -Werte hinaus, zeigt auch die Untersuchung der Differenzen der gemessenen und der aus den Modellen berechneten Strukturfaktoren $\Delta F = |F_o - F_c|$ jeweils des sphärischen und asphärischen Modells, ein eindeutiges Bild zugunsten des Invariomformalismus: Signifikanztests (siehe Abschn. 3.1.2) zu den Hypothesen $H_0 : \Delta F(iam) = \Delta F(mul)$ und $H_1 : \Delta F(iam) > \Delta F(mul)$ ergeben im t -Test und Wilcoxon-Test jeweils p -Werte nahe Null. Gleiches gilt für den Vergleich der Modelle zu den beschränkten Daten *iam0.83* und *mul0.83*. Die Hypothese, daß das sphärische und asphärische Modell gleich wären, kann zugunsten der Alternativhypothese statistisch hochsignifikanter höherer Differenzen in den Modellen *iam* und *iam0.83* zurückgewiesen werden.

Primär zu Vergleichszwecken mit den Makromolekülen zeigt Abb. 5.1 auf der gegenüberliegenden Seite für jeden der gemessenen Reflexe die relative Abweichung von Beobachtung und sphärischem Modell *iam*. In Abhängigkeit von der Auflösung ist dazu der Quotient F_o^2/F_c^2 aufgetragen. Die Anzahl besonders stark abweichender Reflexe, deren Verhältnis die Werte ≥ 2 oder $\leq 1/2$ erreicht oder überschreitet, beträgt 1230 oder 8,8 % der 13992 Reflexe. Hiervon betroffen sind fast ausschließlich Daten mit einer höheren Auflösung als $d = 0,6 \text{ \AA}$. Innerhalb der mit roten Linien gekennzeichneten Grenzen verbleiben 12762 Strukturfaktoren. Betrachtet man nur starke Reflexe $F_o^2 > 2\sigma$, liegen bloß 2,4 % außerhalb der Schwellenwerte.

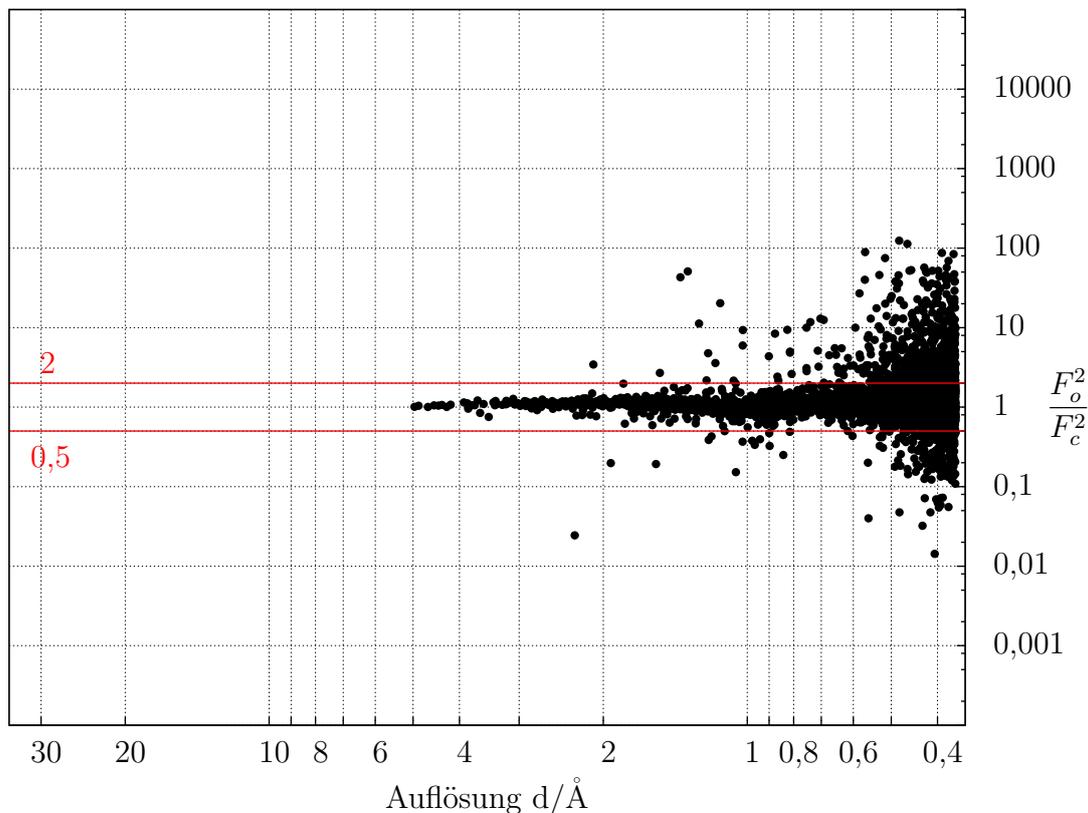


Abbildung 5.1: Quotient beobachteter und berechneter Intensitäten und ihre Auflösungsabhängigkeit im Modell *iam*.

5.2 Geometrieänderungen

Aus den Verfeinerungen in den drei Modellen resultieren unterschiedliche Atomkoordinaten. Die Unterschiede der drei Geometrien sollen zum einen in Form der quadratisch gemittelten Abweichung der Atompositionen voneinander und zum anderen im Sinne der Bindungslängenunterschiede diskutiert werden.

Die mittlere Abweichung der Gesamtstruktur im Modell *mul* beträgt beim Vergleich mit der Verfeinerung *iam* $0,096 \text{ \AA}$, bei Gegenüberstellung der Koordinaten aus *iam_{Tb}* nur $0,051 \text{ \AA}$. In Abb. 5.2(a)–(b) auf der nächsten Seite sind diese Differenzen pro Atom dargestellt, man erkennt, daß die Verschiebungen der Nichtwasserstoffatome mit $0,000 \text{ \AA}$ bis maximal $0,002 \text{ \AA}$ sehr gering sind, letztere betrifft die Sauerstoffatompositionen. Durch die Verlängerung der Bindungen zu den Wasserstoffatomen werden größere Unterschiede hervorgerufen, die gegenüber dem Modell *iam* $0,1\text{--}0,2 \text{ \AA}$ betragen und beim Vergleich mit *iam_{Tb}* noch $0,03$ bis $0,1 \text{ \AA}$ erreichen, da auch Winkeländerungen dazu beitragen.

Analog bietet Abb. 5.2(c)–(d) eine Illustration der Geometrieunterschiede von

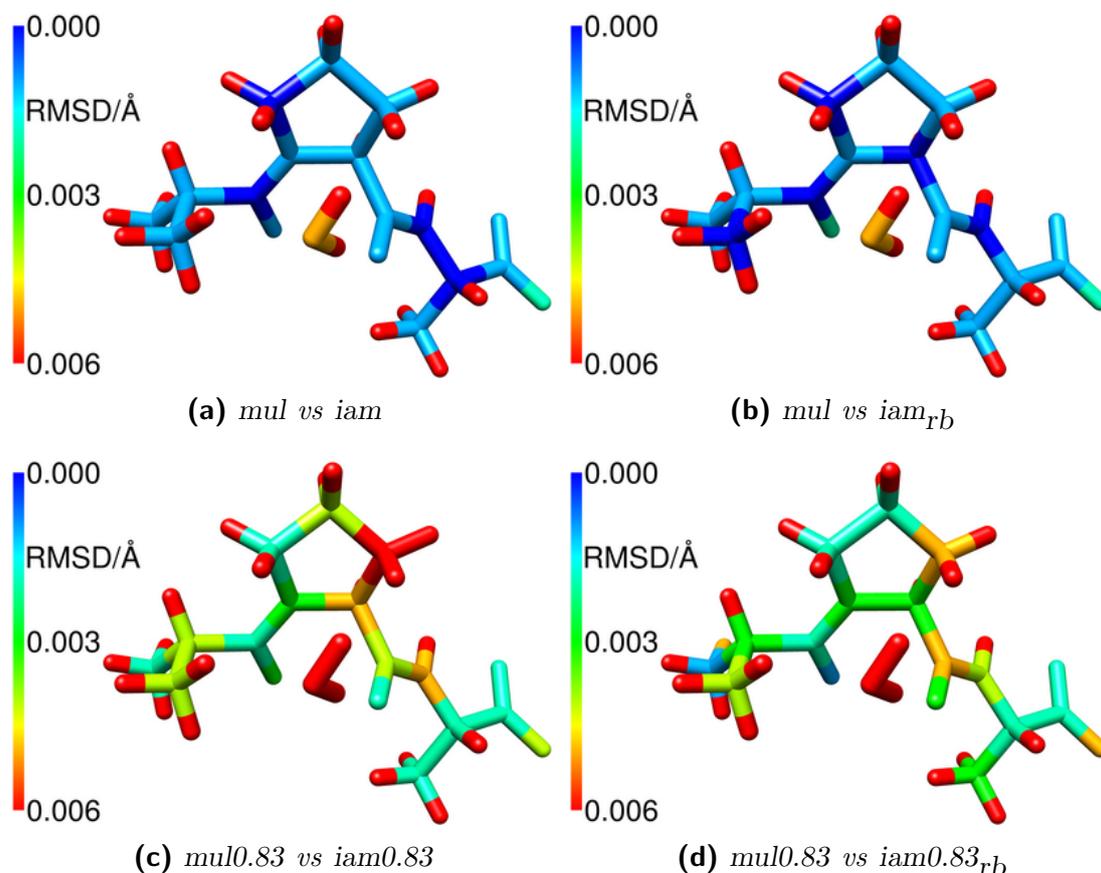


Abbildung 5.2: Quadratisch gemittelte Strukturunterschiede durch asphärische Verfeinerung von APA und APA0.83 – Farbcode: Abweichung in Å.

APA0.83. Hier fallen die über das Molekül gemittelten Veränderungen zwischen den Verfeinerungen *mul0.83* und *iam0.83* mit $0,106 \text{ \AA}$ etwas höher als zuvor aus, der Vergleich mit *iam0.83_{rb}* mit $0,041 \text{ \AA}$ etwas niedriger. Die Unterschiede einzelner Schweratomkoordinaten variieren zwischen $0,002$ und $0,006 \text{ \AA}$, bzw. $0,001$ - $0,005 \text{ \AA}$. Für Wasserstoffatome treten wieder höhere Differenzen von $0,09$ – $0,2 \text{ \AA}$ und $0,007$ – $0,1$ in den Modellen *iam0.83*, bzw. *iam0.83_{rb}* auf. Somit sind die Veränderungen, insbesondere der Nichtwasserstoffatome, bei Verfeinerung gegen die beschränkten Daten ausgeprägter als beim vollen Datensatz. Dazu tragen sicherlich die höheren Parameterverschiebungen und Koordinatenfehler von APA0.83 ursächlich bei. Sowohl für APA0.83 als auch für APA sind die Positionsunterschiede von einigen tausendstel Ångström in der Größenordnung der jeweiligen mittleren Koordinatenunsicherheiten von $0,0006$, bzw. $0,0022 \text{ \AA}$.

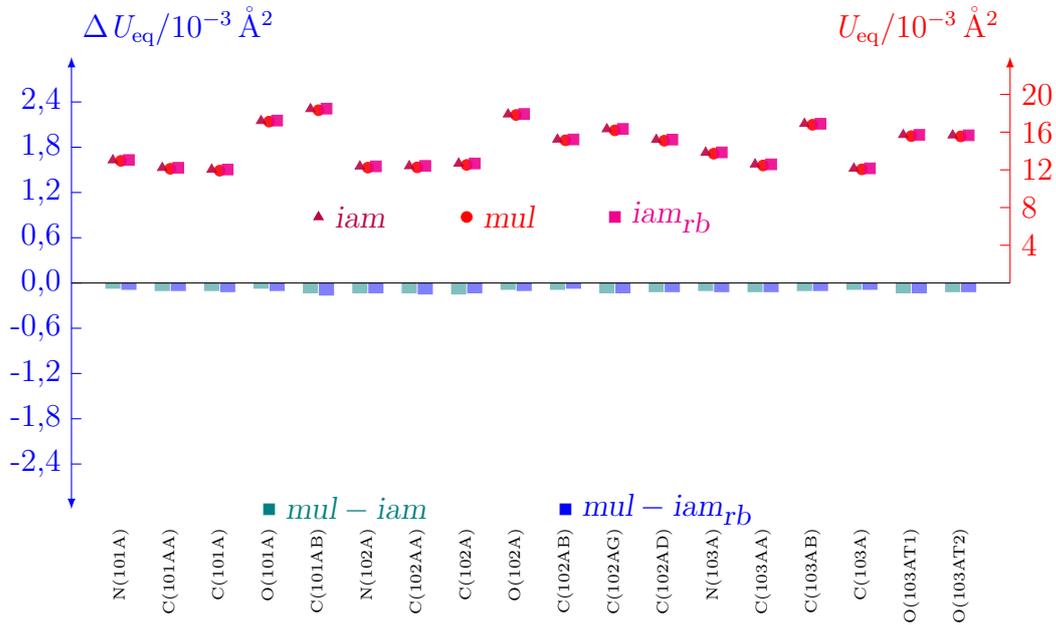
Eine detaillierte Auflistung der Differenzen der einzelnen Atome und Aminosäuren in den Verfeinerungen von APA und APA0.83 findet sich in den Tab. A.1 und A.2 auf den Seiten 319 und 320.

Auch die Änderung der Bindungslängen (Tab. A.3 auf Seite 321) zwischen Nichtwasserstoffatomen durch die unterschiedlichen Verfeinerungen von APA ist ausgesprochen gering, die maximal auftretenden Differenzen zwischen mul und iam , bzw. iam_{rb} , betragen $+0,0010$ und $-0,0013$ Å (siehe auch Tab. A.4 auf Seite 322). Die durchschnittliche Veränderung liegt bei nur $0,0005$ Å. Da sich die Bindungslängen in den Modellen iam und iam_{rb} kaum unterscheiden, sind auch die Differenzen zu mul sehr ähnlich. Im Mittel werden die Bindungen im Modell mul kürzer, wobei insbesondere die Kohlenstoff–Sauerstoffbindungen abnehmen. Dies ist nicht unerwartet, da die Nichtberücksichtigung der freien Elektronenpaare am Sauerstoff bei einer sphärischen Verfeinerung zu einer irrigen Verlängerung der Bindung führt. Auch Kohlenstoff–Stickstoffbindungen sind in den sphärischen Modellen mehrheitlich länger, die Verkürzung durch mul ist aber sehr gering. Die größte Variabilität weisen Kohlenstoff–Kohlenstoffbindungen auf, hier treten auch häufiger geringfügige Bindungsverlängerungen im asphärischen Modell auf.

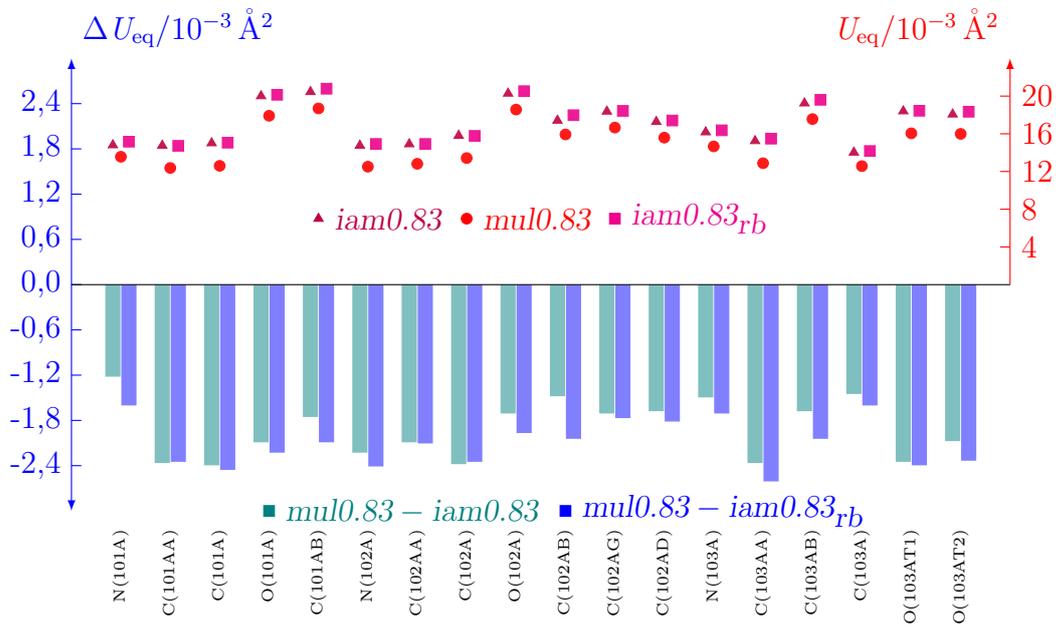
Für die Verfeinerung von APA0.83 gegen die beschränkten Daten, sind sowohl höhere Bindungslängendifferenzen als auch eine größere Schwankungsbreite zu konstatieren. Der Durchschnitt der Beträge der Bindungslängendifferenzen von $iam0.83$ und $iam0.83_{rb}$ zu $mul0.83$ liegt bei $0,0018$ und $0,0019$ Å, wobei insgesamt mehr Bindungen eine Verkürzung als eine Verlängerung erfahren. Zu beiden sphärischen Verfeinerungen treten maximale Differenzen von ungefähr $\pm 0,004$ Å auf. Somit sind die Änderungen viermal so hoch wie bei APA, allerdings ist die Unsicherheit in den Bindungslängen ebenfalls viermal höher. Da die Bindungslängen in den Modellen $iam0.83$ und $iam0.83_{rb}$ recht unterschiedlich ausfallen, sind auch die Differenzen zu $mul0.83$ verschieden und uneinheitlich. Folgten zuvor bei APA die Unterschiede zu den beiden sphärischen Modellen jeweils dem gleichen Trend, kann dies nun nicht mehr beobachtet werden.

5.3 Verschiebungsparameter

Die 18 Nichtwasserstoffatome in APA haben niedrige äquivalente isotrope Temperaturfaktoren zwischen 12 und $20 \cdot 10^{-3} \text{ Å}^2$, und zwar sowohl in den sphärischen Modellen wie nach der asphärischen Verfeinerung. Dies ist in Abb. 5.3(a) auf der folgenden Seite durch purpurne und magentafarbene Drei- und Vierecke für iam und iam_{rb} , sowie rote Kreise für mul illustriert. Darunter sind jeweils die Unterschiede von sphärischem zu asphärischem Modell durch Balken eingezeichnet, die Differenzen zwischen $0,08$ bis $0,15 \cdot 10^{-3} \text{ Å}^2$ darstellen. Wie der Statistik in Tab. A.5 auf Seite 322 entnommen werden kann, liegt der Durchschnitt der 18 U_{eq} bei $14,56$, $14,55$ und $14,44 \cdot 10^{-3} \text{ Å}^2$ in den Modellen iam_{rb} , iam und mul , der der Differenzen $mul - iam$ und $mul - iam_{rb}$ bei $-0,11$ und $-0,12 \cdot 10^{-3} \text{ Å}^2$. Ausgehend von ohnehin niedrigen Verschiebungsparametern, wird im Modell mul nochmal eine kleine Verbesserung von ca. 1% erzielt.



(a) APA



(b) APA0.83

Abbildung 5.3: U_{eq} der Verfeinerung von APA und APA0.83, sowie Differenz der U_{eq} zwischen asphärischem und sphärischen Modellen für die 18 Nichtwasserstoffatome.

Eine analoge Statistik für die Verfeinerung APA0.83 der Daten bis $d = 0,83 \text{ \AA}$ (Tab. A.6 auf Seite 323) ergibt ein etwas höheres durchschnittliches U_{eq} in den Modellen $iam0.83_{rb}$, $iam0.83$ und $mul0.83$ von 17,13, 16,94 und $15,02 \cdot 10^{-3} \text{ \AA}^2$, was auch im dazugehörigen Diagramm 5.3(b) deutlich wird. Auch gibt es nun erkennbare Unterschiede zwischen den beiden sphärischen Verfeinerungen und dem asphärischen Modell. Die Verbesserung in $mul0.83$ liegt zwischen $-1,2$ bis $-2,6 \cdot 10^{-3} \text{ \AA}^2$, der Durchschnitt bei etwa $2 \cdot 10^{-3} \text{ \AA}^2$. Somit liegen die Temperaturfaktoren nach asphärischer Modellierung um ca. 12 % unter denen der sphärischen Behandlung und erreichen fast das Niveau der Verfeinerung gegen alle Daten.

Die geringe Änderung in APA, liegt an den bei hoher Auflösung, auch im sphärischen Fall, besonders gut bestimmten Verschiebungsparametern. Die Streubeiträge zu den Hochwinkeldaten rühren primär von den atomnahen inneren Schalen her, wodurch die Atomlage, und damit ihre Schwingung, besonders gut bestimmt ist. Konsequenterweise nehmen, bei Abwesenheit dieser Information, die U_{eq} in den sphärischen Modellen zu, da nun die Schwingung schlechter von der durch Bindungseffekte verschmierten Elektronendichte getrennt werden kann.

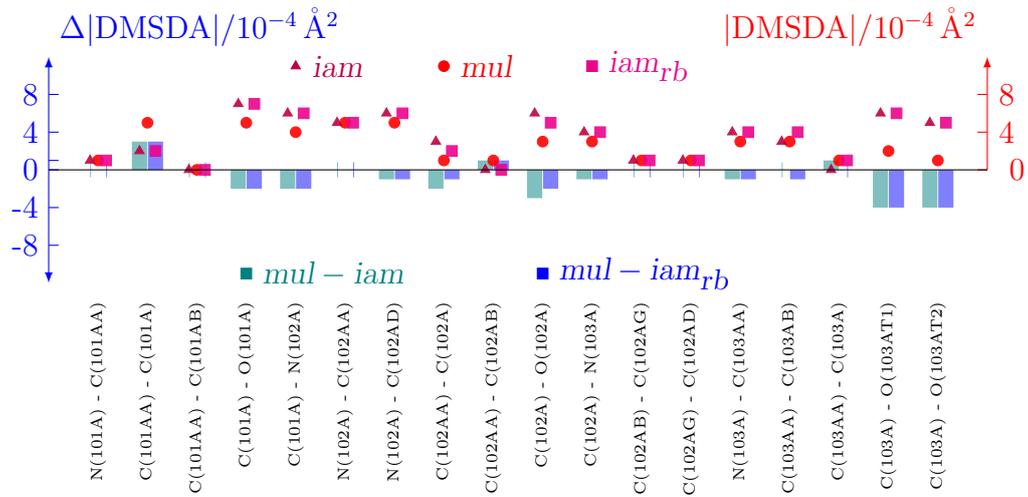
Bei APA ist die U_{eq} -Abnahme an den Kohlenstoffatomen etwas ausgeprägter als an Stickstoff- oder Sauerstoffatomen, wie eine nach Elementen aufgeschlüsselte Statistik (Tab. A.7 auf Seite 324) zeigt. Im Fall APA0.83 ergibt die gleiche Betrachtung (Tab. A.8 auf Seite 324) die größten Änderungen für Sauerstoffatome, gefolgt von Kohlenstoff und Stickstoff.

5.4 DMSDA-Werte

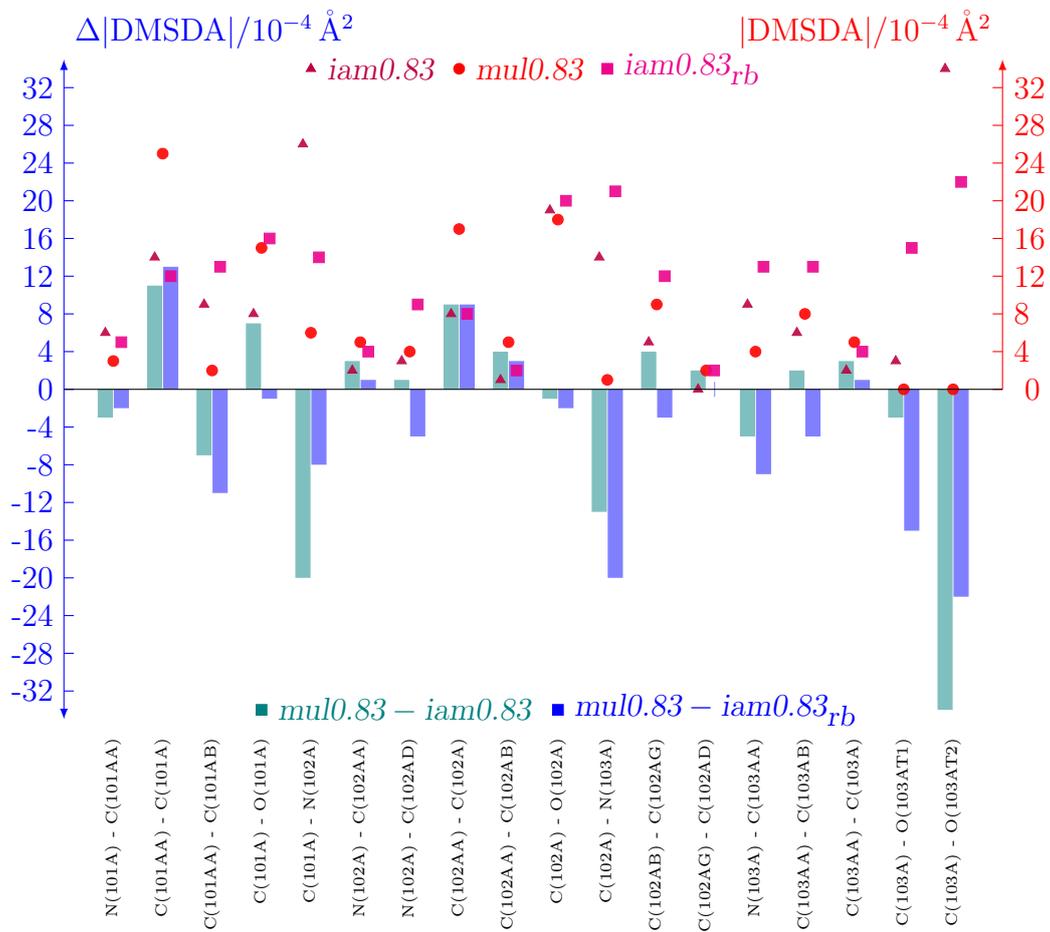
In den Modellen von APA liegen die DMSDA-Werte (siehe Absch. 1.1 auf Seite 18) der 18 Bindungen zwischen Nichtwasserstoffatomen bei $0 - 8 \cdot 10^{-4} \text{ \AA}^2$, sowie bei jeweils ähnlichen Werten pro betrachteter Bindung und Modell. Dementsprechend gering sind die in Abb. 5.4(a) auf der folgenden Seite gezeigten Differenzen ΔDMSDA für die Vergleiche $mul - iam$ und $mul - iam_{rb}$. Betrag der Mittelwert der DMSDA-Beträge $1/n \sum_i^n |DMSDA_i|$ aller $n = 18$ Bindungen in beiden sphärischen Verfeinerungen noch $3,3 \cdot 10^{-4} \text{ \AA}^2$, so wird in mul ein ca. 25 % niedrigerer Wert von $2,5 \cdot 10^{-4} \text{ \AA}^2$ erreicht (siehe Tab. A.9 auf Seite 325).

Dabei verringert sich der DMSDA-Wert von neun Bindungen, bleibt in sechs Bindungen unverändert und verschlechtert sich in drei Fällen (siehe auch Tab. A.10 auf Seite 326). Polare CN- und CO-Bindungen erfahren fast durchweg eine Verbesserung der DMSDA-Werte bis zu $\Delta\text{DMSDA} = -4 \cdot 10^{-4} \text{ \AA}^2$, während nur bei CC-Bindungen eine Zunahme auftritt, die im Extremfall $\Delta\text{DMSDA} = 3 \cdot 10^{-4} \text{ \AA}^2$ beträgt. Eine Abhängigkeit der Differenzen von der Höhe der jeweiligen DMSDA-Beträge kann nicht festgestellt werden.

In Abb. 5.4(b) ist das Verhalten der Verfeinerungen von APA0.83 gezeigt, das besonders durch höhere und volatilere DMSDA-Beträge und -Differenzen gekenn-



(a) APA



(b) APA0.83

Abbildung 5.4: DMSDA-Beträge und ihre Veränderung der 18 Nichtwasserstoffbindungen von APA und APA0.83.

zeichnet ist. Die Modelle bei verringerter Auflösung zeigen drei- bis viermal so hohe DMSDA-Werte zwischen 0 und $34 \cdot 10^{-4} \text{ \AA}^2$, wobei insbesondere CO- und CN-Bindungen betroffen sind. Durchschnittlich werden in den drei Modellen *iam0.83*, *iam0.83_{rb}* und *mul0.83*, Werte von 9,4, 11,4 und $7,2 \cdot 10^{-4} \text{ \AA}^2$ erreicht. Mit der asphärischen Verfeinerung ist also wiederum eine Verbesserung verbunden, die Abnahme der Mittelwerte beläuft sich auf 23 %, bzw. 37 %. Gerade die polaren Bindungen mit den höchsten DMSDA-Beträgen erfahren auch die höchste Verbesserung um bis zu $-34 \cdot 10^{-4} \text{ \AA}^2$. Insgesamt 5–10 Bindungen, je nach Modell, verschlechtern ihre DMSDA-Werte im Modell *mul0.83*, wobei es sich wieder fast ausschließlich um CC-Bindungen handelt. Auch hier kann kein Zusammenhang zwischen Betrag im sphärischen Modell und Höhe der Differenz zur asphärischen Verfeinerung erkannt werden.

5.5 Restdichten

Eine Schlüsselstellung in der Begutachtung der Wirkung unterschiedlicher Streufaktoren nehmen die Restdichten nach den Verfeinerungen ein. Diese Unterschiede zwischen Beobachtung und Modell bestimmen maßgeblich den Erfolg der Beschreibung des Experiments.

Die Restdichten wurden wie in Absch. 3.4 auf Seite 44 beschrieben berechnet. Für die voll aufgelöste Struktur APA, wurden in den drei Modellen jeweils ein die asymmetrische Einheit umfassendes Gitter von $76 \times 60 \times 46 = 209760$ Punkten erstellt, für den beschränkten Datensatz APA0.83 ein gleich großes mit $24 \times 36 \times 21 = 18144$ Punkten. Das Volumen der Voxel beträgt damit $0,0016 \text{ \AA}^3$, bzw. $0,0194 \text{ \AA}^3$.

Die auf das Voxelvolumen normierte Gesamtrestelektronendichte $\sum |\Delta\rho_{res}|V_{vox}$ von APA ist im asphärischen Modell *mul* mit 12,57 e ein Viertel bis ein Drittel niedriger als in den sphärischen Modellen *iam* und *iam_{rb}* mit 17,02 e, bzw. 18,95 e. Bezieht man die integrierte Restdichte auf das Zellvolumen, um eine mit anderen Strukturen vergleichbare Größe zu erzeugen, so erhält man Werte von 0,0375, 0,0508 und $0,0565 \text{ e\AA}^{-3}$. Noch niedrigere Werte liegen bei APA0.83 vor, die Modelle *iam0.83_{rb}*, *iam0.83* und *mul0.83* weisen Restdichten von 10,12, 9,31 und 3,97 e auf, die pro Volumeneinheit der Zelle 0,0302, 0,0278 und $0,0118 \text{ e\AA}^{-3}$ betragen. Hier reduziert sich die Restdichte mit der asphärischen Verfeinerung gegenüber dem Modell *iam0.83* um 57 %, im Vergleich zu *iam0.83_{rb}* sogar um 61 %.

Anhand der Abbildungen 5.5 auf der folgenden Seite und 5.6 auf Seite 125 wird die räumliche Verteilung der Restdichten illustriert, dargestellt sind die Isooberflächen zu verschiedenen Dichtewerten bis $1,4 \text{ \AA}$ um das Molekül.

Nach den sphärischen Verfeinerungen *iam* und *iam_{rb}* von APA (Abb. 5.5(c) und (d)), treten deutliche Konzentrationen nicht beschriebener Dichte auf den Bindungen und an den freien Elektronenpaaren der Sauerstoffatome auf. Zudem wird durch die sphärische Beschreibung zu viel Dichte neben den Bindungen modelliert,

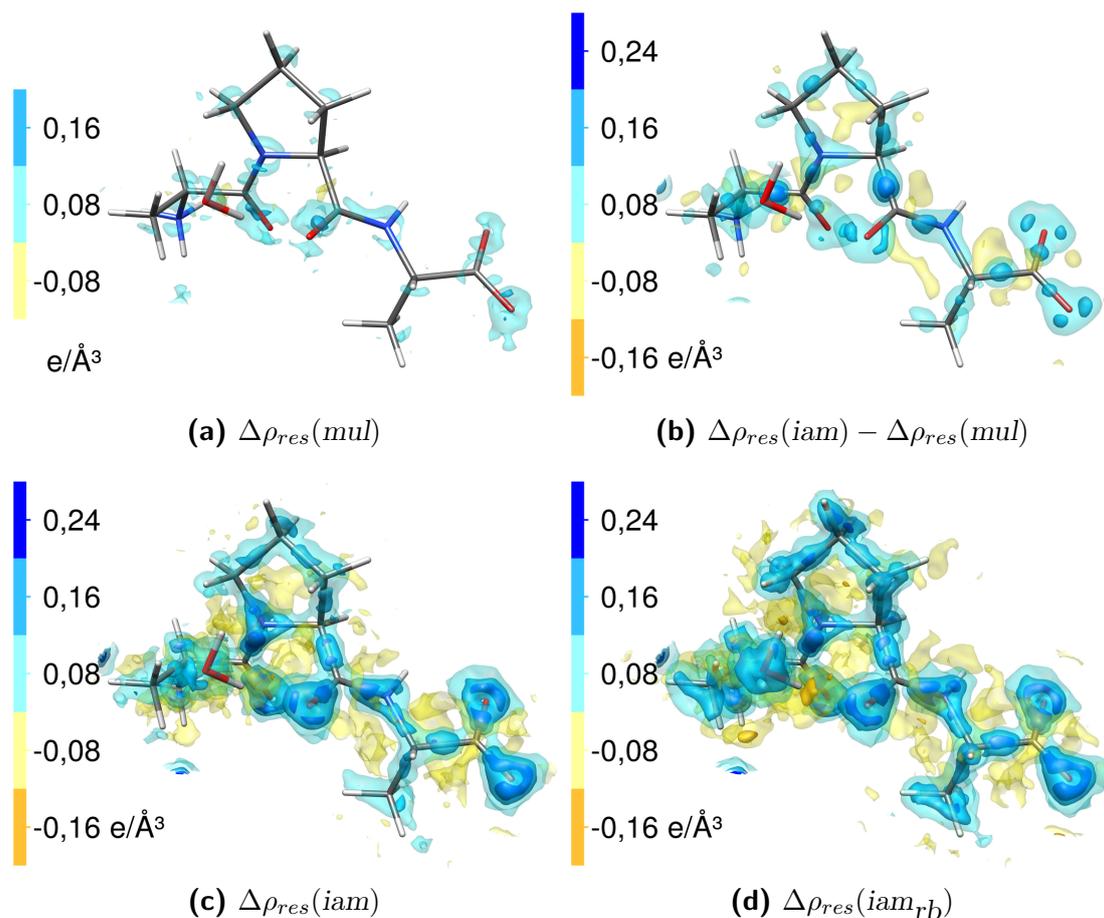


Abbildung 5.5: Restdichten der Verfeinerungen von APA. Farbcodierte Isooberflächen für mehrere Restdichtewerte. Darstellung bis 1,4 Å um das Molekül.

weshalb dort nun negative Restdichten auftreten. Es ist ebenfalls deutlich, daß das Modell iam_{rb} ein schlechteres Ergebnis als iam liefert. Im Resultat der asphärischen Verfeinerung mul (Abb. 5.5(a)) aber sind die Restdichten drastisch verringert. Es verbleiben nur geringe positive Reste in Nähe der Atome. Der Unterschied zur sphärischen Behandlung iam wird in Abb. 5.5(b) durch die Differenz der Restdichten in beiden Modellen $\Delta\rho_{res}(iam) - \Delta\rho_{res}(mul)$ hervorgehoben. Offensichtlich konzentrieren sich die Änderungen auf die Bindungen und freien Elektronenpaare, wo in mul weniger Restdichte als in iam zu finden ist. Dort beschreibt das Invariommodell, im Gegensatz zum sphärischen Modell, die tatsächlich vorhandene Bindungsdichte, die damit in mul inkorporiert ist und keinen Anlaß mehr zu Restdichten gibt. Sehr schön ist auch zu erkennen, daß die Verlängerung der Wasserstoffabstände in iam_{rb} weitere Restdichten auf den Bindungsachsen provoziert, welche im Modell iam durch Positionierung der Wasserstoffatome näher an der Bindungsdichte vermieden wurde.

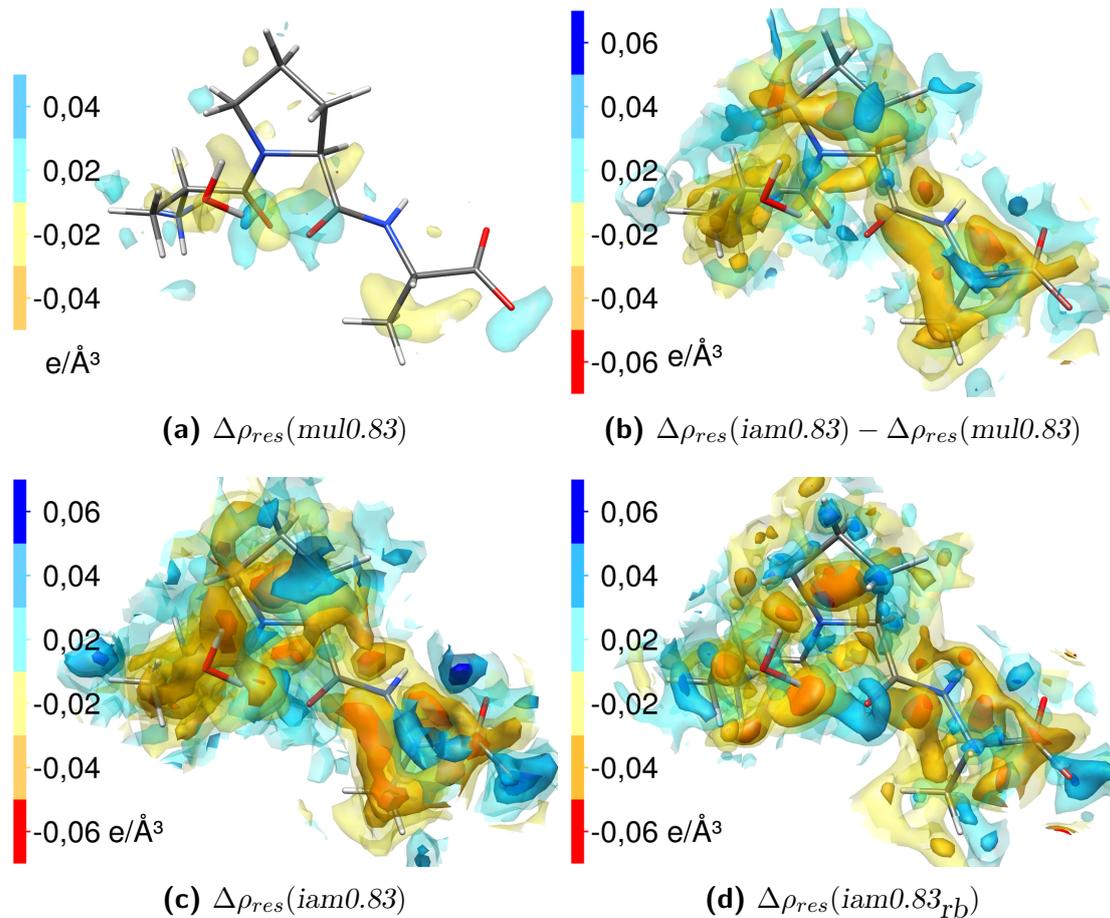


Abbildung 5.6: Restdichten der Verfeinerungen von APA0.83. Farbcodierte Isooberflächen für mehrere Restdichtewerte. Darstellung bis 1,4 Å um das Molekül.

Die nur halb so hohen Restdichten von APA0.83 spiegeln sich in den geringeren Isooberflächenwerten der Abb. 5.6 wieder. Auffällig ist die diffusere Verteilung der Restdichten in allen drei Modellen, dennoch ist auch hier die Restdichteakkumulation auf Bindungen und an Sauerstoffatomen in den sphärischen Modellen ersichtlich. Wie der Vergleich des Modells *mul0.83* in Abb. 5.6(a) mit den Darstellungen 5.6(c) und (d) für *iam0.83* und *iam0.83_{rb}* zeigt, ist mit der asphärischen Modellierung eine besonders starke Abnahme der Restdichten verbunden.

Die punktweise Differenzbildung der Restdichtegitter der Modelle *iam0.83* und *mul0.83* wird durch Abb. 5.6(b) visualisiert. Dabei entsprechen die positiven Isooberflächen der erreichten Verringerung der Restdichten durch den Invariomeinsatz und verdeutlichen die bessere Beschreibung insbesondere der Bindungen. Die Oberflächen zu negativen Werten entsprechen nicht etwa einer Verschlechterung in *mul0.83*, sondern sind die gleichen, die in der Darstellung 5.6(c) zu finden sind. Andernfalls müßten in Abb. 5.6(a) an diesen Orten Restdichten vorhanden sein.

Tabelle 5.3: Statistik der Restdichten von APA und APA0.83.

	$x_{min/max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
$d = 0,37 \text{ \AA} \quad n = 209760$							
<i>mul</i>	-241,3/335,9	-0,04(47,22)	-31,2	-1,4	28,9	60,1	44,6
<i>iam</i>	-258,6/513,4	-0,12(67,72)	-42,2	-5,3	33,5	75,7	56,2
<i>iam_{rb}</i>	-288,3/519,4	-0,22(75,92)	-47,2	-8,2	34,4	81,6	60,3
$d = 0,83 \text{ \AA} \quad n = 18144$							
<i>mul0.83</i>	-80,6/83,3	-0,01(14,69)	-9,0	0,1	9,3	18,3	13,5
<i>iam0.83</i>	-144,2/165,1	0,04(33,79)	-20,8	2,2	22,9	43,7	32,2
<i>iam0.83_{rb}</i>	-196,2/192,1	-0,26(38,01)	-22,8	0,7	22,5	45,3	33,4

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der Restdichte;

[‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten 10^{-3} e/\AA^3 .

Eine statistische, statt graphischer Auswertung der gesamten Restdichte der asymmetrischen Einheit von APA und APA0.83 bietet Tab. 5.3. Besonders für APA ist die Abstufung des Verfeinerungserfolgs in der Reihenfolge *mul*, *iam* und *iam_{rb}* deutlich. Mit der asphärischen Verfeinerung wird der geringste Mittel- und Zentralwert erreicht, auch liegen die Extremwerte unter denen der sphärischen Modelle. Zudem wird die Verteilung der Restdichten um den Idealwert Null enger, was sich in einer Abnahme der Standardabweichung, des Interquartilabstandes und des Medians der absoluten Abweichungen vom Zentralwert äußert. Die sphärischen Modelle, und hier insbesondere *iam_{rb}*, zeigen höhere Maximal- und Minimalwerte der Restelektronendichte, ebenso sind die Mittelwerte vier- bis sechsmal weiter von Null entfernt. Gegenüber *mul* sind auch die Streuungsmaße um ein bis zwei Drittel erhöht, wobei die Verteilung eher in den negativen Bereich verschoben ist, wie man es an den unsymmetrischen Lagen der 0,25- und 0,75-Quartile ersieht.

Dagegen scheint auf den ersten Blick die Häufigkeitsverteilung der Restdichten in $0,01 \text{ e\AA}^{-3}$ breiten Intervallen in Abb. 5.7 auf der nächsten Seite zu sprechen, da die sphärischen Modelle eine breite Schulter im Bereich positiver Restdichte zeigen. Bei Beachtung der logarithmischen Teilung der Ordinate aber umfaßt dieser Bereich recht wenige Werte. Durch die Überlagerung der Histogramme für *mul* und *iam*, bzw. *iam_{rb}*, wird die Restdichtereduktion besonders deutlich. Es treten weniger hohe Werte auf, der Anteil sehr niedriger Restdichten nahe Null nimmt zu und die Verteilung von *mul* nimmt bei Entfernung vom Nullpunkt sehr viel schneller ab als in den sphärischen Verfeinerungen. So liegen in den besonders guten Intervallen $(-0,01; 0]$ und $(0; 0,01]$ bei *mul* 18 % aller Restdichtewerte, während im Fall *iam* und *iam_{rb}* nur noch 14 % und 13 % erreicht werden. Umgekehrt ist das 99 % aller Restdichte einschließende Intervall bei *mul* mit $(-0,116; 0,149]$ viel schmalere als in den anderen Verfeinerungen mit $(-0,154; 0,273]$ und $(-0,166; 0,299]$.

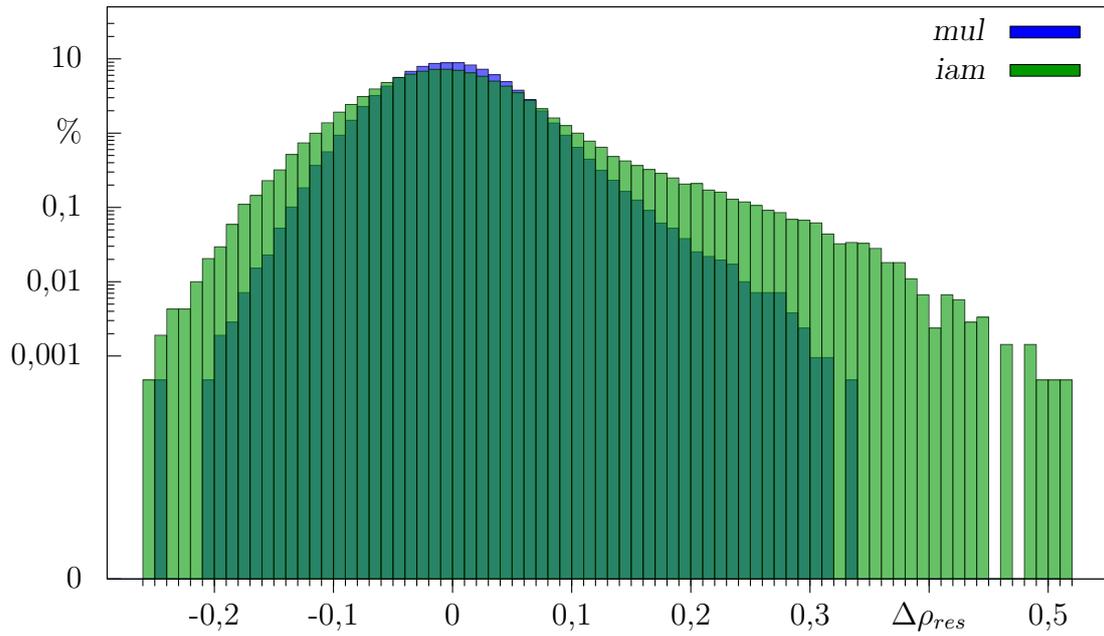
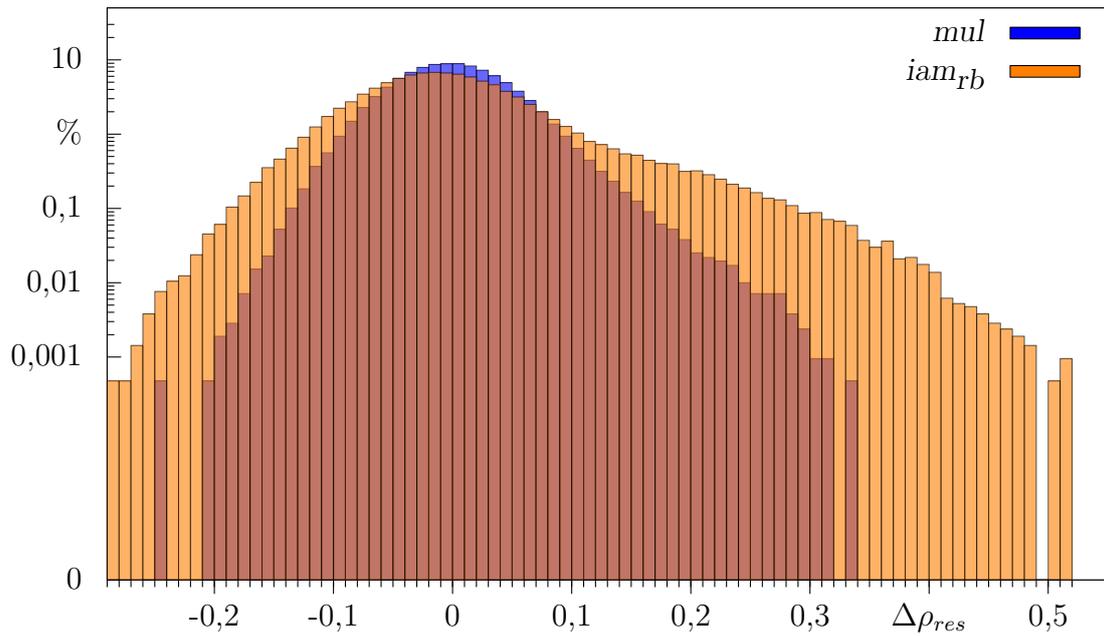
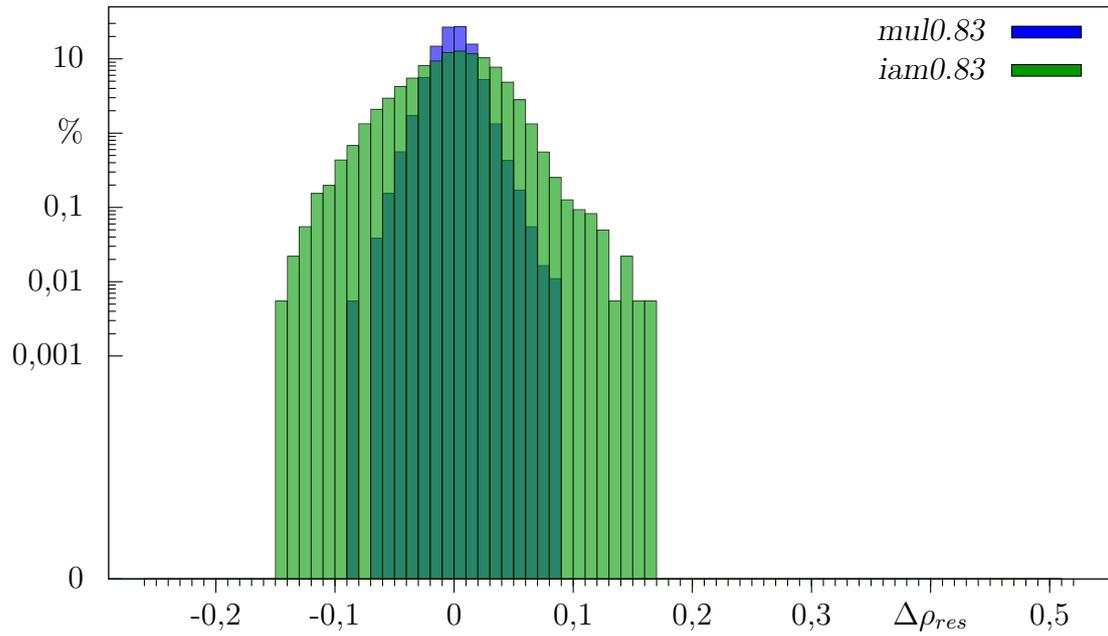
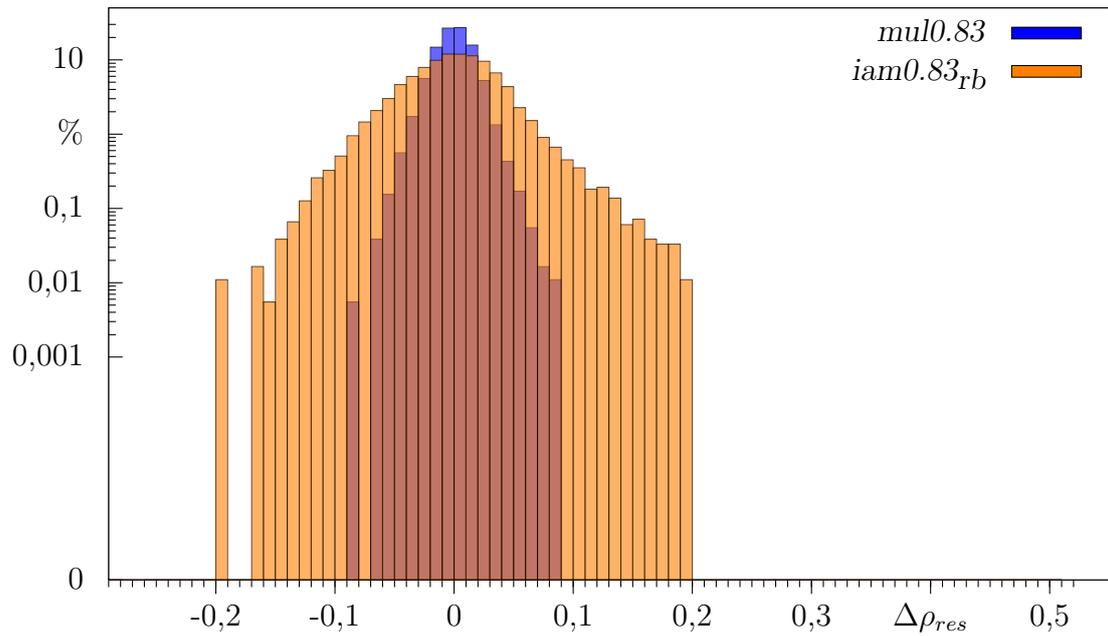
(a) Restdichten APA, *mul* vs *iam*.(b) Restdichten APA, *mul* vs *iam_{rb}*.

Abbildung 5.7: Histogramm der APA-Restdichtegitter $\sum(F_o - F_c) \exp(i\phi_c)$ des asphärischen Modells (blau) und der sphärischen Modelle (grün und orange). Intervalle $\Delta(i; i + 0,01]$, Einheiten $e\text{\AA}^{-3}$.



(a) Restdichten APA0.83, *mul0.83* vs *iam0.83*.



(b) Restdichten APA0.83, *mul0.83* vs *iam0.83_{rb}*.

Abbildung 5.8: Histogramm der APA0.83-Restdichtegitter $\sum(F_o - F_c) \exp(i\phi_c)$ des asphärischen Modells (blau) und der sphärischen Modelle (grün und orange). Intervalle $\Delta(i; i + 0,01]$, Einheiten $\text{e}\text{\AA}^{-3}$.

Mit der asphärischen Modellierung von APA0.83 ist eine noch deutlichere Ausprägung der genannten Effekte verbunden: Die ohnehin nahe Null liegenden Restdichtemittelwerte der sphärischen Verfeinerungen sinken besonders stark und die Extremwerte und Streuungsmaße erfahren eine Halbierung in *mul0.83*. Die Statistik der sphärischen Modelle *iam0.83* und *iam0.83_{rb}* in Tab. 5.3 weist doppelt so gute Deskriptoren wie die der Verfeinerung mit allen Daten auf, allerdings sind keine so deutlichen Unterschiede mehr zwischen ihnen ersichtlich.

Das generell niedrigere Restdichteniveau von APA0.83 im Vergleich zu APA wird auch in den Diagrammen 5.8(a) und (b) deutlich. Anhand dieser Verteilungen, sowie den unteren und oberen Quartilsgrenzen, ist die zum Nullpunkt symmetrische Restdichteverteilung in nunmehr allen Modellen ersichtlich. Die eklatante Wirkung der Invariommodellierung äußert sich zudem in den Unterschieden der Populierung der Intervalle $(-0,01; 0]$ und $(0; 0,01]$, in denen in der Summe bei *mul0.83* 54 % aller Restdichten liegen, aber nur 25 % und 24 % der unbeschriebenen Elektronendichte der Modelle *iam0.83* und *iam0.83_{rb}*. Fast die gesamte Restdichte, 99 %, liegt in der asphärischen Modellierung im Intervall $(-0,043; 0,043]$, bei den sphärischen Beschreibungen in weiteren Grenzen von $(-0,099; 0,085]$ und $(-0,111; 0,123]$.

5.6 Deformationsdichten

Die unterschiedliche Beschreibung der Elektronendichte in asphärischem und sphärischem Modell läßt sich gut anhand der Deformationselektronendichte darstellen, welche in Abb. 5.9 auf der folgenden Seite gezeigt wird. Im Multipolmodell findet eine gerichtete Modellierung statt, um der Abweichung der Valenzelektronendichte von der Kugelform entlang der Bindungen und an den freien Elektronenpaaren zu entsprechen. Da die Elektronendichte nur unverteilt wird, sind benachbarte Raumbereiche geringer als das sphärische Modell populiert, so daß sich positive und negative Deformationsdichten gegenseitig aufheben.

Diese Größe wurden auf einem Gitter mit $0,11 \text{ \AA}$ weiten Abständen berechnet, womit das Voxelvolumen von $0,0013 \text{ \AA}^3$ dem der Restelektronendichte ähnlich ist. Die Summation $\sum |\Delta\rho_{def}|V_{vox}$ des normierten Betrags der Deformationsdichte von APA ergibt $15,12 e$, um die das Invariommodell vom sphärischen verschieden ist. Damit werden im Mittel $0,38 e$ pro Atom verschoben. Das Verhältnis von Deformations- und Restdichte beträgt somit $\Delta\rho_{def}/\Delta\rho_{res} = 15,12 e/12,57 e = 1,203$ – es gibt 20 % mehr Deformationsdichte als verbleibende, vom Modell *mul* unbeschriebene Dichte. Mit der geringeren Restdichte von APA0.83 beträgt dieses Verhältnis sogar $15,12 e/3,97 e = 3,810$, hier übersteigt die Deformationsdichte die Restdichte von *mul0.83* um ein Vielfaches.

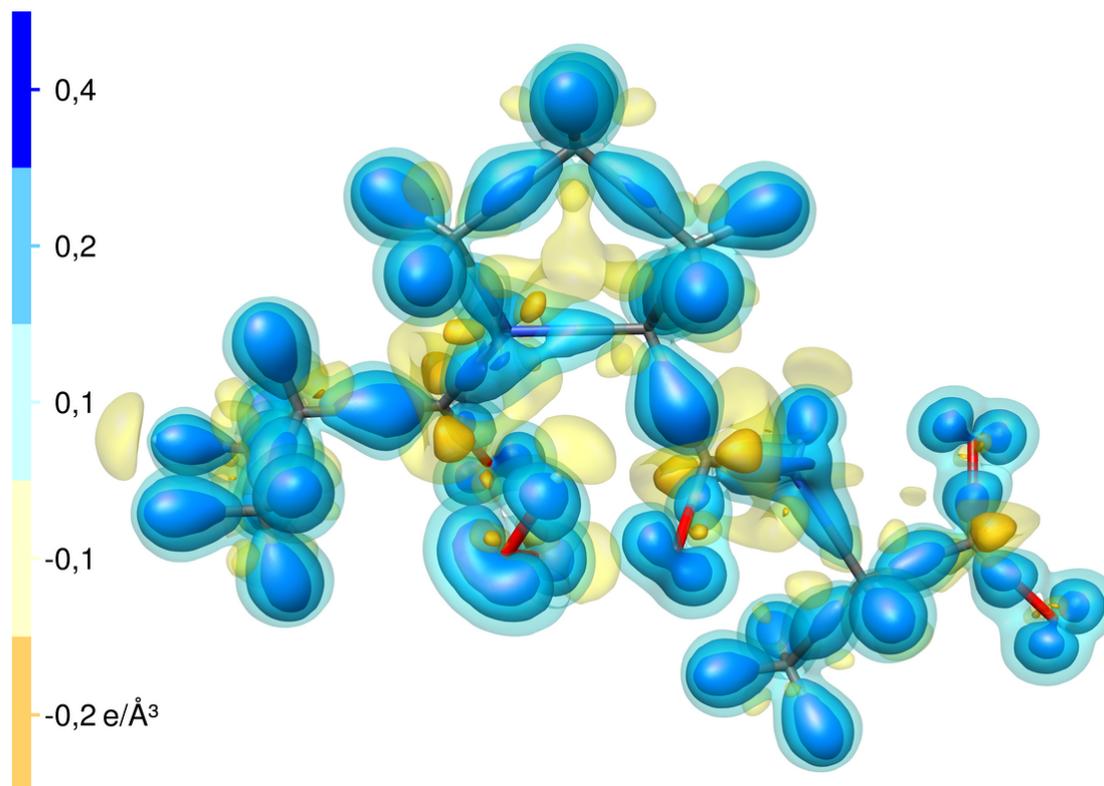


Abbildung 5.9: Deformationsdichte APA.

5.7 Abgeleitete Eigenschaften

Stellvertretend für etliche aus einer hochwertigen Elektronendichteverteilung ableitbare Größen, wurde das elektrostatische Potential (ESP) von APA aus der Multipolelektronendichte *mul* bestimmt. Zur Abschätzung der Qualität dieses ESPs, wurde es ebenfalls aus mehreren *ab initio* und semiempirischen Rechnungen bestimmt. Für die kleine Struktur APA sind diese auch auf hohem theoretischen Niveau zugänglich, was für Makromoleküle nur sehr eingeschränkt zutrifft. Die Unterschiede im Potential und der Elektronendichteverteilung werden quantitativ ausgewertet, wodurch die Resultate aus *mul* eingeordnet werden können. Unter der Annahme der Übertragbarkeit dieser Ergebnisse auf die makromolekularen Strukturen, ergibt sich eine indirekte Beurteilung der Güte des dort aus dem Invariommodell abgeleiteten Potentials. Des weiteren erfolgte die Berechnung durch Lösung der Poisson-Boltzmann-Gleichung für die Ladungsverteilung aus *amber03*-Partiellladungen. Auch für Makromoleküle sind solche Punktladungs-ESPs schnell erhältlich, hier können sie höherwertigen Rechnungen gegenübergestellt werden.

Abb. 5.10 auf der nächsten Seite zeigt das aus *mul* bestimmte elektrostatische Potential, als Farbcode in Einheiten von e/a_0 auf der Isooberfläche der ebenfalls aus

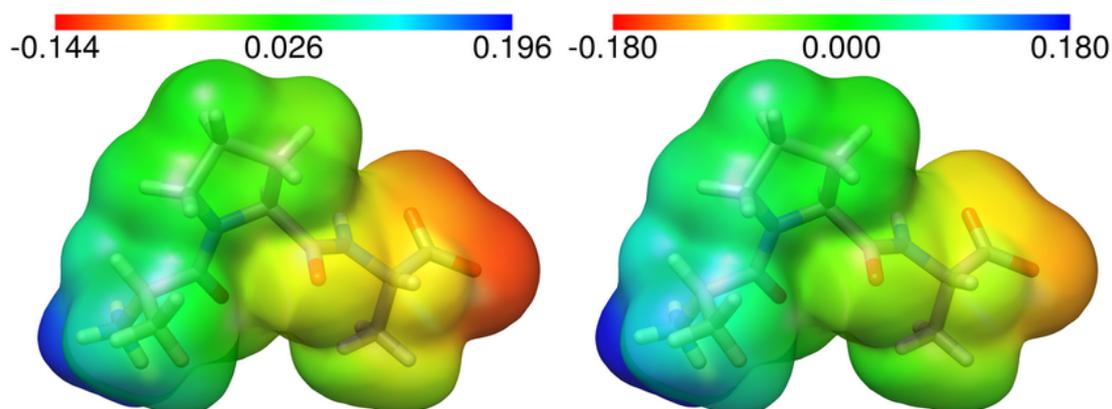


Abbildung 5.10: Elektrostatistisches Potential in e/a_0 von APA im Multipolmodell *mul* auf $0,001 e/a_0^3$ -Oberfläche, voller Wertebereich (links) und beschränkt auf $\pm 0,18 e/a_0$ (rechts).

mul bestimmten Elektronendichte bei einem Wert von $0,001 e/a_0^3$. Diese entspricht ungefähr dem van-der-Waals-Volumen des Moleküls. Auf der linken Seite der Abbildung umfaßt die Kolorierung den vollständigen ESP-Wertebereich von $-0,144$ bis $0,196 e/a_0$. Rechterhand ist die Farbcodierung auf $\pm 0,180 e/a_0$ festgelegt, um den visuellen Vergleich mit nachfolgenden Abbildungen zu erleichtern. Das der Erwartung entsprechende Ergebnis zeigt einen stark elektropositiven Bereich um die N-terminale Ammoniumgruppe, die stark elektronegative Carboxylatgruppe des C-Terminus und weitestgehend elektroneutrale Regionen an den aliphatischen Strukturteilen.

Die zur Darstellung und weiteren Auswertung benutzten Elektronendichte- und ESP-Gitter wurden mit einer recht groben Auflösung von $0,3 \text{ \AA}$, entsprechend einem Voxelvolumen von $0,027 \text{ \AA}^3$, erstellt, um den Auswertungen der Makromoleküle vergleichbare Randbedingungen sicherzustellen.

5.7.1 Vergleich *ab initio*-ESP mit Invariom-ESP

Für vier theoretische Näherungen wurden, wie im vorherigen Fall *mul*, kubische Elektronendichte- und ESP-Gitter berechnet, mit jeweils gleichem Mittelpunkt ($x_0, y_0, z_0 = 3,87, 3,18, 17,51$ im kartesischen Koordinatensystem) und jeweils gleicher Anzahl von 51 Gitterpunkten in allen drei Raumrichtungen, bei identischen Abständen von $0,3 \text{ \AA}$. Damit wird ein Würfel von 15 \AA Seitenlänge erzeugt, bei dem für alle Dichten und Potentiale die Koordinaten jedes Gitterpunktes übereinstimmen. Somit können die ρ - und ϕ -Gitter der verschiedenen Rechnungen voneinander subtrahiert werden, wodurch jeweils ein neues Gitter mit den quantitativen Unterschieden in der dreidimensionalen Verteilung der beiden Eigenschaften zwischen den Methoden erzeugt wird.

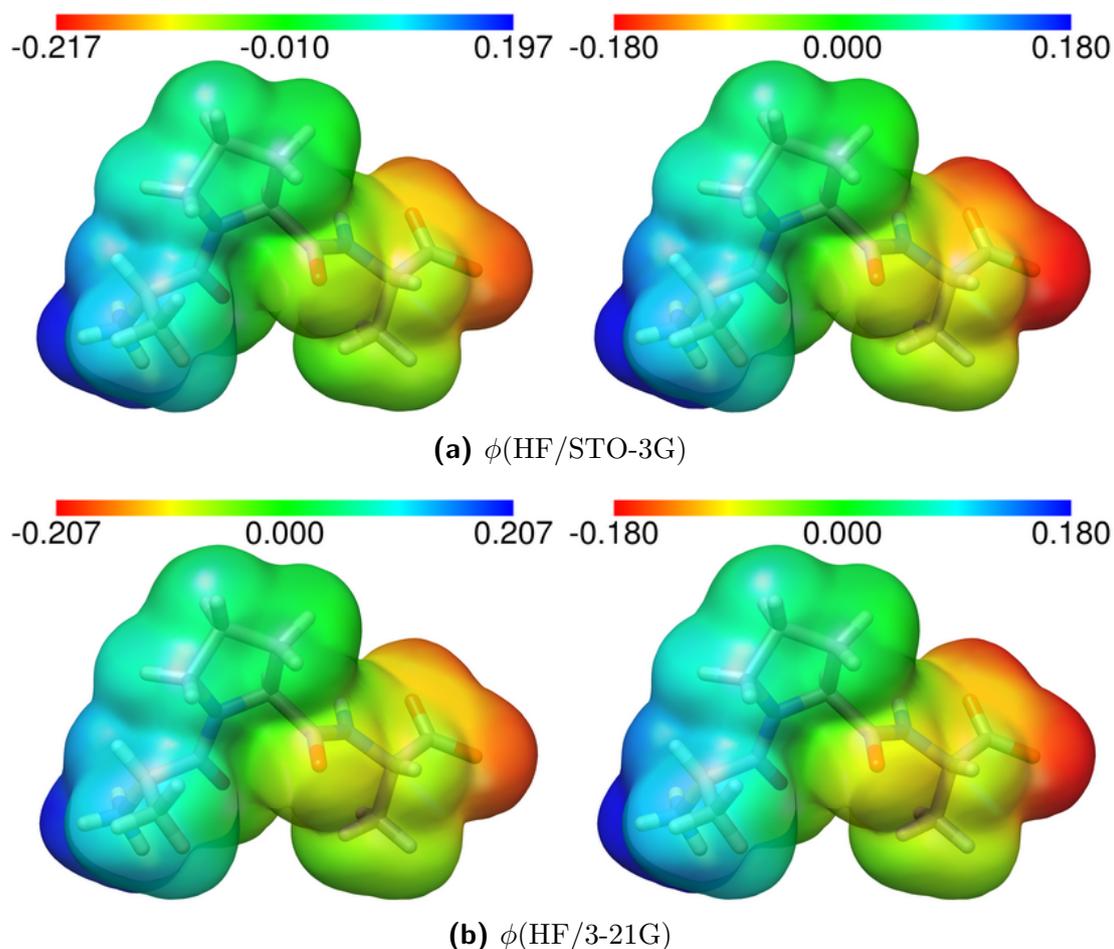


Abbildung 5.11: Elektrostatisches Potential in e/a_0 von APA aus verschiedenen *ab initio*-Methoden auf $0,001 e/a_0^3$ -Oberfläche, voller Wertebereich (links) und beschränkt auf $\pm 0,18 e/a_0$ (rechts).

Die elektrostatischen Potentiale aus *ab initio*-Näherungen sind in den Abb. 5.11(a) bis (d) in der gleichen Weise wie in Abb. 5.10 aufgetragen. Sie wurden aus g03-Rechnungen mit zwei Methoden (HF und B3LYP) und drei Basissätzen (STO-3G, 3-21G und 6-311++G(3df,3pd)) erzeugt. Mit der um der Vergleichbarkeit willen gewählten Gitterauflösung entfallen sicherlich einige Unterschiede zwischen den Rechnungen. Dennoch zeigt sich mit steigender Güte der theoretischen Näherung ein eindeutiger Trend des ESPs: Von 5.11(a) zu 5.11(d) wird dieses immer schwächer, was zum einen an den abnehmenden Extremwerten linkerhand und zum anderen an der weniger intensiven Farbgebung in den auf $\pm 0,18 e/a_0$ beschränkten Illustrationen erkennbar wird.

Eine Aussage zur Ähnlichkeit mit dem Ergebnis in Abb. 5.10 auf der vorherigen Seite fällt noch schwer, wiewohl bereits jetzt klar ist, daß es sich mehr von den

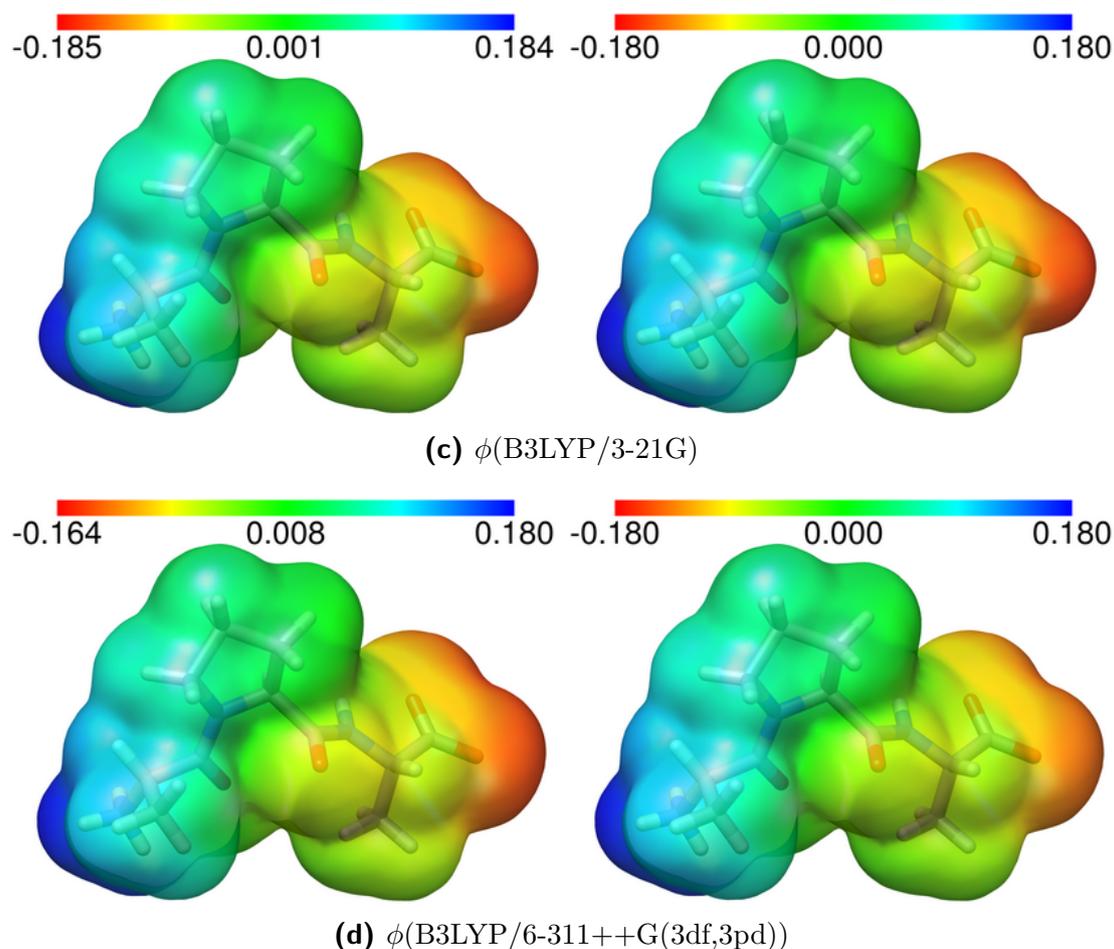
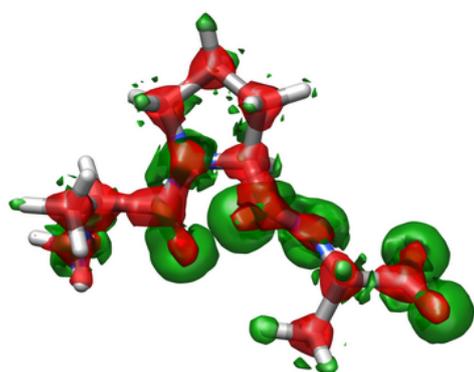


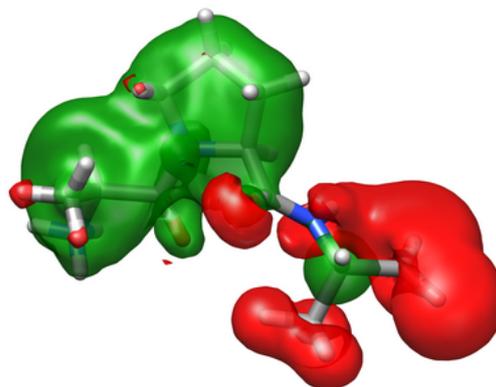
Abbildung 5.11: Elektrostatistisches Potential in e/a_0 von APA aus verschiedenen *ab initio*-Methoden auf $0,001 e/a_0^3$ -Oberfläche, voller Wertebereich (links) und beschränkt auf $\pm 0,18 e/a_0$ (rechts).

Varianten HF/STO-3G und HF/3-21G unterscheidet als von den höherwertigen Rechnungen B3LYP/3-21G und B3LYP/6-311++G(3df,3pd). Deshalb wurden Differenzgitter der Elektronendichte $\Delta\rho$ und des elektrostatischen Potentials $\Delta\phi$ aus den vier *ab initio*-Rechnungen und dem Invariommodell *mul* gebildet. Zur Auswertung werden sie in Abb. 5.12 auf der nächsten Seite visualisiert, in Tab. 5.4 auf Seite 137 statistisch ausgewertet, sowie in den Histogrammen 5.13(a) bis (d) auf Seite 138 weiter untersucht.

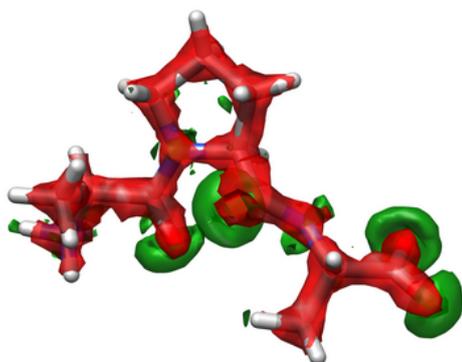
Die Darstellung der Differenzen als Isooberflächen ermöglicht durchaus, das Ausmaß an Abweichung abzuschätzen, allerdings eingeschränkt durch die Auswahl und Anzahl dargestellter Oberflächen. Zusätzlich gibt sie aber einen Eindruck vom Ort der Abweichungen, ob z. B. Unterschiede bevorzugt an einem bestimmten Element auftreten oder ob eher Bindungen, denn kernnahe Bereiche betroffen sind.



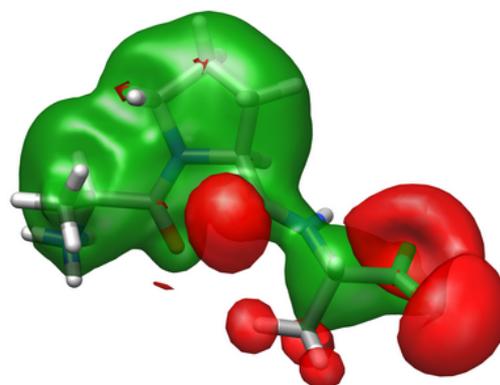
(a) $\Delta\rho(\text{HF}/\text{STO-3G} - \text{mul})$



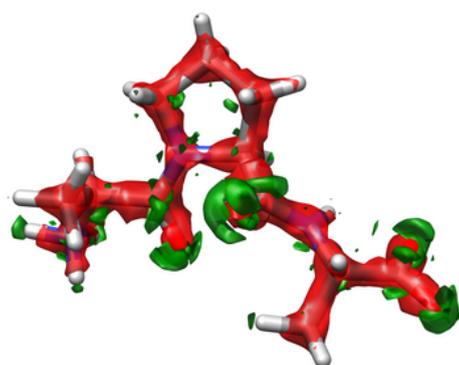
(b) $\Delta\phi(\text{HF}/\text{STO-3G} - \text{mul})$



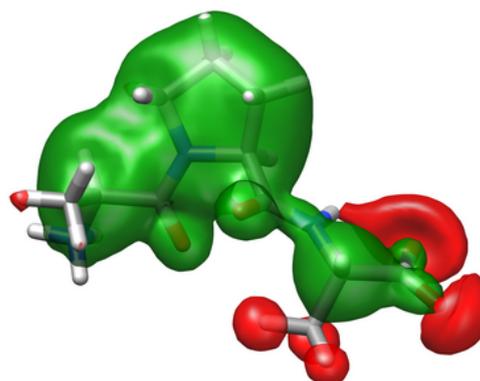
(c) $\Delta\rho(\text{HF}/3\text{-21G} - \text{mul})$



(d) $\Delta\phi(\text{HF}/3\text{-21G} - \text{mul})$

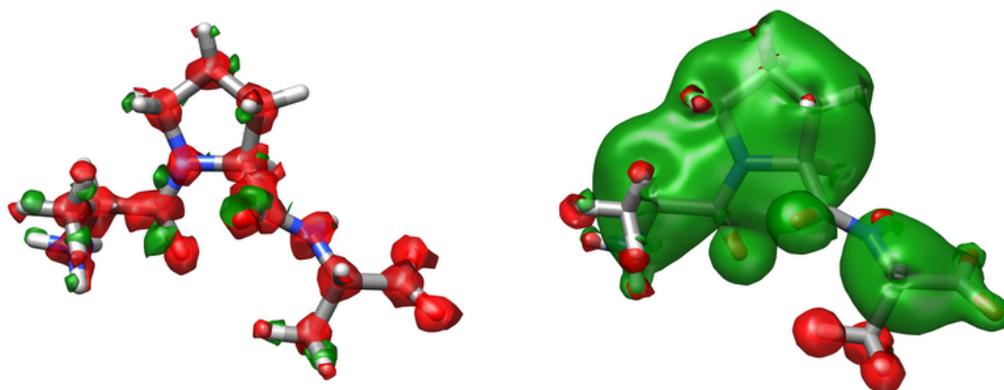


(e) $\Delta\rho(\text{B3LYP}/3\text{-21G} - \text{mul})$

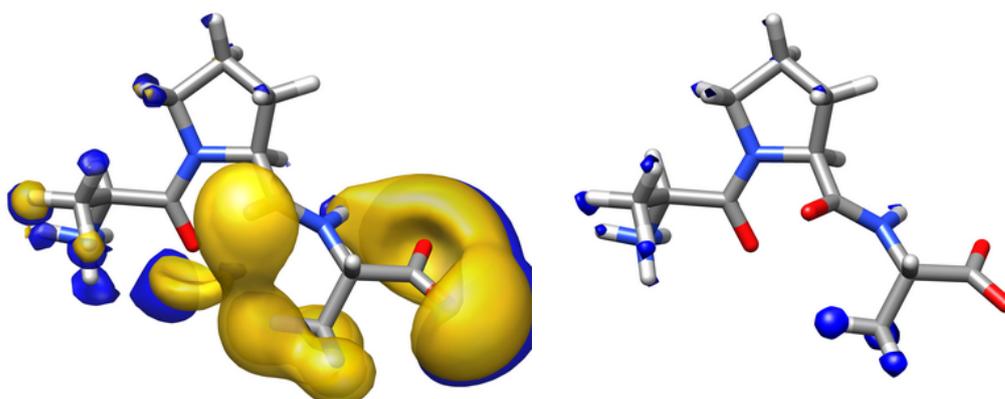


(f) $\Delta\phi(\text{B3LYP}/3\text{-21G} - \text{mul})$

Abbildung 5.12: (a)-(h) Isooberflächen der Differenzen verschiedener Theorie-*mul*-Kombinationen bei $\Delta\rho = \pm 0,01 \text{ e}/\text{a}_0^3$ und $\Delta\phi = \pm 0,05 \text{ e}/\text{a}_0$ (positiv: grün, negativ: rot). (i)-(j) Weitere Isooberflächen für ESP-Differenzen aus B3LYP/3-21G-*mul* (gelb) und B3LYP/6-311++G(3df,3pd)-*mul* (blau) bei $-0,025 \text{ e}/\text{a}_0$ und $-0,1 \text{ e}/\text{a}_0$.



(g) $\Delta\rho(\text{B3LYP}/6\text{-}311++\text{G}(3\text{df},3\text{pd}) - \text{mul})$ (h) $\Delta\phi(\text{B3LYP}/6\text{-}311++\text{G}(3\text{df},3\text{pd}) - \text{mul})$



(i) $\Delta\phi(\text{B3LYP}/6\text{-}311++\text{G}(3\text{df},3\text{pd}) - \text{mul})$, (j) $\Delta\phi(\text{B3LYP}/6\text{-}311++\text{G}(3\text{df},3\text{pd}) - \text{mul})$,
 $\Delta\phi(\text{B3LYP}/3\text{-}21\text{G} - \text{mul})$ $\Delta\phi(\text{B3LYP}/3\text{-}21\text{G} - \text{mul})$

Abbildung 5.12: (a)-(h) Isooberflächen der Differenzen verschiedener Theorie-*mul*-Kombinationen bei $\Delta\rho = \pm 0,01 \text{ e}/\text{a}_0^3$ und $\Delta\phi = \pm 0,05 \text{ e}/\text{a}_0$ (positiv: grün, negativ: rot). (i)-(j) Weitere Isooberflächen für ESP-Differenzen aus B3LYP/3-21G - *mul* (gelb) und B3LYP/6-311++G(3df,3pd) - *mul* (blau) bei $-0,025 \text{ e}/\text{a}_0$ und $-0,1 \text{ e}/\text{a}_0$.

Wie man an den kleinen Isooberflächenwerten erkennt, sind die Unterschiede in der Verteilung der Elektronendichte gering. An den Kernorten und teilweise entlang der Bindungen zeigen die roten Oberflächen bei $-0,01 \text{ e}/\text{a}_0^3$ eine geringfügig höhere Ladungsdichte im Modell *mul* an. Umgekehrt weisen die grünen Oberflächen bei $+0,01 \text{ e}/\text{a}_0^3$ Orte aus, an denen in den *ab initio*-Modellen mehr Dichte zu finden ist. Dies trifft insbesondere auf die Umgebung der Sauerstoffatome zu. Am geringsten erscheinen die Unterschiede, zumindest auf dieser Isooberfläche, zwischen *mul* und B3LYP/6-311++G(3df,3pd).

Die Beurteilung der $\Delta\phi$ -Darstellungen ist insofern komplizierter, als die vonein-

ander subtrahierten Gitter bereits positive und negative Werte aufweisen können, womit vier Arten von Differenzen („+“−, „+“=, „±“, „+“−, „−“=, „+“, „−“−, „+“=, „−“, „−“−, „−“=, „±“) auftreten könnten. Ob der Ähnlichkeit der ESPs in Abb. 5.10 und 5.11 aber kann ein gleiches Vorzeichen der Werte an verschiedenen Orten angenommen werden, was zudem anhand der Isooberflächen der ursprünglichen Gitter überprüft wurde.

Somit zeigen alle *ab initio*-Rechnungen entlang fast des gesamten Atomgerüsts, insbesondere aber im aliphatischen und N-terminalen Bereich, ein höheres positives elektrostatisches Potential (grüne $+0,05 e/a_0$ Oberfläche). An der der Carboxylatfunktion benachbarten Methylgruppe liegt in *mul* ein etwas höheres elektropositives Potential (rote $-0,05 e/a_0$ Oberfläche) vor. Um die Carboxylatfunktion und an einem Carbonylsauerstoff wiederum, zeigt *mul* ein geringeres elektronegatives Potential als die meisten *ab initio*-Modelle (ebenfalls rote $-0,05 e/a_0$ Oberflächen). Mit der Einschränkung, daß hier nur zwei Isooberflächen bei $\pm 0,05 e/a_0$ betrachtet wurden, erscheint die beste Übereinstimmung wieder mit B3LYP/6-311++G(3df,3pd) zu bestehen, gefolgt von B3LYP/3-21G.

Dies relativiert sich mit Betrachtung der Abb. 5.12(i) und (j), in denen für die beiden letztgenannten Differenzen jeweils zwei weitere Isooberflächen bei $-0,025 e/a_0$ und $-0,1 e/a_0$ gegenübergestellt werden. Je nach betrachteter Molekülregion, kann auch der Unterschied in B3LYP/3-21G geringer ausfallen, was beim höchsten Isooberflächenwert $-0,1 e/a_0$ besonders deutlich wird, da hier nur noch Differenzen zur Rechnung B3LYP/6-311++G(3df,3pd) sichtbar sind. Somit wäre das ESP aus B3LYP/3-21G, dem aus *mul* am ähnlichsten.

Abschließend soll die Übereinstimmung von Theorie- und Multipolergebnissen anhand der $\Delta\rho$ - und $\Delta\phi$ -Statistiken über die gesamten Differenzgitter in Tab. 5.4 und den Histogrammen in Abb. 5.13 auf Seite 138 beurteilt werden.

Die $\Delta\rho$ -Statistik zeigt für die Extrema aller Verteilungen aus der Subtraktion des Modells *mul* von den *ab initio*-Ergebnissen, sehr hohe negative (ca. $-106 e/a_0^3$) und kleine positive ($0,03-0,14 e/a_0^3$) Werte. Allerdings liegen die Mittelwerte bei $-0,00254 e/a_0^3$, die Mediane unter $10^{-5} e/a_0^3$ und der Interquartilabstand ebenfalls. Die extremen negativen Ausreißer rühren von einer einstelligen Zahl von fehlerhaft berechneten Werten des XD-Gitters in unmittelbarer Nähe der Atompositionen her und sind für das Gesamtgitter irrelevant. Daraus folgt, daß die Elektronendichte aus *mul* den *ab initio*-Rechnungen bei der aktuellen Gitterfeinheit äquivalent ist.

Unterstützt wird dieser Befund durch die Histogramme 5.13(a) bis (d), die auf der Abszisse die Differenzen in Intervallen von $(i; i+0,1]$ (e/a_0^3 für $\Delta\rho$, e/a_0 für $\Delta\phi$) angeben und auf der Ordinate die prozentualen relativen Häufigkeiten der Werte logarithmisch darstellen. Fast alle $\Delta\rho$ -Werte liegen nahe Null, beispielsweise beträgt für das Gitter $\Delta\rho(\text{HF}/\text{STO-3G} - \text{mul})$ die relative Häufigkeit von Werten im Bereich $(-0,1; 0,0] e/a_0^3$ 97,07%. Daß heißt, von 132651 Gitterpunkten haben 128758 einen Wert in diesem Intervall. Im nachfolgenden Intervall $(0,0; 0,1]$ befinden sich 3771 oder 2,84% der Werte, nur die restlichen 122 Differenzgitterpunkte verteilen sich

Tabelle 5.4: Statistik der Differenzgitter aus *ab initio*-Rechnungen und dem Modell *mul*.

	Δ_{min}^\dagger	Δ_{max}^\dagger	$\bar{\Delta} \cdot 10^{3\dagger}$	$Q1 \cdot 10^{3\dagger}$	$Q2 \cdot 10^{3\dagger}$	$Q3 \cdot 10^{3\dagger}$	$IQR \cdot 10^{3\dagger}$
(a) HF/STO-3G – <i>mul</i>							
$\Delta\rho$	-106,60	0,14	-2,54	0,00	0,00	0,00	0,00
$\Delta\phi$	-525,40	0,72	-24,63	-4,90	2,60	7,80	12,70
(b) HF/3-21G – <i>mul</i>							
$\Delta\rho$	-106,07	0,03	-2,54	-0,00	0,00	0,00	0,00
$\Delta\phi$	-525,34	0,72	-22,50	-5,29	2,56	8,96	14,24
(c) B3LYP/3-21G – <i>mul</i>							
$\Delta\rho$	-105,95	0,03	-2,54	-0,00	0,00	0,00	0,00
$\Delta\phi$	-525,35	0,66	-23,38	-4,03	1,52	6,50	10,53
(d) B3LYP/6-311++G(3df,3pd) – <i>mul</i>							
$\Delta\rho$	-105,68	0,04	-2,54	0,00	0,00	0,00	0,00
$\Delta\phi$	-525,53	0,73	-22,27	-4,26	2,02	7,49	11,75

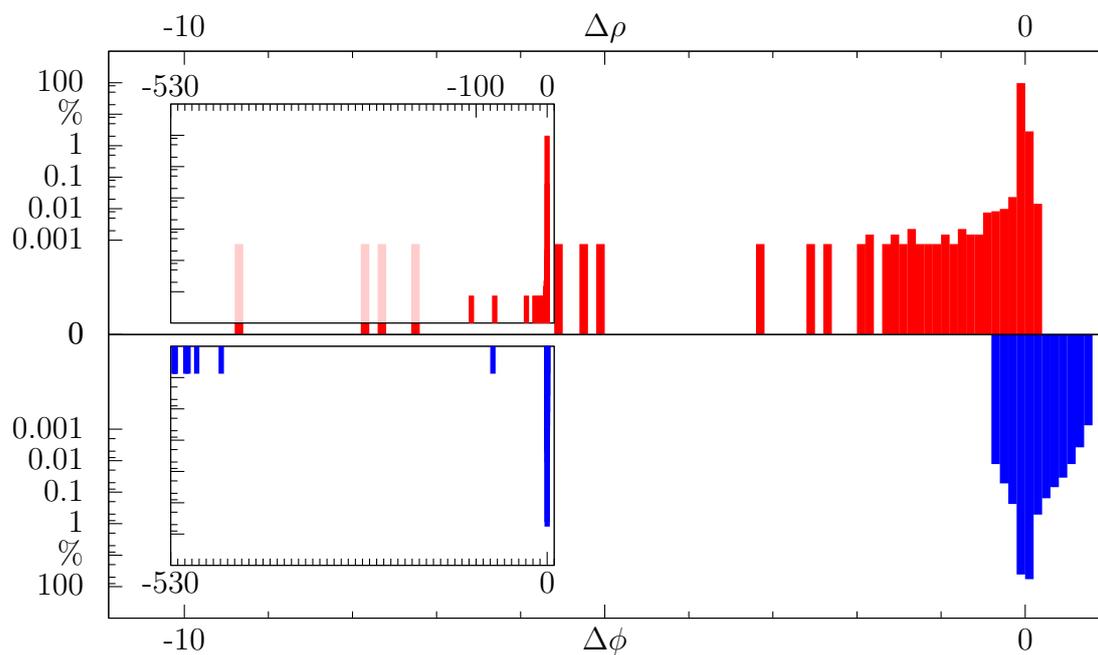
† Minimum, Maximum und Mittelwert der Gitterpunkte; ‡ 0,25-Quartil, Median, 0,75-Quartil und Interquartilabstand; ρ in e/a_0^3 , ϕ in e/a_0 .

über weiter von Null entfernte Intervalle.

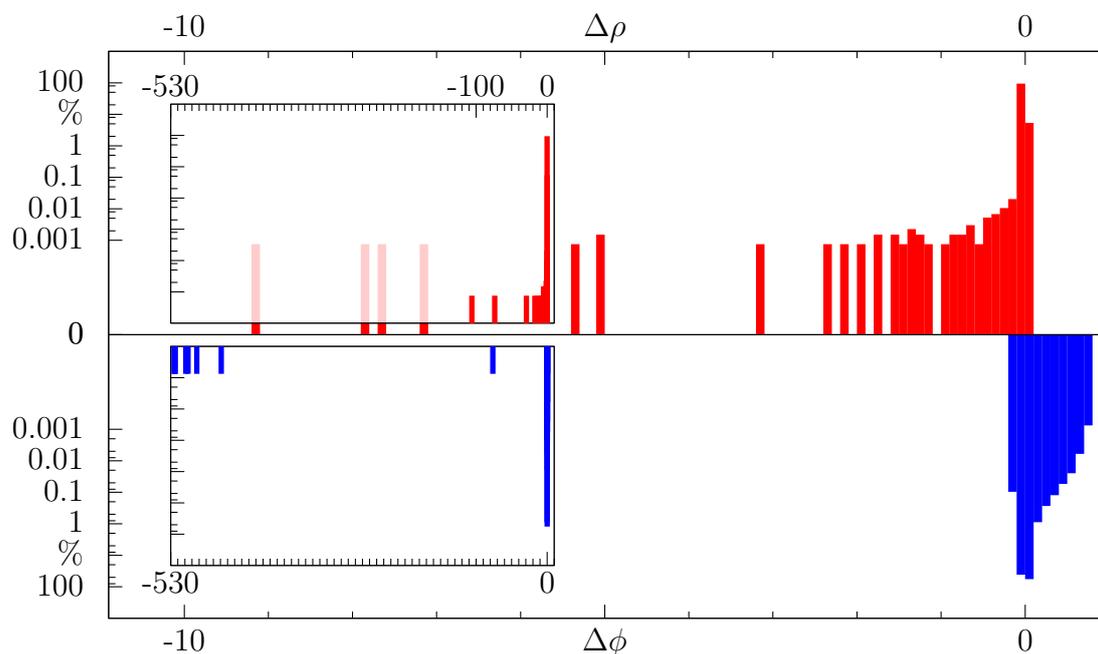
Anders das Verhalten der Differenzen des elektrostatischen Potentials. Es zeigen sich wieder sehr hohe negative (ca. $-525 e/a_0$) und kleine positive ($0,66-0,73 e/a_0$) Extremwerte. Die Mittelwerte sind, durch die starken Ausreißer beeinflusst, zehnmal so hoch wie im Fall der Elektronendichten und liegen nun bei ca. $-0,023 e/a_0$. Aber auch die dafür unempfindlichen Mediane differieren um $1,5-2,6 \cdot 10^{-3} e/a_0$ vom Idealwert Null, bei dem kein Unterschied zwischen *ab initio*-Ergebnissen und *mul* vorläge. Die unteren und oberen Quartile liegen in den vier Vergleichen bei ungefähr $-5 \cdot 10^{-3}$ und $8 \cdot 10^{-3} e/a_0$, die Verteilung ist also leicht ins Positive verschoben. Im Gegensatz zu $\Delta\rho$ sind nun deutliche Unterschiede zwischen den theoretischen Näherungen zu finden. So nimmt der Median der *ab initio-mul*-Differenzen von HF/STO-3G, über HF/3-21G zu B3LYP/3-21G ab, um bei B3LYP/6-311++G(3df,3pd) wieder anzusteigen. Ebenso hat der Vergleich (B3LYP/3-21G – *mul*) den kleinsten Interquartilabstand mit $0,011 e/a_0$, was hier die engste Verteilung um die Differenz Null bedeutet. Die HF-Rechnungen, aber auch die Rechnung mit B3LYP/6-311++G(3df,3pd) liegen wieder darüber.

Auch legt das Histogramm 5.13(c) die am wenigsten von Null verschiedene Verteilung von $\Delta\phi$ nahe. Dies äußert sich unter anderem in der mit 98,92% höchsten Häufigkeitsverteilung im Bereich $(-0,1; 0,1] e/a_0$. Alle anderen Differenzen weisen hier etwas niedrigere Werte auf und somit breitere Verteilungen.

Unter Beachtung aller Befunde wird das elektrostatische Potential aus B3LYP/3-21G dem aus *mul* am ähnlichsten befunden.

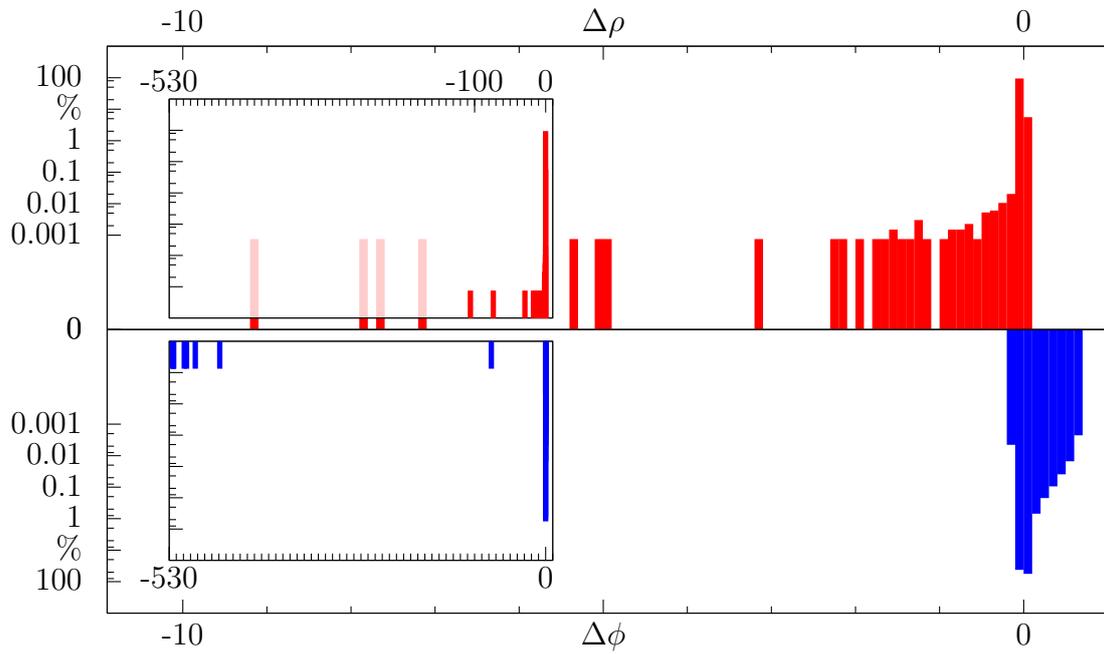


(a) HF/STO-3G - XD

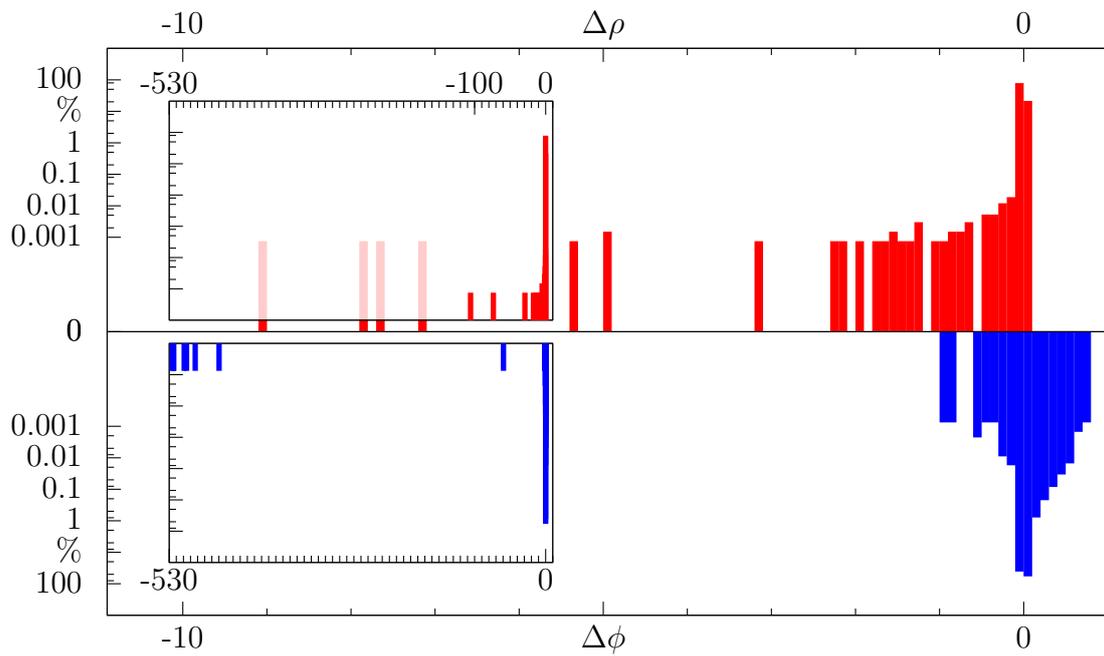


(b) HF/3-21G - XD

Abbildung 5.13: Histogramm der Differenzen $\Delta\rho$ (rot, in e/a_0^3) und $\Delta\phi$ (blau, in e/a_0) zwischen *ab initio*-Modellen und *mul.* Intervalle $(i; i + 0,1]$. Gesamtbereich in jeweils kleinerer Darstellung.



(c) B3LYP/3-21G – XD



(d) B3LYP/6-311++G(3df,3pd) – XD

Abbildung 5.13: Histogramm der Differenzen $\Delta\rho$ (rot, in e/a_0^3) und $\Delta\phi$ (blau, in e/a_0) zwischen *ab initio*-Modellen und *mul.* Intervalle ($i; i + 0,1$]. Gesamtbereich in jeweils kleinerer Darstellung.

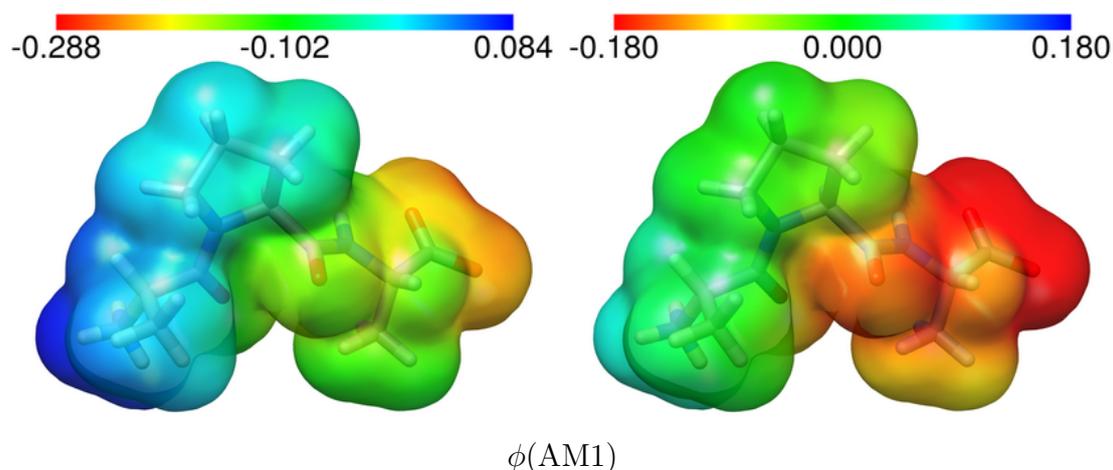


Abbildung 5.14: AM1-ESP (e/a_0) auf $0,001 e/a_0^3$ -Oberfläche, vollständiger (links) und auf $\pm 0,18 e/a_0$ beschränkter (rechts) Wertebereich.

5.7.2 Vergleich semiempirisches ESP mit Invariom-ESP

Im Gegensatz zu den *ab initio*-Methoden ist es möglich, wenn auch mit mehreren Monaten langwierig, für Makromoleküle eine semiempirische Rechnung durchzuführen und so das elektrostatische Potential zu gewinnen (siehe Abschn. 6.7 auf Seite 183). Zur Beurteilung dieser theoretischen Näherung wurden auch für APA die ESPs aus AM1 und PM3 gewonnen. Hier werden nur die AM1-Ergebnisse diskutiert, PM3 ergibt ganz Ähnliches. Wie sich herausstellen wird, sind die Resultate stark von den bisherigen Befunden abweichend. Die Rechnungen wurden mit zwei Versionen des *gaussian*-Programmpaketes (*g03/RevD.01* und *g98/RevA.11.3*) durchgeführt. Weil die *g03*-ESP als unbrauchbar zu erkennen waren, beschränkt sich die nachfolgende Diskussion auf die *g98*-Ergebnisse.

Wie Abb. 5.14 entnommen werden kann, ist zumindest die Verteilung des ESPs auf der $0,001 e/a_0^3$ -Oberfläche in Übereinstimmung mit den Erwartungen und den bisherigen Befunden. Allerdings sind die Potentialwerte ins Negative verschoben, sie reichen von $-0,288$ bis nur $0,084 e/a_0$ auf dieser Oberfläche. In der auf $\pm 0,18 e/a_0$ beschränkten Darstellung, äußert sich das durch einen wesentlich größeren rotgefärbten Bereich und kleineren blauen Bereich als in den vorherigen Illustrationen.

Abb. 5.15(a) und (b) zeigen die Differenz zu *mul* in ρ und ϕ . Hier konnte für $\Delta\rho$ der gleiche Isooberflächenwert von $\pm 0,01 e/a_0^3$ wie in Abb. 5.12 auf Seite 134 gewählt werden, für $\Delta\phi$ aber mußte ein doppelt so hoher Wert von $\pm 0,1 e/a_0$ angesetzt werden, um die Darstellung übersichtlich zu halten. Daran erkennt man bereits, daß die Unterschiede der Elektronendichte, in Übereinstimmung mit den *ab initio*-Methoden, gering sind, sie entsprechen ungefähr den Resultaten aus HF/STO-3G. Das elektrostatische Potential aber zeigt gravierendere Abweichungen.

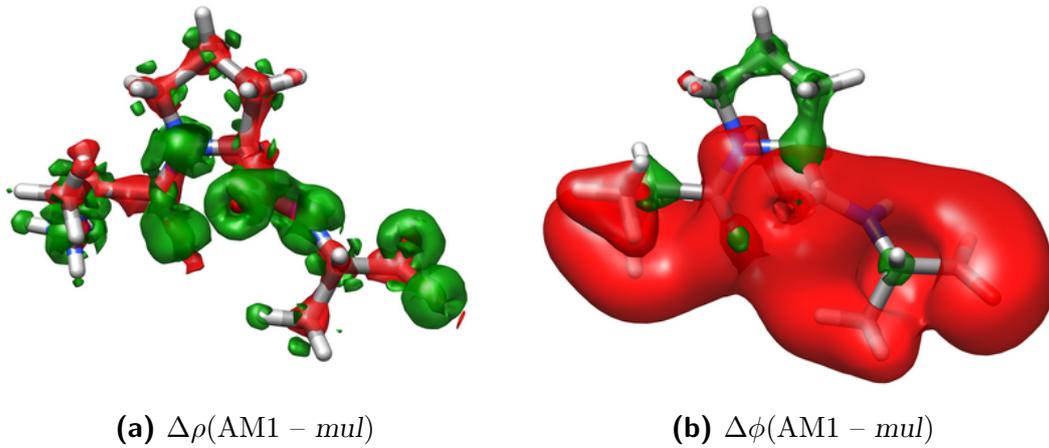


Abbildung 5.15: Isooberflächen der Differenz AM1 – *mul* bei $\pm 0,01\rho e/a_0^3$ und $\pm 0,1\phi e/a_0$ (positiv: grün, negativ: rot).

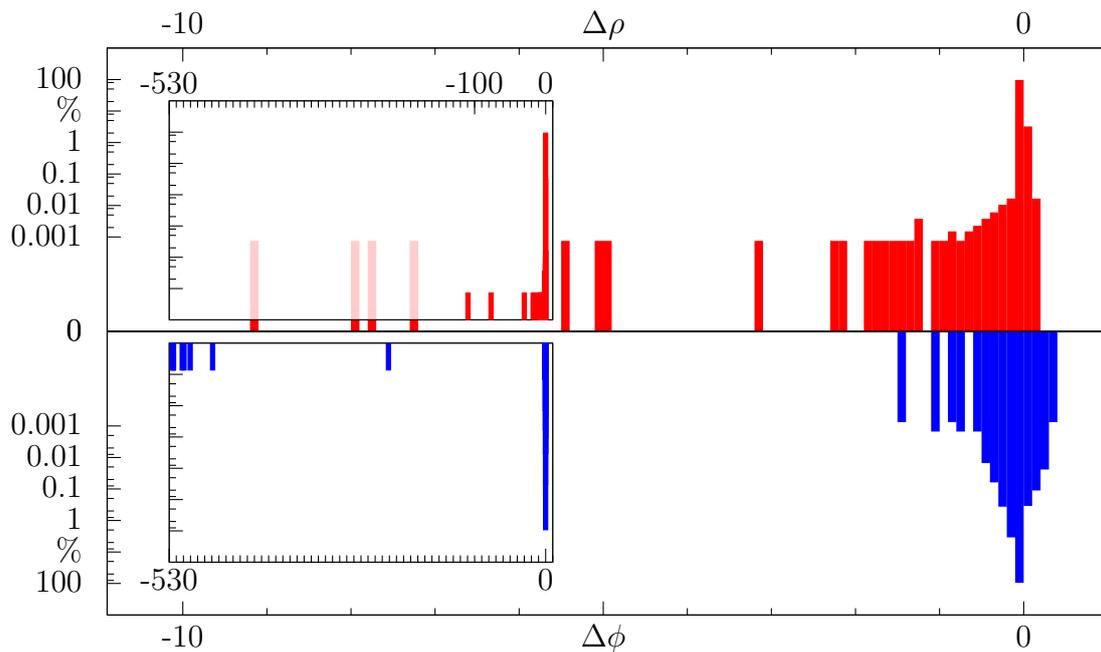


Abbildung 5.16: Histogramm der Differenzen $\Delta\rho$ (rot, in e/a_0^3) und $\Delta\phi$ (blau, in e/a_0) zwischen AM1-Ergebnissen und *mul*. Intervalle $(i; i + 0,1]$. Gesamtbereich in jeweils kleinerer Darstellung.

Tabelle 5.5: Statistik der Differenz AM1 – *mul*.

	Δ_{min}^\dagger	Δ_{max}^\dagger	$\bar{\Delta} \cdot 10^{3\dagger}$	$Q1 \cdot 10^{3\ddagger}$	$Q2 \cdot 10^{3\ddagger}$	$Q3 \cdot 10^{3\ddagger}$	$IQR \cdot 10^{3\ddagger}$
	AM1 – <i>mul</i>						
$\Delta\rho$	-109,36	0,13	-2,55	-0,00	0,00	0,00	0,00
$\Delta\phi$	-525,50	0,33	-72,77	-51,04	-37,15	-28,35	22,69

† Minimum, Maximum und Mittelwert der Gitterpunkte; ‡ 0,25-Quartil, Median, 0,75-Quartil und Interquartilabstand; ρ in e/a_0^3 , ϕ in e/a_0 .

Dies wird durch eine Tab. 5.4 analoge Statistik in Tab. 5.5 unterstützt. Die Unterschiede in der Elektronendichte aus AM1 und *mul* sind sehr gering, die des ESPs aber sehr ausgeprägt. Der Mittelwert, der Median und der Interquartilabstand sind deutlich von Null verschieden und höher als in den *ab initio*-Fällen. In den negativen 0,25- und 0,75-Quartilen spiegelt sich die bereits erwähnte Verschiebung zu elektronegativerem Potential wieder.

Das Histogramm über $\Delta\rho$ in Abb. 5.16 auf der vorherigen Seite zeigt nach wie vor eine enge Verteilung um den Nullpunkt, wie die mit 99,84 % hohe relative Häufigkeit von Werten im Intervall $(-0,1; 0,0] e/a_0^3$ belegt. Sogar im zehnmal engeren Intervall $(-0,01; 0]$ befinden sich noch 97,36 % aller Differenzgitterpunkte. Im Fall der Häufigkeitsverteilung von $\Delta\phi$ liegen in $(-0,1; 0,0] e/a_0$ nur 95,98 % der Werte, alle Differenzen zu *ab initio*-Potentialen zeigten höhere Werte.

Es wurde versucht, das ESP durch Skalierung auf die Werte aus B3LYP/3-21G zu verbessern. Dazu wurden mehrere Regressionsanalysen, sowohl über alle ϕ -Werte beider Gitter als auch nur über eine Auswahl an Punkten, durchgeführt. Aus der Anpassung der *ab initio*- und AM1-Potentialwerte, welche auf der $0,001 e/a_0^3$ -Oberfläche liegen, resultierte mit $0,0731(13) + 1,0334(150) \cdot \phi_{AM1,alt}$ die Funktion, mit der die Transformation der AM1-ESP-Werte ein annehmbares Ergebnis lieferte. Natürlich nur auf der betrachteten Isooberfläche, in anderen Raumbereichen liegt die Differenz zu *mul* nach wie vor viel höher. Eine Darstellung des skalierten ESPs und der Differenz-ESP ist in Abb. A.1 auf Seite 327 im Anhang gegeben. Zudem sind diese und andere hier bestimmte Skalierungsfunktionen, nicht auf das analoge Problem des Makromoleküls übertragbar.

Ob der Probleme mit dem AM1-ESP aus **gaussian**, wurde die Berechnung in **mopac** wiederholt. Es wird programmbedingt kein kubisches Gitter wie in **gaussian** oder **XD** berechnet, sondern das Potential nur an Koordinaten relativ zur van-der-Waals-Oberfläche ausgegeben. Deshalb wurde die Statistik über die Differenzen der ESPs aus den bisherigen und der neuen theoretischen Näherung zu *mul* an ebendiesen Punkten wiederholt und in Tab. 5.6 zusammengefasst. Zusätzlich erfolgt die Darstellung des elektrostatischen Potentials in Abb. 5.17 auf der nächsten Seite durch einen Farbcode auf diese Koordinaten symbolisierenden Kugeln, statt einer Isooberfläche.

Tabelle 5.6: Statistik der ESP-Differenz theoretischer Näherungen und *mul* an 2926 Punkten nahe der vdW-Oberfläche.

	Δ_{min}^\dagger	Δ_{max}^\dagger	$\bar{\Delta} \cdot 10^{3\dagger}$	$Q1 \cdot 10^{3\dagger}$	$Q2 \cdot 10^{3\dagger}$	$Q3 \cdot 10^{3\dagger}$	$IQR \cdot 10^{3\dagger}$
AM1 – <i>mul</i>							
$\Delta\phi$	-0,18	-0,01	-67,88	-86,22	-69,52	-42,64	43,59
MOPAC/AM1 – <i>mul</i>							
$\Delta\phi$	-0,11	0,07	-0,45	-17,79	2,22	17,58	35,36
HF/STO-3G – <i>mul</i>							
$\Delta\phi$	-0,08	0,12	3,00	-18,92	8,69	19,85	38,77
HF/3-21G – <i>mul</i>							
$\Delta\phi$	-0,09	0,14	3,49	-19,30	6,12	23,68	42,97
B3LYP/3-21G – <i>mul</i>							
$\Delta\phi$	-0,06	0,12	2,50	-15,11	1,43	18,02	33,13
B3LYP/6-311++G(3df,3pd) – <i>mul</i>							
$\Delta\phi$	-0,04	0,12	5,79	-13,48	4,03	22,10	35,57

† Minimum, Maximum und Mittelwert der Gitterpunkte; ‡ 0,25-Quartil, Median, 0,75-Quartil und Interquartilabstand; ϕ in e/a_0 .

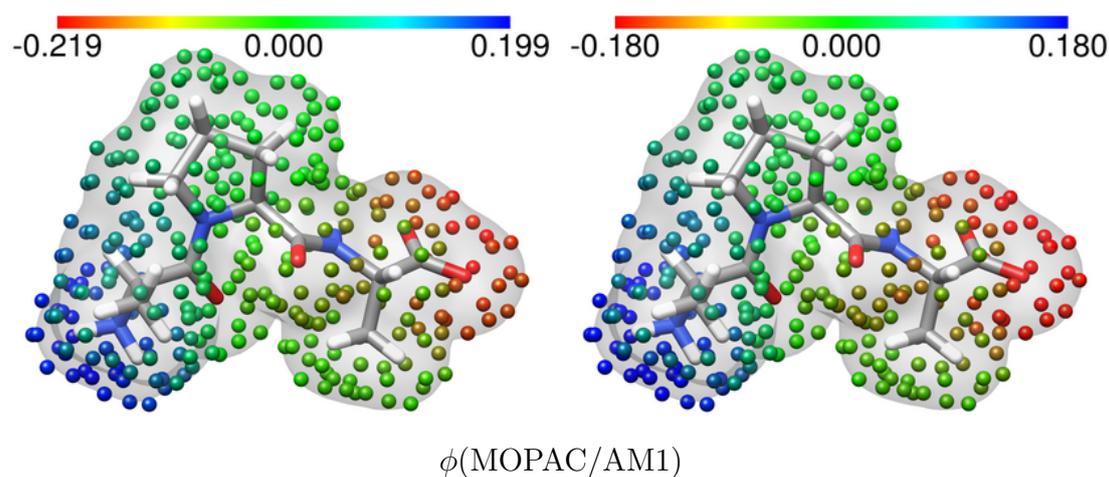


Abbildung 5.17: Elektrostatisches Potential (e/a_0) aus MOPAC/AM1 an 283 Punkten nahe der van-der-Waals-Oberfläche, vollständiger (links) und auf $\pm 0,18 e/a_0$ beschränkter (rechts) Wertebereich. Orientierungshilfe: $0,001 e/a_0^3$ -Oberfläche aus G03/B3LYP/3-21G.

Offensichtlich ist die Statistik des Differenzpotentials an 2926 Koordinaten konsistent zu den bereits durchgeführten Analysen der Gesamtdifferenzgitter: $\Delta\phi(\text{B3LYP}/3\text{-}21\text{G} - \text{mul})$ weist, mit dem am nächsten bei Null liegenden Median und kleinsten Interquartilabstand bei geringen Minimal- und Maximalwerten, die besten Werte auf. Während die Resultate der bisherigen AM1-Rechnung die erkennbar schlechtesten sind, zeigen sich nur geringe Unterschiede zwischen *mul* und der *mopac*-AM1-Rechnung, was durch die zweitkleinsten Mittel- und Streuungswerte belegt wird. Ebenso sind die Absolutwerte dieses ESPs, wie Abb. 5.17 auf der vorherigen Seite zeigt, durchaus denen aus *ab initio*-Rechnungen, und damit dem aus *mul*, vergleichbar. Die bisherigen Probleme sind also auf einen Fehler in *gaussian/cubegen* zurückzuführen.

Schlußfolgernd ist das elektrostatische Potential von APA im Modell *mul* dem aus *ab initio*- und semiempirischen Rechnungen als ähnlich zu bezeichnen, die Abweichungen liegen in der Größenordnung von $0,01 e/a_0$.

5.7.3 Punktladungs-ESP

Abb. 5.18 auf der nächsten Seite ermöglicht einen Vergleich des aus *mul* bestimmten ESPs ϕ_{mul} mit dem aus Punktladungen berechneten ϕ_{PC} (siehe Abschn. 3.4 auf Seite 47). Die Auftragung erfolgte zum einen auf der $0,001 e/a_0^3$ -Isooberfläche von *mul* und zum anderen der van-der-Waals-Oberfläche (siehe Abschn. 3.4 auf Seite 46). Letztere benötigt, wie das Punktladungs-ESP selbst, keine Informationen aus dem Modell *mul*. Die Einheiten wurden in kT/e gewählt, der Umrechnungsfaktor ist $1 e/a_0 = 1059,59 \text{kT}/e$. Somit sind die Abbildungen zu $\pm 191 \text{kT}/e$ den bisherigen Darstellungen zu $\pm 0,18 e/a_0$ äquivalent.

Zumindest auf der $0,001 e/a_0^3$ -Oberfläche sind die Potentiale sehr ähnlich, so der optische Eindruck aus dem Vergleich von Abb. 5.18(a), einer Wiederholung von Abb. 5.10, und dem Punktladungsanalogon in Illustration 5.18(b). Auch der volle Umfang der Werte von ϕ_{PC} in Abb. 5.18(e) ist mit $-159\text{--}197 \text{kT}/e$ ($-0,150\text{--}0,184 e/a_0$) in einem nur etwas elektronegativeren Bereich als ϕ_{mul} mit $-0,144\text{--}0,196 e/a_0$. In dieser Beziehung ähnelt es den *ab initio*-Potentialen, zeigt allerdings nicht ganz so hohe negative Werte wie diese.

Auf der zumeist näher an den Atomen liegenden van-der-Waals-Oberfläche A_{vdw} , sind die Differenzen ausgeprägter, wie die Gegenüberstellung von Abb. 5.18(c) und (d) verdeutlicht. Hier verstärkt sich beim Punktladungs-ESP der Trend des im Verhältnis zu *mul* elektronegativeren Potentials. Das wird durch Abb. 5.18(f) hervorgehoben, hier umfassen die Werte des ESPs einen Bereich von insgesamt $-209\text{--}207 \text{kT}/e$ ($-0,197\text{--}0,195 e/a_0$).

Wiewohl nur Extrem- und Farbwerte untersucht wurden, und unter Beachtung der Abhängigkeit der Ergebnisse von z. B. der Gitterfeinheit und der gewählten Oberfläche, scheint ϕ_{PC} den *ab initio*-Ergebnissen ähnlicher als ϕ_{mul} zu sein.

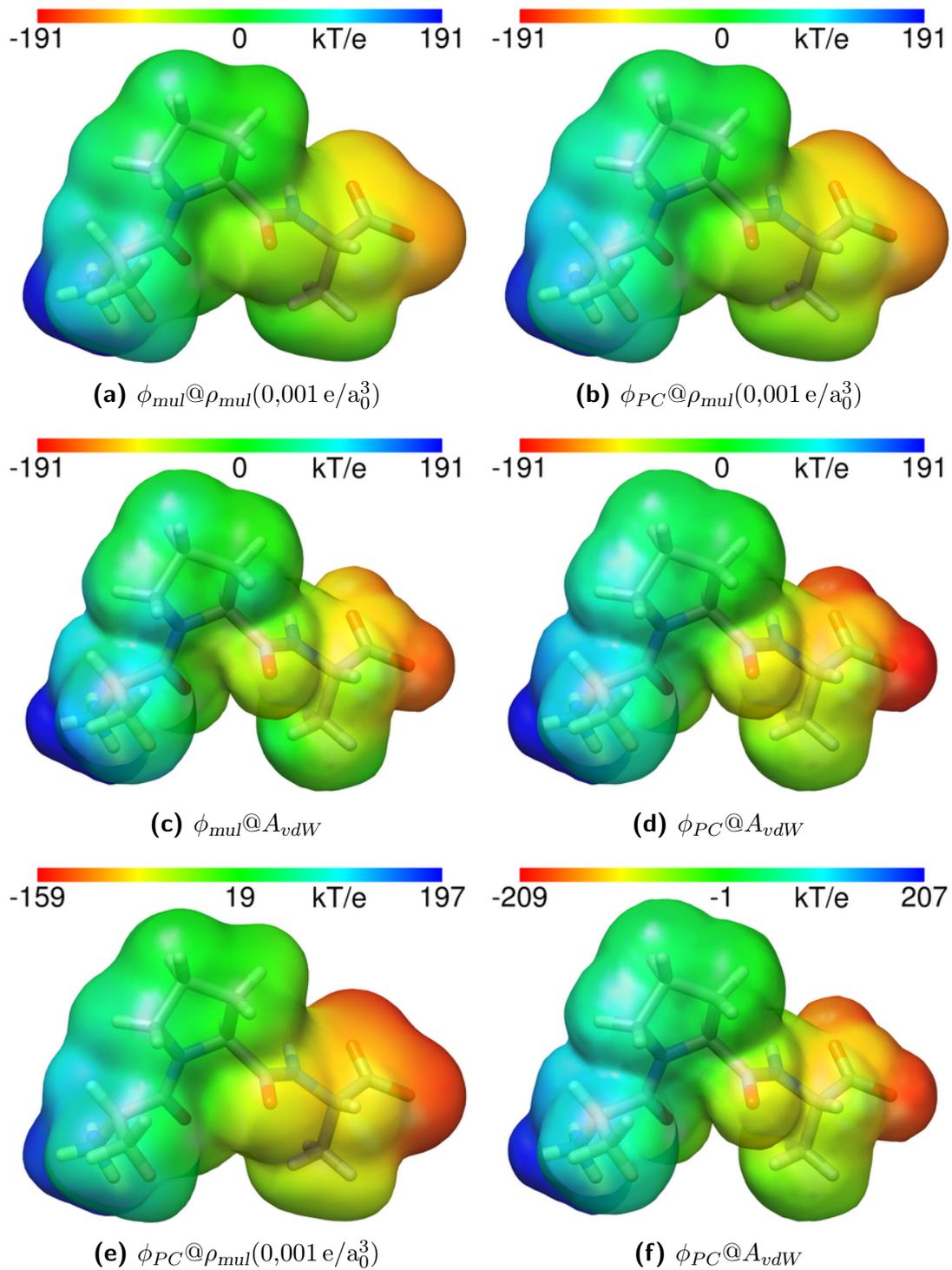


Abbildung 5.18: Multipol-ESP (ϕ_{mul}) und Punktladungs-ESP (ϕ_{PC}) auf $0,001 \text{ e}\text{\AA}^{-3}$ Isooberfläche der Multipolelektronendichte (ρ_{mul}) und van-der-Waals-Oberfläche (A_{vdW}). (a)-(d): Fester Wertebereich $\pm 191 \text{ kT/e} = \pm 0,18 \text{ e}/a_0$. (e)-(f): Voller Wertebereich.

6 Triklines Lysozym

6.1 Gütefaktoren

Wie Tab. 6.1 zeigt, unterscheiden sich die drei Modelle weder bezüglich der Reflexe noch der Parameteranzahl, wohl aber in den R -Werten: Gegenüber der sphärischen Verfeinerung iam sinkt der R -Wert für das asphärische Modell mul um 0,12 Prozentpunkte, gegenüber dem sphärischen Modell mit verlängerten Wasserstoffabständen iam_{rb} um 0,26 Punkte. Somit liegt die relative Verringerung dieses Gütefaktors bei 1,5 % bzw. 3,1 %. Die Änderungen des R_{all} und R_w betragen 0,11 und 0,15 beim Vergleich von iam und mul , in Relation zu iam_{rb} immerhin 0,25 und 0,44 Prozentpunkte. Der *Goodness of fit* schwankt zwischen 1,3 und 1,5. Anhand der mittleren auf die jeweilige Standardabweichung bezogenen Parameterverschiebungen in der Größenordnung von 10^{-6} erkennt man die Verfeinerungen als konvergent, wobei die asphärische Verfeinerung noch etwas höhere Verschiebungen exhibierte.

Bei gleichbleibender Parameterzahl wurden also die Gütefaktoren durch Anwendung der Invariome verbessert. Allerdings sind die Unterschiede recht gering, je nach verwendetem R -Faktor liegt das Verhältnis zwischen sphärischen und asphärischen Werten bei 1,015 bis 1,007. Um eine Aussage zur Signifikanz dieser Änderung treffen zu können, wurden zum einen entsprechende Tests auf die Modellunterschiede in den die R -Werte konstituierenden Differenzen zwischen beobachteten und berechneten Strukturparametern durchgeführt und zum anderen das R -Wertverhältnis nach Hamilton (Hamilton 1965) auf Signifikanz geprüft.

Tabelle 6.1: Gütefaktoren der sphärischen und asphärischen Verfeinerung von P1LYSO.

	iam	iam_{rb}	mul
Reflexe $F_o^2 > 2\sigma(F_o^2)$	174992	174992	174992
Parameter	5752	5752	5752
N_{ref}/N_{par}	30,423	30,423	30,423
$R(F)$	0,0830	0,0844	0,0818
$R_{all}(F)$	0,0859	0,0873	0,0848
$R_w(F^2)$	0,2236	0,2265	0,2221
GoF_w	1,0402	1,0542	1,0326
Mittl.(Versch./esd)	$3,75 \cdot 10^{-6}$	$6,24 \cdot 10^{-6}$	$2,98 \cdot 10^{-5}$

Als Parameter der Verteilung $\Delta F(iam) = |F_{o,iam} - F_{c,iam}|$ ergeben sich ein arithmetisches Mittel von 1,7130 und ein Zentralwert von 0,7917 bei einer maximalen Differenz von 757. Die Betrachtung für $\Delta F(mul) = |F_{o,mul} - F_{c,mul}|$ ergibt 1,6900 und 0,7866 als Mittel und Median mit einem Extremwert von 747,8. Werden die Differenzen mit $\sigma(F_o)^{-1}$ gewichtet, verringern sich die Unterschiede zwischen F_o und F_c sowohl für *iam* als auch *mul*: Der Mittelwert und Median im sphärischen Fall betragen nun 0,4206 und 0,3110, im Maximum 209,8. Die Abweichungen nach der asphärischen Verfeinerung liegen bei maximal 208,1, im arithmetischen Mittel und im Median bei 0,4174 und 0,3081.

Die Unterschiede in den Mittelwerten der beiden Modelle sind gering: Ungefähr 0,04 im ungewichteten und 0,003 im gewichteten Fall. Durch einen gepaarten *t*-Test und Wilcoxon-Test (siehe auch Absch. 3.1.2 auf Seite 36) wurden die Unterschiede der beiden Differenzen $\Delta F_w(iam) = |F_{o,iam} - F_{c,iam}|\sigma(F_{o,iam})^{-1}$ und $\Delta F_w(mul) = |F_{o,mul} - F_{c,mul}|\sigma(F_{o,mul})^{-1}$ untersucht. Mit den Hypothesen $H_0 : \Delta F_w(iam) = \Delta F_w(mul)$ und $H_1 : \Delta F_w(iam) > \Delta F_w(mul)$ erhält man im *t*-Test eine Überschreitungswahrscheinlichkeit von $p \approx 0$, ebenso im Wilcoxon-Test. Die Nullhypothese kann also statistisch signifikant abgelehnt werden, was bedeutet, daß die Verringerung der Abweichung zwischen Modell und Messung in der Verfeinerung *mul* nicht zufällig ist, sondern trotz des geringen Ausmaßes relevant ist. Auch bei Verwendung der ungewichteten Differenzen, liegen die Irrtumswahrscheinlichkeiten bei $p \approx 0$, womit die Veränderungen im asphärischen Fall wieder von einer der Nullhypothese zugrunde liegenden Zufallsverteilung zu unterscheiden sind.

Im Hamilton-Test wird das *R*-Wertverhältnis mit einem den Freiheitsgraden der Verfeinerung entsprechendem Signifikanzniveau verglichen. Der Austausch der Streufaktoren wurde als eindimensionales Problem angenommen, der zweite Freiheitsgrad ist $N_{ref} - N_{par} = 169240$. Damit wurde der Signifikanzpunkt zu einer Irrtumswahrscheinlichkeit von 5% als $\mathcal{R}_{1,169240,0,005} = 1,0000241$ extrapoliert. Da das gefundene *R*-Wertverhältnis weit darüber liegt, kann die Verbesserung auch nach Hamilton als signifikant betrachtet werden.

Effekt des Beschneidens der Daten

Die in der Verfeinerung verwendeten Intensitäten stammen, wie eingangs gezeigt, aus einer ausgesprochen guten und sorgfältigen Messung und ebensolchen Datenintegration und Verfeinerung. Dennoch konnten bisher nur geringe Verbesserungen durch das asphärische Modell erzielt werden, ganz im Gegensatz zu kleinen Molekülen (siehe Absch. 5 auf Seite 115). Um zu testen, ob günstigere Resultate bei künstlich erhöhter Datenqualität erreichbar sind, können die am stärksten abweichenden Beobachtungen von der Verfeinerung ausgeschlossen werden. Natürlich ist das eine sehr grobe Methode der „Datenverbesserung“, denn die Unterschiede zwischen beobachteten und berechneten Strukturfaktoren sind zum einen Teil auf Meßfehler und zum anderen Teil auf Modellungenauigkeiten zurückzuführen. Wäh-

Tabelle 6.2: Gütefaktoren der sphärischen und asphärischen Verfeinerung von PLYSO gegen beschnittene Daten.

	<i>iam</i>	$\frac{1}{2} \leq \frac{F_o}{F_c} \leq 2$ <i>iam_{rb}</i>	<i>mul</i>
Reflexe $F_o^2 > 2\sigma(F_o^2)$	170868	170868	170868
Parameter	5752	5752	5752
N_{ref}/N_{par}	29,706	29,706	29,706
$R(F)$	0,0760	0,0774	0,0750
$R_{all}(F)$	0,0763	0,0778	0,0754
$R_w(F^2)$	0,1959	0,1993	0,1959
GoF_w	0,9112	0,9278	0,9112
Mittl.(Versch./esd)	$6,40 \cdot 10^{-7}$	$8,65 \cdot 10^{-7}$	$2,05 \cdot 10^{-5}$

rend man sich bei einer Kleinmolekülstruktur mit einer gewissen Berechtigung auf den Standpunkt stellen kann, daß das Strukturmodell richtig ist und die Diskrepanzen auf experimentelle Fehler zurückzuführen sind, ist bei einer Proteinstruktur das Strukturmodell kaum als vollständig und ideal anzunehmen.

Dazu wurden alle F_o^2 aus dem Datensatz entfernt, die in der sphärischen Verfeinerung einen Quotienten $F_o^2/F_c^2 \geq 2$ oder $\leq 1/2$ erreichen bzw. über- oder unterschreiten. Zur Verdeutlichung dient Abb. 6.1 auf der nächsten Seite, welche den Quotienten in Abhängigkeit von der Auflösung eines jeden Reflexes zeigt. Es handelt sich um eine doppellogarithmische Auftragung, um den umfangreichen Wertebereich in gut erkennbarer Weise abzubilden. Die fast 190000 Datenpunkte überschneiden sich natürlich trotzdem in der Abbildung. Die stärksten Abweichungen vom Idealwert 1 treten bei besonders geringer und hoher Auflösung in Erscheinung. Mit dem Beschneiden mindert sich die Reflexzahl um 10602 Beobachtungen oder 5,7%, von 187165 auf 176563 Reflexe. Der Anteil starker Reflexe $> 2\sigma(F_o^2)$ steigt von 93,5% auf nunmehr 96,8%, auch das mittlere I/σ -Verhältnis wird geringfügig erhöht. Andererseits nimmt die Vollständigkeit auf 92% ab.

Sowohl die Gütefaktoren des sphärischen als auch des asphärischen Modells verringern sich in dieser Verfeinerung, wie ein Vergleich von Tab. 6.2 mit der vorherigen Auflistung Tab. 6.1 auf Seite 147 zeigt. Dabei verbessern sich die Modelle *iam* und *iam_{rb}* mit 0,7 Prozentpunkten sogar stärker als das Multipolmodell *mul*, das eine R -Wertdifferenz von 0,68 Punkten aufweist. Auf diese Weise kann der Effekt der Invariomverfeinerung also nicht näher an die, durch die Verfeinerung kleiner Moleküle begründete, Erwartungshaltung geführt werden. Dies betrifft darüberhinaus auch strukturelle Parameter wie z. B. die Temperaturfaktoren, welche analog den R -Werten sinken, allerdings wieder etwas stärker im sphärischen Modell.

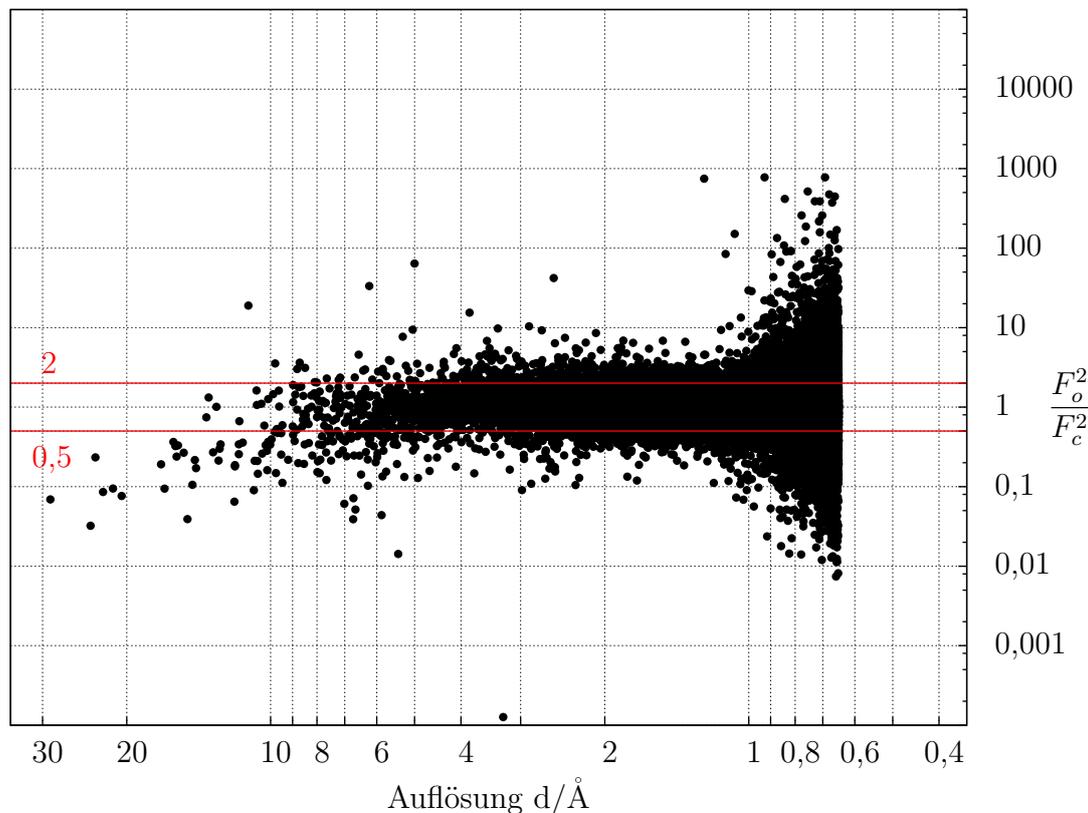


Abbildung 6.1: Quotient beobachteter und berechneter Intensitäten von P1LYSO und ihre Auflösungsabhängigkeit im Modell *iam*.

6.2 Geometrie

Zunächst soll die Geometrieänderung anhand der quadratisch gemittelten Abweichung (RMSD) der Atompositionen der Verfeinerung *mul* zu den sphärischen Modellen untersucht werden. Diese beträgt für die Gesamtstruktur inklusive Wasserstoffatomen lediglich 0,095 Å in Relation zu *iam* und 0.003 Å zu *iam_{rb}*. Die geringen Werte sind kaum verwunderlich, sind doch viele Koordinaten fixiert. An den verfeinerten Nichtwasserstoffpositionen werden maximale Abweichungen im Modell *mul* von 0,027 Å zum Modell *iam* und 0,033 Å zu *iam_{rb}* gefunden. Im gleichen Rahmen unterscheiden sich die Koordinaten der Wasserstoffatome zwischen *mul* und *iam_{rb}*. Mit den kurzen Wasserstoffabständen in *iam*, ergeben sich Diskrepanzen bis zu 0,163 Å. In diese Werte gehen sowohl Bindungslängenveränderungen als auch eventuelle Winkelunterschiede ein.

In den Abbildungen Abb. 6.2 auf Seite 152 und Abb. 6.3 auf Seite 153 sind die Strukturunterschiede graphisch dargestellt: Einem jeden Atom ist hierbei die Positionsänderung als Farbcode zugeordnet. Die hervorgehobenen Strukturabschnitte

(dick gezeichnet) entsprechen den tatsächlich verfeinerten Atomen. Die von der Verfeinerung ausgeschlossenen Bereiche sind nur dünn gezeichnet und tragen naturgemäß den blauen Farbcode für null Änderung. Die Ausnahme sind dabei die Wasserstoffatome beim Vergleich von *mul* mit *iam*, da im asphärischen Modell alle X-H-Abstände auf theoretische Werte gesetzt wurden. Um die Unterschiede zwischen den Strukturen deutlicher werden zu lassen, ist in beiden Darstellungen bereits ab einem RMSD-Wert von 0,03 Å der die maximale Abweichung kennzeichnende Farbcode rot zugeordnet, wodurch der obengenannte Verschiebungsrahmen für Schweratome differenziert dargestellt werden kann. Liegen, wie im Fall der Wasserstoffatompositionen in der Verfeinerung *iam*, größere Verschiebungen vor, tragen diese dennoch den gleichen Farbcode.

Somit ist das hervorstechendste Merkmal der Illustration 6.2 die Färbung der Wasserstoffatome, im Modell *iam* sind deren Koordinaten besonders verschieden zum Modell *mul*. Aber auch etliche der Nicht-H-Atome weisen Unterschiede in der Größenordnung 0,01-0,02 Å auf, und zwar um so stärker, je randständiger diese Atome liegen. Die weniger volatilen Koordinaten des Molekülzentrums weisen kaum eine Lageverschiebung auf. Dies wird in der Darstellung 6.3 für die Unterschiede der Modelle *mul* und *iam_{rb}* noch deutlicher. Weil hier in beiden Verfeinerungen die gleichen Wasserstoffatomabstände verwendet wurden, sind nur wenige Bereiche als voneinander abweichend zu erkennen. Auch hier gilt, daß außenliegende Atome, eher als innenliegende, höhere Verschiebungen zeigen.

Die geringen Geometrieveränderungen müssen nicht zwingend auf die unterschiedlichen Streumodelle zurückgeführt werden, da hauptsächlich die beweglicheren und schlechter bestimmten Randbereiche von P1LYSO betroffen sind und die Koordinatenunsicherheit von gleicher Größenordnung ist. Es wurde aber nochmal klar vor Augen geführt, daß die realitätsnäheren Wasserstoffbindungslängen im Multipolmodell deutlich von einer üblichen sphärischen Verfeinerung verschieden sind. Erreicht man dies mit einem sphärischen Streumodell wie in *iam_{rb}*, tut man dies um den Preis verschlechterter Gütefaktoren.

Um eine Aussage über die Richtung und physikalische Bedeutung der Koordinatenunterschiede zu erhalten, bietet es sich an, die Unterschiede in den Bindungslängen näher zu betrachten. In „Nächster Nachbar Näherung“ hat P1LYSO 54 verschiedene Bindungstypen (siehe Tab. B.1 auf Seite 339), während die in der Proteinkristallographie verwendeten Engh&Huber-Abstandsrestraints nur 46 Bindungen unterscheiden. Von den insgesamt 1378 Bindungen wurden 723 verfeinert, enthielten also wenigstens ein nicht fixiertes Atom.

Die Untersuchung beschränkt sich auf die allen Aminosäuren, exklusive Prolin und Glycin, gemeinsamen Peptidbindungen C–O, C–N, C_α–C und C_α–N der Hauptkette. Unter Beachtung der „Nächsten Nachbarn“ sind diese als C[onc]–O[c], C[onc]–N[cch], C[ncch]–C[onc] und C[ncch]–N[cch] zu bezeichnen. Insgesamt gibt es 607 solcher Bindungen, von denen 379 zur Verfeinerung freigegeben waren und zugleich nur Atome der Hauptpopulation miteinander verknüpfen.

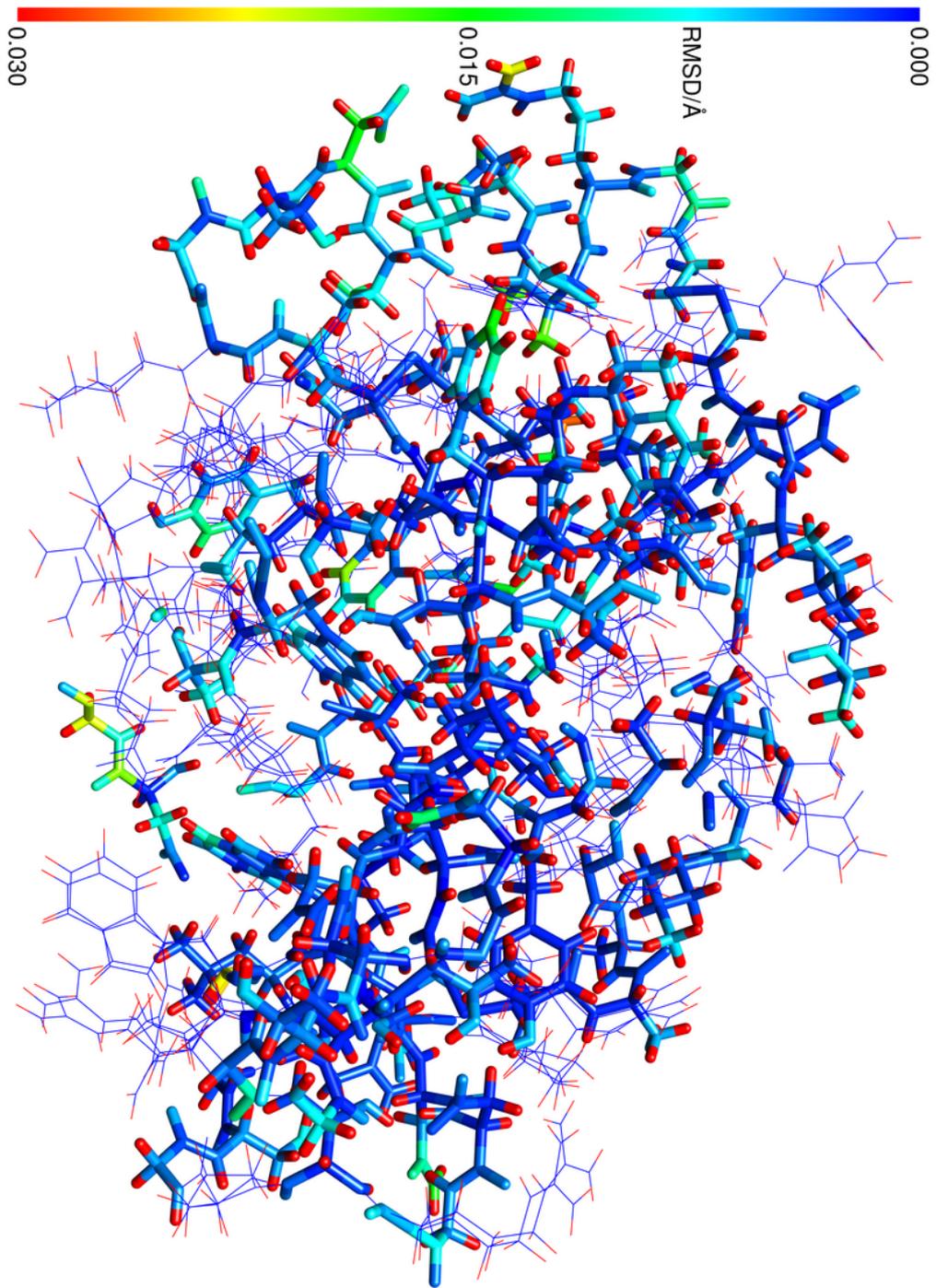


Abbildung 6.2: Quadratisch gemittelter Strukturunterschied zwischen den Modellen *ml1* und *iam* von P1LYSO. Verfeinerte Koordinaten hervorgehoben.

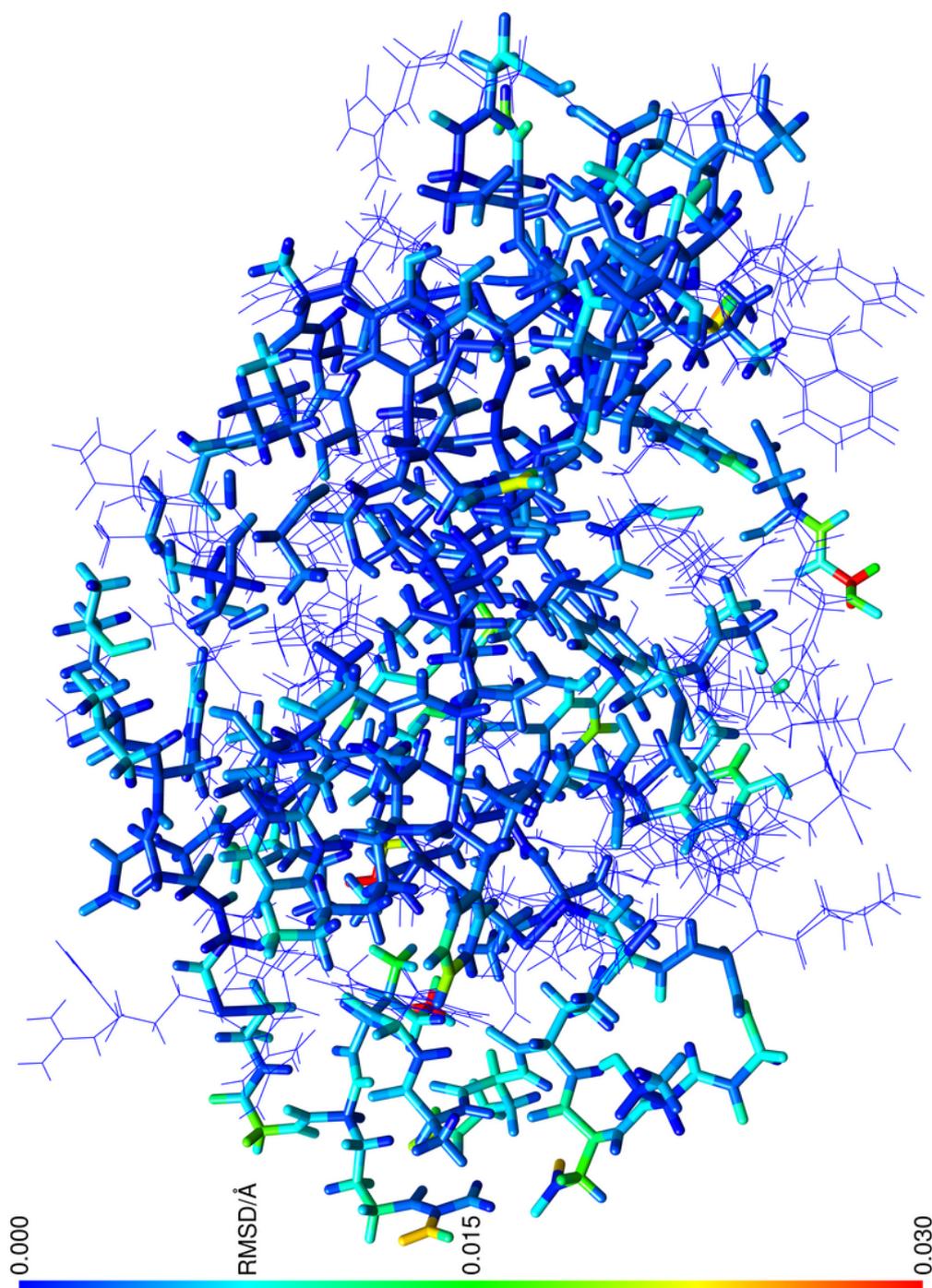


Abbildung 6.3: Quadratisch gemittelter Strukturunterschied zwischen den Modellen mul und iam_{rb} von PLYSO. Verfeinerte Koordinaten hervorgehoben.

Tabelle 6.3: Verfeinerte Peptidbindungslängen der P1LYSO-Hauptpopulation.

	$d_{E\&H}^*$	$d_{min/max}^\dagger$	$\bar{d}(\sigma_d)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–O ($n = 100$)								
<i>mul</i>		1,195/1,262	1,229(09)	1,225	1,230	1,233	0,009	0,006
<i>iam</i>	1,231(20)	1,200/1,277	1,233(10)	1,229	1,233	1,237	0,008	0,006
<i>iam_{rb}</i>		1,200/1,274	1,233(10)	1,228	1,233	1,238	0,009	0,007
C–N ($n = 107$)								
<i>mul</i>		1,252/1,380	1,337(16)	1,333	1,338	1,344	0,011	0,008
<i>iam</i>	1,329(14)	1,251/1,378	1,337(17)	1,333	1,338	1,344	0,012	0,009
<i>iam_{rb}</i>		1,252/1,378	1,337(17)	1,332	1,339	1,345	0,013	0,009
C_α–C ($n = 90$)								
<i>mul</i>		1,508/1,616	1,533(16)	1,525	1,529	1,539	0,014	0,010
<i>iam</i>	1,525(21)	1,501/1,608	1,531(16)	1,523	1,529	1,537	0,014	0,010
<i>iam_{rb}</i>		1,498/1,607	1,530(16)	1,523	1,529	1,537	0,014	0,009
C_α–N ($n = 82$)								
<i>mul</i>		1,370/1,496	1,451(13)	1,448	1,452	1,456	0,008	0,006
<i>iam</i>	1,458(19)	1,385/1,494	1,452(12)	1,450	1,453	1,457	0,007	0,006
<i>iam_{rb}</i>		1,386/1,492	1,453(12)	1,450	1,454	1,458	0,008	0,006

*Engh&Huber-Bindungslängen mit Standardabweichung; [†]Minimum, Maximum und Mittelwert mit Standardabweichung der Bindungslängen; [‡]0,25-Quantil, Median, 0,75-Quantil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in Å.

Für die drei Modelle sind Mittelwerte der betreffenden Bindungen in Tab. 6.3 und weitere Merkmale der Statistik zusammen mit den mittleren Bindungslängen nach Engh u. Huber (1991) angegeben. Innerhalb einer Standardabweichung stimmen die aus der aktuellen Struktur ermittelten Bindungslängenmittelwerte mit diesen überein. Der Mittelwert der Carbonylbindung in den sphärischen Modellen liegt etwas über der Referenzbindungslänge, der der asphärischen Verfeinerung darunter. Die C–N- und C_α–C-Mittelwerte aller Modelle liegen über, die der C_α–N-Bindung unter dem $d_{E\&H}$ -Vergleichswert. Die Unterschiede belaufen sich aber auf nur $5 \cdot 10^{-3}$ bis $8 \cdot 10^{-3}$ Å, also ungefähr eine halbe Standardabweichung.

Zwischen den kürzesten und längsten Bindungen des gleichen Typs können nicht unerhebliche 0,07 bis 0,13 Å liegen, mithin das bis zu Zehnfache der Standardabweichungen. Trotzdem liegen die Mediane der Bindungslängen recht nahe an den arithmetischen Mittelwerten, ein Indiz für normalverteilte Abstände. Während die verschiedenen Bindungstypen neben merklich unterschiedlichen Mittelwerten auch Streuungsmaße verschiedener Höhe aufweisen, trifft dies auf die drei Verfeinerungen einer jeden Bindung nicht zu: Sowohl die Mittelwerte als auch Streuungsmaße, wie die Standardabweichung, der Interquartilabstand oder die mittlere absolute

Tabelle 6.4: Differenzen der verfeinerten Peptidbindungslängen der P1LYSO-Hauptpopulation.

	$\Delta d_{min/max}^\dagger$	$\overline{\Delta d}(\sigma_\Delta)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–O ($n = 100$)							
<i>mul-iam</i>	–14,70/5,90	–3,50(3,26)	–4,30	–2,85	–1,90	2,40	1,70
<i>mul-iam_{rb}</i>	–16,90/4,90	–3,38(3,05)	–4,50	–3,20	–1,98	2,52	1,93
C–N ($n = 107$)							
<i>mul-iam</i>	–11,90/23,20	–0,17(3,94)	–2,00	–0,60	1,00	3,00	2,22
<i>mul-iam_{rb}</i>	–11,80/13,20	–0,82(3,30)	–2,25	–0,90	0,90	3,15	2,22
C_α–C ($n = 90$)							
<i>mul-iam</i>	–5,10/12,90	2,36(3,60)	0,30	1,40	4,00	3,70	2,45
<i>mul-iam_{rb}</i>	–5,60/14,10	2,51(3,72)	0,10	1,40	4,00	3,90	2,45
C_α–N ($n = 82$)							
<i>mul-iam</i>	–14,60/6,10	–0,95(2,64)	–2,08	–0,70	0,40	2,48	1,93
<i>mul-iam_{rb}</i>	–15,70/4,80	–1,49(2,95)	–2,77	–1,15	0,28	3,05	2,30

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der Bindungslängendifferenzen; [‡]0,25-Quartil, Median, 0,75-Quartil; alle Einheiten in 10^{-3} Å.

Abweichung vom Median, sind in den drei Verfeinerungen sehr ähnlich – das Modell hat kaum Einfluß auf die Breite der Verteilung, zwingt die Struktur also nicht enger an einen bestimmten Wert.

Die Modellunterschiede können vorteilhafter durch die in Tab. 6.4 gezeigte Statistik der Bindungslängendifferenzen zwischen dem asphärischen Modell *mul* und den beiden sphärischen Verfeinerungen *iam* und *iam_{rb}* studiert werden. Anhand der Maximalunterschiede $\Delta d_{min/max}$ wird das Ausmaß der Veränderungen deutlich. Die höchste absolute Abweichung liegt bei 0,023 Å, im Mittel betragen die Unterschiede aber nur 0,0002 bis maximal 0,0035 Å, je nach betrachteter Bindung und Modell. Tendentiell verkürzen sich Kohlenstoff–Heteroatom-Bindungen, was sich in negativen Werten für $\overline{\Delta d}$ und den Median im Falle der C–O-, C–N- und C_α–N-Bindungen äußert, während der Abstand in der C_α–C-Bindung zunimmt.

Die mittlere Verkürzung der Bindungen zu Stickstoff ist besonders klein, die Quartilgrenzen fallen sowohl in den negativen wie positiven Bereich. Nur CC- und CO-Bindungen ändern sich merklich, dabei ist die C–O-Verkürzung in Übereinstimmung mit den Beobachtungen in Abschnitt 5.2 auf Seite 117 und zudem die deutlichste der hier betrachteten Veränderungen. Dieser Befund deckt sich mit den Erwartungen, da die freien Elektronenpaare des Sauerstoffs im asphärischen Modell Berücksichtigung finden und somit keine künstliche Bindungsverlängerung wie im sphärischen Modell mehr bewirken.

6.3 Verschiebungsparameter

Die 2901 Atomlagen von P1LYSO haben einen durchschnittlichen äquivalenten isotropen Temperaturfaktor U_{eq} von $97 \cdot 10^{-3} \text{ \AA}^2$. In diesem Mittel sind auch extreme Verschiebungsparameter, wie der Maximalwert von $981 \cdot 10^{-3} \text{ \AA}^2$ eines Wassers der Solvenssphäre, enthalten. Aber auch im Lysozymmolekül selbst, treten für terminale Atome noch Extrema bis zu $607 \cdot 10^{-3} \text{ \AA}^2$ auf, wohingegen der Minimalwert $25 \cdot 10^{-3} \text{ \AA}^2$ beträgt. Abzüglich der Solvensmoleküle und Wasserstoffatome beträgt der Mittelwert noch $77 \cdot 10^{-3} \text{ \AA}^2$ für 1323 Atomlagen, wobei die Hälfte der Werte in einem $38 \cdot 10^{-3} \text{ \AA}^2$ breiten Bereich um einen Median von nur $59 \cdot 10^{-3} \text{ \AA}^2$ streut. In Abb. 6.4 auf der gegenüberliegenden Seite sind die Temperaturfaktoren und ihre Änderung durch die asphärische Verfeinerung für diese 1323 Atome aufgetragen. Offensichtlich ist der periodische Anstieg der Temperaturfaktoren in schlecht definierten Bereichen der Struktur, wie insbesondere Seitenketten und dem C-Terminus. Betrachtet man nur die 1001 Schweratome der Hauptpopulation verringert sich das mittlere U_{eq} auf $69 \cdot 10^{-3} \text{ \AA}^2$. Die davon zur Verfeinerung ausgewählten 639 Nichtwasserstoffatome haben einen durchschnittlichen Verschiebungsparameter von $58 \cdot 10^{-3} \text{ \AA}^2$ oder – als realistischeren Schätzer – einen Median von $52 \cdot 10^{-3} \text{ \AA}^2$. Von den zur Verfeinerung freigegebenen Atomen, haben 50 % ein U_{eq} zwischen $42 \cdot 10^{-3}$ und $67 \cdot 10^{-3} \text{ \AA}^2$. Eine vollständige Statistik der U_{eq} -Werte für verschiedene Untermengen des Moleküls ist in Tab. B.2 auf Seite 340 aufgeführt, Abb. 6.6 auf Seite 161 illustriert die anisotropen Temperaturfaktoren des verfeinerten Strukturteils.

Die Temperaturfaktoren nach der asphärischen Verfeinerung *mul* sind um durchschnittlich $0,3 \cdot 10^{-3} \text{ \AA}^2$ (0,5 %) kleiner als im Modell *iam*. Anhand von *p*-Werten nahe Null in einem *t*- und Wilcoxon-Test, wird diese Verringerung als statistisch signifikant erkannt. Die stärksten Verbesserungen betragen um $-6 \cdot 10^{-3} \text{ \AA}^2$. Allerdings tritt an einigen Atomen auch eine Erhöhung um maximal $11 \cdot 10^{-3} \text{ \AA}^2$ auf, wobei eine solche Verschlechterung nur bei 91 (*mul vs iam*), bzw. 77 (*mul vs iam_{TB}*) Atomen vorliegt und im Mittel mit $0,8 \cdot 10^{-3} \text{ \AA}^2$ moderat ausfällt. Im Endeffekt verbessert sich der U_{eq} -Wert für über 85 % der verfeinerten Atome von P1LYSO. Anhand Abb. 6.5 auf Seite 158 ist ersichtlich, daß bei kleinen Ausgangsverschiebungsparametern fast durchgängig eine Verbesserung erzielt wurde. Nur bei den Atomen, die von vornherein einen hohen Temperaturfaktor von ca. $> 85 \cdot 10^{-3} \text{ \AA}^2$ aufweisen, treten die Verschlechterungen auf. Diese Teile der Struktur sind zu schlecht bestimmt, um verlässlich verfeinert zu werden.

Schlüsselt man die Verschiebungsparameter der 639 verfeinerten Atome und ihre Veränderung nach Elementen auf, verhält es sich ähnlich der vorherigen Beobachtung: Die zehn Schwefelatome in der Struktur haben die niedrigsten Temperaturfaktoren, der U_{eq} -Zentralwert beträgt $48 \cdot 10^{-3} \text{ \AA}^2$, und verbessern sich mit durchschnittlich $-0,6 \cdot 10^{-3} \text{ \AA}^2$ fast doppelt so stark wie Kohlenstoffatome. Diese zeigen einen U_{eq} -Median von $51 \cdot 10^{-3} \text{ \AA}^2$ bei einer Verringerung der Temperaturfaktoren von

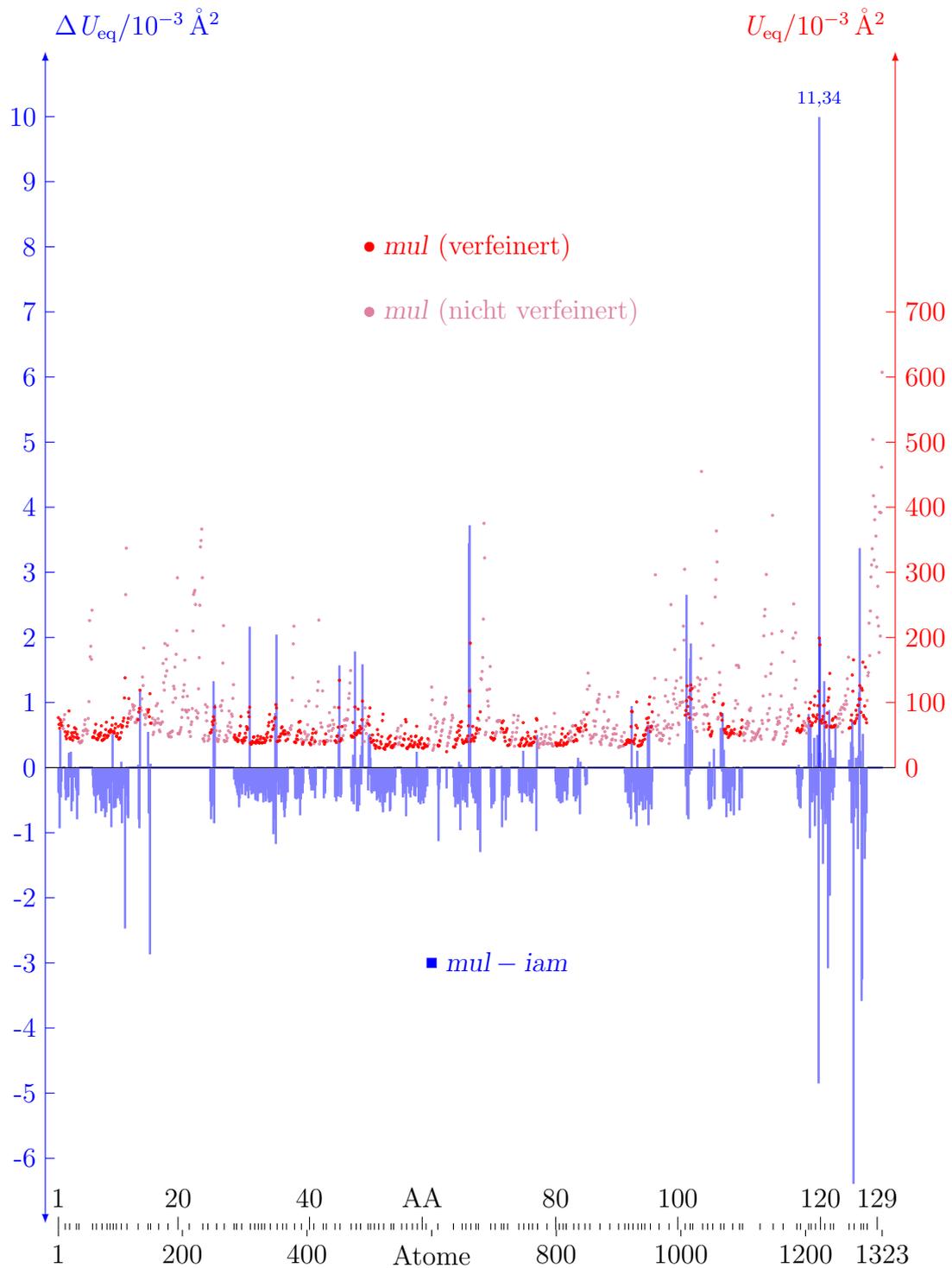


Abbildung 6.4: Veränderung der U_{eq} -Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die Nichtwasserstoffatome von P1LYSO.

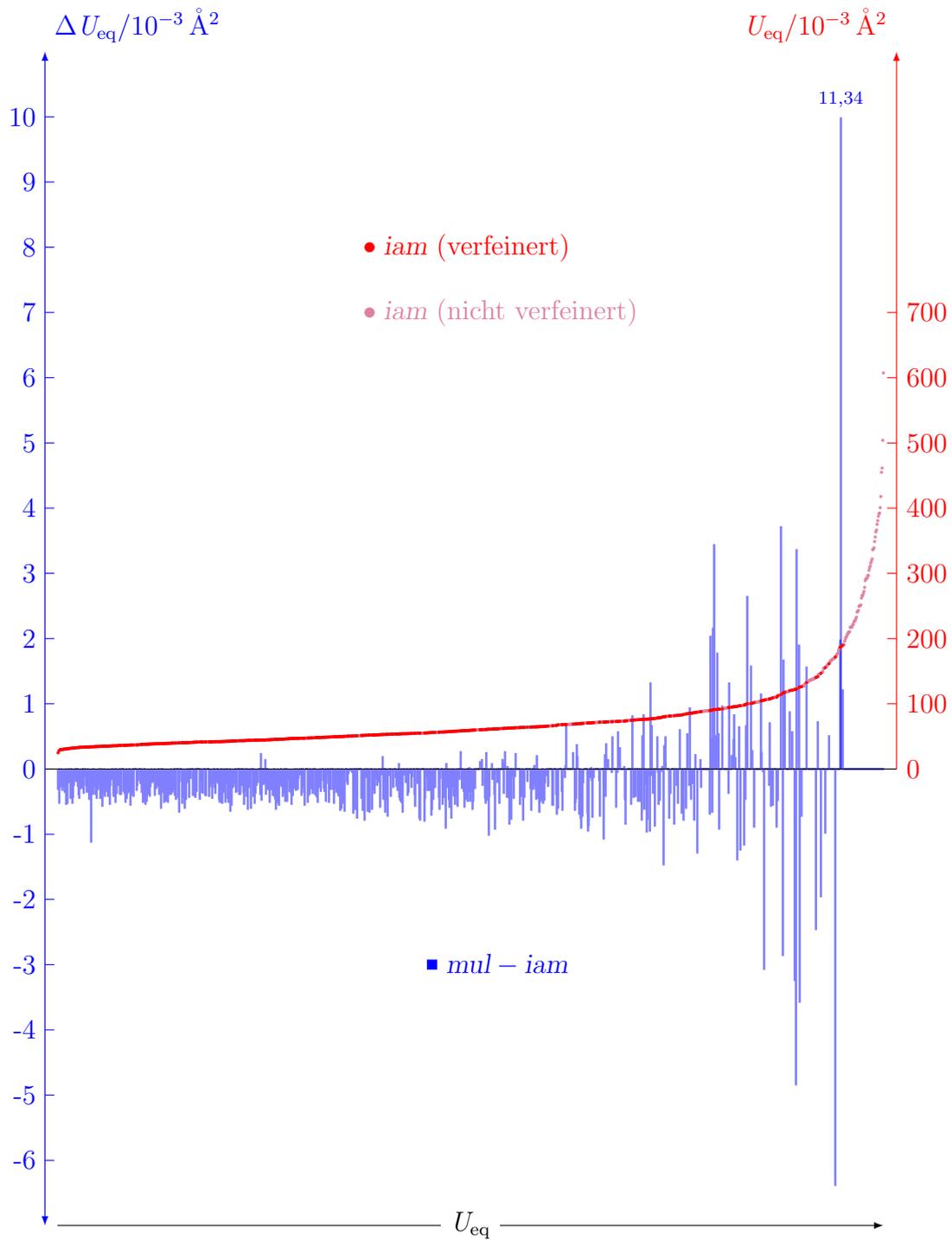


Abbildung 6.5: Veränderung der U_{eq} -Werte im asphärischen Modell mul gegenüber dem sphärischen Modell iam für die Nichtwasserstoffatome von P1LYSO. Auftragung nach $U_{eq}(iam)$.

$-0,4 \cdot 10^{-3} \text{ \AA}^2$. Noch höhere U_{eq} der Sauerstoffatome, im Median $56 \cdot 10^{-3} \text{ \AA}^2$, bedingen eine abermals geringere Verbesserung von $-0,3 \cdot 10^{-3} \text{ \AA}^2$. Die Stickstoffatome wiederum, zeigen zwar mit $49 \cdot 10^{-3} \text{ \AA}^2$ einen recht geringen Median der Temperaturfaktoren, aber gleichzeitig die geringste Abnahme von $-0,2 \cdot 10^{-3} \text{ \AA}^2$. Im Detail ist dies in Tab. B.3 auf Seite 341 aufgeführt, anhand der auch deutlich wird, daß die Verbesserungen im Modell *mul* immer ausgeprägter gegenüber *iam_{TB}* als im Verhältnis zu *iam* sind.

In Tab. 6.5 auf der nächsten Seite ist eine solche Statistik für vier der insgesamt 30 Atomtypen gezeigt, welche die zuvor erwähnten Peptidbindungen der Hauptpopulation bilden. Die Auswahl umfaßt 372 der 639 verfeinerten Atome, ihre mittleren Verschiebungsparameter liegen unter denen der Gesamtheit aller verfeinerten Atome des entsprechenden Elements und die Temperaturfaktorveränderung folgt dem bereits beschriebenen Verhalten. Hierbei sind die Extremwerte der Differenzen geringer ausgeprägt. Bezogen auf arithmetisches Mittel und Standardabweichung beträgt die Streuung der U_{eq} -Werte ca. $\pm 30\%$ des Mittelwertes und unterscheidet sich kaum zwischen den verschiedenen Modellen. Abhängig vom betrachteten Maß der Verteilungsbreite, liegt dieses für einen Atomtyp im asphärischen Modell mal über und mal unter dem der sphärischen Modelle. Obwohl mit dem Modell *mul* keine Verringerung der Streuungsmaße einher geht, was in sehr geringem Umfang noch für alle 639 Atome galt, ist der Anteil der verbesserten U_{eq} mit 87–91 % höher als in den vorherigen Betrachtungen.

Abb. 6.6 auf Seite 161 zeigt die anisotropen Temperaturfaktoren des verfeinerten Strukturteils, die festgehaltenen Atome sind in Strichdarstellung gezeigt. Zur Verdeutlichung ist jedes Atom mit einem dem äquivalenten isotropen Temperaturfaktor U_{eq} entsprechendem Farbcode für den Bereich $0-240 \cdot 10^{-3} \text{ \AA}^2$ gekennzeichnet – blaue Bereiche der Struktur weisen kleine Verschiebungsparameter auf, rote besonders hohe. Hieran wird nochmals deutlich, daß Atome mit besonders hohem U_{eq} von der Verfeinerung ausgeschlossen wurden. Eine Darstellung der Verschiebungsparameter des gesamten Moleküls findet sich in Abb. B.1 auf Seite 342.

Mit der abgebildeten magentafarbenen Kugel wird die Lage des massengewichteten Zentrums der Struktur angegeben, es berechnet sich (mit $x_i = \text{kart. Koord.}$, $m_i = \text{Atomgewicht der } i \text{ Atome}$) zu

$$\begin{aligned} x_{MZ} &= \frac{\sum_i^n x_i m_i}{\sum_i^n m_i} = -1.8681 \\ y_{MZ} &= \frac{\sum_i^n y_i m_i}{\sum_i^n m_i} = 14.8638 \\ z_{MZ} &= \frac{\sum_i^n z_i m_i}{\sum_i^n m_i} = 24.5546 \end{aligned} \quad (6.1)$$

und der Abstand d_{MZ} eines Atoms mit den Koordinaten (x_i, y_i, z_i) ist der Betrag

Tabelle 6.5: Statistik der U_{eq} der verfeinerten Atome der Peptidbindungen von P1LYSO.

	$x_{\min/\max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C_α (n = 82)							
<i>mul</i>	29,13/87,73	48,01(14,23)	38,03	42,88	58,01	19,98	12,75
<i>iam</i>	29,42/88,02	48,39(14,24)	38,26	43,26	58,35	20,09	12,80
<i>iam_{rb}</i>	29,41/87,73	48,42(14,28)	38,30	43,33	58,35	20,05	12,79
<i>mul-iam</i>	-1,13/0,10	-0,38(0,19)	-0,47	-0,37	-0,26	0,21	0,16
<i>mul-iam_{rb}</i>	-1,18/-0,00	-0,41(0,21)	-0,46	-0,37	-0,29	0,17	0,13
C (n = 100)							
<i>mul</i>	28,15/118,80	50,35(18,46)	36,60	44,52	59,66	23,07	14,42
<i>iam</i>	28,68/122,00	50,76(18,50)	36,95	44,51	59,90	22,95	13,68
<i>iam_{rb}</i>	28,66/119,00	50,75(18,41)	37,06	44,48	60,13	23,07	13,65
<i>mul-iam</i>	-3,25/2,66	-0,41(0,49)	-0,53	-0,41	-0,29	0,24	0,17
<i>mul-iam_{rb}</i>	-2,05/2,12	-0,40(0,40)	-0,54	-0,43	-0,27	0,28	0,22
N (n = 90)							
<i>mul</i>	29,60/118,60	49,32(16,19)	38,09	44,70	56,64	18,55	13,98
<i>iam</i>	29,83/116,90	49,45(16,01)	38,35	44,78	56,82	18,48	14,04
<i>iam_{rb}</i>	29,86/118,70	49,52(16,10)	38,44	44,93	56,77	18,34	13,99
<i>mul-iam</i>	-0,77/1,69	-0,13(0,33)	-0,28	-0,17	-0,08	0,20	0,16
<i>mul-iam_{rb}</i>	-0,68/0,72	-0,20(0,25)	-0,34	-0,23	-0,10	0,24	0,18
O (n = 100)							
<i>mul</i>	32,57/161,90	62,01(25,72)	44,00	53,40	72,83	28,83	17,11
<i>iam</i>	32,86/161,40	62,23(25,65)	44,24	53,71	72,49	28,26	17,13
<i>iam_{rb}</i>	32,95/161,20	62,28(25,69)	44,24	53,69	72,45	28,22	17,09
<i>mul-iam</i>	-2,86/1,59	-0,22(0,50)	-0,40	-0,25	-0,08	0,32	0,24
<i>mul-iam_{rb}</i>	-2,81/1,91	-0,27(0,55)	-0,46	-0,33	-0,15	0,31	0,24

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der U_{eq} und ΔU_{eq} ;

[‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-3}\AA^2 .

der Differenz zum Schwerpunktsvektor (x_{MZ}, y_{MZ}, z_{MZ}) :

$$d_{MZ} = \sqrt{(x_i - x_{MZ})^2 + (y_i - y_{MZ})^2 + (z_i - z_{MZ})^2} \quad (6.2)$$

Tendentiell scheinen weiter außen liegende Molekülteile einen höheren Temperaturfaktor aufzuweisen, dies gilt insbesondere für in die Solvenssphäre ragende Seitenketten (siehe z.B. die auffällige ARG 128-Seitenkette links oben im Bild). Ein solches Verhalten ist nicht unerwartet, da zum einen die inneren Reste mehr sie stabilisierende Nachbarn haben und zum anderen man sich eine Librationsbewegung des Gesamtmoleküls vorstellen kann.

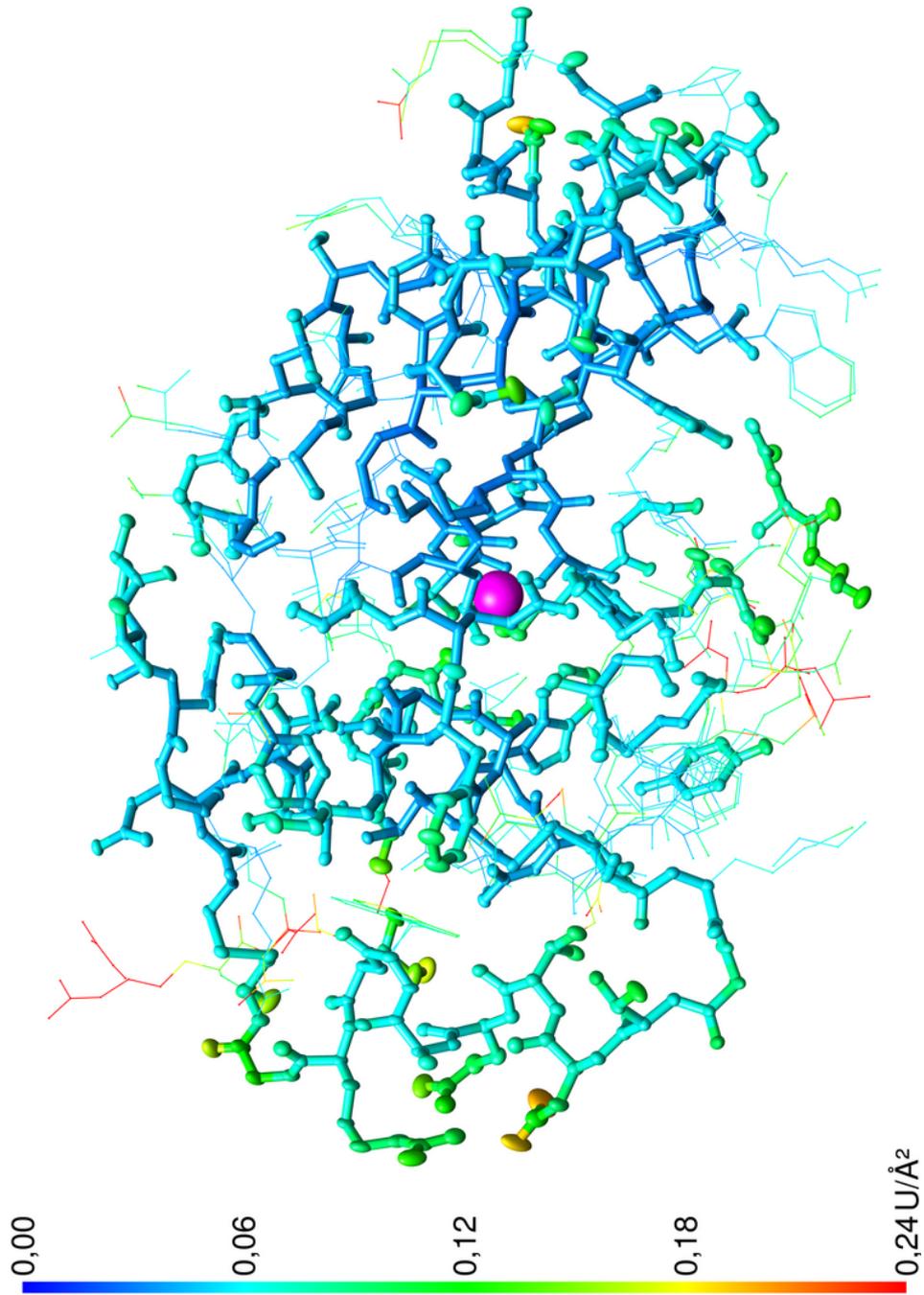


Abbildung 6.6: Anisotrope Verschiebungsparameter des verfeinerten Teils von P1LYSO im Modell *mul*. Farbcode: U_{eq} . Darstellung des massengewichteten Zentrums als magentafarbene Kugel.

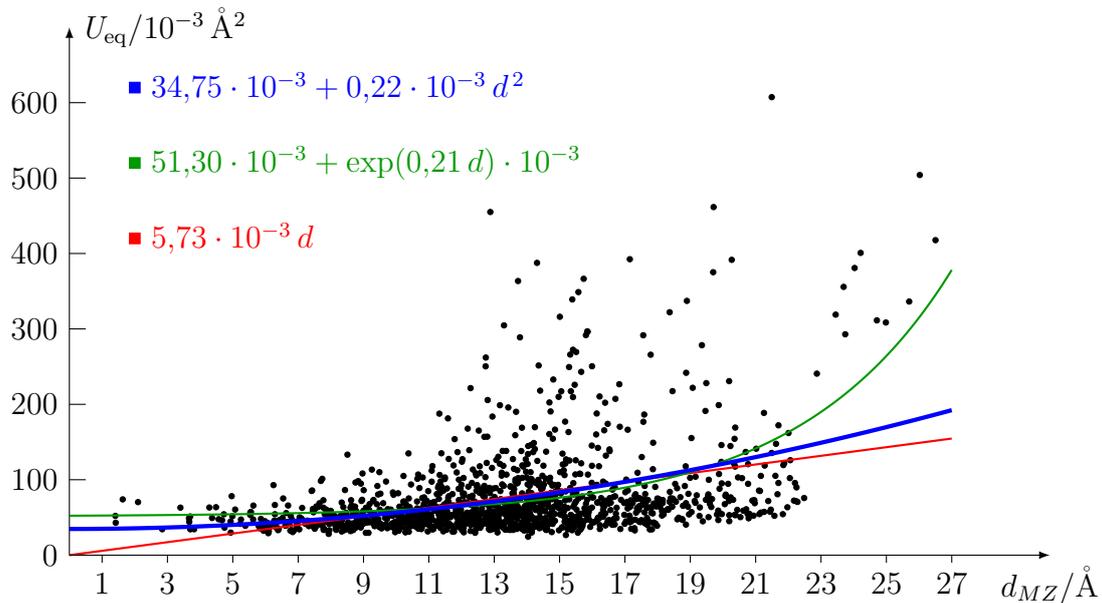


Abbildung 6.7: U_{eq} (*mul*) der Nichtwasserstoffatome von P1LYSO in Abhängigkeit vom Abstand zum Massenzentrum.

Zur Überprüfung wurde das Diagramm 6.7 angefertigt, das den isotropen Temperaturfaktor jedes Nichtwasserstoffatoms in Abhängigkeit von der Entfernung des jeweiligen Atoms vom Massenzentrum zeigt. An die Punktwolke wurden lineare, quadratische und exponentielle Ausgleichsfunktionen angepaßt. Die Unsicherheit der angegebenen Koeffizienten der Gradenfunktion (rot) und der Exponentialfunktion (grün) liegen zwischen $\pm 2\%$ und $\pm 4\%$. Höhere Fehlermargen weisen die Konstanten der quadratischen Funktion (blau) mit $\pm 9\%$ und $\pm 6\%$ auf. Während die exponentielle Regressionsfunktion bei zu hohen Werten ansetzt, um bei steigendem Abstand den Anstieg zu überschätzen, hat die Geradengleichung den Nachteil, die U_{eq} am oder nahe des Zentrums gegen Null zu extrapolieren. Versucht man eine Gerade $a + b * x$ mit freiem Ordinatenabschnitt an die Daten anzupassen, verbleibt der Parameter a mit 2,1 nahe dem Ursprung und der Parameter b verfeinert sich auf 5,59. Dabei steigen die Fehler allerdings stark an, so daß der obigen Ursprungsgrade der Vorzug gegeben wurde. Die 1323 Datenpunkte werden am besten durch die quadratische Ausgleichsfunktion beschrieben, welche von $35 \cdot 10^{-3} \text{ \AA}^2$ ausgehend, einen moderaten Anstieg von $0,44 \cdot 10^{-3} d \text{ \AA}$ mit der Entfernung vom massengewichteten Zentrum angibt. Trotz dieser moderaten Steigung kommen, den Randbereichen entsprechend, besonders hohe Werte bei Entfernungen $> 12 \text{ \AA}$ vor. Allerdings machen Atome mit einem $100 \cdot 10^{-3} \text{ \AA}^2$ überschreitendem Temperaturfaktor nur etwas mehr als ein Sechstel der betrachteten Daten aus.

6.4 DMSDA-Werte

Einen ersten Überblick über die zwischen anisotrop verfeinerten Atomen von P1LY-SO auftretenden „Differenzen der mittleren quadratischen Verschiebungsamplitude“ (DMSDA, siehe auch Absch. 1.1 auf Seite 18) geben die Abbildungen 6.8 und 6.9 auf den Seiten 164 und 165. Dabei sind an der rechten, roten y2-Achse die DMSDA-Werte der 1378 Bindungen aufgetragen. Angeordnet sind die Bindungen auf der Abszisse einmal in der Reihenfolge der beitragenden Atome in der Peptidsequenz, man erhält also einen ähnlichen Eindruck vom Gang der Werte wie in Abb. 6.4, und zum anderen nach steigenden DMSDA-Werten des sphärischen Modells. Bis auf einen einzelnen Ausreißer (kleingedruckte Zahlen), liegen die Werte zwischen 0 und ungefähr $300 \cdot 10^{-4} \text{ \AA}^2$. Die Werte streuen in allen Modellen sehr stark, der Mittelwert aber liegt bei moderaten $38 \cdot 10^{-4} \text{ \AA}^2$, der Median bei $22 \cdot 10^{-4} \text{ \AA}^2$ und damit dem Doppelten des Richtwertes nach dem *rigid bond*-Postulat.

Für die 723 verfeinerten Bindungen liegen die DMSDA-Werte wesentlich niedriger, nämlich bei $28,94 \cdot 10^{-4} \text{ \AA}^2$ (Median $15 \cdot 10^{-4} \text{ \AA}^2$) im Modell *iam* und $28,25 \cdot 10^{-4} \text{ \AA}^2$ (Median $13 \cdot 10^{-4} \text{ \AA}^2$) im Modell *mul*. Mit dem Übergang vom sphärischen zum asphärischen Modell ist also eine Verringerung der DMSDA-Werte verbunden. Obwohl die Abnahme im arithmetischen Mittel und im Median mit $-0,7 \cdot 10^{-4}$ bzw. $-2 \cdot 10^{-4} \text{ \AA}^2$ gering ist und dabei die Streubreite im asphärischen Modell sogar geringfügig zunimmt, ist diese Verringerung nach dem Wilcoxon-Test signifikant (p -Wert 0,0005431 für den Test $H_0 : \mu_{mul} = \mu_{iam}$ vs $H_1 : \mu_{mul} < \mu_{iam}$).

Die modellabhängigen Veränderungen sind in die Abb. 6.8 und 6.9 als Differenz $|\text{DMSDA}(mul)| - |\text{DMSDA}(iam)|$, in Form blauer Balken, eingetragen, entsprechend einer ebenfalls blau gekennzeichneten Ordinate (links). In positive wie negative Richtung treten starke Veränderungen auf. Da diese aber dennoch geringer als die auftretenden DMSDA-Werte sind, weist die y1-Achse die dreifache Skalierung der y2-Achse auf.

Versucht man die ΔDMSDA -Werte mit den entsprechenden DMSDA-Beträgen des sphärischen Ausgangsmodells wie in Abb. 6.9 zu korrelieren, so zeigt sich, daß sowohl für die niedrigsten ($< 7 \cdot 10^{-4} \text{ \AA}^2$) als auch die höchsten ($> 140 \cdot 10^{-4} \text{ \AA}^2$) DMSDA-Werte eher eine Verschlechterung im asphärischen Modell eintritt, erkenntlich an den überwiegend positiven Differenzen.

Eine detailliertere, nach Bindungstypen aufgeschlüsselte Statistik der DMSDA-Werte der bereits in Absch. 6.2 auf Seite 151 angeführten 379 Peptidbindungen, ist in Tab. 6.6 auf Seite 166 zu finden. Die hier besonders interessierenden Effekte der Multipolbehandlung werden durch die Statistik über die Differenzen $|\text{DMSDA}|(mul) - |\text{DMSDA}|(iam)$ und $|\text{DMSDA}|(mul) - |\text{DMSDA}|(iam_{rb})$ nochmal hervorgehoben.

In der sphärischen Verfeinerung erreicht zumindest das Mittel der C_α -N-Bindung mit $11 \cdot 10^{-4} \text{ \AA}^2$ Werte nahe am Hirshfeldlimit. Zugleich ist es auch der Bindungstyp mit der geringsten Streuung und den geringsten Maximalwerten. Die Verbesserung

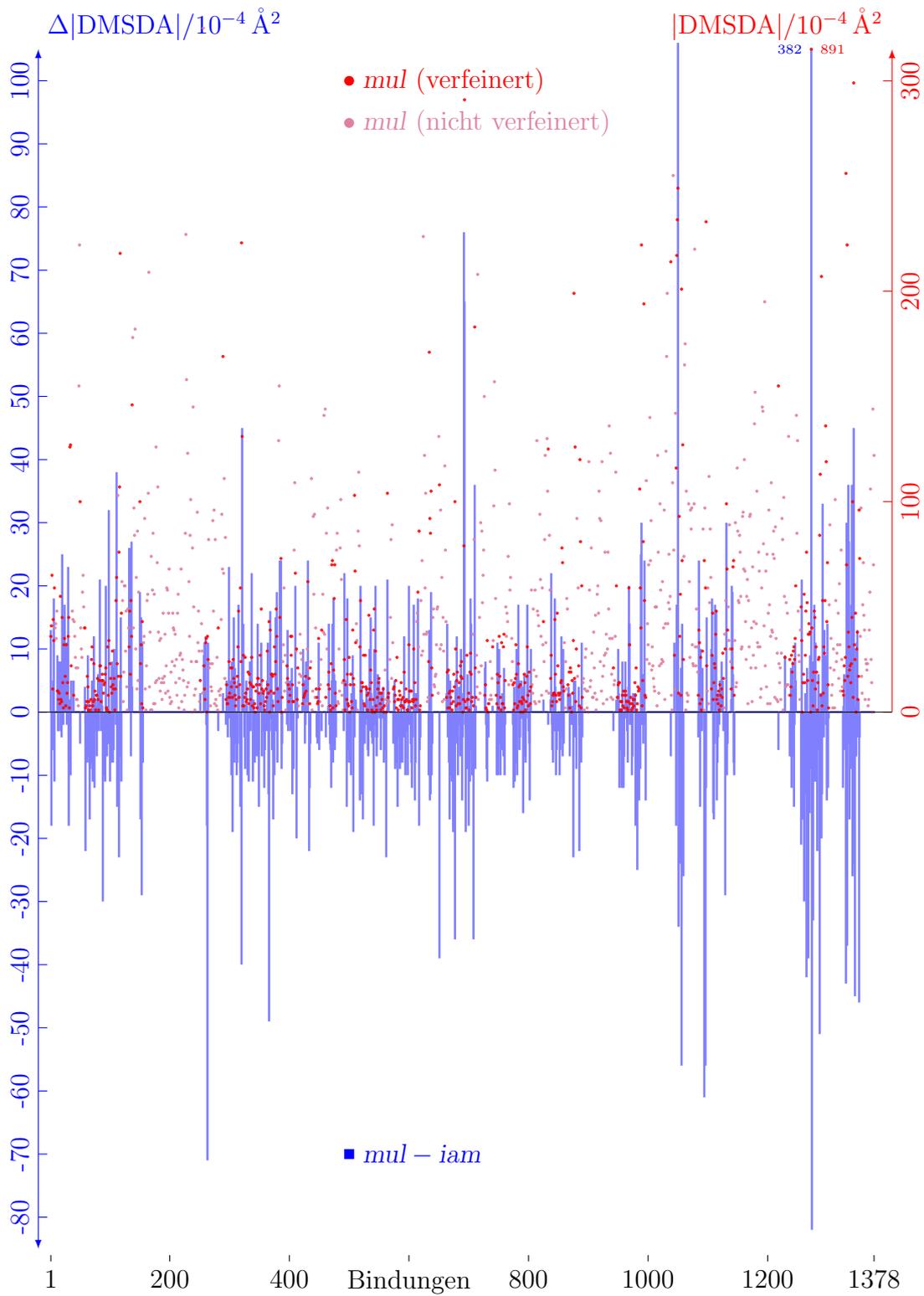


Abbildung 6.8: Veränderung der DMSDA-Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die Nichtwasserstoffatome von P1LYSO.

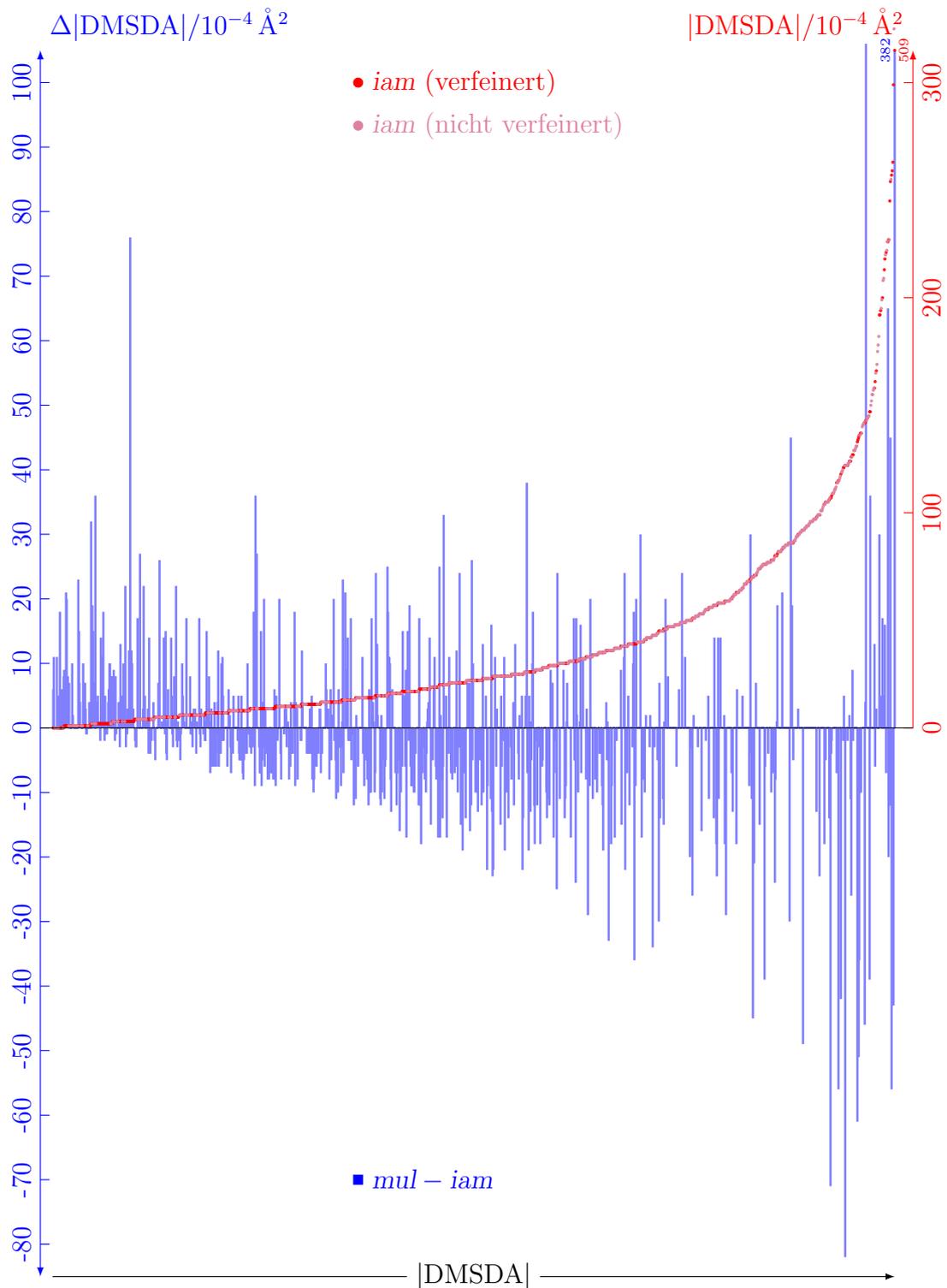


Abbildung 6.9: Veränderung der DMSDA-Werte im asphärischen Modell mul gegenüber dem sphärischen Modell iam für die Nichtwasserstoffatome von PLYSO. Auftragung nach steigenden $|DMSDA|$ -Werten des Modells iam .

Tabelle 6.6: DMSDA-Beträge der verfeinerten Peptidbindungen der P1LYSO-Hauptpopulation und deren Differenzen.

	$x_{min/max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–O ($n = 100$)							
<i>mul</i>	0/146	17,31(19,69)	7,00	12,50	21,25	14,25	11,86
<i>iam</i>	0/142	22,21(21,66)	10,00	17,00	26,00	16,00	11,86
<i>iam_{rb}</i>	1/139	21,52(20,17)	9,75	17,00	25,00	15,25	11,86
<i>mul-iam</i>	-46/15	-4,90(8,68)	-8,00	-5,00	-1,00	7,00	4,45
<i>mul-iam_{rb}</i>	-43/31	-4,21(8,79)	-7,00	-5,00	-1,75	5,25	4,45
C–N ($n = 107$)							
<i>mul</i>	0/249	24,03(42,68)	5,00	9,00	24,00	19,00	10,38
<i>iam</i>	1/245	25,72(38,49)	8,00	15,00	27,00	19,00	13,34
<i>iam_{rb}</i>	0/215	25,23(36,67)	8,00	14,00	28,00	20,00	13,34
<i>mul-iam</i>	-18/106	-1,69(14,06)	-10,00	-4,00	3,50	13,50	8,90
<i>mul-iam_{rb}</i>	-19/104	-1,21(13,86)	-9,00	-4,00	4,00	13,00	8,90
C_α–C ($n = 90$)							
<i>mul</i>	0/217	26,21(29,41)	9,00	19,00	33,50	24,50	16,31
<i>iam</i>	1/200	23,96(32,35)	5,25	13,00	25,75	20,50	13,34
<i>iam_{rb}</i>	0/188	24,09(31,32)	5,25	13,00	26,50	21,25	13,34
<i>mul-iam</i>	-61/36	2,26(19,28)	-12,00	7,00	17,00	29,00	19,27
<i>mul-iam_{rb}</i>	-44/35	2,12(19,52)	-11,00	8,00	17,75	28,75	19,27
C_α–N ($n = 82$)							
<i>mul</i>	0/116	13,87(15,50)	5,00	11,00	17,75	12,75	9,64
<i>iam</i>	1/134	14,84(16,92)	5,00	11,00	19,75	14,75	10,38
<i>iam_{rb}</i>	0/125	15,17(16,06)	5,25	12,00	20,75	15,50	10,38
<i>mul-iam</i>	-18/26	-0,98(6,51)	-4,00	-0,50	2,00	6,00	5,19
<i>mul-iam_{rb}</i>	-18/22	-1,31(6,62)	-5,00	-1,00	1,00	6,00	4,45

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der DMSDA-Beträge;

[‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-4}Å^2 .

in *mul* ist allerdings sehr klein und beträgt, gemessen an den Differenzen zwischen sphärischem und asphärischem Modell, ungefähr $-1 \cdot 10^{-4}\text{Å}^2$. Andererseits nimmt die Streubreite der DMSDA-Werte am deutlichsten ab.

Die zweite Stickstoffbindung C–N wiederum, verbessert sich wesentlich. Ausgehend von $15 \cdot 10^{-4}\text{Å}^2$, sinkt der Median um 40%. Die Differenzen der DMSDA-Werte zwischen *mul* und den sphärischen Modellen haben einen Zentralwert von $-4 \cdot 10^{-4}\text{Å}^2$. Im asphärischen Modell wird nun das genannte Limit im Mittel unterschritten. Dieser Bindungstyp zeigt aber auch die drastischsten Verschlechterungen

einiger Bindungen, die Varianz der DMSDA-Beträge bleibt in allen Modellen ungefähr gleich hoch.

Eine Verringerung der durchschnittlichen DMSDA-Werte, von 17 auf $12,5 \cdot 10^{-4} \text{Å}^2$, zeigt auch die dritte Heteroatombindung C–O, was einer 26 %-igen Abnahme entspricht. Besonders deutlich zeigen die Differenzen zwischen den Modellen die Verbesserung, sie haben einen Zentralwert von $-5 \cdot 10^{-4} \text{Å}^2$ und die Quartilsgrenzen liegen vollständig im negativen Bereich. Hier wird mit der Behandlung *mul*, trotz höherer Maximalwerte, eine geringere Varianz erreicht.

Ein gegenläufiges Verhalten bringt die C_α –C-Bindung hervor, der Zentralwert der DMSDA-Beträge steigt um ca. 46 %. Auch das arithmetische Mittel erhöht sich, wenn auch weniger stark. Die Differenzen zwischen den Modellen liegen in einem besonders breiten Bereich und sind im Mittel positiv. Gegenüber den sphärischen Modellen, erhöht sich auch die Streuung der DMSDA-Werte, ebenso wie die Maximalbeträge.

Bei allen Bindungstypen erreichen die Minima und Maxima der Differenzen Werte, die um ein Vielfaches größer als die Mittelwerte sind, sowie außerhalb der dreifachen Standardabweichung und auch außerhalb des 1,5-fachen Interquartilabstands liegen. Solche Ausreißer sind nicht mehr physikalisch sinnvoll zu erklären, sie unterstreichen vielmehr die unzureichende Stabilität des Strukturmodells.

Alles in allem besteht der allgemeine Trend des Hirshfeld-Tests für die Peptidbindungen in einer Verbesserung der heteroatomaren Bindungen und Verschlechterung der homoatomaren Kohlenstoff-Kohlenstoff-Bindungen.

Die nachfolgenden Tafeln 6.10 bis 6.13 zeigen für alle 1378 Bindungen, aufgeschlüsselt nach CO- ($n = 230$), CC- ($n = 660$), CN- ($n = 472$) und CS/SS-Bindungen ($n = 12/4$), die entsprechenden DMSDA-Werte des Modells *mul* zusammen mit der Differenz zum sphärischen Modell, aufgetragen über der jeweiligen Bindungslänge. Dabei sind die Beträge nach verfeinert/nicht verfeinert und der Beteiligung/Abwesenheit von Atomen in Alternativpositionen unterteilt. Da Differenzen nur für die 723 verfeinerten Bindungen gebildet werden können, sind diese nur nach Abwesenheit/Vorkommen fehlgeordneter Atome sortiert, wobei letzteres nur bei 38 Bindungen der Fall ist, in denen ein verfeinertes Atom an ein fehlgeordnetes gebunden ist. Wieder sind die Differenzen auf der linken Skala (blau) und die Beträge auf der rechten Achse (rot) aufgetragen. Zur Orientierung bezüglich der vorliegenden Bindungslängen sind im Abbildungsfuß die Engh&Huber-Referenzwerte jeweils als Striche angedeutet und teilweise mit zugehörigem Zahlenwert versehen.

In Abb. 6.10 auf Seite 169 läßt sich zum einen die generelle Verkürzung von CO-Bindungslängen ablesen, da sie sich eher links der Engh&Huber-Werte gruppieren. Die um $1,225 \text{Å}$ streuenden Bindungen gehören zu Carbonyl- oder Carboxylfunktionen, die Werte bei $1,42 \text{Å}$ zu Hydroxylgruppen in Serin und Threonin. Verfeinert wurden 135 Bindungen (DMSDA-Werte als rote Punkte gekennzeichnet), keine davon enthielt fehlgeordnete Atome (violett). Die nichtverfeinerten Bindungen exklusive Fehlordnung (grün) umfassen 50 Werte ($\tilde{x} \approx 16 \cdot 10^{-4} \text{Å}^2$), die nichtver-

feinerten inklusive Fehlordnung (orange) 45 Beträge ($\tilde{x} \approx 48 \cdot 10^{-4} \text{Å}^2$). Besonders abweichende Bindungslängen kommen sowohl bei Abwesenheit fehlgeordneter Atome als auch bei Beteiligung selbiger vor und waren auch teilweise in die Verfeinerung einbezogen, mit dem Resultat drastischer Verschlechterungen des DMSDA-Wertes. Die meisten der verfeinerten Bindungen liegen allerdings in einem recht schmalen Längenfenster bei kleinen DMSDA-Werten und erfahren überwiegend eine moderate Abnahme des DMSDA-Betrages (von $\tilde{x} \approx 17 \cdot 10^{-4} \text{Å}^2$ um $\approx -5 \cdot 10^{-4} \text{Å}^2$ auf $\approx 12 \cdot 10^{-4} \text{Å}^2$), erkenntlich an den negativen Differenzen.

Die analoge Auftragung für CC-Bindungen 6.11 auf Seite 170 zeigt zwei nach Bindungslängen unterscheidbare Gruppen. Die um $1,39 \text{Å}$ streuenden Werte gehören zu den Ringsystemen in Histidin, Tryptophan, Tyrosin und Phenylalanin, alle anderen CC-Bindungen sind in der zweiten Gruppe. Die nicht verfeinerten Bindungen (grün und orange) haben mittlere DMSDA-Werte von $\tilde{x} \approx 32$ und $52 \cdot 10^{-4} \text{Å}^2$ und umfassen 171 und 157 der insgesamt 660 Werte. Die 311 verfeinerten Bindungen (rote Punkte/blau Balken) zwischen wohlgeordneten Atomen verschlechtern ihren DMSDA-Wert von einem Zentralwert $13 \cdot 10^{-4} \text{Å}^2$ auf $14 \cdot 10^{-4} \text{Å}^2$. Dabei gibt es sehr wohl Bindungen, die eine negative Differenz aufweisen, aber ebenso sich verschlechternde Bindungen. Abb. 6.11 macht dies durch die gleichförmig nach unten und oben weisenden Balken unmittelbar deutlich. Die 21 verfeinerten Bindungen (violett), deren einer Bindungspartner fehlgeordnet ist, verschlechtern sich in gleichem Maße wie die zuvor genannten, von $80 \cdot 10^{-4} \text{Å}^2$ auf $81 \cdot 10^{-4} \text{Å}^2$. Tendentiell weisen Bindungen inklusive fehlgeordneter Atome weiter streuende Abstände und höhere DMSDA-Werte auf, bezüglich der Änderungen durch das sphärische Modell gibt es aber keinen Unterschied zu den geordneten Bindungen.

Für die CN-Bindungen in Abb. 6.12 auf Seite 171 gibt es wieder zwei Gruppierungen von Bindungslängen. Bei ca. $1,34 \text{Å}$, den C-N-Peptidbindungen, mesomeren Bindungen in Heterozyklen und Arginin, sowie Amidgruppen zuzuordnen und um $1,45 \text{Å}$, welche primär C $_{\alpha}$ -N-Bindung enthält, sowie die Einfachbindungen in Arginin, Prolin und Lysin. Die Längenverteilung der vier Bindungskategorien gestaltet sich ähnlich den vorherigen Fällen uneindeutig. Der DMSDA-Median verringert sich von 14 auf $11 \cdot 10^{-4} \text{Å}^2$ bei Betrachtung der 223 verfeinerten, nicht fehlgeordneten Verknüpfungen und verschlechtert sich von 58 auf $60 \cdot 10^{-4} \text{Å}^2$ im Fall der 17 Bindungen mit Fehlordnungsanteil. Die nicht verfeinerten 119 Bindungen ohne bzw. die 113 mit Partnern in Alternativpositionen, bleiben bei mittleren DMSDA-Werten von $29 \cdot 10^{-4} \text{Å}^2$ und $39 \cdot 10^{-4} \text{Å}^2$.

Nur 12 Kohlenstoff-Schwefel-Bindungen und 4 Disulfidbrücken sind in der letzten Tafel 6.13 auf Seite 172 eingetragen, die alle verfeinert und fehlordnungsfrei sind. Die Bindungen haben vergleichsweise niedrige DMSDA-Werte, die bei den CS-Verknüpfungen einen Anstieg vom sphärischen ($\tilde{x} = 9 \cdot 10^{-4} \text{Å}^2$) zum sphärischen ($\tilde{x} = 13 \cdot 10^{-4} \text{Å}^2$) Streumodell erfahren. Jede der vier Disulfidbrücken verbessert sich hingegen beim Übergang von *iam* zu *mul*, kenntlich an einem Median von $4,5 \cdot 10^{-4} \text{Å}^2$ nach zuvor $6 \cdot 10^{-4} \text{Å}^2$.

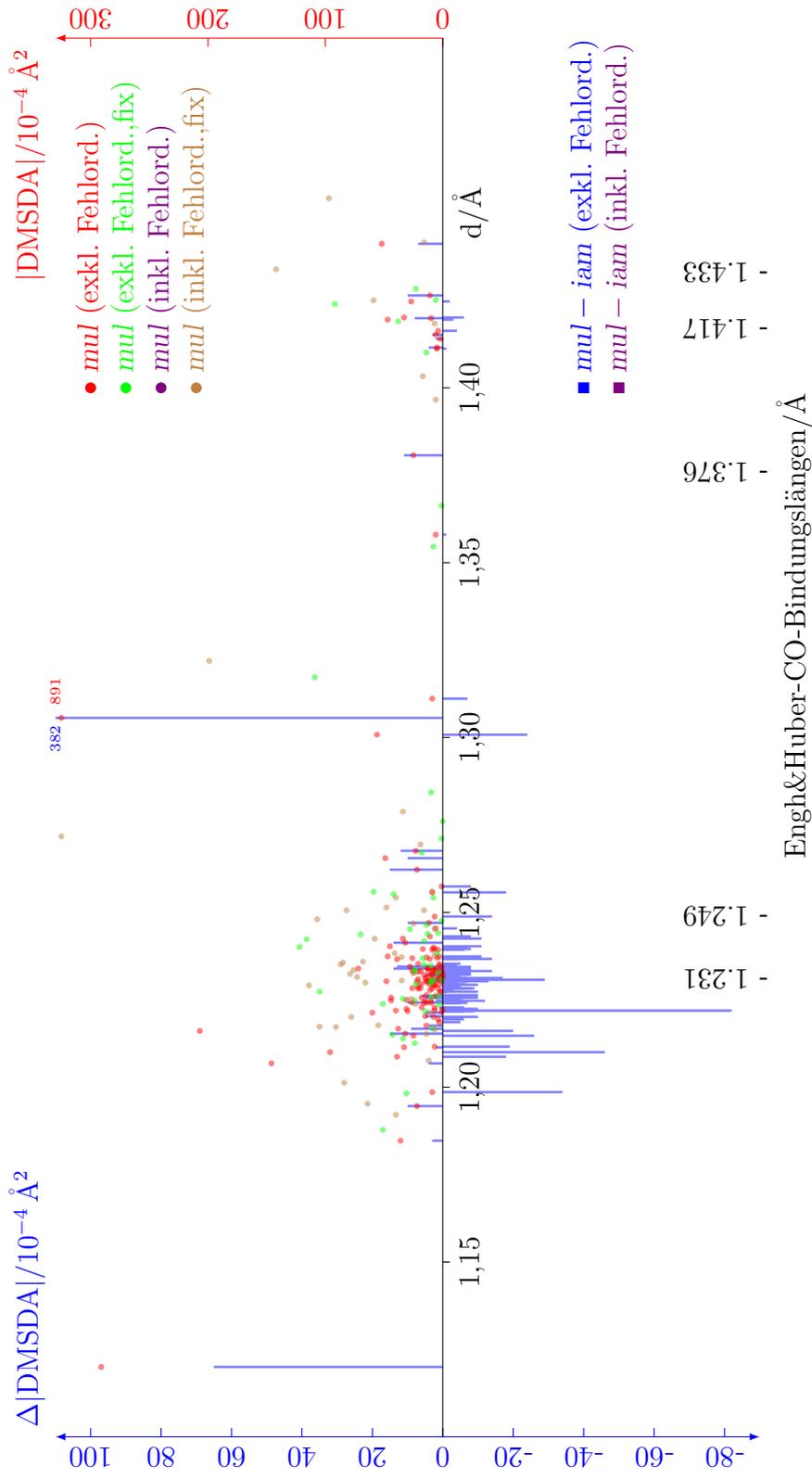


Abbildung 6.10: Veränderung der DMSDA-Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die CO-Bindungen von P1LYSO.

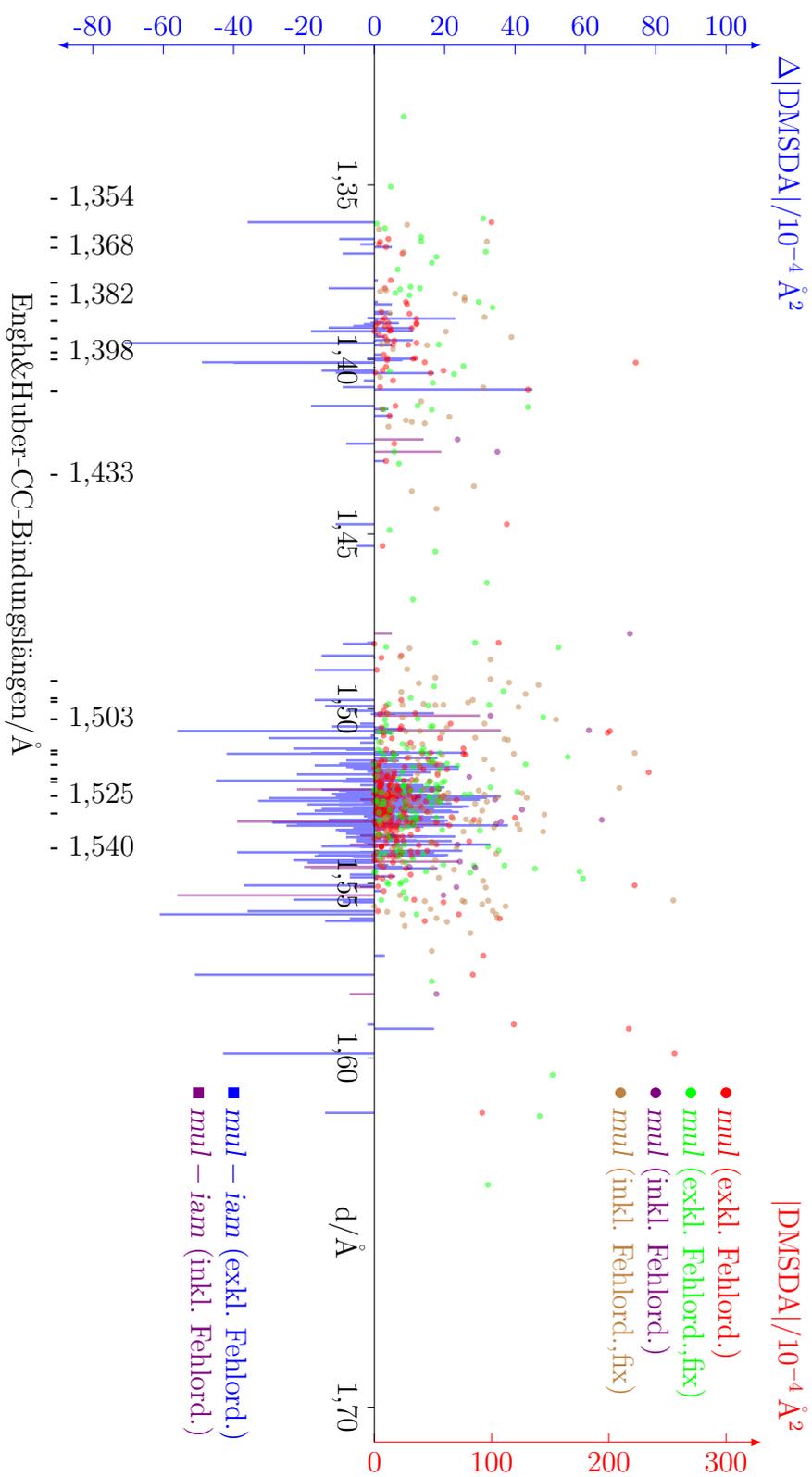


Abbildung 6.11: Veränderung der DMMSDA-Werte im asphärischen Modell *mml* gegenüber dem sphärischen Modell *iam* für die CC-Bindungen von P1LYSO.

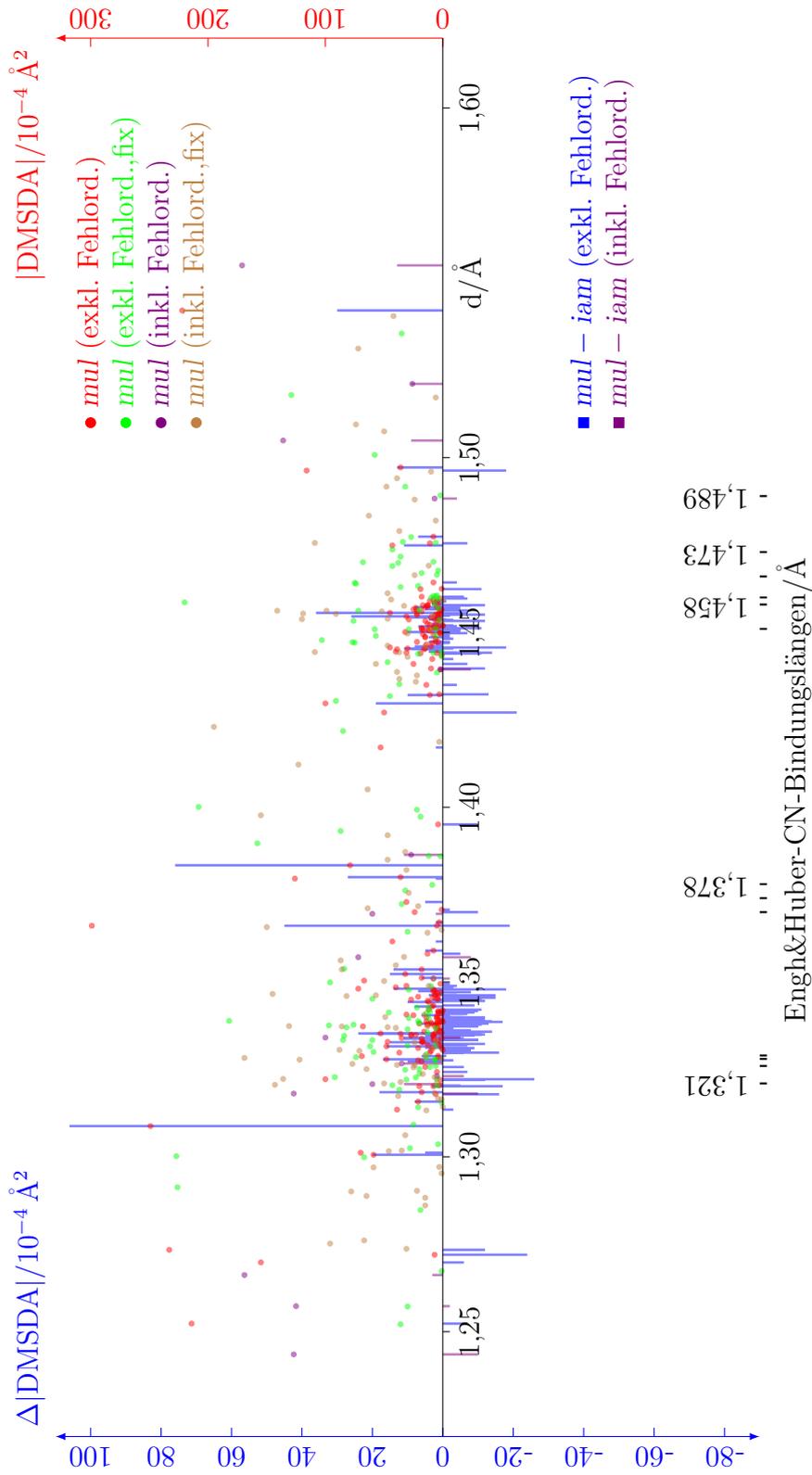


Abbildung 6.12: Veränderung der DMSDA-Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die CN-Bindungen von P1LYSO.

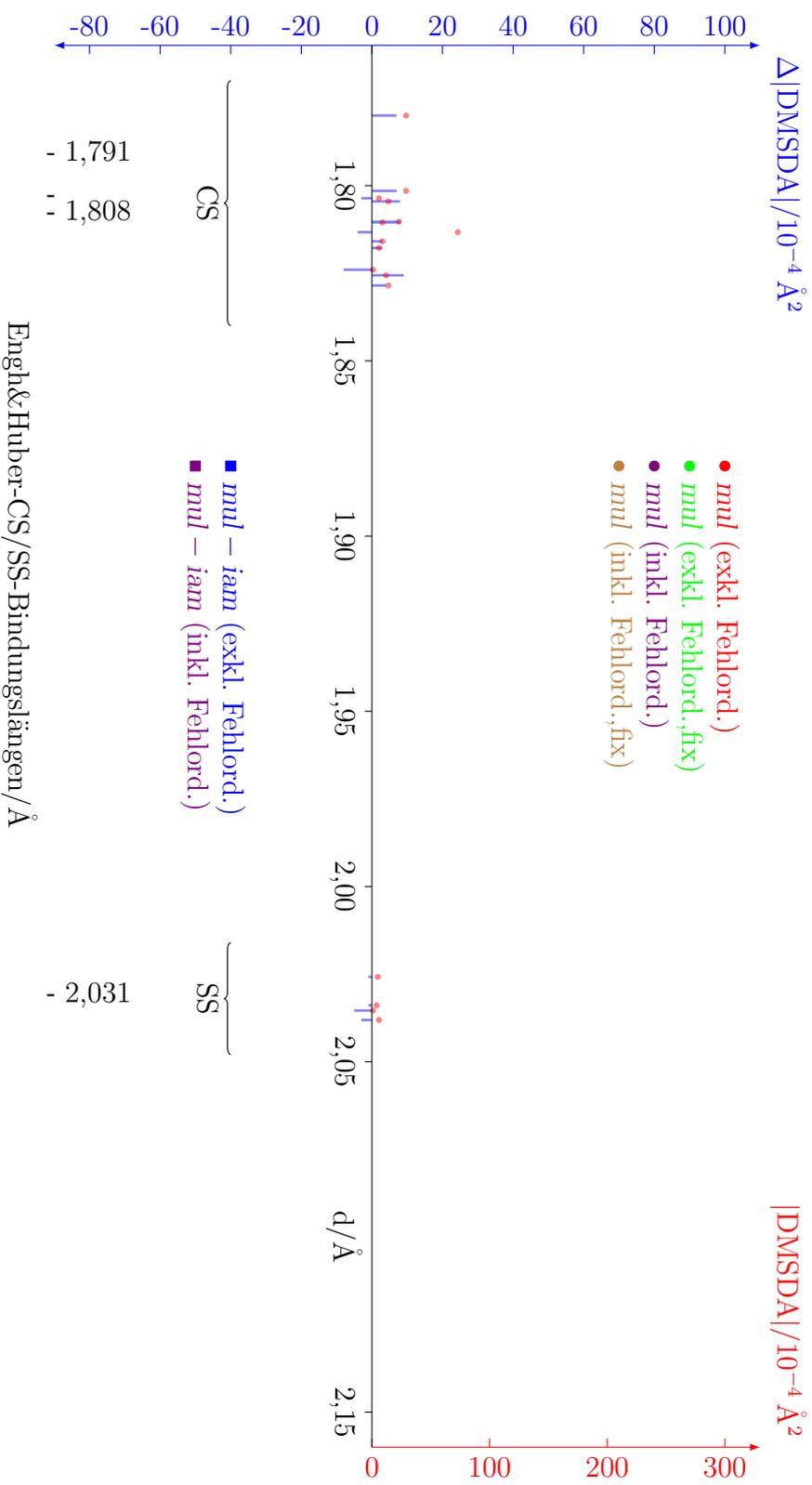


Abbildung 6.13: Veränderung der DMSDA-Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die CS-Bindungen und SS-Bindungen von P1LYSO.

6.5 Restdichten

Die vom Modell *mul* nicht beschriebene, in den Beobachtungen aber enthaltene Restelektronendichte $(F_o^2 - F_c^2)\phi_c$ von P1LYSO zeigt Abb. 6.14 auf der nächsten Seite. Dazu wurden im Umkreis von 4 \AA um alle Atomlagen des Moleküls, Isooberflächen bei $\pm 0,40$, $\pm 0,25$ und $\pm 0,10 \text{ e\AA}^{-3}$ dargestellt. Hierbei wird augenfällig, daß der Großteil der Restdichten in der Solvenssphäre zu finden ist. Eine analoge Überblicksdarstellung der Ergebnisse der Verfeinerung *iam* in Abb. 6.15 ist von der Restdichtekarte des asphärischen Modells kaum zu unterscheiden. Gleiches gilt für die Abb. B.2 auf Seite 343 zum Modell *iam_{rb}*.

Mit Abb. 6.16 auf Seite 176 wurde deshalb zusätzlich eine Darstellung der Differenz der Restelektronendichten der Modelle *iam* und *mul* angefertigt, indem die Werte der asphärischen Verfeinerung von denen der sphärischen subtrahiert wurden. Daran werden zwei Sachverhalte unmittelbar deutlich: Zum einen sind die Unterschiede der Restdichten der Modelle wesentlich geringer als die subtrahierten Dichten selbst, weshalb die abgebildeten Oberflächen auf ungefähr einem Drittel der in den vorherigen Illustrationen verwendeten Skalen gezeichnet wurden. Zum anderen sind die Differenzen hauptsächlich auf das Lysozymmolekül beschränkt sind, was zu erwarten war, denn nur dort wurden die Atome asphärisch behandelt. Fehlgeordnete Bereiche zeigen hierbei eine Anhäufung sowohl positiver wie negativer Restdichtedifferenzen. Der außerhalb des Moleküls liegende Solvensbereich wurde in beiden Modellen gleichartig behandelt, wodurch dort kaum eine Änderung zu verzeichnen ist.

Die gezeigten Isooberflächenwerte in Abb. 6.16 sind schwieriger zu interpretieren. Positive (blaue) Isooberflächen kennzeichnen Bereiche, in denen das Modell *mul* reduzierte Restdichten aufweist, was sowohl ein Weniger an positiven als auch ein Mehr an negativen Restdichten bedeuten kann. Umgekehrt beinhaltet das Modell *mul* geringere negative oder stärker positive Restdichten an Orten mit negativen (gelbrotten) Differenzisooberflächen. Bei Berücksichtigung der ursprünglichen Restdichten zeigt sich eine Reduktion der im Modell *iam* noch vorhandenen Restdichten auf den Bindungen, allerdings bewirkt die asphärische Behandlung eine Verstärkung negativer Restdichten neben den Bindungen.

Exemplarisch zeigen die Abbildungen Abb. 6.17 auf Seite 177 und Abb. 6.18 auf Seite 178 die Restdichten und Restdichtedifferenzen an den Aminosäureresten Tryptophan 108 und Glutaminsäure 35 bis zu einer Entfernung von $1,4 \text{ \AA}$ um die Fragmente. Die jeweiligen oberen Abbildungen (a) und (b) zeigen das Ergebnis der asphärischen Verfeinerung sowie, analog zur Darstellung 6.16, die Differenz zum Modell *iam*. In den jeweils zwei unteren Bildern (c) und (d) sind die Restdichten der beiden sphärischen Modelle gezeigt. Gegenüber den Abbildungen der Gesamtstruktur, konnten die darzustellenden Konturwerte und -intervalle verringert werden, da die auftretenden Restdichten weniger stark in positive oder negative Richtung ausgeprägt sind.

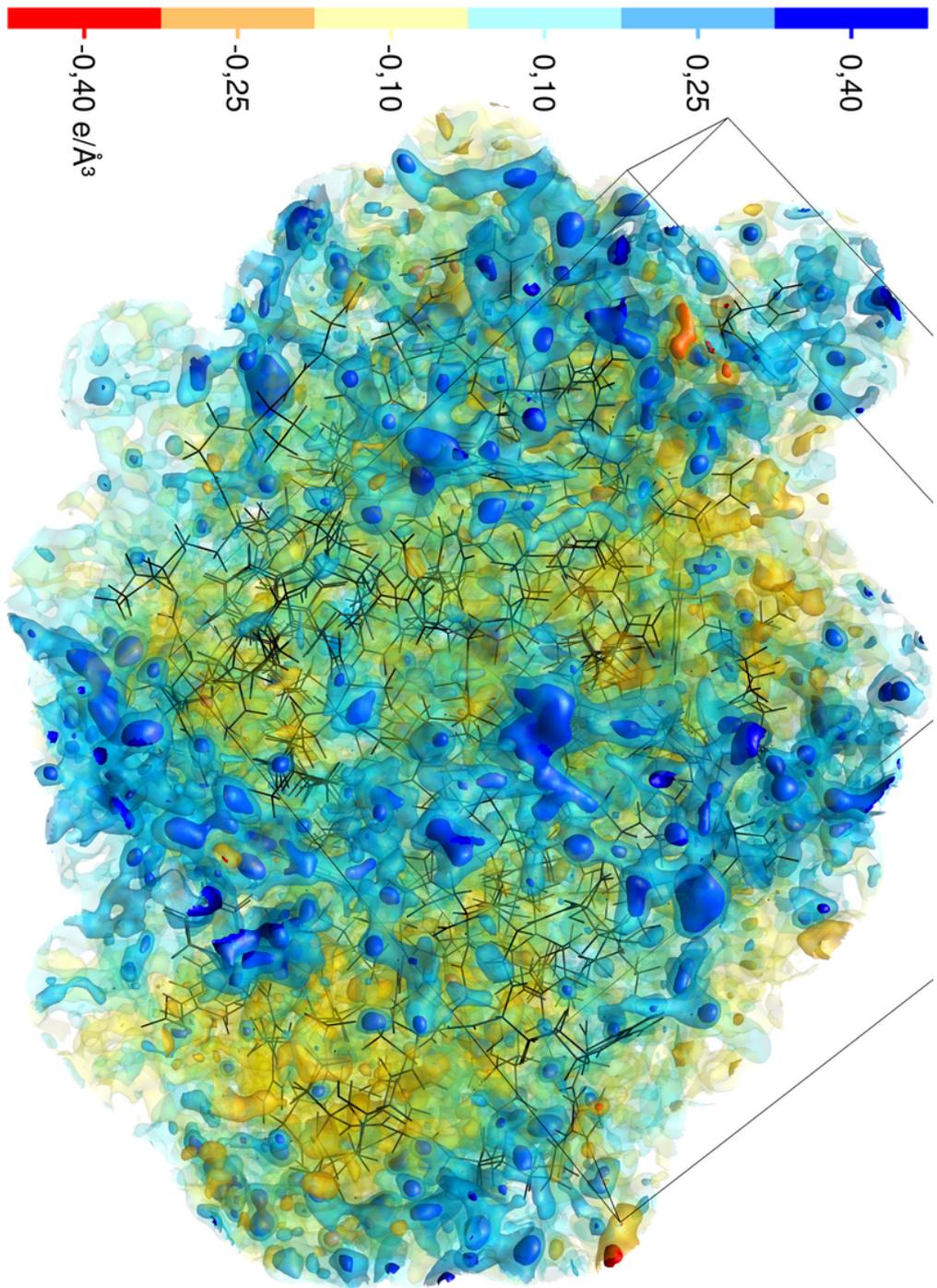


Abbildung 6.14: Restdichten von P1LYSO im Modell *mul*, Darstellung bei sechs Isooberflächenwerten bis zu 4 Å um die Struktur.

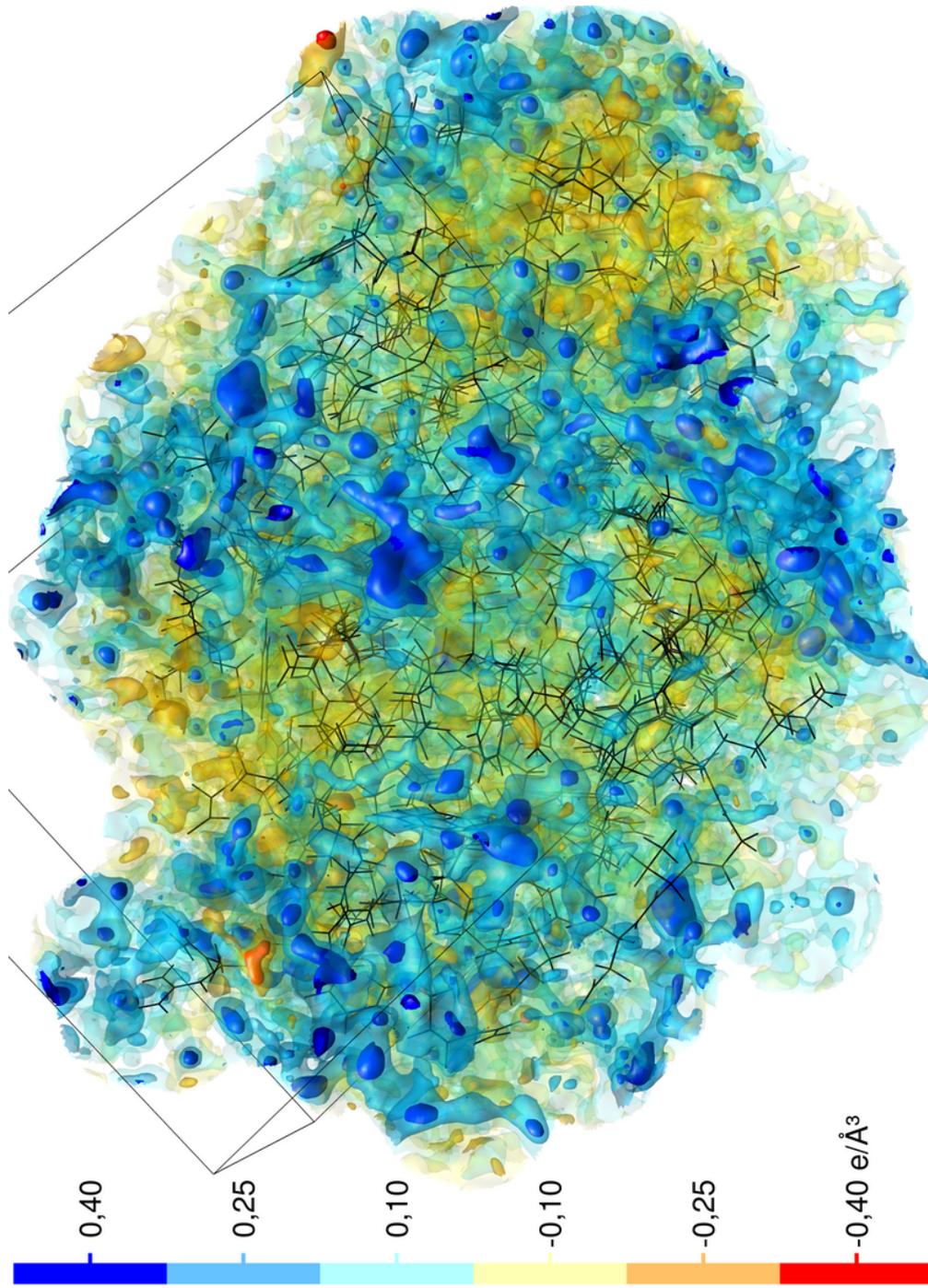


Abbildung 6.15: Restdichten von P1LYSO im Modell iam, Darstellung bei sechs Isooberflächenwerten bis zu 4\AA um die Struktur.

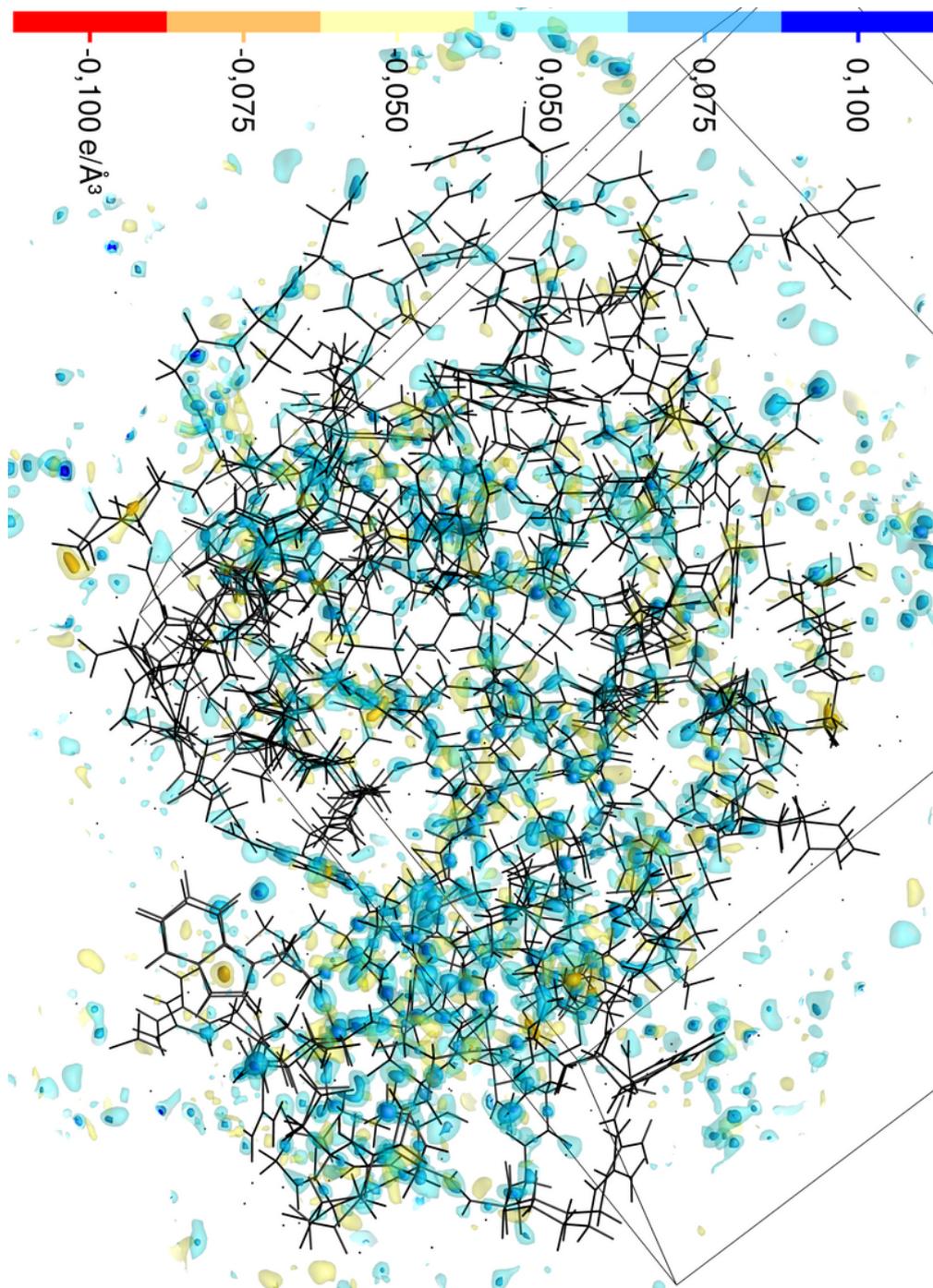


Abbildung 6.16: Restdichtedifferenzen $\text{iam-mul } 4 \text{ \AA}$ um P1LYSO.

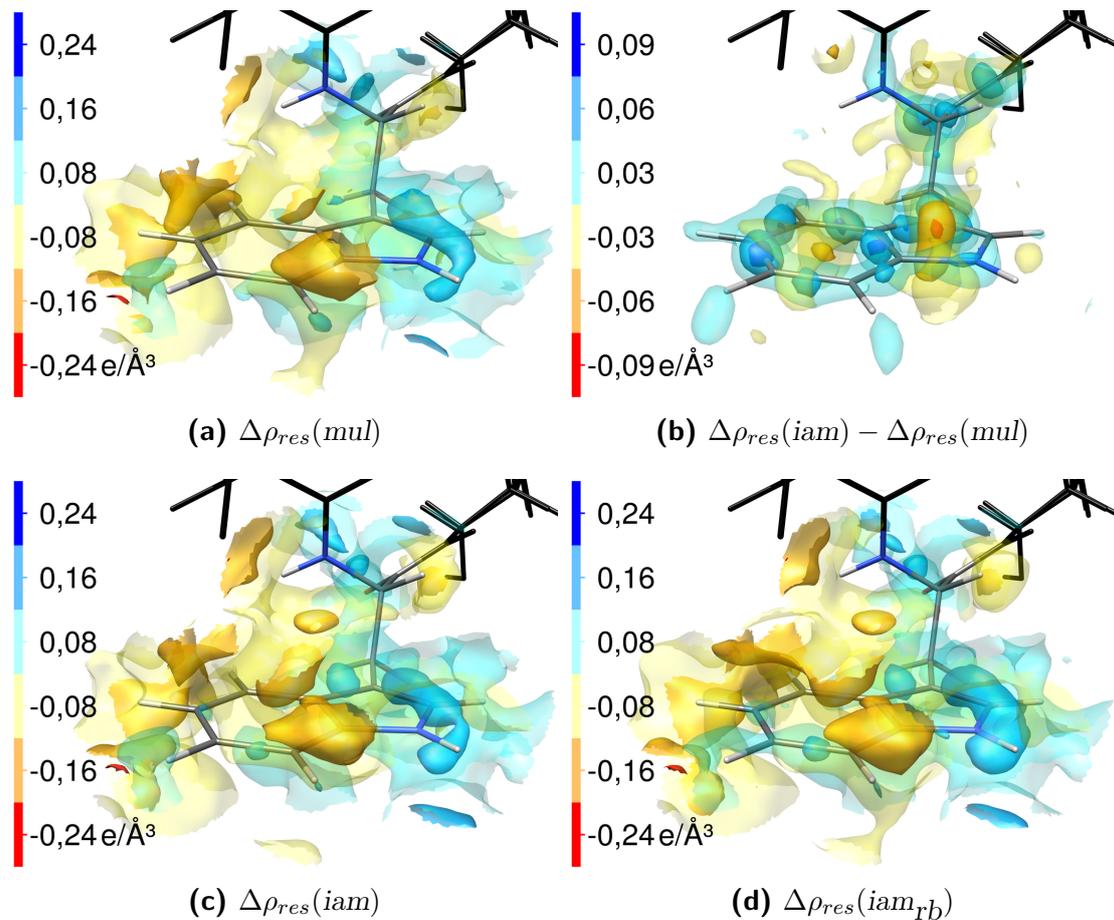


Abbildung 6.17: Restdichten und Restdichtedifferenzen bis $1,4\text{ \AA}$ um den P1LYSO-Tryptophanrest 108 für sphärische und asphärische Verfeinerungen.

Die Restdichten des Tryptophanrests liegen in der Größenordnung von $0,1\text{--}0,2\text{ e}\text{\AA}^{-3}$. Wie die Abb. 6.17(c) und (d) zeigen, sind die positiven Restdichten der sphärischen Verfeinerungen hauptsächlich am Indolzyklus ausgeprägt, dort besonders am Pyrrolring und auf vielen Bindungen. Die negativen Restdichten verteilen sich ober- und unterhalb der Ringebene und sind noch schlechter als die positiven lokalisiert. Als Resultat der asphärischen Verfeinerung verringert sich die unbeschriebene Bindungsdichte des Heterozyklus, wie in 6.17(a) gezeigt. In der Differenz 6.17(b) von *iam* und *mul* tritt dies deutlicher zutage: Alle Bindungen weisen nun eine geringere Restdichtekonzentration auf. Dafür sind in den Ringmitten im Modell *mul* um ca. $0,06\text{ e}\text{\AA}^{-3}$ höhere negative Restdichten als im Modell *iam* zu finden – aus den Ringmitten hat das asphärische Modell Elektronendichte zu Gunsten der Bindungen abgezogen, obwohl die beobachteten Strukturformfaktoren auch dort Dichte nahelegen. An den ausgeprägten negativen Restdichten im Umfeld des Fragments ändert sich in allen Modellen wenig.

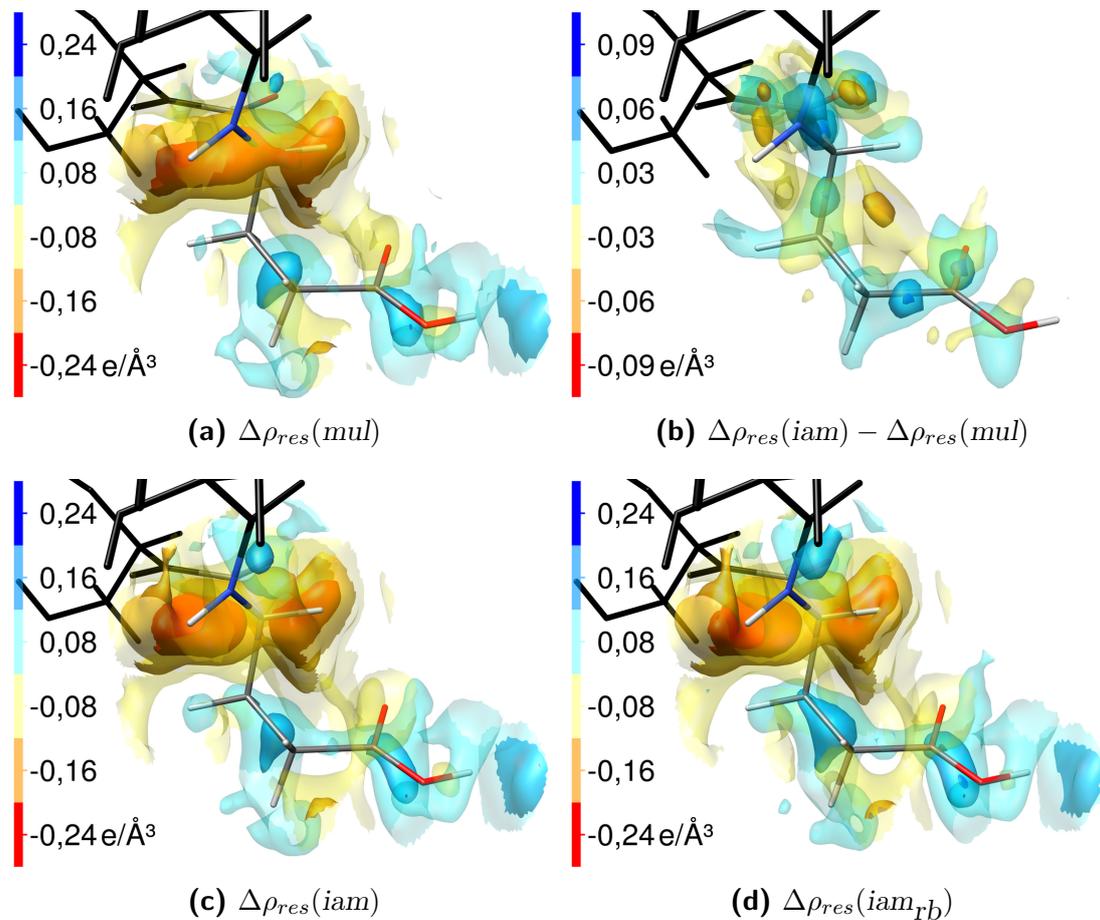


Abbildung 6.18: Restdichten und Restdichtedifferenzen bis $1,4 \text{ \AA}$ um den P1LYSO-Glutaminsäurerest 35 für sphärische und asphärische Verfeinerungen.

Ein ähnliches Bild läßt sich für das zweite Beispiel, den Aminosäurerest Glutaminsäure 35, zeichnen. Die Darstellungen werden durch hohe negative Restdichten an der Peptidbindung dominiert, was auf eine Überschätzung, evtl. durch eine nichtbeachtete Fehlordnung, der vorhandenen Elektronendichte durch die Modelle hindeutet. Dazu kommen positive Reste auf und neben der Seitenkette. Betrachtet man die Restdichtedifferenz in Abb. 6.18(b), ist die Abnahme der Restdichte auf den Bindungen $C_\alpha-C_\beta$, $C_\gamma-C_\delta$ und $C_\delta-O_\epsilon$ deutlich. Diese fällt für $C_\beta-C_\gamma$ geringer aus, da die Restdichte eher neben der Bindung lokalisiert ist. Wieder sind die Veränderungen wesentlich geringer als die absoluten Restdichten, was sich auch in der engeren Skalenteilung ausdrückt. Auffällig in Abb. 6.18(b) ist der ausgeprägt positive Bereich auf der Peptidbindung. Da in 6.18(a) und den sphärischen Resultaten 6.18(c) und (d) dort negative Restdichten lokalisiert sind, bedeutet die blaue Kontour um die Peptidbindung nicht etwa eine Abnahme der Restdichte in *mul*, sondern eine Zunahme negativer Restdichte gegenüber der sphärischen Verfeinerung *iam*.

Tabelle 6.7: Statistik der Restdichten der asymmetrischen Einheit von P1LYSO.

	$x_{min/max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
<i>mul</i>	-649,2/1450,0	0,0(134,1)	-90,0	-11,5	76,7	166,7	122,9
<i>iam</i>	-638,8/1455,0	0,0(135,9)	-91,4	-11,5	78,2	169,6	124,8
<i>iam_{rb}</i>	-642,4/1446,0	0,0(136,8)	-92,3	-11,5	79,4	171,7	126,6

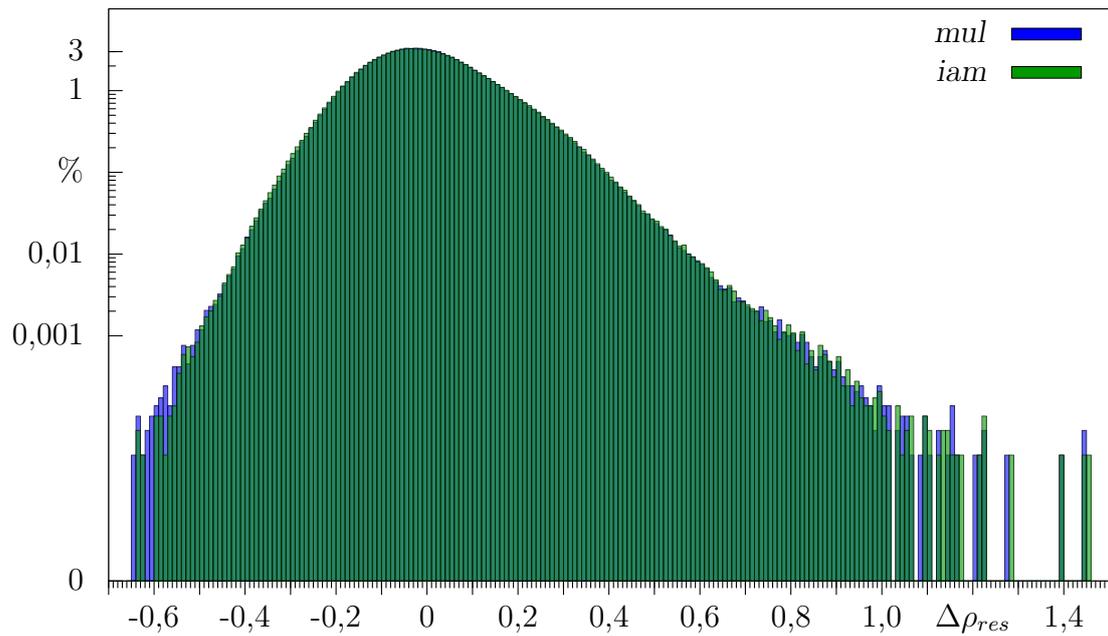
[†]Minimum, Maximum und Mittelwert mit Standardabweichung der Restdichte;

[‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in $10^{-3} \text{ e}/\text{\AA}^3$.

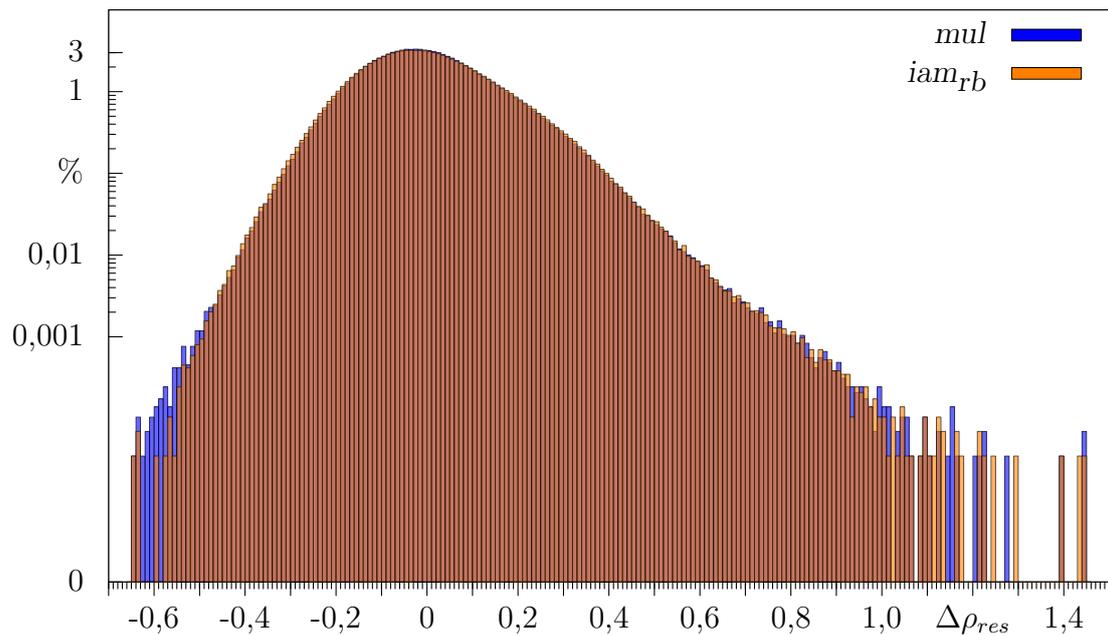
Scheinbar werden physikalisch sinnvolle und plausible Restdichteverringerungen auf den Bindungen durch Zunahmen der Restdichtebeträge andernorts konterkariert. Eine Betrachtung der Diskrepanzen zwischen Modell und Beobachtungen, im Sinne einer Statistik über die gesamte asymmetrische Einheit, soll zur Klärung beitragen.

Die Restdichten $\sum(F_o - F_c) \exp(i\phi_c)$ der Modelle *iam*, *iam_{rb}* und *mul* in der asymmetrischen Einheit wurden auf einem Gitter mit $128 \times 144 \times 156 = 2875392$ Punkten berechnet. Das Volumen eines Voxels beträgt $0,009 \text{ \AA}^3$, normiert darauf errechnet sich die summierte Gesamtrestelektronendichte $\sum |\Delta\rho_{res}|$ in absteigender Reihenfolge zu 2663 e im Modell *iam_{rb}*, 2640 e für *iam* und 2601 e bei *mul*; bezogen auf das Zellvolumen, sind dies Restdichten von 0,1063, 0,1054 und $0,1038 \text{ e}\text{\AA}^{-3}$. Insgesamt nimmt die Diskrepanz zwischen Beobachtungen und Modell im asphärischen Fall um ca. 2 % ab. Positive und negative Beiträge zu diesen Summen halten sich ungefähr die Waage, wie Tab. 6.7 zeigt, denn die mittlere Restelektronendichte liegt nahe bei Null. Die Quartilsgrenzen zeigen, daß die Hälfte der Restdichten zwischen $-0,09$ und $0,08 \text{ e}\text{\AA}^{-3}$ liegt. Außerdem läßt die Statistik eine marginal engere Streuung um den Nullwert im Invariommodell erkennen, wiewohl die Maximal- und Minimalwerte gegenüber den sphärischen Modellen zunehmen.

Einen Eindruck von der Häufigkeitsverteilung der Restdichten in $0,01 \text{ e}\text{\AA}^{-3}$ breiten Intervallen geben die Histogramme in Abb. 6.19 auf der folgenden Seite. Auch dort sind die höheren Extremwerte des Modells *mul* kenntlich, aber ebenso, daß sie weniger als ein zehntausendstel des betrachteten Volumens ausmachen. Ebenso wird deutlich, daß in allen Modellen der Hauptteil der Restdichte Werte um $\pm 0,1 \text{ e}\text{\AA}^{-3}$ annimmt. Die Unterschiede zwischen den Verfeinerungen sind allerdings gering, sodaß die Verteilungen fast aufeinander liegen. Die Ähnlichkeit äußert sich bspw. in der Häufigkeit von Voxeln im Intervall $(0;0,01]$ von 3,17 % im Modell *mul* und 3,13 % bzw. 3,08 % in den Modellen *iam* und *iam_{rb}*. Im Intervall $(-0,01;0]$ liegen bei *mul* 3,24 % und bei den sphärischen Modellen jeweils 3,20 % und 3,16 % der Werte. Das Maximum der Häufigkeitsverteilung nach der asphärischen Verfeinerung liegt mit 3,32 % im Intervall $(-0,03;0,02]$, nach der Modellierung *iam* bei 3,26 % zwischen $(-0,04;0,03]$ und bei 3,21 % im gleichen Intervall für das Modell *iam_{rb}*.



(a) Restdichten P1LYSO, *mul* vs *iam*.



(b) Restdichten P1LYSO, *mul* vs *iam_{rb}*.

Abbildung 6.19: Histogramm der Restdichtegitter $\sum(F_o - F_c) \exp(i\phi_c)$ der sphärischen (grün/orange) und asphärischen (blau) Modelle. Intervalle $\Delta(i; i+0,01]$, Einheiten $e\text{\AA}^{-3}$.

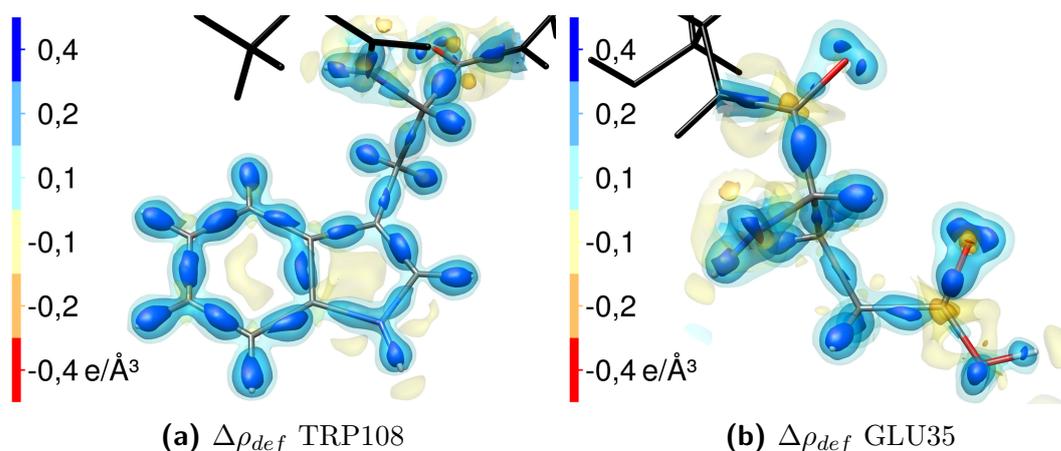


Abbildung 6.20: Deformationsdichten der P1LYSO-Reste Tryptophan 108 und Glutaminsäure 35

6.6 Deformationsdichten

Die gegenüber dem sphärischen Modell deformierte Valenzdichte $\Delta\rho_{def}$, wurde auf einem 301^3 Punkte großen Gitter berechnet. Jede Volumeneinheit umfaßte dabei $0,011 \text{ \AA}^3$. Auf diese Gitterfeinheit normiert, summiert sich die Gesamtdeformationsdichte $\sum |\Delta\rho_{def}| V_{vox}$ auf $998 e$ – im Mittel wird bei jedem Atom ein halbes Elektron asphärisch verteilt.

Die das Multipolmodell von der sphärischen Beschreibung unterscheidenden Deformationsdichten, zeigt Abb. 6.20 exemplarisch für die bereits angeführten Aminosäurereste Tryptophan 108 und Glutaminsäure 35. Eine Darstellung für das gesamte Lysozymmolekül gibt Abb. 6.21 auf der folgenden Seite. Es wird deutlich sichtbar wie das Invariommodell Bindungsdichten und freie Elektronenpaare berücksichtigt, indem Valenzdichte dorthin verschoben wird, wohingegen es andernorts geringere Elektronendichte als das sphärische Modell aufweist. Zudem erklären diese Illustrationen die geringen Restdichteänderungen. Während letztere recht diffus verteilt sein können, sind die das asphärische Modell charakterisierenden Deformationsdichten auf einen engen Raumbereich konzentriert. Nicht unmittelbar auf Bindungen oder nahe Atomen lokalisierte Merkmale der Elektronendichte, können nicht direkt beeinflußt werden.

Da diese gerichtete Beschreibung von Elektronendichte den Unterschied zum sphärischen Modell ausmacht, stellt sich die Frage nach dem Verhältnis der Deformationsdichte zur gesamten Restdichte der Struktur: Mit den Werten aus dem vorherigen Abschnitt, ergibt sich ein Quotient $\Delta\rho_{def}/\Delta\rho_{res} = 998,25 e/2600,97 e$ von $0,384$. Die im Modell nicht beschriebene Restdichte übersteigt die eingeführte Deformationsdichte um das 2,6-fache. Somit wird unmittelbar deutlich, daß keine großen Veränderungen zur sphärischen Verfeinerung erwartet werden konnten.

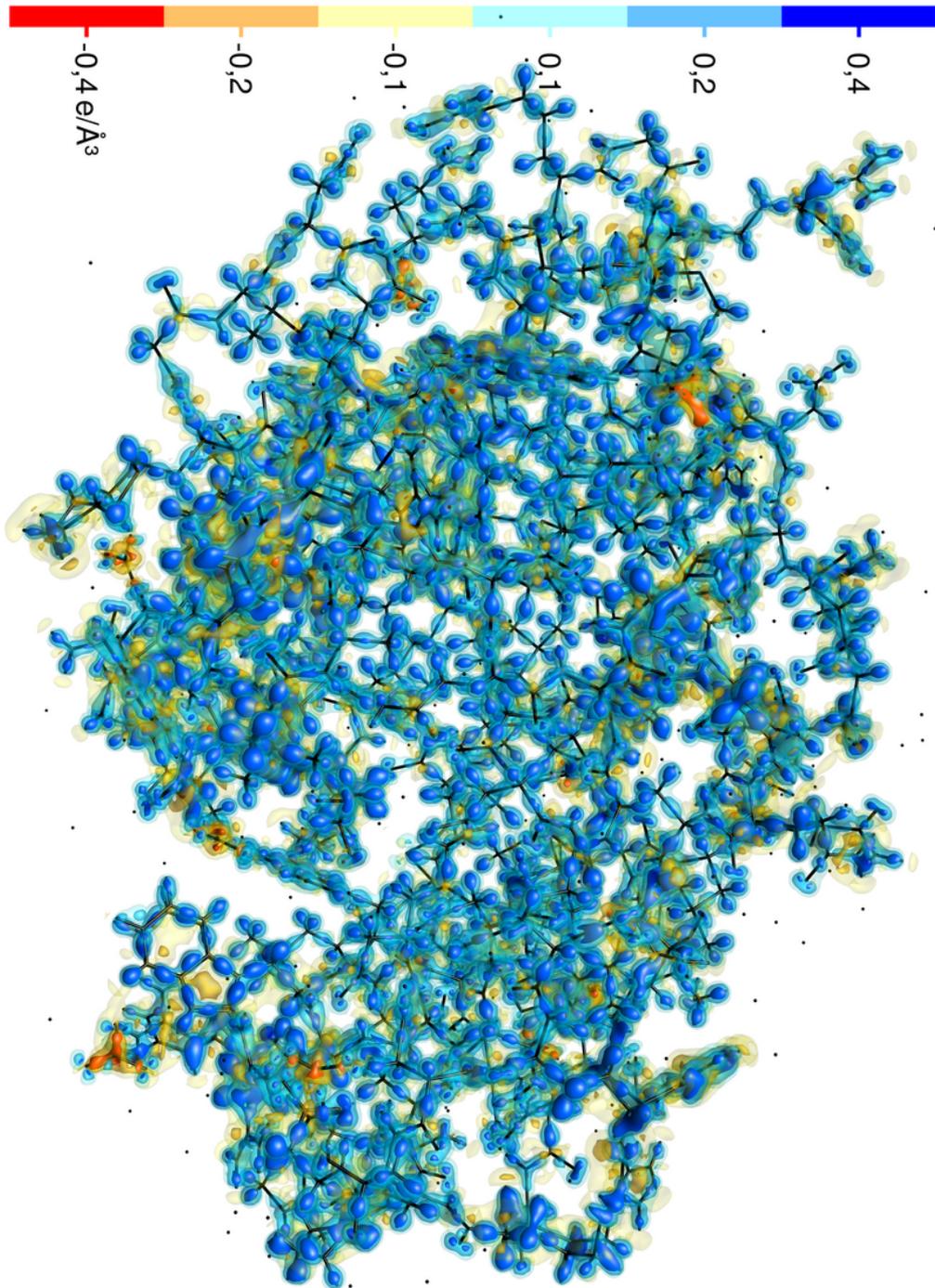


Abbildung 6.21: Deformationsdichte $\Delta\rho_{def}$ der Gesamtstruktur P1LYSO.

6.7 Abgeleitete Eigenschaften

Das aus den Invariomen des Modells *mul* für P1LYSO abgeleitete elektrostatische Potential auf der ebenfalls dem Invariommodell entstammenden Isooberfläche der Elektronendichte bei $0,001 \text{ e}/\text{a}_0^3$ ist in Abb. 6.22 auf der nächsten Seite gezeigt. Bei der Berechnung wurden nur die 1961 Atome der Hauptpopulation des Lysozymmoleküls berücksichtigt, Atome der Alternativkonformation und Solvensmoleküle tragen nicht zum ESP bei. Die für die katalytische Wirkung des Enzyms besonders relevanten Aminosäurereste Glutaminsäure 35 und Asparaginsäure 52, sind in der Darstellung mittels Beschriftung hervorgehoben.

Fast die gesamte Moleküloberfläche zeigt ein moderates Potential von $\pm 0,2 \text{ e}/\text{a}_0$. Nur die Tasche weist um die deprotonierte Asparaginsäure 52 ein verstärkt negatives ESP auf, dieses erreicht Werte bis zu $-0,325 \text{ e}/\text{a}_0$. Auf der gegenüberliegenden Seite herrscht um GLU 35 ein annähernd neutrales Potential. Während der Katalyse wird das Proton der Carboxylgruppe der Glutaminsäure 35 auf ein glykosidisches Sauerstoffatom des Oligosaccharids übertragen, indes das entstehende Oxocarbeniumkation durch die negative Ladung von ASP 52 stabilisiert oder gar in einem kovalenten Intermediat abgefangen wird.

In Abb. 6.23 auf Seite 185 wird das Resultat der *g98/AM1*-Kalkulation dargestellt. Wie schon in Absch. 5.7.2 auf Seite 140 für APA ausgeführt wurde, ergeben sich daraus fehlerhafte, ins Negative verschobene Ergebnisse, während *g03* auch hier ein offensichtlich unbrauchbares Potential lieferte. So nimmt der Wertebereich des AM1-ESPs, auf der aus der gleichen Rechnung erhaltenen $0,001 \text{ e}/\text{a}_0^3$ -Oberfläche, ausschließlich negative Werte an. Die Verteilung der negativsten bis positivsten Bereiche ist jedoch in guter Übereinstimmung mit den Resultaten aus *mul*.

Für den Testfall APA, konnten mit *mopac* zu *ab initio*-Rechnungen konsistente ESPs erzeugt werden, was auch für dieses Molekül versucht wurde. Diese Rechnung konnte leider nicht erfolgreich beendet werden, da die Bestimmung der Oberflächenpunkte scheiterte.

Deswegen wird als Alternativpotential nur das mit *delphi* berechnete Punktladungs-ESP ϕ_{PC} in Abb. 6.24 auf Seite 186 aufgeboden und dem bereits gezeigten Multipol-ESP ϕ_{mul} gegenübergestellt. Die jedem Atom zugeordneten Ladungen entsprechen der Standardparametrisierung des *amber03*-Kraftfeldes, die Atomradien den in *vvv* vorgesehenen Werten. Die Potentiale werden in beschränktem und vollem Umfang, sowohl auf der aus dem Multipolmodell stammenden $0,001 \text{ e}/\text{a}_0^3$ -Elektronendichteoberfläche ρ_{mul} als auch auf der mit *vvv* bestimmten van-der-Waals-Oberfläche A_{vdW} gezeigt. Mit letzterer Variante wird zum einen eine vollkommen invariomfreie ESP-Beschreibung geliefert und zum anderen werden die Potentialunterschiede aufgrund unterschiedlich gewählter Oberflächen verdeutlicht. Die Einheiten wurden in kT/e gewählt, der Umrechnungsfaktor zu den bisher gezeigten atomaren Einheiten beträgt $1 \text{ e}/\text{a}_0 = 1059,59 \text{ kT}/\text{e}$.

Während im Fall APA noch eine gute Übereinstimmung von ϕ_{mul} und ϕ_{PC}

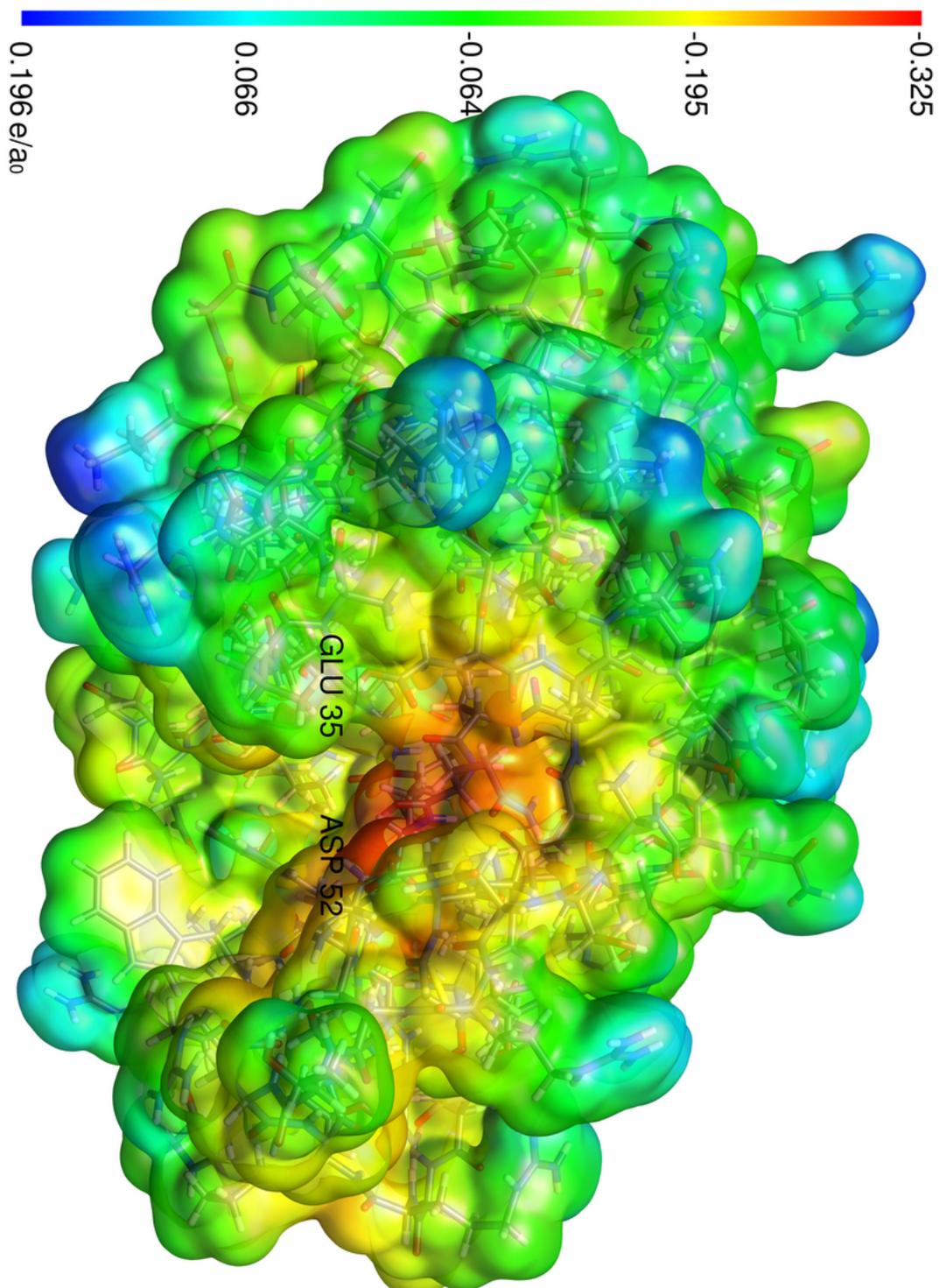


Abbildung 6.22: Elektrostatisches Potential des Multipolmodells von P1LYSO auf der $0,001 \text{ e}/a_0^3$ -Isooberfläche.

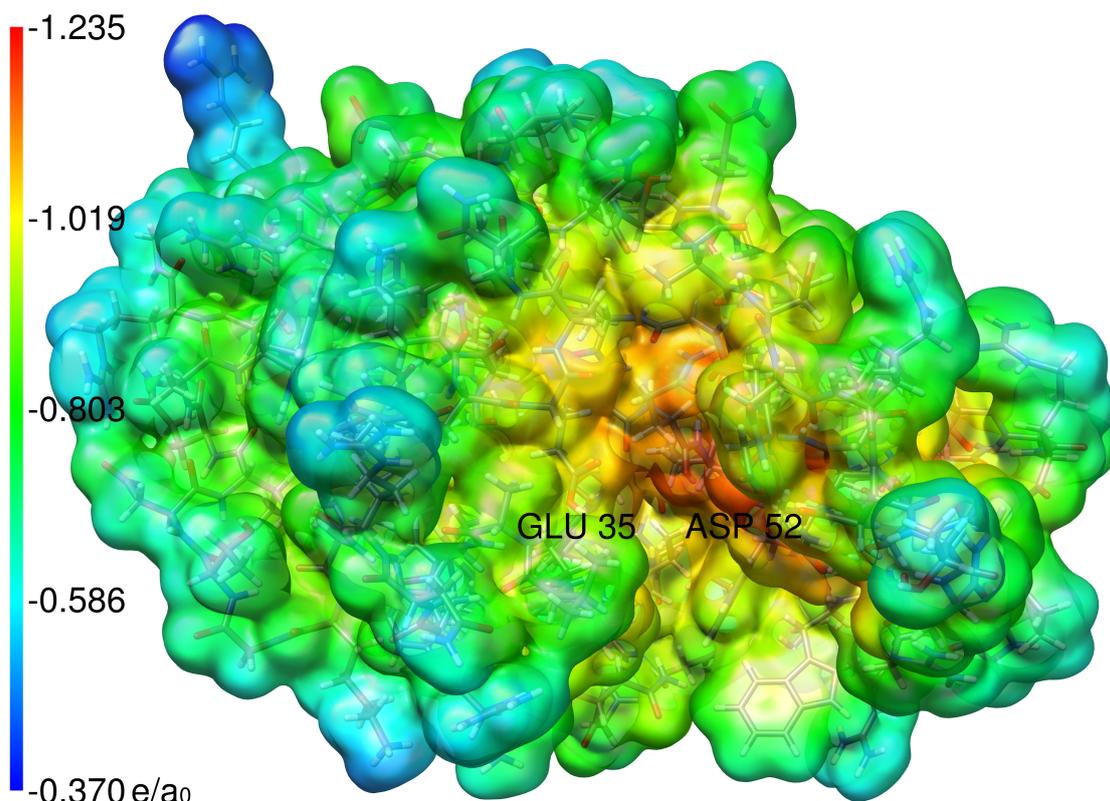


Abbildung 6.23: Semiempirisches elektrostatisches Potential von P1LYSO aus g98/AM1-Rechnung auf der $0,001 e/a_0^3$ -Oberfläche.

festgestellt wurde, sind im aktuellen Fall die Unterschiede besonders eklatant. Das Invariompotential hat auf der $\rho_{mul} = 0,001 e\text{\AA}^{-3}$ -Oberfläche Minimal- und Maximalwerte von $-345,4$ und $205,3 \text{ kT/e}$, auf der vdW-Oberfläche von $-382,0$ und $392,6 \text{ kT/e}$. Da die vdW-Oberfläche meist näher am Molekül liegt, sind die extremen Werte nicht überraschend. Aus den Punktladungen resultiert hingegen ein Wertebereich von nur -129 – 252 kT/e bzw. -160 – 289 kT/e auf den genannten Oberflächen.

Die stärkere Polarisierung von ϕ_{mul} wird durch den direkten Vergleich mit ϕ_{PC} verdeutlicht. Abb. 6.24(a) und (b) zeigen dies auf der $0,001 e/a_0^3$ -Oberfläche bei fixem Wertebereich von $\pm 191 \text{ kT/e} = \pm 0,18 e/a_0$. Offensichtlich nimmt ϕ_{mul} stärker negative Werte an, die sich über einen größeren Teil der Oberfläche, insbesondere im Bereich der Tasche, erstrecken. Gleiches zeigen Abb. 6.24(c) und (d) für die van-der-Waals-Oberfläche.

Läßt man das Punktladungs-ESP seinen vollen Wertebereich auf der jeweiligen Oberfläche annehmen, ergeben sich die beiden unteren Abbildungen 6.24(e) und (f): Die Verteilung elektronegativer und -positiver Bereiche ist den bisherigen Ergebnissen ausgesprochen vergleichbar.

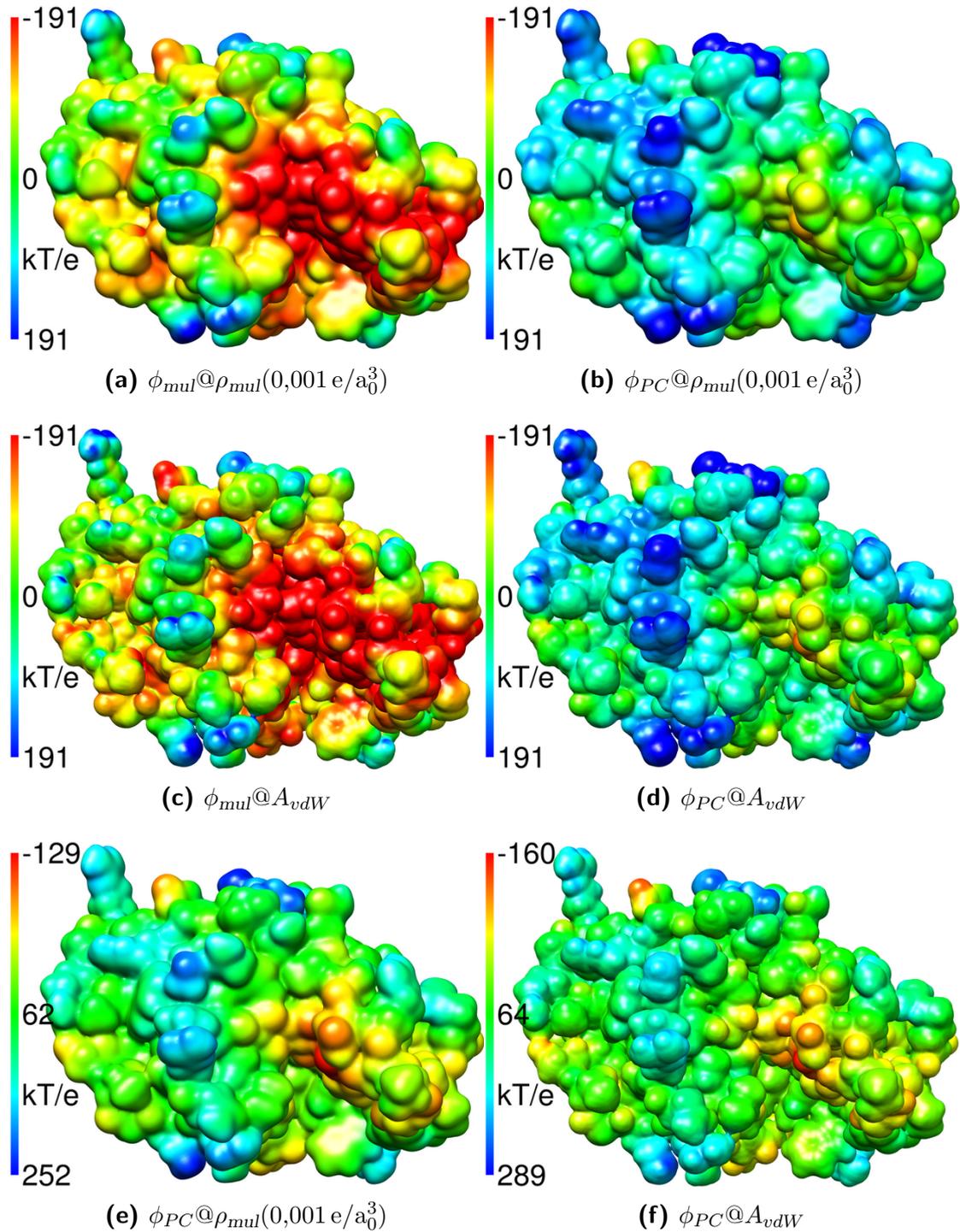


Abbildung 6.24: Multipol-ESP (ϕ_{mul}) und Punktladungs-ESP (ϕ_{PC}) auf $0,001 e/a_0^3$ -Isooberfläche der Multipolelektronendichte (ρ_{mul}) und van-der-Waals-Oberfläche (A_{vdW}). (a)-(d): Fester Wertebereich ± 191 kT/e = $\pm 0,18 e/a_0$. (e)-(f): Voller Wertebereich.

7 Tetragonales Lysozym

7.1 Gütefaktoren

Die in Abschn. 4.4 auf Seite 83 beschriebenen sphärischen und asphärischen Modelle verfeinern sich zu den in Tab. 7.1 gelisteten Gütefaktoren. Bei gleichbleibender Parameterzahl erzielt das asphärische Modell *mul* einen um 0,23 Prozentpunkte niedrigeren $R1$ als die sphärische Verfeinerung *iam_{rb}* und einen um 0,11 Prozentpunkte niedrigeren als die Wiederholung der klassischen Verfeinerung *iam*. Somit wird durch das Invariommodell die Beschreibung von TETRALYSO um 1–2 % gegenüber dem sphärischen Modell verbessert.

Als komplementäre Betrachtung der Unterschiede in den Abweichungen der Modelle von den Beobachtungen, sollen direkt die gewichteten Differenzen $\Delta F = |F_o - F_c| \sigma(F_o)^{-1}$ zwischen beobachteten und aus dem Modell berechneten Struktur Faktoren der Verfeinerungen *iam* und *mul* miteinander verglichen werden: Die Abweichungen $\Delta F(\textit{iam})$ liegen im arithmetischen Mittel bei 0,7315, der Median beträgt 0,4888, die höchste Abweichung 226. Analoge Parameter der Verteilung der Differenzen $\Delta F(\textit{mul})$, bestimmen sich zu einem Mittel von 0,7269 und einem Median von 0,4848, mit einer Maximalabweichung von 218,7. Ohne $\sigma(F_o)$ -Gewichtung werden zehnmal höhere Unterschiede beobachtet. In beiden Fällen entspricht das Ergebnis der Erwartung kleinerer Abweichungen im Modell *mul*, wie die um 0,004 und ohne Gewichtung um etwa 0,04 geringeren Mediane belegen.

Da die Unterschiede sehr klein sind, stellt sich die Frage nach der Signifikanz

Tabelle 7.1: Gütefaktoren der sphärischen und asphärischen Verfeinerung von TETRALYSO.

	<i>iam</i>	<i>iam_{rb}</i>	<i>mul</i>
Reflexe $F_o^2 > 2\sigma(F_o^2)$	107618	107618	107618
Parameter	6760	6760	6760
N_{ref}/N_{par}	15,920	15,920	15,920
$R(F)$	0,1248	0,1260	0,1237
$R_{all}(F)$	0,1310	0,1323	0,1299
$R_w(F^2)$	0,2987	0,3015	0,2983
GoF_w	2,7284	2,7558	2,7225
Mittl.(Versch./esd)	$1,12 \cdot 10^{-4}$	$1,21 \cdot 10^{-3}$	$2,97 \cdot 10^{-3}$

Tabelle 7.2: Gütefaktoren der sphärischen und asphärischen Verfeinerung von TETRALY-
SO gegen beschnittene Daten.

	<i>iam</i>	$\frac{1}{2} \leq \frac{F_o}{F_c} \leq 2$ <i>iam_{rb}</i>	<i>mul</i>
Reflexe $F_o^2 > 2\sigma(F_o^2)$	99818	99818	99818
Parameter	6760	6760	6760
N_{ref}/N_{par}	14,766	14,766	14,766
$R(F)$	0,0974	0,0988	0,0970
$R_{all}(F)$	0,0978	0,0992	0,0975
$R_w(F^2)$	0,2375	0,2410	0,2380
GoF_w	2,1792	2,2126	2,1833
Mittl.(Versch./esd)	$3,75 \cdot 10^{-6}$	$3,81 \cdot 10^{-5}$	$3,19 \cdot 10^{-5}$

dieser Verbesserung durch die asphärische Verfeinerung. Mit Bezugnahme auf das in Absch. 3.1.2 auf Seite 36 zu statistischen Tests Gesagte, ergibt sich für die zwei Verteilungen $\Delta F_w(iam) = |F_{o,iam} - F_{c,iam}| \sigma(F_{o,iam})^{-1}$ und $\Delta F_w(mul) = |F_{o,mul} - F_{c,mul}| \sigma(F_{o,iam})^{-1}$ mit den Hypothesen $H_0 : \Delta F_w(iam) = \Delta F_w(mul)$ und $H_1 : \Delta F_w(iam) > \Delta F_w(mul)$ sowohl im t -Test als auch im Wilcoxon-Rangsummentest mit $p < 2 \cdot 10^{-16}$ eine statistisch signifikante Ablehnung der Nullhypothese – die Strukturfaktordifferenzen sind durch die Verfeinerung *mul* signifikant kleiner als im sphärischen Fall *iam*. Gleiches gilt für die ungewichteten Daten.

Effekt des Beschneidens der Daten

Von 122388 Reflexen haben 17486 ein Verhältnis $F_o^2/F_c^2 \geq 2$ oder $\leq 1/2$ nach der sphärischen Verfeinerung, wovon Daten höchster und niedrigster Auflösung anteilig am stärksten betroffen sind, was Abb. 7.1 auf der gegenüberliegenden Seite verdeutlicht. Nimmt man diese 14,3% aus dem Datensatz heraus, so verbleiben 104902 Reflexe, deren Vollständigkeit im Auflösungsbereich $\infty - 0,8 \text{ \AA}$ auf nur noch 86% sinkt, nach vorherigen 100%. Es fallen insbesondere schwache Reflexe in den höchsten Auflösungsschalen heraus – die Vollständigkeit sinkt dort sogar um 30% – weshalb der prozentuale Anteil von Beobachtungen $> 2\sigma$ ebenso zunimmt wie die durchschnittliche auf den Fehler normierte Intensität.

Die Verfeinerung der drei eingangs genannten Modelle gegen auf diese Weise beschnittene Daten resultiert in um ca. 2,7 Prozentpunkte verringerten konventionellen R -Werten. Wie Tab. 7.2 entnommen werden kann, verzeichnet die sphärische Verfeinerung eine stärkere Verbesserung als die asphärische Modellierung. Dadurch wird der zuvor gezeigte Vorsprung des Modells *mul* fast aufgehoben.

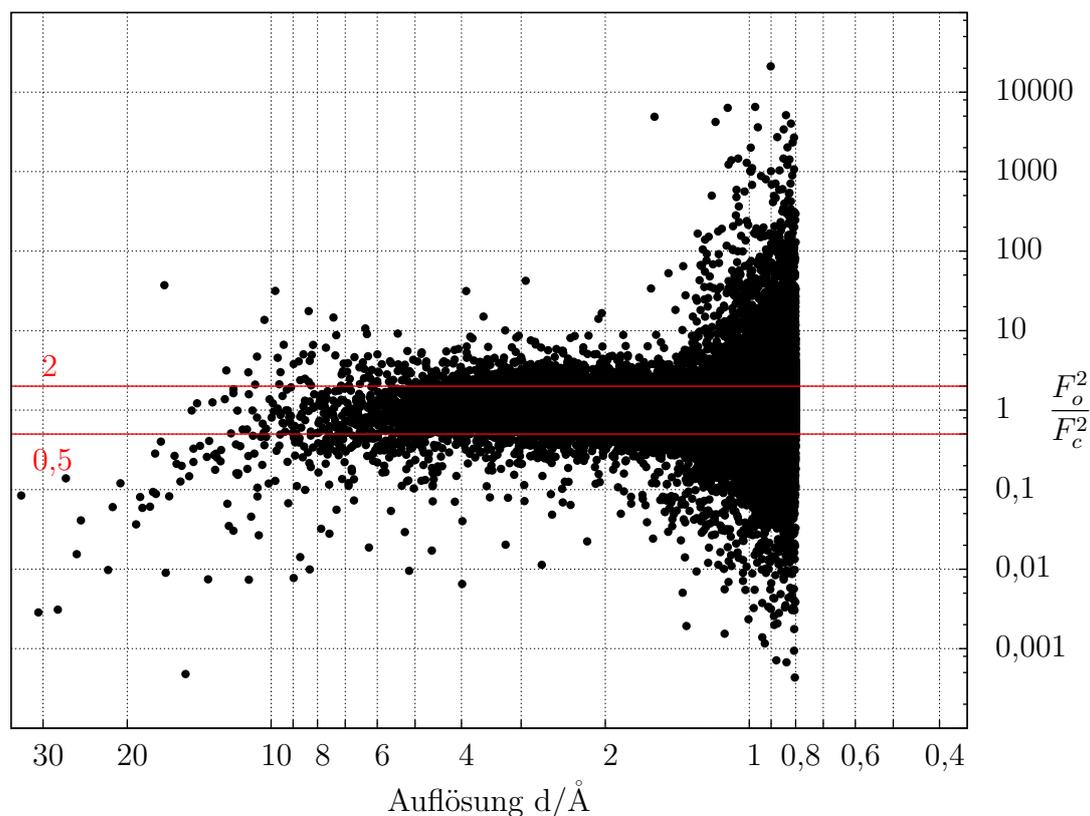


Abbildung 7.1: Quotient beobachteter und berechneter Intensitäten von TETRALYSO und ihre Auflösungsabhängigkeit im Modell *iam*.

7.2 Geometrie

Für den quadratisch gemittelten Unterschied aller Atompositionen des Lysozymmoleküls im Modell *mul*, ergibt sich gegenüber der Verfeinerung *iam* eine Abweichung von 0,095 Å. Im Vergleich zur Variante *iam_{rb}* beträgt der Unterschied nur 0,006 Å. Allerdings beeinhalteten diese vernachlässigbar kleinen Differenzen auch die fixierten Koordinaten.

Bezüglich einzelner Atomlagen belaufen sich die Positionsunterschiede der Nichtwasserstoffatome im Modell *iam* auf bis zu 0,031 Å, die der Wasserstoffatome betragen 0,083–0,190 Å. Letzteres ist den verlängerten XH-Abständen des Modells *mul* zuzuschreiben. In Abb. 7.2 auf der folgenden Seite sind diese Abweichungen der Koordinaten in den beiden Verfeinerungen farblich gekennzeichnet. Abgesehen von den Wasserstoffatomen, zeigen nur wenige randständige Atome eine merkliche Verschiebung.

Die Gegenüberstellung der Verfeinerungen *mul* und *iam_{rb}* zeigt für die Nicht-H-Positionen ganz ähnliche Unterschiede von bis zu 0,045 Å. Ob der in beiden Modellen

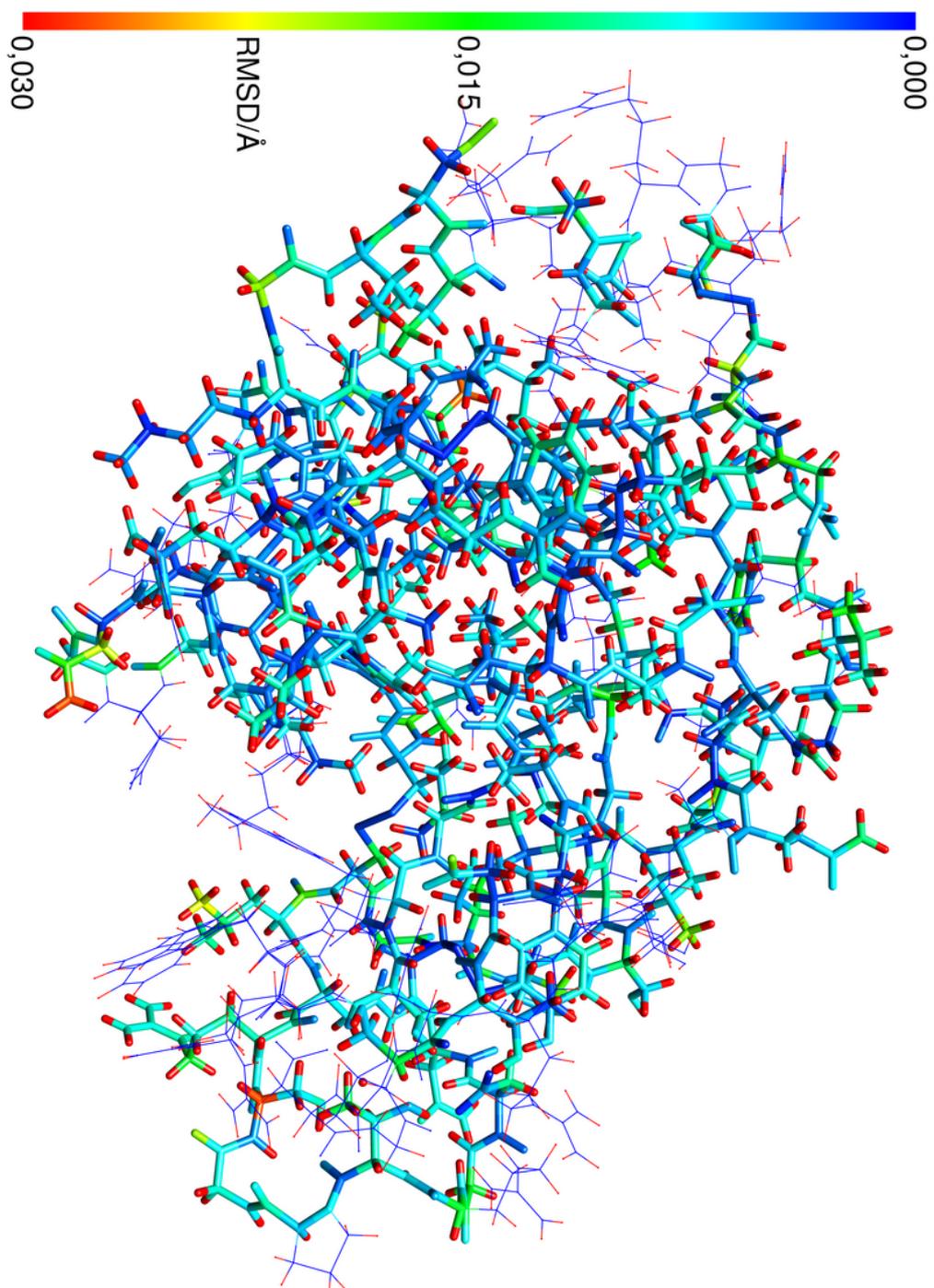


Abbildung 7.2: Quadratisch gemittelte Abweichung zwischen den Strukturen *ml* und *iam* von TETRALYSO. Verfeinerte Koordinaten hervorgehoben.

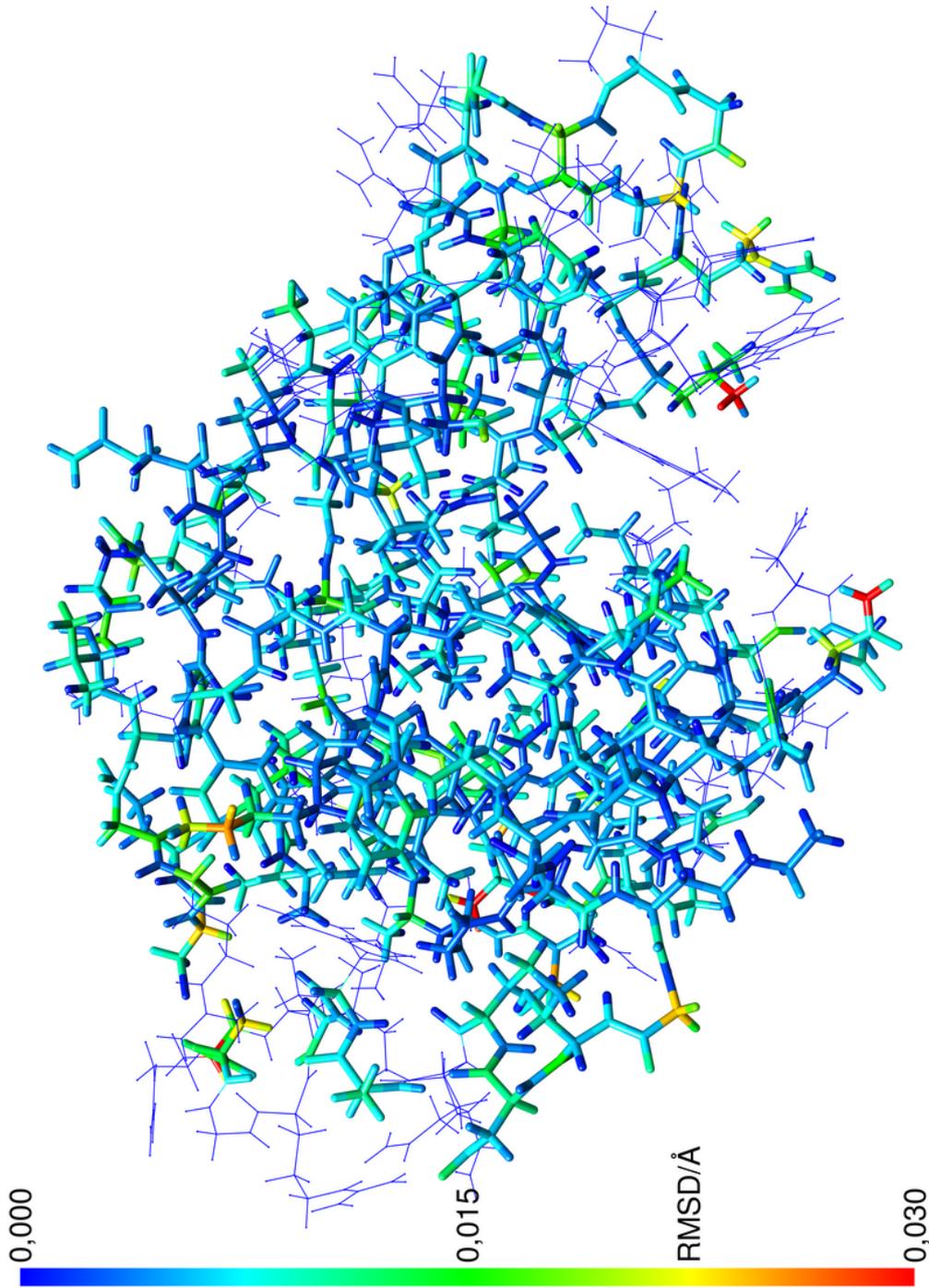


Abbildung 7.3: Quadratisch gemittelte Abweichung zwischen den Strukturen *mul* und *iam_{rb}* von TETRALYSO. Verfeinerte Koordinaten hervorgehoben.

Tabelle 7.3: Verfeinerte Bindungslängen der Peptidbindungen in der Hauptpopulation von TETRALYSO.

	$d_{E\&H}^*$	$d_{min/max}^\dagger$	$\bar{d}(\sigma_d)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–O ($n = 111$)								
<i>mul</i>		1,159/1,293	1,224(20)	1,214	1,223	1,234	0,019	0,014
<i>iam</i>	1,231(20)	1,166/1,302	1,230(21)	1,219	1,228	1,240	0,021	0,016
<i>iam_{rb}</i>		1,161/1,304	1,230(21)	1,219	1,228	1,241	0,022	0,015
C–N ($n = 115$)								
<i>mul</i>		1,242/1,419	1,336(22)	1,327	1,339	1,346	0,020	0,016
<i>iam</i>	1,329(14)	1,244/1,413	1,334(23)	1,324	1,336	1,347	0,023	0,018
<i>iam_{rb}</i>		1,248/1,420	1,336(23)	1,325	1,338	1,348	0,023	0,018
C_α–C ($n = 101$)								
<i>mul</i>		1,445/1,626	1,537(22)	1,528	1,536	1,547	0,018	0,013
<i>iam</i>	1,525(21)	1,410/1,612	1,529(24)	1,520	1,530	1,540	0,020	0,015
<i>iam_{rb}</i>		1,421/1,619	1,530(24)	1,519	1,531	1,542	0,023	0,017
C_α–N ($n = 98$)								
<i>mul</i>		1,374/1,515	1,454(21)	1,446	1,456	1,466	0,020	0,015
<i>iam</i>	1,458(19)	1,361/1,513	1,456(23)	1,449	1,458	1,468	0,020	0,015
<i>iam_{rb}</i>		1,368/1,521	1,459(23)	1,452	1,463	1,471	0,019	0,014

*Engh&Huber-Bindungslängen mit Standardabweichung; [†]Minimum, Maximum und Mittelwert mit Standardabweichung der Bindungslängen; [‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in Å.

verlängerten Wasserstoffabstände, liegen die Veränderungen der Wasserstofflagen nun bei maximal 0,033 Å. Abb. 7.3 auf der vorherigen Seite macht die geringen Unterschiede deutlich. Außer wenigen, wiederum randständigen Atomen zeigen die meisten Koordinaten den Farbcode für eine nur 0,01 Å betragende Abweichung.

Die Standardabweichung der Koordinaten liegt in ebendieser Größenordnung, weshalb nun die Auswirkungen auf die Bindungslängen anhand einer Auswahl von Peptidbindungen (siehe dazu auch Absch. 6.2 auf Seite 151) näher untersucht werden sollen. Wie die Tabellen 7.3 für diese Abstände und insbesondere 7.4 auf der nächsten Seite für deren Differenzen zeigen, liegen die Unterschiede zwischen den Verfeinerungsmodellen unter 0,01 Å. Allerdings sind einige Trends zu beobachten.

Zunächst ist festzustellen, daß das arithmetische Mittel der Bindungslängen des Modells *iam* um maximal 0,005 Å von den Engh&Huber-Referenzwerten verschieden ist. Im Modell *iam_{rb}* sind die Abweichungen marginal höher. Die asphärische Verfeinerung *mul* bewirkt nun etwas stärkere Unterschiede zu den Referenzmittelwerten und zwar jeweils in Richtung der bereits in den sphärischen Modellen bestehenden Abweichungen. Dadurch scheint diese Entwicklung durch reale Gegebenheiten in

Tabelle 7.4: Differenzen der verfeinerten Peptidbindungsängen der TETRALYSO-Hauptpopulation.

	$\Delta d_{min/max}^\dagger$	$\overline{\Delta d}(\sigma_\Delta)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–O ($n = 110$)							
<i>mul-iam</i>	–29,80/7,80	–5,85(5,09)	–8,25	–5,20	–2,90	5,35	4,00
<i>mul-iam_{rb}</i>	–40,40/2,80	–5,90(5,85)	–8,00	–4,70	–1,85	6,15	4,74
C–N ($n = 107$)							
<i>mul-iam</i>	–9,70/14,00	2,01(4,82)	–1,20	2,00	5,60	6,80	5,04
<i>mul-iam_{rb}</i>	–11,90/16,70	0,25(5,18)	–3,10	0,10	3,55	6,65	5,04
C_α–C ($n = 90$)							
<i>mul-iam</i>	–10,60/35,60	8,19(8,89)	2,30	7,60	13,50	11,20	8,01
<i>mul-iam_{rb}</i>	–12,20/31,50	7,16(8,65)	1,70	5,90	12,20	10,50	8,01
C_α–N ($n = 82$)							
<i>mul-iam</i>	–27,10/16,30	–1,93(6,38)	–4,88	–2,35	2,17	7,05	5,78
<i>mul-iam_{rb}</i>	–36,30/12,10	–5,08(6,46)	–7,97	–5,50	–2,12	5,85	4,52

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der Bindungslängendifferenzen; [‡]0,25-Quartil, Median, 0,75-Quartil; alle Einheiten in 10^{-3}Å .

dieser Struktur verursacht und nicht das Resultat zufälliger Schwankungen zu sein.

Insbesondere stimmen die Effekte durch die Multipolverfeinerung meist mit den für P1LYSO und APA gemachten Beobachtungen überein. So tritt auch hier eine Verkürzung der C–O-Bindung ein, die mit $-5 \cdot 10^{-3}\text{Å}$ auch zu den deutlichsten Modifikationen durch die Invariombeschreibung zählt. Die geringsten Unterschiede zwischen sphärischer und asphärischer Bindungslänge, zeigen die Bindungen zu Stickstoffatomen. Hierbei erfährt die C–N-Bindung eine Verlängerung um durchschnittlich $2 \cdot 10^{-3}\text{Å}$ beim Vergleich von *mul* und *iam*, jedoch fast gar keine Änderung beim Vergleich mit *iam_{rb}*, dem Modell, welches gleichfalls eine verlängerte Bindung zum amidischen Wasserstoffatom an diesem Stickstoff hat. Umgekehrt kommt es bei der C_α–N-Bindung zu einer Verkürzung selbiger, die gegenüber *iam_{rb}* mit $-5,5 \cdot 10^{-3}\text{Å}$ ausgeprägter, als die Abnahme um $-2,4 \cdot 10^{-3}\text{Å}$ verglichen mit *iam* ist. Auch dieser Befund ist im Einklang mit den vorherigen Untersuchungen. Die C_α–C-Bindung schließlich weist die deutlichsten Unterschiede zwischen asphärischem und sphärischem Modell auf. Sie verlängert sich im Modell *mul* um durchschnittlich 8 bzw. $6 \cdot 10^{-3}\text{Å}$ gegenüber den Werten aus *iam* und *iam_{rb}*.

Bei allen Bindungstypen nimmt in der asphärischen Verfeinerung die Streuung um das Mittel etwas ab. Die Spannweite zwischen Minimal- und Maximallängen einer Bindung liegt mit etwa $0,16\text{Å}$ aber nach wie vor in einem die Modellunterschiede bei weitem überschreitenden Bereich.

7.3 Verschiebungsparameter

Der Median der äquivalenten isotropen Verschiebungsparameter U_{eq} aller 2393 Atomlagen von TETRALYSO beträgt ca. $0,134 \text{ \AA}^2$, wobei die Extremwerte der Verteilung bei $0,06$ und 2 \AA^2 liegen. Für die 751 zur Verfeinerung ausgewählten Nichtwasserstoffatome, ergeben sich moderatere Werte (Tab. C.1 auf Seite 355 enthält Statistiken über weitere Untermengen dieser Struktur). Zwar beläuft sich das maximale U_{eq} auf immerhin noch $0,386$ und $0,380 \text{ \AA}^2$ in den sphärischen Modellen und $0,370 \text{ \AA}^2$ nach der Verfeinerung *mul*, aber der Median von *iam* und *iam_{rb}* liegt bei nur noch $97 \cdot 10^{-3} \text{ \AA}^2$, der des Modells *mul* bei etwas niedrigeren $96 \cdot 10^{-3} \text{ \AA}^2$. Ebenfalls geringer in *mul*, ist das Minimal- U_{eq} $58 \cdot 10^{-3} \text{ \AA}^2$, verglichen mit $59 \cdot 10^{-3} \text{ \AA}^2$ in den sphärischen Modellen. In Abb. 7.6 auf Seite 198 sind die anisotropen Verschiebungsparameter dieser Atome illustriert. Zusätzlich tragen sie einen dem jeweiligen U_{eq} -Wert entsprechenden Farbcode. Die Ellipsoide aller Nichtwasserstoffatome zeigt Abb. C.1 auf Seite 356. Mit diesen Abbildungen, wird die Auswahl verfeinerter Atome graphisch wiedergegeben, sowie hervorgehoben, daß primär Strukturteile mit zu hohen Verschiebungsparametern fixiert wurden. Durch die Statistik über die Differenzen $\Delta U_{\text{eq}} = U_{\text{eq}}(\textit{mul}) - U_{\text{eq}}(\textit{iam})$, bzw. $U_{\text{eq}}(\textit{mul}) - U_{\text{eq}}(\textit{iam}_{\text{rb}})$, wird die Verbesserung der Temperaturfaktoren bekräftigt: Die Unterschiede zur klassischen sphärischen Verfeinerung weisen einen Median von $-1,18 \cdot 10^{-3} \text{ \AA}^2$ auf, gegenüber dem Modell *iam_{rb}* von $-1,33 \cdot 10^{-3} \text{ \AA}^2$. Dabei liegen die Quartilsgrenzen komplett im negativen Bereich, mit einem recht schmalen Abstand von $\approx 1 \cdot 10^{-3} \text{ \AA}^2$.

Diesen Sachverhalt verbildlichen die Tafeln 7.4 auf der nächsten Seite und 7.5 auf Seite 196, durch Auftragung der isotropen Temperaturfaktoren aller Nichtwasserstoffatome des Lysozymmoleküls mit den jeweiligen Unterschieden zwischen *mul* und *iam*. Hierbei wird offenbar, daß 665 oder 88,5% aller verfeinerten Atome nach asphärischer Behandlung ein kleineres U_{eq} aufweisen, während für nur 86 (11,5%) Verschiebungsparameter eine Verschlechterung eintritt. Letzteres ist bei höheren U_{eq} häufiger als bei niedrigen. Je größer der Temperaturfaktor bereits nach sphärischer Verfeinerung war, desto drastischer sind die auftretenden Veränderungen. Sowohl die stärkste Verkleinerung von $-9,5 \cdot 10^{-3} \text{ \AA}^2$ als auch die größte Zunahme eines U_{eq} -Wertes von $10,5 \cdot 10^{-3} \text{ \AA}^2$ korrespondieren mit hohen Ausgangstemperaturfaktoren. Beim Vergleich der Modelle *mul* und *iam_{rb}* liegen die Extrema der U_{eq} -Differenzen doppelt so hoch.

Eine Aufschlüsselung nach Elementen bietet Tab. C.2 auf Seite 357. Den kleinsten durchschnittlichen Temperaturfaktor von im Median $89 \cdot 10^{-3} \text{ \AA}^2$ haben Schwefelatome, welcher sich mit *mul* um $-2 \cdot 10^{-3} \text{ \AA}^2$ am stärksten verbessert. Mit $-1,4$ und $-1,6 \cdot 10^{-3} \text{ \AA}^2$ ist die Abnahme für Kohlenstoffatomen in den Modellen *iam* und *iam_{rb}* geringer ausgeprägt, deren mittleres U_{eq} bei $96 \cdot 10^{-3} \text{ \AA}^2$ liegt. Die nächstniedrigere Verbesserung von, je nach Modell, $-1,0$ und $-1,3 \cdot 10^{-3} \text{ \AA}^2$

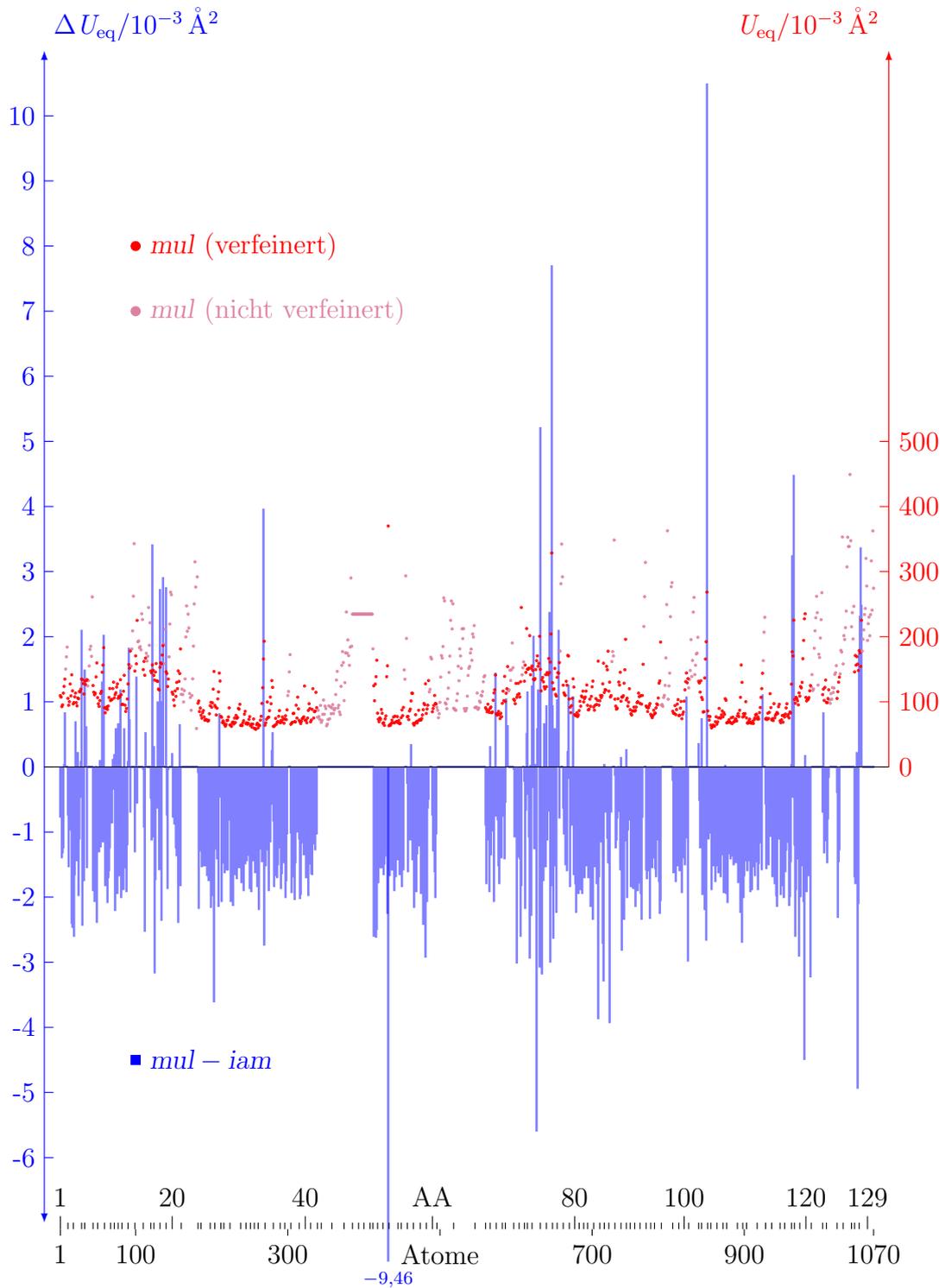


Abbildung 7.4: Veränderung der U_{eq} -Werte im asphärischen Modell mul gegenüber dem sphärischen Modell iam für die Nichtwasserstoffatome von TETRALYSO.

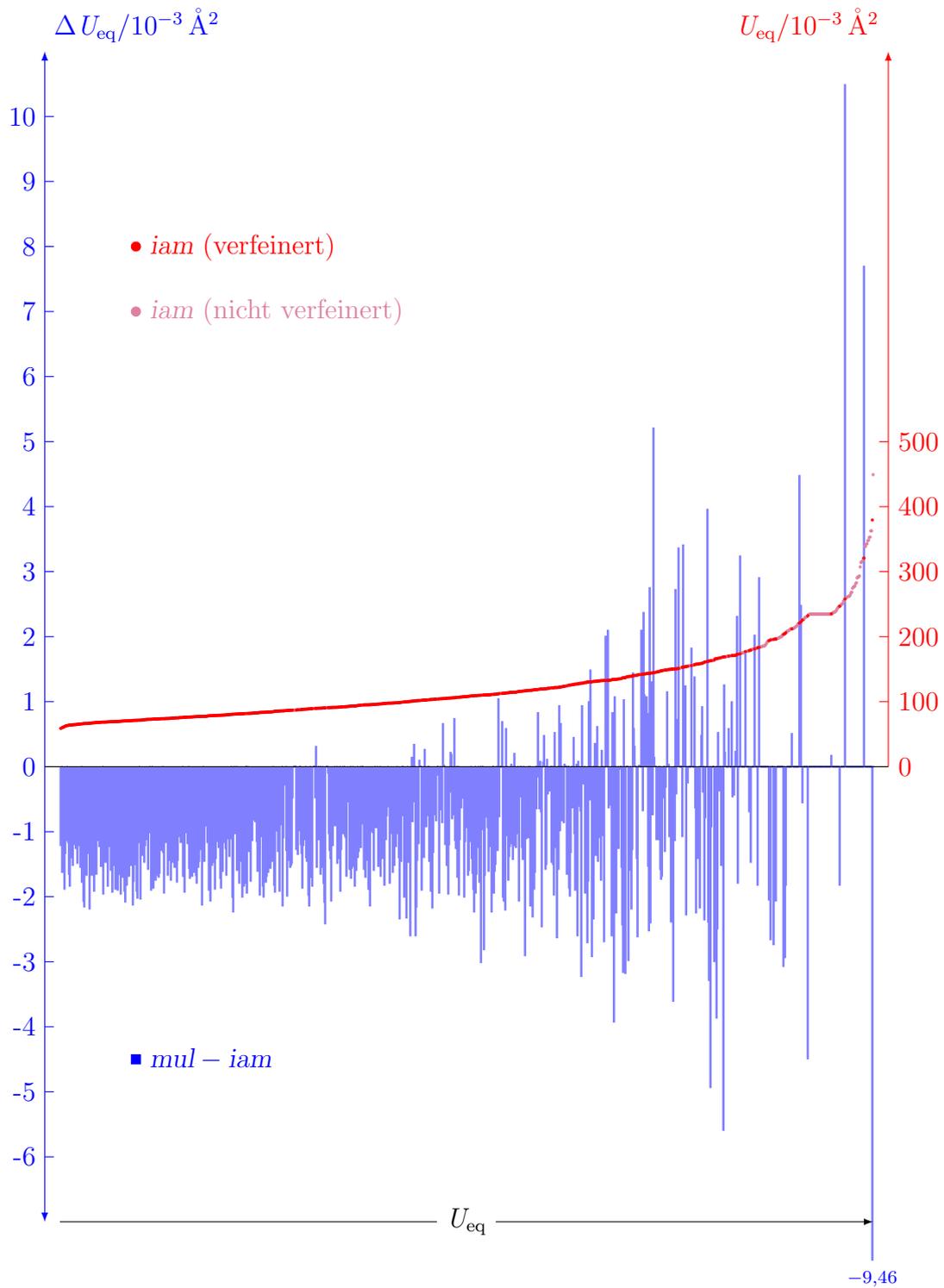


Abbildung 7.5: Veränderung der U_{eq} -Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die Nichtwasserstoffatome von TETRALYSO. Auftragung nach U_{eq} (*iam*).

Tabelle 7.5: Statistik der U_{eq} der verfeinerten Peptidbindungsatome von TETRALYSO.

	$x_{\min/\max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C_α (n = 101)							
<i>mul</i>	58,54/175,20	97,08(28,13)	74,77	91,57	111,60	36,86	27,12
<i>iam</i>	60,17/172,90	98,56(27,71)	76,34	93,34	112,40	36,08	27,09
<i>iam_{rb}</i>	60,43/173,90	98,78(27,58)	76,65	93,54	112,50	35,88	27,07
<i>mul-iam</i>	-3,93/2,32	-1,48(0,92)	-1,93	-1,66	-1,27	0,66	0,44
<i>mul-iam_{rb}</i>	-3,87/2,06	-1,70(1,12)	-2,21	-1,99	-1,44	0,77	0,42
C (n = 111)							
<i>mul</i>	57,73/170,40	97,07(27,92)	73,39	91,70	112,40	38,98	28,20
<i>iam</i>	58,95/170,90	98,12(27,54)	74,74	93,28	113,10	38,36	28,44
<i>iam_{rb}</i>	59,22/169,20	98,30(27,55)	74,79	93,74	113,40	38,62	29,00
<i>mul-iam</i>	-3,19/3,41	-1,05(1,00)	-1,57	-1,21	-0,84	0,73	0,54
<i>mul-iam_{rb}</i>	-4,47/4,72	-1,23(1,09)	-1,77	-1,33	-1,10	0,67	0,55
N (n = 107)							
<i>mul</i>	59,40/178,70	96,56(27,13)	74,46	93,48	107,80	33,31	26,44
<i>iam</i>	60,67/177,00	96,89(26,54)	75,00	95,06	107,90	32,92	27,20
<i>iam_{rb}</i>	60,92/177,00	96,89(26,37)	75,34	94,80	107,90	32,53	26,36
<i>mul-iam</i>	-1,91/5,22	-0,32(0,99)	-0,84	-0,59	-0,12	0,73	0,54
<i>mul-iam_{rb}</i>	-2,21/4,33	-0,33(1,14)	-1,07	-0,66	0,00	1,07	0,68
O (n = 111)							
<i>mul</i>	64,32/244,90	110,80(38,86)	83,34	99,77	131,10	47,74	30,90
<i>iam</i>	65,54/246,70	111,50(38,46)	84,36	100,20	131,60	47,23	30,00
<i>iam_{rb}</i>	65,69/255,30	111,90(38,81)	84,56	100,50	132,50	47,91	29,96
<i>mul-iam</i>	-2,94/4,48	-0,73(0,89)	-1,24	-0,90	-0,49	0,75	0,57
<i>mul-iam_{rb}</i>	-10,39/8,37	-1,12(1,45)	-1,45	-1,20	-0,85	0,60	0,48

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der U_{eq} und ΔU_{eq} ;

[‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-3}Å^2 .

erfahren Sauerstoffatome, ihr mittleres U_{eq} beträgt ca. $103 \cdot 10^{-3}\text{Å}^2$. Am geringsten ist die Veränderung bei Stickstoffatomen, deren U_{eq} -Mittel sich, ausgehend von $95 \cdot 10^{-3}\text{Å}^2$, um etwa $-0,65 \cdot 10^{-3}\text{Å}^2$ verringert.

Anhand Tab. 7.5 ist für die nach bekanntem Schema ausgewählten Atome der Peptidbindung ein ähnliches Muster erkennbar – ein geringeres sphärisches U_{eq} bedingt eine stärkere Verbesserung, welche bei Stickstoffatomen aber unerwartet gering ist. Im asphärischen Modell erhöht sich die Streuung der U_{eq} -Verteilung aller Atome leicht, bei C_α und N geht dies mit größeren Maximalwerten einher. Die Schwankungsbreite der Verbesserungen ist bei jedem Atomtyp vergleichbar.

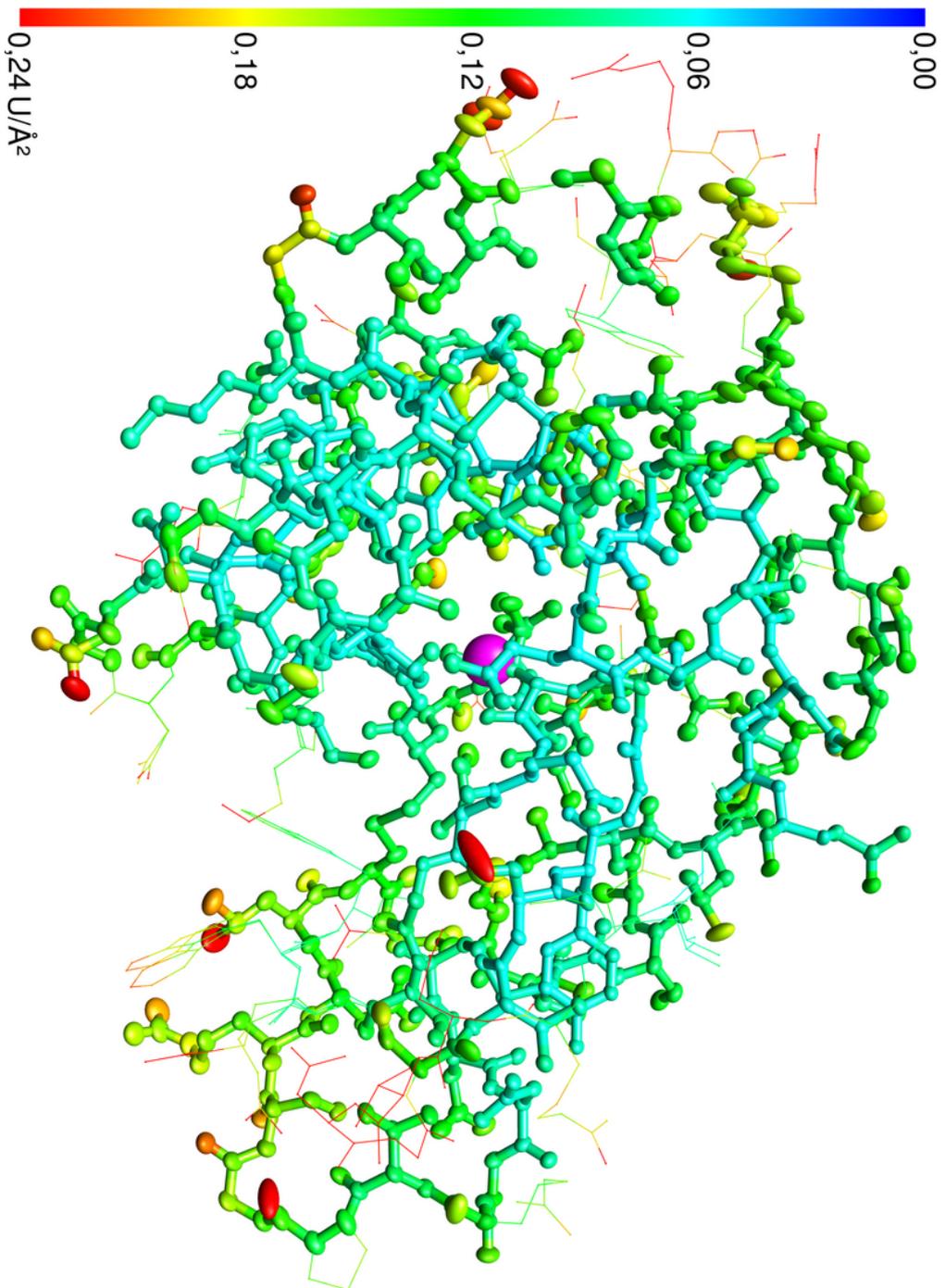


Abbildung 7.6: Anisotrope Verschiebungsparameter des verfeinerten Teils von TETRALYSSO im Modell *ml*.
Farbcode: U_{eq} . Darstellung des massengewichteten Zentrums als magentafarbene Kugel.

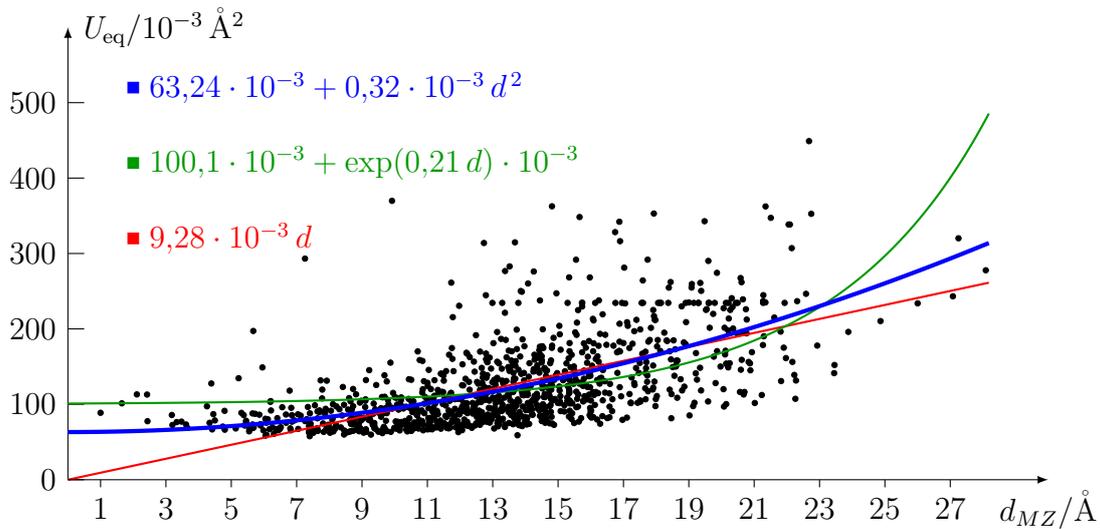


Abbildung 7.7: U_{eq} (*mul*) der Nichtwasserstoffatome von TETRALYSO in Abhängigkeit vom Abstand zum Massenzentrum.

Die analog Gl. 6.1 auf Seite 159 berechneten Koordinaten des massengewichteten Zentrums liegen bei $(58,42; 78,51; 28,53)$ Å, in Abb. 7.6 ist es als magentafarbene Kugel symbolisiert. Wie der Differenzvektor aller 1070 Nichtwasserstoffatome nach Gl. 6.2 angibt, befinden sie sich in Abständen von 1 bis 28 Å dazu. Die U_{eq} dieser Atome sind in Abhängigkeit von dieser Distanz in Abb. 7.7 aufgetragen.

Offenbar nehmen die Verschiebungsparameter mit der Entfernung vom Massenzentrum zu. Drei Typen von Ausgleichsfunktion sollen das Verhalten der Verteilung approximieren. Die Unsicherheit der Koeffizienten der linearen (rot) und exponentiellen (grün) Funktion beträgt dabei ca. 1-2%, die der quadratischen (blau) 3-4%. Wie schon im vorherigen Kapitel (siehe Seite 162) ergibt sich der Eindruck, daß die Beschreibung durch eine quadratisch mit dem Abstand d_{MZ} vom Schwerpunkt steigende Funktion das Verhalten der U_{eq} am besten widerspiegelt. Die Steigung ist mit $0,64 \cdot 10^{-3} d \text{ Å}$ ein Drittel höher als im Fall des triklinen Lysozyms, beginnend bei einem extrapolierten Temperaturfaktor im Zentrum des Moleküls von $63 \cdot 10^{-3} \text{ Å}^2$. Damit werden die Daten besser wiedergegeben als durch die Exponentialfunktion, welche einen zu starken Anstieg bei großem Abstand und zu geringen nahe des Zentrums beschreibt. Problematisch an der linearen Beschreibung ist die Extrapolation auf Null im Zentrum der Struktur. Passt man stattdessen eine lineare Funktion $a + bx$ an die Daten an, resultiert $11,70(4,85) \cdot 10^{-3} + 8,48(0,35) \cdot 10^{-3} d$. Die hohen Parameterunsicherheiten von $\pm 42\%$ und $\pm 4\%$ in Kombination mit dem ohnehin niedrigen Parameter a führten zu dem Entschluß, diesen aus der Regression auszuschließen.

7.4 DMSDA-Werte

Zwecks weiterer Untersuchung der Temperaturfaktoren zeigen Abb. 7.8 auf der nächsten Seite und Abb. 7.9 auf Seite 202 die DMSDA-Beträge der 1102 Nichtwasserstoffatombindungen in den Modellen *mul* und *iam*, sowie die Veränderung durch die unterschiedliche Verfeinerung von 795 Bindungen. Sowohl im sphärischen als auch im asphärischen Modell treten sehr niedrige und sehr hohe Werte auf, die stark schwankend über die gesamte Sequenz verteilt sind. Um einen in allen Verfeinerungen ähnlichen Zentralwert von $\approx 41 \cdot 10^{-4} \text{ \AA}^2$ gibt es Beträge von bis zu $973 \cdot 10^{-4} \text{ \AA}^2$ im Modell *iam* und $941 \cdot 10^{-4} \text{ \AA}^2$ im Fall *mul*.

Gleiches gilt für die Differenzen $\Delta|\text{DMSDA}| = |\text{DMSDA}|(\textit{mul}) - |\text{DMSDA}|(\textit{iam})$, nur bei ungefähr einem Viertel der vorherigen Größenordnung. Hier liegt das Mittel über alle Bindungen nahe Null, die stärksten Veränderungen zu *iam* betragen $\pm 146 \cdot 10^{-4} \text{ \AA}^2$, beim Vergleich mit *iam_{rb}* sogar -197 und $166 \cdot 10^{-4} \text{ \AA}^2$. Tendentiell verschlechtern sich niedrige und hohe DMSDA-Werte des sphärischen Modells, während mittlere eher eine Verbesserung erfahren. Zudem ist das Ausmaß der Unterschiede bei hohen Ausgangswerten ebenfalls besonders hoch.

Wie Tab. C.3 auf Seite 358 detailliert, weisen die frei gegebenen Bindungen einen Median von 44, 42 und $40 \cdot 10^{-4} \text{ \AA}^2$ für die DMSDA-Beträge in den Modellen *iam_{rb}*, *iam* und *mul* auf. Die 795 Differenzen haben einen Zentralwert von $-1 \cdot 10^{-4} \text{ \AA}^2$, sowohl für die Unterschiede zum Modell *iam* als auch zu *iam_{rb}*. Dazu haben die nach Elementen diskriminierten fünf Bindungstypen unterschiedliche Beiträge: Auf den vier Disulfidbrücken nehmen die Extremwerte im Modell *mul* ab, womit ein etwa $2 \cdot 10^{-4} \text{ \AA}^2$ geringerer Median von $42 \cdot 10^{-4} \text{ \AA}^2$ verbunden ist. Hingegen kommt es auf den zwölf C–S-Bindungen zu einer durchschnittlichen Zunahme des DMSDA-Betrages um $6 \cdot 10^{-4} \text{ \AA}^2$ auf $36 \cdot 10^{-4} \text{ \AA}^2$. Die 148 C–O-Bindungen verbessern ihren Schnitt in *mul* deutlich, von ca 59 auf nun $47 \cdot 10^{-4} \text{ \AA}^2$. Eine weniger ausgeprägte Verringerung, von ca. 41 auf $37 \cdot 10^{-4} \text{ \AA}^2$, ist für 258 C–N-Verknüpfungen zu beobachten. Fast unverändert zeigen sich die 373 C–C-Bindungen, deren asphärischer Median bei 40, der des Modells *iam* bei 39 und im Fall *iam_{rb}* bei $42 \cdot 10^{-4} \text{ \AA}^2$ liegt.

Ein davon abweichendes Bild ergibt die Untersuchung der insgesamt 425 Peptidbindungen des Rückgrats der Hauptpopulation von TETRALYSO, exklusive terminaler Reste und Glycin, sowie Prolin, deren Statistik in Tab. 7.6 auf Seite 203 aufgeführt ist. Die stärkste Verbesserung erfährt im Median die C–N-Bindung, von 40 auf $35 \cdot 10^{-4} \text{ \AA}^2$. Auch auf den C_α–N-Bindungen sinken im Durchschnitt die DMSDA-Beträge, von 37,5 in den sphärischen Modellen auf nur noch $32 \cdot 10^{-4} \text{ \AA}^2$ im Modell *mul*. Allerdings nehmen die Beträge auf einigen Bindungen stark zu, das Mittel der Differenzen ist deshalb nahe bei Null. In geringerem Ausmaß verbessert sich auch die C–O-Bindung, deren ohnehin höheres sphärisches Mittel von $56\text{--}59 \cdot 10^{-4} \text{ \AA}^2$ auf nur $53 \cdot 10^{-4} \text{ \AA}^2$ sinkt. Andererseits sind die Extremwerte der

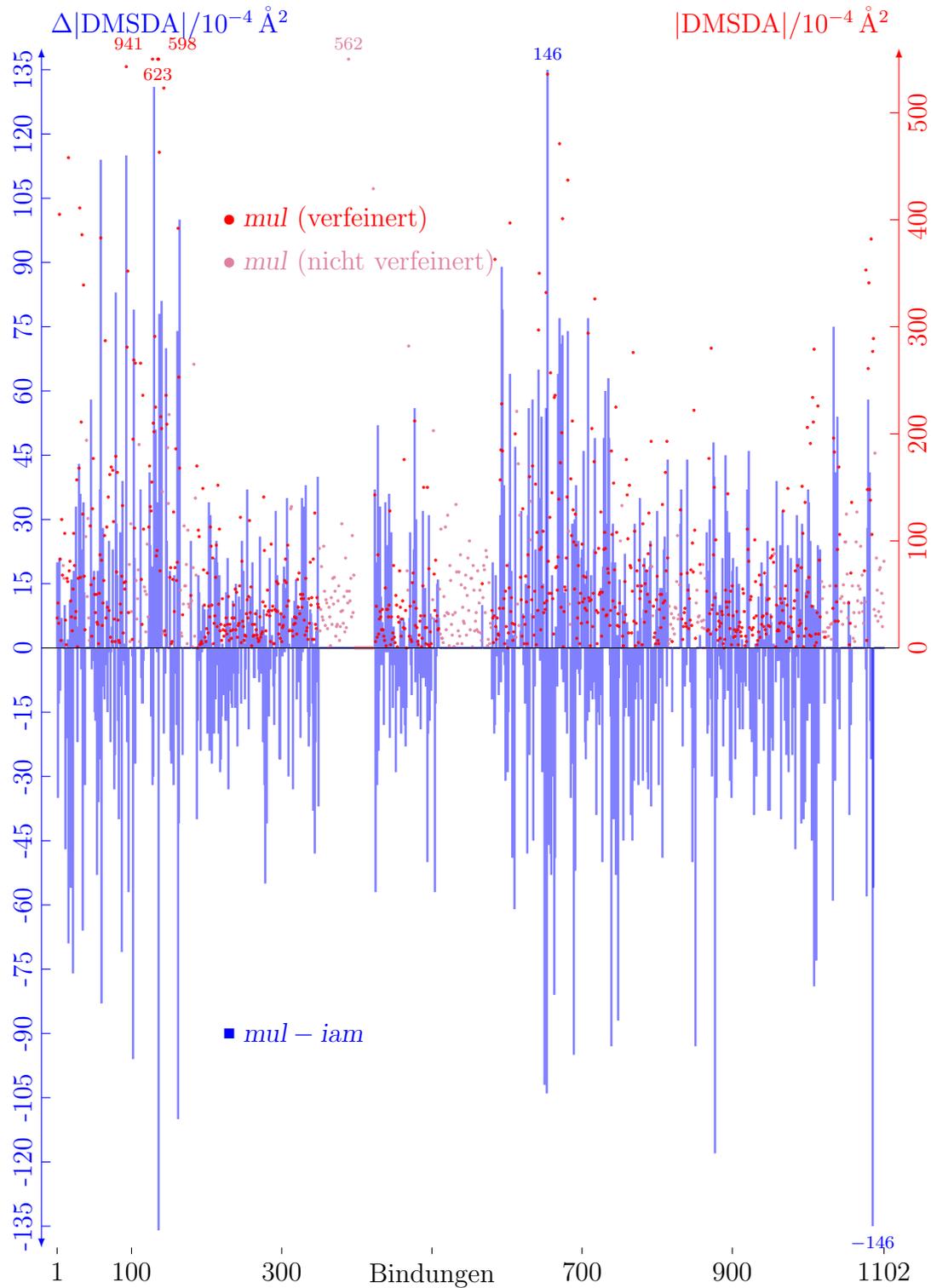


Abbildung 7.8: Veränderung der DMSDA-Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die Nichtwasserstoffatome von TETRALYSO.

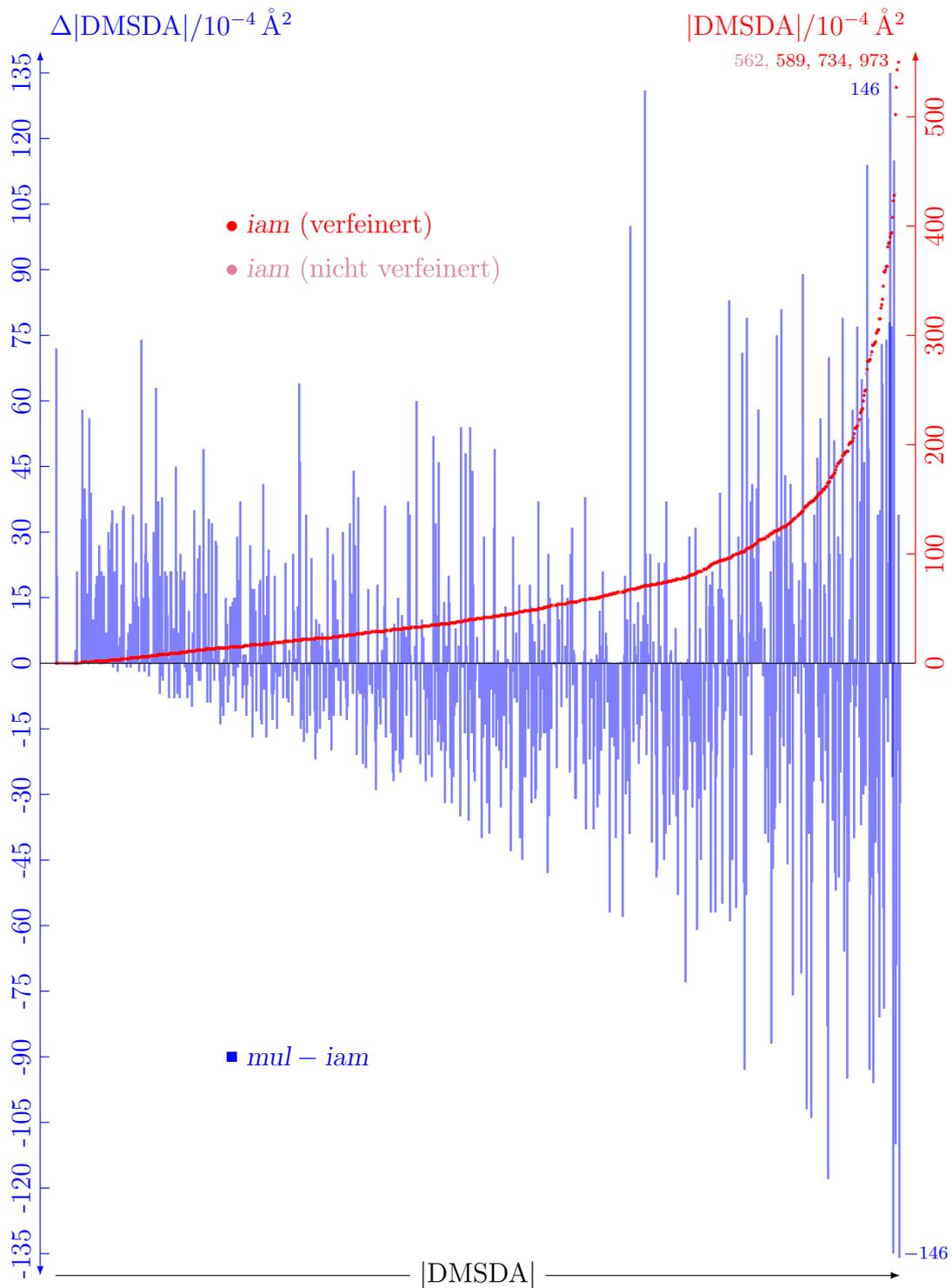


Abbildung 7.9: Veränderung der DMSDA-Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die Nichtwasserstoffatome von TETRALYSO. Auftragung nach steigenden $|DMSDA|$ -Werten des Modells *iam*.

Tabelle 7.6: DMSDA-Beträge der verfeinerten Peptidbindungen der TETRALYSO-Hauptpopulation und deren Differenzen.

	$x_{min/max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–O ($n = 111$)							
<i>mul</i>	0/437	88,61(94,95)	20,00	53,00	115,00	95,00	54,86
<i>iam</i>	0/381	90,27(91,09)	26,00	56,00	125,50	99,50	54,86
<i>iam_{rb}</i>	0/396	90,26(91,54)	24,50	58,00	121,50	97,00	57,82
<i>mul-iam</i>	-95/74	-1,66(25,11)	-14,00	-1,00	9,00	23,00	16,31
<i>mul-iam_{rb}</i>	-107/95	-1,65(29,09)	-12,00	-2,00	10,00	22,00	17,79
C–N ($n = 115$)							
<i>mul</i>	4/405	64,02(73,56)	20,50	35,00	74,00	53,50	31,13
<i>iam</i>	0/423	65,65(76,00)	21,00	40,00	74,50	53,50	38,55
<i>iam_{rb}</i>	0/444	67,02(78,47)	19,00	40,00	78,50	59,50	38,55
<i>mul-iam</i>	-146/131	-1,64(32,01)	-20,00	-5,00	17,50	37,50	26,69
<i>mul-iam_{rb}</i>	-167/131	-3,00(36,93)	-25,50	-6,00	17,00	42,50	31,13
C_α–C ($n = 101$)							
<i>mul</i>	2/941	82,67(128,97)	21,00	44,00	88,00	67,00	44,48
<i>iam</i>	1/973	84,47(136,69)	14,00	42,00	100,00	86,00	48,93
<i>iam_{rb}</i>	0/947	86,34(142,73)	11,00	43,00	92,00	81,00	53,37
<i>mul-iam</i>	-136/115	-1,79(41,84)	-33,00	-1,00	30,00	63,00	47,44
<i>mul-iam_{rb}</i>	-197/110	-3,66(50,83)	-35,00	0,00	31,00	66,00	50,41
C_α–N ($n = 98$)							
<i>mul</i>	1/623	75,28(121,25)	12,25	32,00	66,00	53,75	34,84
<i>iam</i>	1/589	72,84(110,10)	15,00	37,50	71,50	56,50	36,32
<i>iam_{rb}</i>	0/629	75,13(112,90)	16,25	37,50	71,50	55,25	35,58
<i>mul-iam</i>	-59/146	2,44(25,05)	-10,75	1,00	10,00	20,75	14,83
<i>mul-iam_{rb}</i>	-66/159	0,14(26,18)	-10,00	-2,00	7,00	17,00	12,60

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der DMSDA-Beträge;
[‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-4}Å^2 .

Differenzen weniger stark ausgeprägt. Die homoatomare C_α–C-Bindung schließlich zeigt einen leicht steigenden DMSDA-Mittelwert im asphärischen Modell. Lag dieser im Modell *iam* noch bei $42 \cdot 10^{-4} \text{Å}^2$, sind es nach *mul* $44 \cdot 10^{-4} \text{Å}^2$. Hierbei weisen die recht hohen Extremwerte der Differenzen und z. B. der besonders hohe Interquartilabstand auf ausgeprägte Änderungen der DMSDA-Beträge dieses Bindungstyps hin, wiewohl immerhin der Maximalbetrag mit der Verfeinerung *mul* abnimmt. Auf allen vier Bindungen nimmt die Verteilungsbreite der DMSDA-Beträge im asphärischen Modell etwas ab.

Tabelle 7.7: Statistik der Restdichten der asymmetrischen Einheit von TETRALYSO.

	$x_{min/max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
<i>mul</i>	-903,3/1560,0	-0,1(150,9)	-109,4	-17,8	96,8	206,2	149,9
<i>iam</i>	-915,7/1479,0	-0,1(154,0)	-112,0	-18,7	99,2	211,2	153,2
<i>iam_{rb}</i>	-902,1/1475,0	-0,1(154,9)	-112,9	-18,0	100,6	213,5	155,2

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der Restdichte;

[‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in $10^{-3} \text{ e}/\text{\AA}^3$.

7.5 Restdichten

Zur Auswertung der Restdichten in den drei Modellen *mul*, *iam* und *iam_{rb}*, wurde der Term $\sum(F_o - F_c) \exp(i\phi_c)$ auf einem die asymmetrische Einheit umfassenden Gitter von $300 \times 300 \times 136 = 1620000$ Punkten berechnet, das Volumen eines Voxels ist somit $0,0188 \text{\AA}^3$. Die gesamte Restelextronendichte beläuft sich im Betrag $\sum |\Delta\rho_{res}|V_{vox}$ auf 3688 e im Modell *mul* und auf 3768 e sowie 3797 e in den sphärischen Modellen *iam* und *iam_{rb}*. Zur besseren Vergleichbarkeit mit den restlichen Strukturen auf das Volumen der asymmetrischen Einheit normiert, lauten diese Werte 0,1284, 0,1312 und $0,1322 \text{ e}\text{\AA}^{-3}$. Somit reduziert sich die Gesamtrestdichte in der asphärischen Verfeinerung um 2 bzw. 3 % gegenüber den Modellen *iam* und *iam_{rb}*.

Auch die Statistik über die Restdichten der asymmetrischen Einheit in Tab. 7.7 läßt die, wenn auch geringe, Abnahme der Restdichten erkennen. Hierbei wird deutlich, daß im Modell *mul* von TETRALYSO insbesondere die negativen Restdichten gemildert werden: Der Mittelwert und die Lage der Quartilsgrenzen ist zu positiveren und näher bei Null liegenden Werten verschoben. Anhand der verminderten Indikatoren für die Streuung der Restdichten, wird zudem eine engere Verteilung um den Nullpunkt nahegelegt, bei zugleich erhöhtem Maximum.

Gleiches wird auch in den Histogrammen 7.10(a) und 7.10(b) auf der gegenüberliegenden Seite deutlich. Zwar gibt es in *mul* einige wenige Gitterpunkte zu höheren Extrema, andererseits liegen in diesem Modell 5,0 % im optimalen Intervall $(-0,01; 0,01]$, während die sphärischen Verfeinerungen Werte von 4,9 % und 4,8 % zeigen. Das Verteilungsmaximum erreicht das asphärische Modell mit 2,81 % im Intervall $(-0,06; -0,05]$, das von *iam* liegt bei 2,75 % in $(-0,07; -0,06]$ und mit 2,69 % bei $(-0,08; -0,07]$ hat *iam_{rb}* die schlechtesten Werte. Die geringen Unterschiede der Modelle lassen die Verteilungen fast genau übereinander liegen.

Die Statistiken ergänzend, geben die Abb. 7.11 bis 7.15 auf Seite 206–210 eine Illustration der Restdichte und Restdichtedifferenzen als Isooberflächen zu mehreren Werten, analog denen in Absch. 6.5 auf Seite 173 für P1LYSO gezeigten und erläuterten.

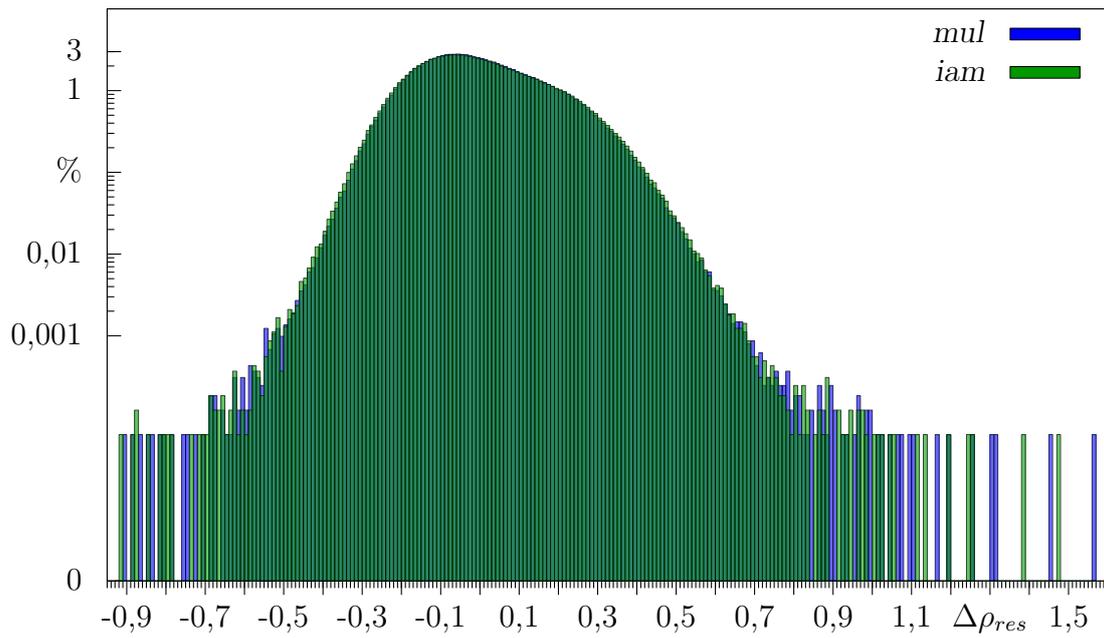
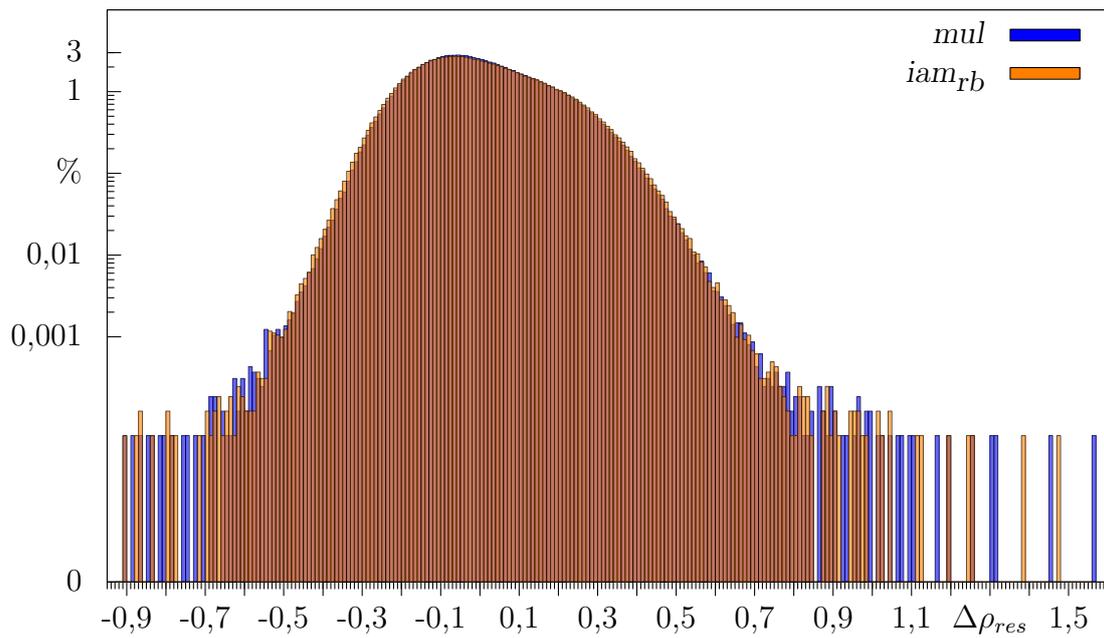
(a) Restdichten TETRALYSO, *mul* vs *iam*.(b) Restdichten TETRALYSO, *mul* vs *iam_{rb}*.

Abbildung 7.10: Histogramm der Restdichtegitter $\sum(F_o - F_c) \exp(i\phi_c)$ der sphärischen (grün/orange) und asphärischen (blau) Modelle. Intervalle $\Delta(i; i+0,01]$, Einheiten $\text{e}\text{\AA}^{-3}$.

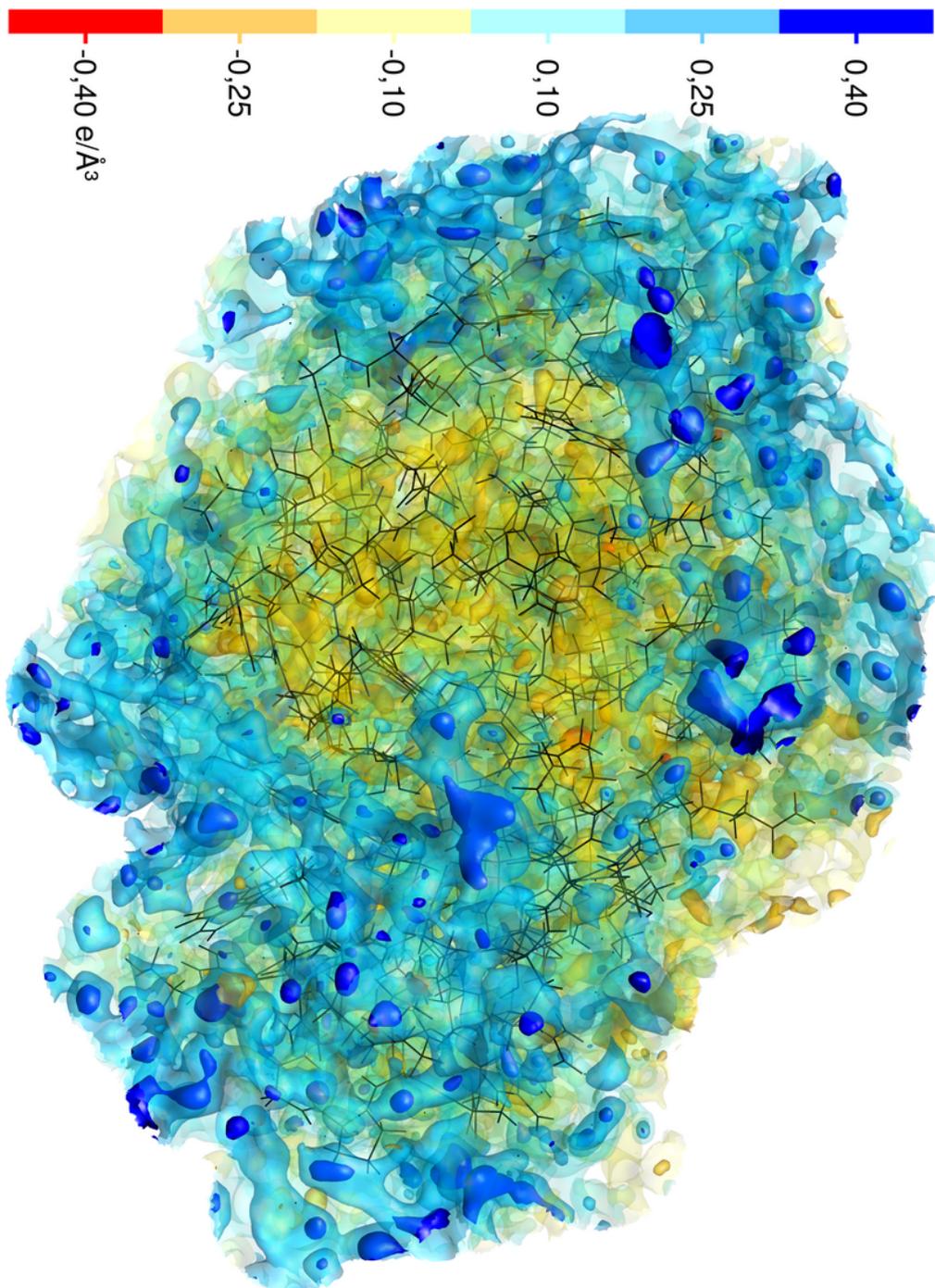


Abbildung 7.11: Restdichten von TETRALYSO im Modell *mul*, Darstellung bei sechs Isooberflächenwerten bis zu 4 Å um die Struktur.

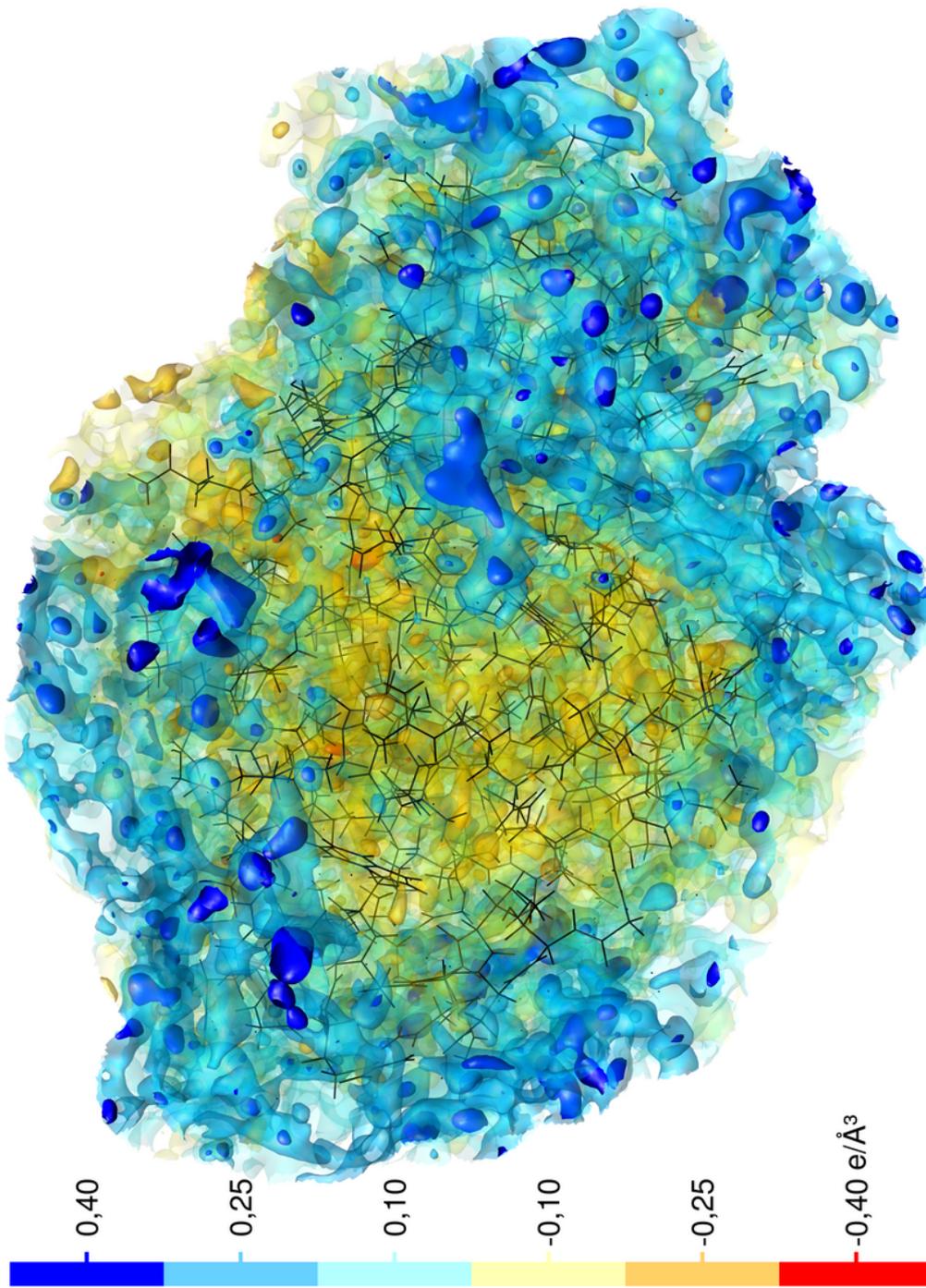


Abbildung 7.12: Restdichten von TETRALYSO im Modell iam, Darstellung bei sechs Isooberflächenwerten bis zu 4 Å um die Struktur.

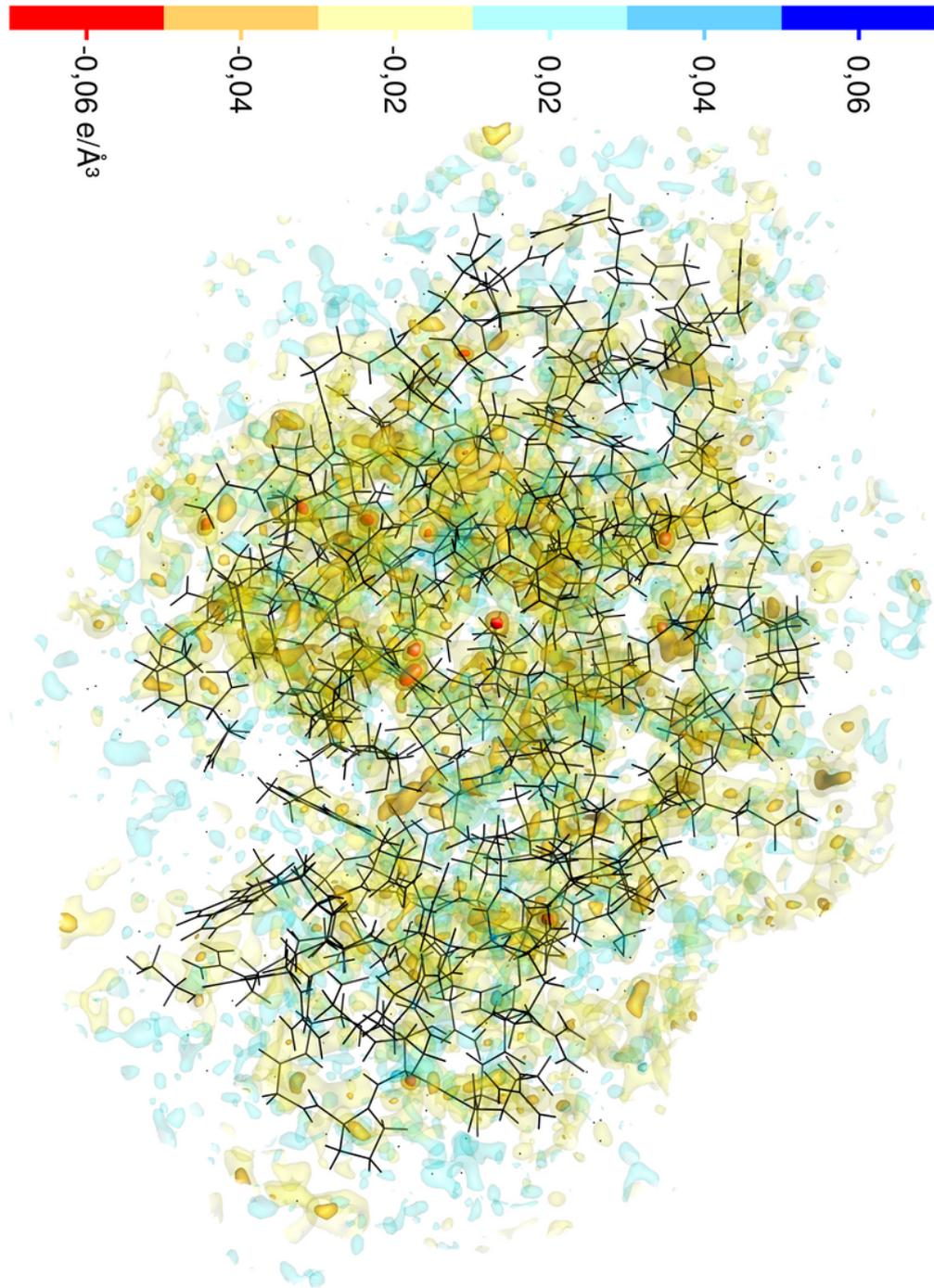


Abbildung 7.13: Restdichtedifferenz iam-mnl 4 Å um TETRALYSO.

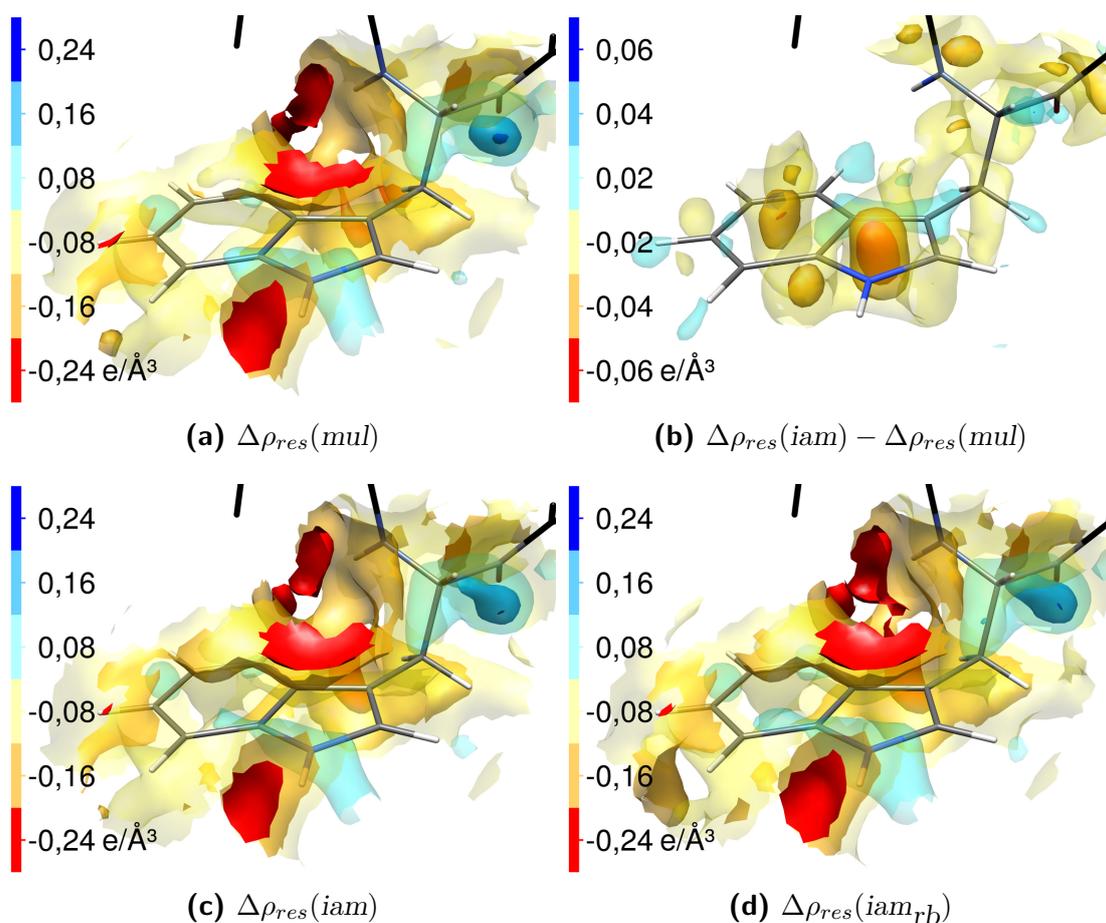


Abbildung 7.14: Restdichten und Restdichtedifferenzen bis $1,4 \text{ \AA}$ um den TETRALYSO-Tryptophanrest 108 für sphärische und asphärische Verfeinerungen.

Die Überblicksdarstellungen Abb. 7.11 und 7.12 zeigen die Modelle *mul* und *iam*, eine fast identische Darstellung für *iam_{fb}* ist in Abb. C.2 auf Seite 359 zu finden. Nur geringe Akkumulationen positiver Restdichten sind in den sphärischen Modellen auf den Atomen oder Bindungen von TETRALYSO zu finden, sie sind eher um das Molekül, im Solvensbereich, lokalisiert. Zumeist sind die Atomlagen von negativer Restdichte umhüllt. Der Einfluß des asphärischen Modells ist ausgesprochen gering, wie allein die Skala der Differenz der Restdichten der Modelle *iam* und *mul* in Abb. 7.13 zeigt – sie wurde zu ungefähr einem Sechstel des Wertebereichs der eigentlichen Restdichten gewählt.

In Abb. 7.14 und in Abb. 7.15 auf der folgenden Seite sind Ausschnitte für die Aminosäurereste Tryptophan 108 und Glutaminsäure 35 zu sehen. Die Verteilung der Restdichte der sphärischen Modelle ist offensichtlich nur noch schwer mit der Erwartung lokalisierter Bindungsphänomene vereinbar. Vielmehr ist die Topologie der Restdichten diffus und nicht deckungsgleich mit der Molekülstruktur.

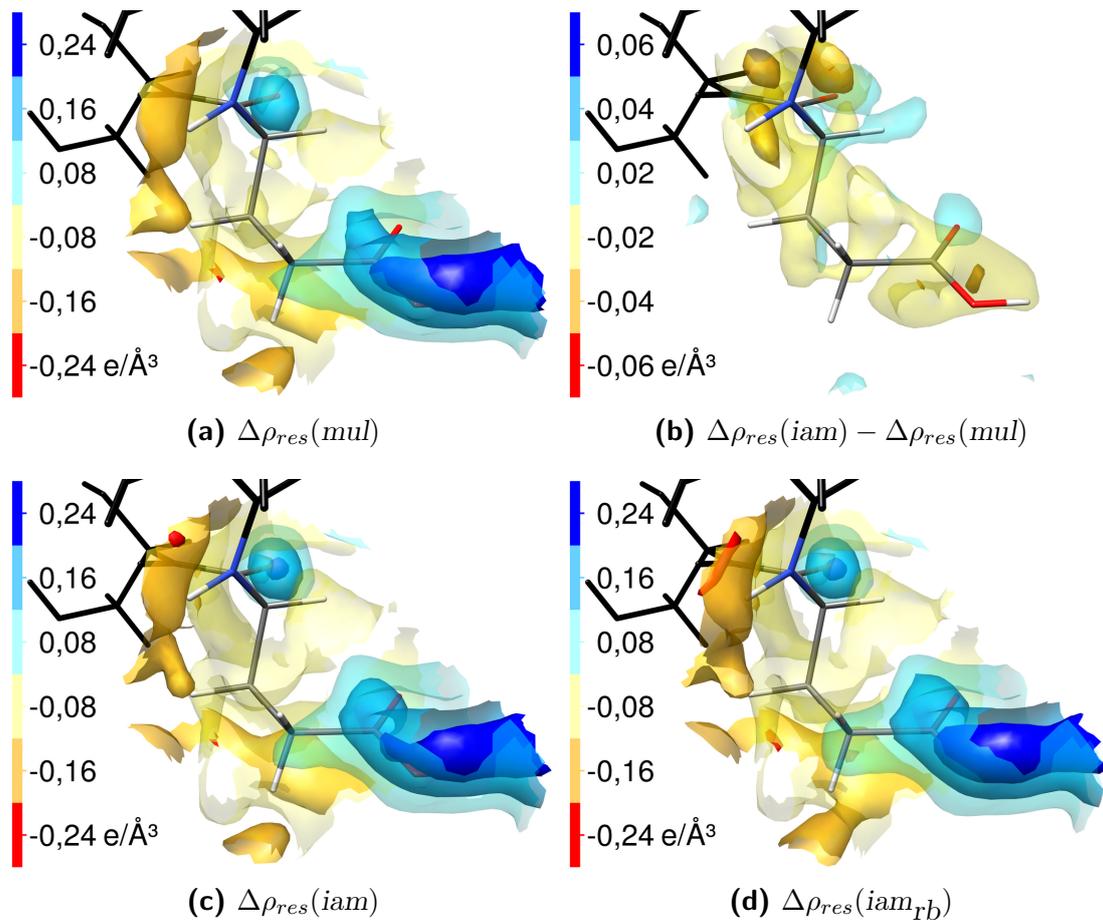


Abbildung 7.15: Restdichten und Restdichtedifferenzen bis $1,4 \text{ \AA}$ um den TETRALYSO-Glutaminsäurerest 35 für sphärische und asphärische Verfeinerungen.

Hierfür ist die ungenau bestimmte Elektronendichte des Experiments und der durch die geringe Datenaufösung von $d = 0,8 \text{ \AA}$ prinzipiell eingeschränkte Detailreichtum der Beobachtungen ursächlich.

Die Hauptmerkmale der Restdichten sind bei beiden Beispielen neben den Atomen und Bindungen lokalisiert, während direkt auf den Bindungen, nur wenig Restdichte, insbesondere kaum positive, befindlich ist. Im Fall des Tryptophanrestes liegen hauptsächlich negative Restdichten in ca. 1 \AA Entfernung vom Molekül. Bei der Carboxylgruppe der Glutaminsäure sind hohe positive Restdichten knapp über dieser lokalisiert, was eher auf eine Fehlordnung hinweist. Zwischen sphärischem und asphärischem Modell betragen die Differenzen nur etwa ein Viertel der beitragenden Restdichten. Die Konzentration von Elektronendichte auf den Bindungen, bedingt eine Abnahme selbiger in der Umgebung. Somit wird die dort vorhandene negative Restdichte etwas reduziert, andererseits positive Restdichte verstärkt.

7.6 Deformationsdichten

In Abb. 7.16 ist die Deformationsdichte $\Delta\rho_{def}$ des asphärischen Modells *mul* gegenüber einer sphärischen Elektronendichteverteilung für die zuvor betrachteten zwei Aminosäurereste gezeigt. Abb. 7.17 auf der folgenden Seite zeigt die gleiche Größe für die Gesamtstruktur TETRALYSO. Hiermit wird unter anderem belegt, daß die Orientierung der atomzentrierten Multipoldichte korrekt erfolgte, denn schließlich liegen die im asphärischen Formalismus gegenüber dem sphärischen erhöhten Elektronendichten auf den Bindungen und an den Orten freier Elektronenpaare des Sauerstoffs.

Die Deformationsdichte wurde auf einem kubischen Gitter mit 231^3 Punkten berechnet, welches die Gesamtstruktur einschließt. Es wurde ein Punktabstand von $0,266 \text{ \AA}$ in alle Raumrichtungen gewählt, um ein ähnlich feines Gitter zu erzeugen, wie es für die Restdichte Verwendung fand. Damit beträgt das Volumen eines jeden Voxels $0,0188 \text{ \AA}^3$. Aus der Summation der Restdichtebeträge, ergibt sich nach Normierung auf die Gittereinheit eine Gesamtmenge $\sum |\Delta\rho_{def}| V_{vox}$ von $830 e$ positiver wie negativer Deformationsdichte. Für jede der 2017 asphärisch beschriebenen Atomlagen, ergibt sich im Mittel $0,41 e$ Valenzdichte, die von einer kugelförmigen Beschreibung abweichend verteilt ist.

Aus dem Quotienten der Deformations- und Restelektronendichte $\Delta\rho_{def}/\Delta\rho_{res} = 830 e/3688 e = 0,225$ ist ersichtlich, daß die vom Modell *mul* nicht beschriebene Dichte fast das Fünffache der das asphärische vom sphärischen Modell unterscheidenden Dichte ausmacht.

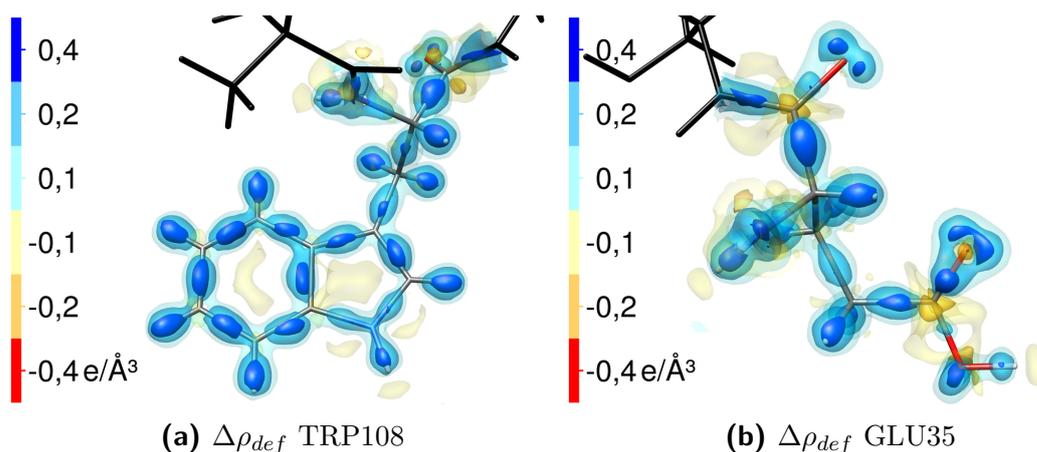


Abbildung 7.16: Deformationsdichten der TETRALYSO-Reste Tryptophan 108 und Glutaminsäure 35

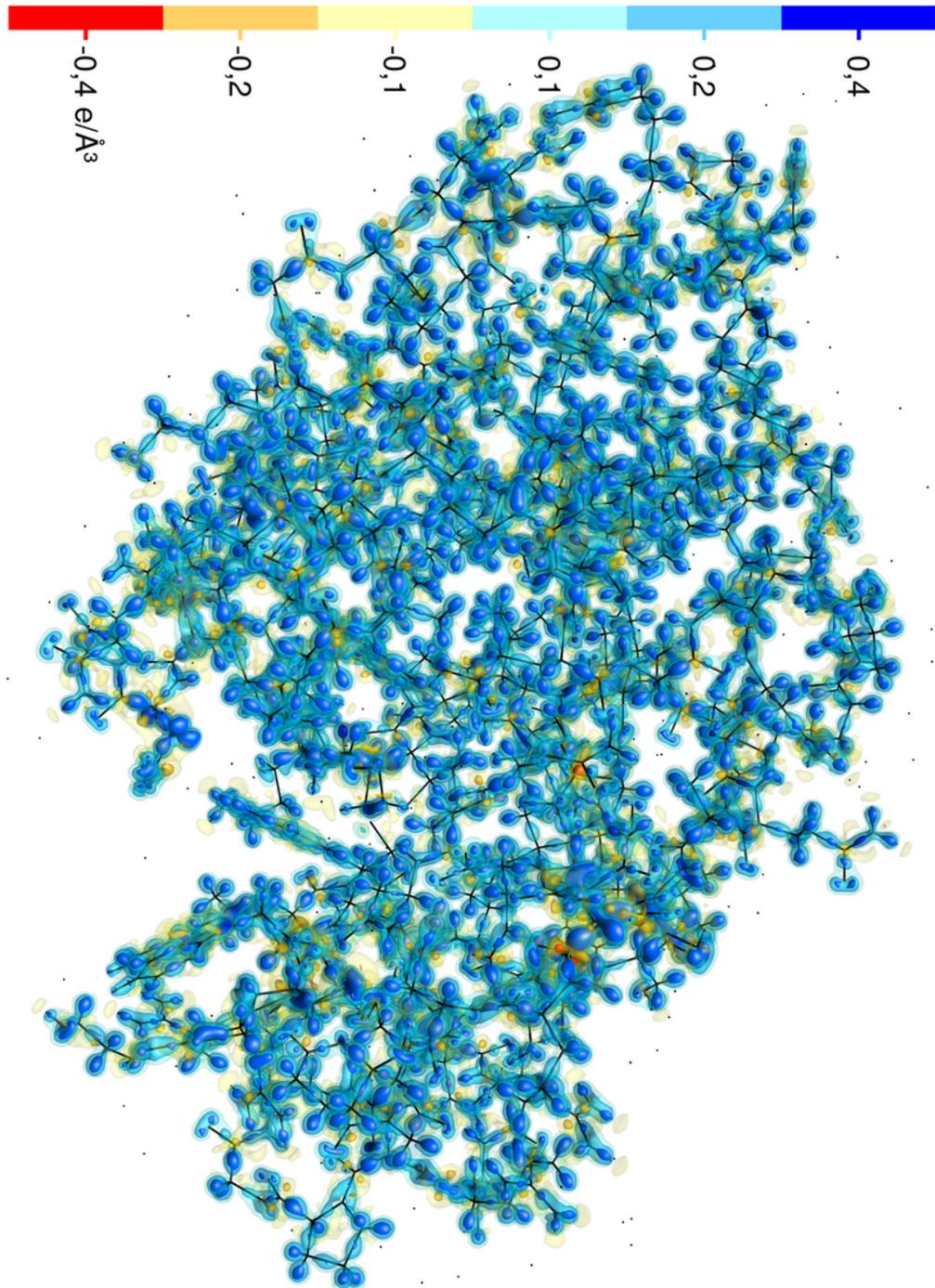


Abbildung 7.17: Deformationsdichte $\Delta\rho_{def}$ von TETRALYSO.

7.7 Abgeleitete Eigenschaften

Abb. 7.18 auf der nächsten Seite zeigt das elektrostatische Potential von TETRALYSO, welches aus der Multipolelektronendichte der Verfeinerung *mul* abgeleitet wurde. Es ist durch Farbwerte auf einer $0,001 e/a_0^3$ -Isooberfläche der Elektronendichte illustriert, die ebenfalls den Invariompopulationen entstammt. Für die Bestimmung der Elektronendichteverteilung und des Potentials wurden ausschließlich die 1961 Atome der Lysozymhauptbesetzung herangezogen. Beiträge minderpopulierter Strukturteile oder der Solvenssphäre wurden von der Rechnung ausgeschlossen.

Auf dieser Oberfläche, welche die van-der-Waals-Oberfläche approximiert, nimmt das ESP Werte zwischen $-0,2545$ und $0,1966 e/a_0$ an. Zum größten Teil wird das Molekül durch ein elektropositives bis schwach negatives Potential charakterisiert. Nur im Bereich der Enzymtasche tritt ein ausgeprägt elektronegatives ESP zutage, insbesondere an der Carboxylatgruppe des Aminosäurerestes Asparaginsäure 52. Dem steht ein weniger negatives bis positives Potential um die Glutaminsäure 35 gegenüber, wie es bereits in den Ausführungen zum P1LYSO-ESP in Absch. 6.7 auf Seite 183 erläutert wurde.

Eine Wiederholung dieses Potentials ϕ_{mul} auf der gleichen Oberfläche stellt Abb. 7.19(a) auf Seite 215 dar, nur bei einem festen Wertebereich von $\pm 191 \text{ kT/e} = \pm 0,18 e/a_0$. Damit kann Abb. 7.19(b) verglichen werden, die auf der gleichen Oberfläche das aus der Lösung der Poisson-Boltzmann-Gleichung mit *delphi* resultierende elektrostatische Potential ϕ_{PC} zeigt. Hierfür wurden jedem Proteinatorom die Standardladungen des *amber03*-Kraftfeldes zugewiesen, um die notwendige Punktladungsverteilung zu erzeugen. Die Gesamtladung des Moleküls betrug $+8,98 e$, was sehr nahe den neun überschüssigen Formalladungen liegt.

Offenbar ist das Multipol-ESP, besonders im Bereich der Tasche, stärker elektronegativ ausgeprägt als das Punktladungs-ESP. Dazu passen auch die stärker elektropositiven Extremwerte von ϕ_{PC} , die bei $-0,1364$ und $0,2526 e/a_0$ (-145 und 268 kT/e) liegen, was Abb. 7.19(e) zeigt. Dort wird aber ebenfalls deutlich, daß die prinzipielle Verteilung positiver und negativer Bereiche von ϕ_{PC} der von ϕ_{mul} in Abb. 7.18 sehr ähnlich ist.

Mit den Abb. 7.19(c), (d) und (f) wird der Vergleich von ϕ_{mul} und ϕ_{PC} auf einer weiteren Oberfläche durchgeführt. Es handelt sich um die van-der-Waals-Oberfläche A_{vdW} , welche aus den Atomradien bestimmt wurde. Auf ihr nehmen beide Potentiale höhere Extremwerte an. So weist das Punktladungs-ESP in 7.19(f) Maxima bei $-0,1812$ und $0,2825 e/a_0$ (-192 und 302 kT/e) auf, die entsprechenden Werte für ϕ_{mul} liegen bei $-0,2809$ und $0,2499 e/a_0$ (-298 und 265 kT/e). Hieran ist die Verschiebung von ϕ_{PC} ins Elektropositivere gegenüber ϕ_{mul} ebenso deutlich erkennbar wie durch den Vergleich der Abb. 7.19(c) und (d).

Auch für TETRALYSO ist die Abweichung beider Potentiale höher, als es die Tests für APA in Absch. 5.7.3 auf Seite 144 nahegelegt hätten, andererseits geringer, als bei P1LYSO beobachtet.

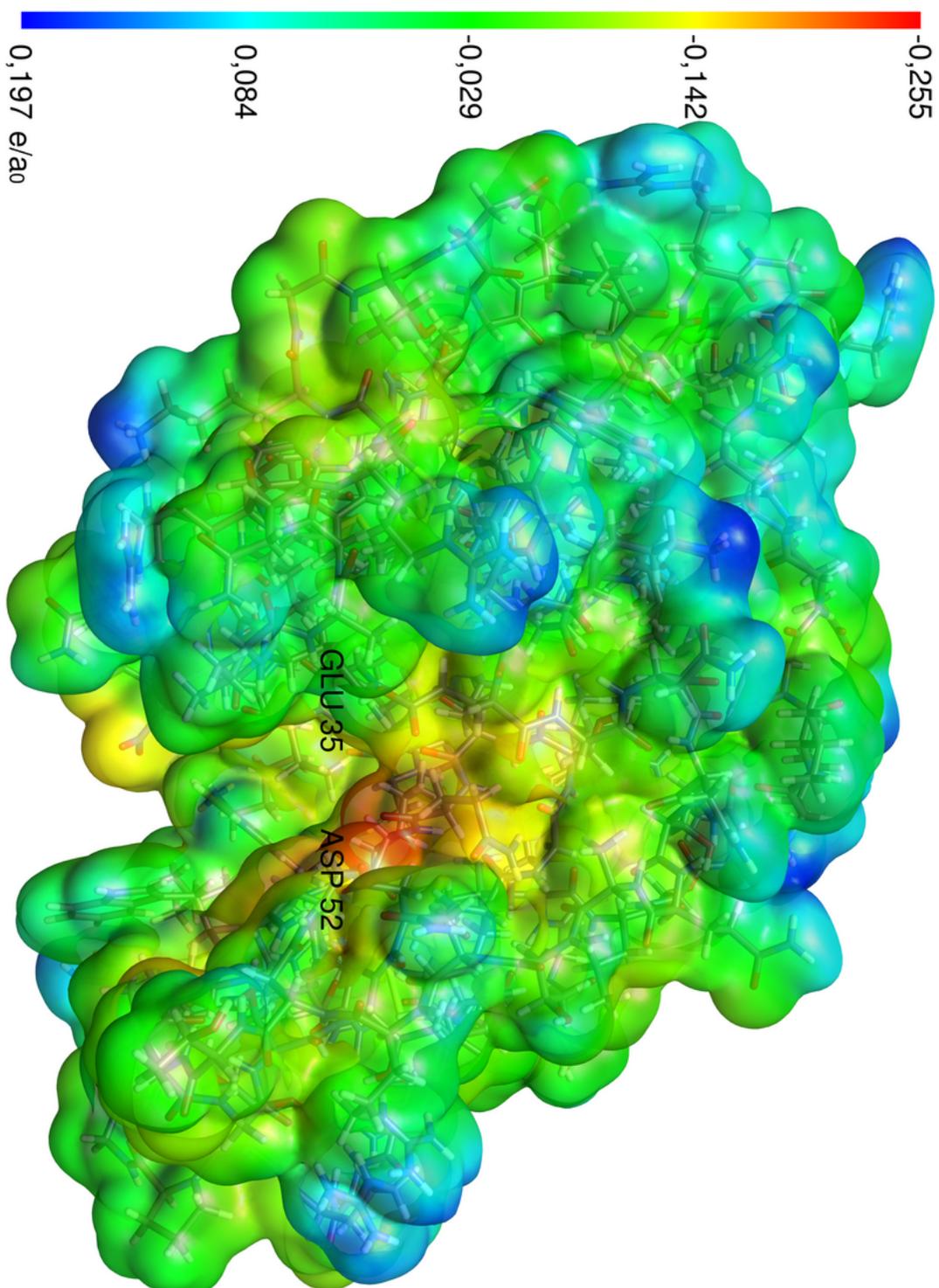


Abbildung 7.18: Elektrostatisches Potential des Multipolmodells von TETRALYSO auf der 0,001 e/a_0^3 -Isooberfläche.

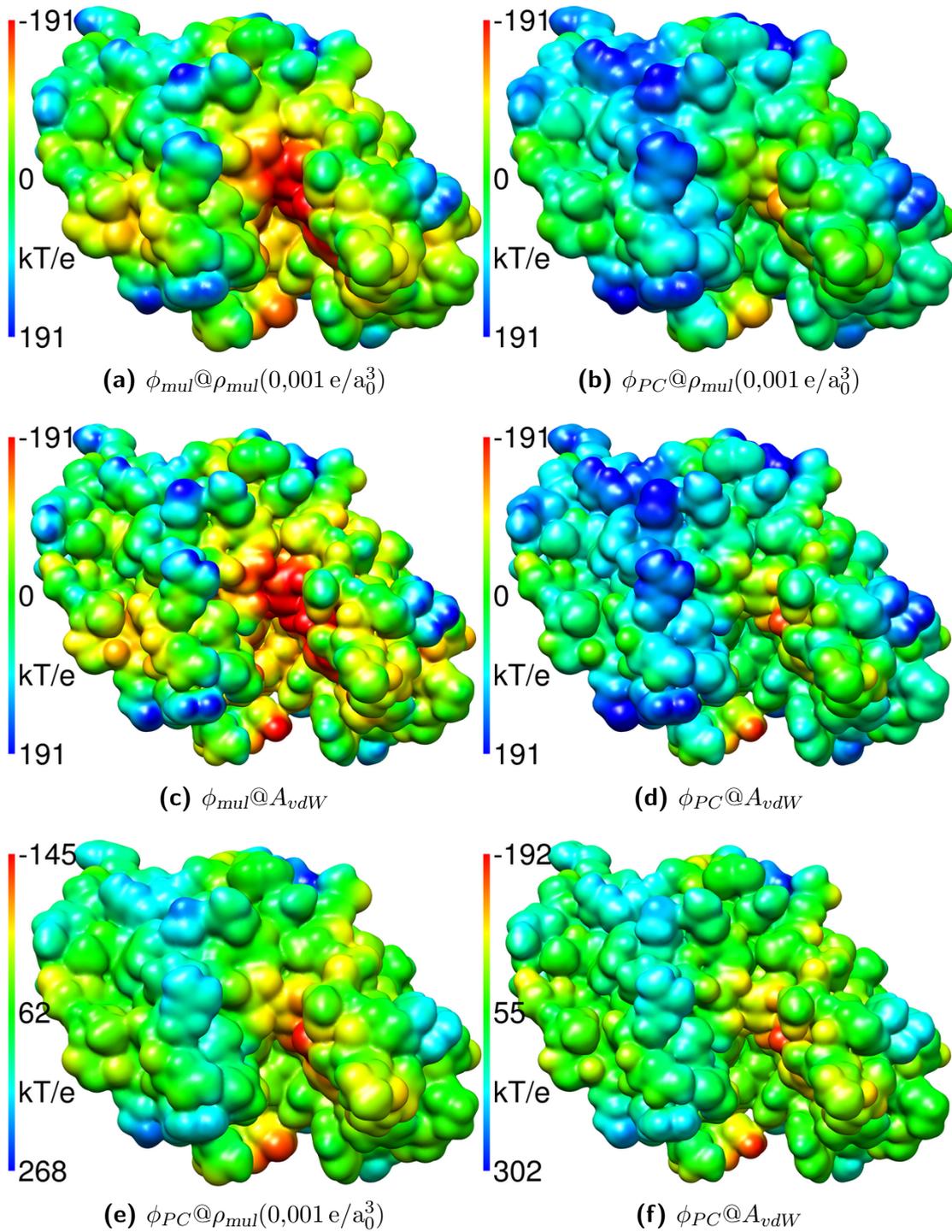


Abbildung 7.19: Multipol-ESP (ϕ_{mul}) und Punktladungs-ESP (ϕ_{PC}) auf $0,001 e/a_0^3$ -Isooberfläche der Multipolelektronendichte (ρ_{mul}) und van-der-Waals-Oberfläche (A_{vdW}). (a)-(d): Fester Wertebereich $\pm 191 \text{ kT/e} = \pm 0,18 e/a_0$. (e)-(f): Voller Wertebereich.

8 Insulin

8.1 Gütefaktoren

Aus der Verfeinerung der in Absch. 4.5 auf Seite 94 beschriebenen Modelle resultieren die in Tab. 8.1 aufgelisteten Gütefaktoren. Die konvergente Anpassung von 3628 identischen Parametern der drei Modelle an 54157 Beobachtungen verlief in allen Fällen konvergent, wie die mittlere, auf die Standardabweichung bezogene Parameterverschiebung von etwa 10^{-5} zeigt, allerdings auch zu fast identischen R -Werten. Zwischen $R(F)$ und $R_{all}(F)$ der sphärischen Modelle und dem asphärischen Modell mul liegen maximal 0,2 % hohe relative Unterschiede zugunsten der Invariomverfeinerung. Der $R_w(F^2)$ verschlechtert sich aber von 27,18 % im Modell iam bzw. 27,28 % im Modell iam_{rb} auf 27,37 % nach der Verfeinerung mul , eine relative Zunahme um bis zu 0,7 %.

Ob tatsächlich eine Verschlechterung im asphärischen Modell vorliegt, soll direkt anhand der gewichteten Differenzen von beobachteten und berechneten Strukturparametern der Modelle iam und mul betrachtet werden. Die Unterschiede $\Delta F_w(iam) = |F_{o,iam} - F_{c,iam}| \sigma(F_{o,iam})^{-1}$ haben ein arithmetisches Mittel von 0,5148 und einen Median von 0,3928, wohingegen die Deskriptoren der Verteilung $\Delta F_w(mul) = |F_{o,mul} - F_{c,mul}| \sigma(F_{o,mul})^{-1}$ bei 0,5179 und 0,3967 liegen. Im asphärischen Modell sind die Differenzen im Mittel höher, was auch ohne Gewichtung gilt. Aus Signifikanztests zu den Hypothesen $H_0 : \Delta F_w(iam) = \Delta F_w(mul)$ und $H_1 : \Delta F_w(iam) > \Delta F_w(mul)$ ergeben sich im t - und Wilcoxon-Test Über-

Tabelle 8.1: Gütefaktoren der sphärischen und asphärischen Verfeinerung von INSULIN.

	<i>iam</i>	<i>iam_{rb}</i>	<i>mul</i>
Reflexe $F_o^2 > 2\sigma(F_o^2)$	54157	54157	54157
Parameter	3628	3628	3628
N_{ref}/N_{par}	14,928	14,928	14,928
$R(F)$	0,1123	0,1124	0,1123
$R_{all}(F)$	0,1179	0,1180	0,1178
$R_w(F^2)$	0,2718	0,2728	0,2737
GoF_w	2,3679	2,3774	2,3829
Mittl.(Versch./esd)	$4,76 \cdot 10^{-5}$	$3,76 \cdot 10^{-5}$	$7,29 \cdot 10^{-5}$

Tabelle 8.2: Gütefaktoren der sphärischen und asphärischen Verfeinerung von INSULIN gegen beschnittene Daten.

	<i>iam</i>	$\frac{1}{2} \leq \frac{F_o}{F_c} \leq 2$ <i>iam_{rb}</i>	<i>mul</i>
Reflexe $F_o^2 > 2\sigma(F_o^2)$	50739	50739	50739
Parameter	3628	3628	3628
N_{ref}/N_{par}	13,985	13,985	13,985
$R(F)$	0,0990	0,0991	0,0992
$R_{all}(F)$	0,0995	0,0997	0,0997
$R_w(F^2)$	0,2359	0,2369	0,2388
GoF_w	2,0837	2,0932	2,1075
Mittl.(Versch./esd)	$1,08 \cdot 10^{-5}$	$8,85 \cdot 10^{-6}$	$2,09 \cdot 10^{-5}$

schreitungswahrscheinlichkeiten von jeweils 1. Die Nullhypothese kann also nicht abgelehnt werden, es gibt keine statistisch signifikante Bevorzugung der Hypothese höherer Strukturfaktordifferenzen im Modell *iam*. Bei gegenteiliger Alternativhypothese $H_1 : \Delta F_w(iam) < \Delta F_w(mul)$ jedoch, lauten die *p*-Werte in beiden Fällen 0. Hier kann die Nullhypothese zugunsten geringerer Differenzen im Modell *iam* abgelehnt werden. Beide Tests bestätigen die statistisch signifikant erhöhten Abweichungen zwischen Beobachtungen und Modell im Fall der Verfeinerung *mul*.

Effekt des Beschneidens der Daten

Tab. 8.2 zeigt die Resultate des Versuchs, die Verfeinerungen durch Ausschluß der am stärksten abweichenden Reflexe zu verbessern. Von 62155 Reflexen, liegen 8066 oder 13,0 % außerhalb des in Abb. 8.1 auf der gegenüberliegenden Seite rot gekennzeichneten Bereichs $F_o^2/F_c^2 \leq 2$ bis $\geq 1/2$, sie wurden für diesen Test aus dem Datensatz entfernt. Dadurch bricht allerdings die Vollständigkeit von 100 % auf 87 % ein, Schalen höchster und niedrigster Auflösung sind besonders betroffen. Mit dieser Behandlung sinkt der *R*-Wert um 1,3 Prozentpunkte, der $R_w(F^2)$ sogar um ca. 3,6 Punkte. Die Verbesserung ist für das Modell *iam* immer ausgeprägter, so daß nun die Verschlechterung der Gütefaktoren im Modell *mul* eindeutig ist.

Die zu obigen Ausführungen für den vollen Datensatz analoge Untersuchung der Strukturfaktordifferenzen $\Delta F = |F_o - F_c|$ der beschnittenen Daten in den Modellen *iam* und *mul*, ergibt für beide Verfeinerungen geringere Mittelwerte als zuvor. Jedoch sind die Differenzen im asphärischen Modell in Relation zum sphärischen erhöht, was die Wiederholung der Signifikanztests bestätigt.

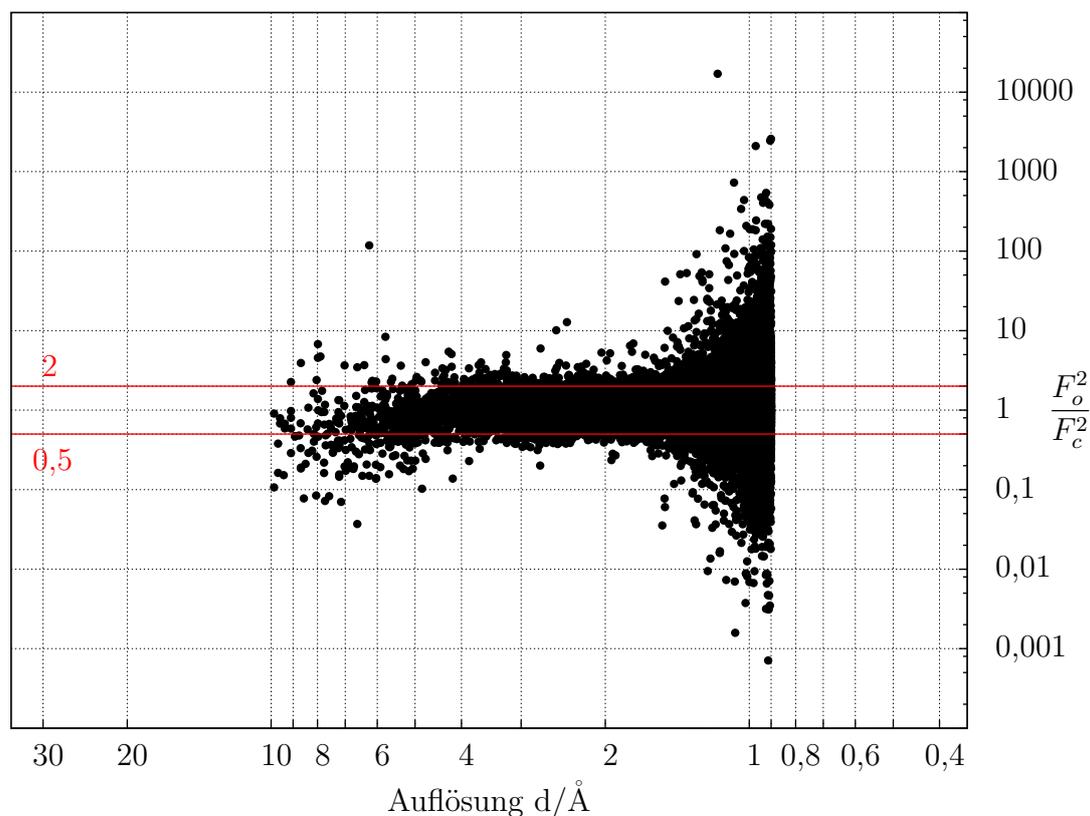


Abbildung 8.1: Quotient beobachteter und berechneter Intensitäten von INSULIN und ihre Auflösungsabhängigkeit im Modell *iam*.

8.2 Geometrie

Anhand der Abb. 8.2 auf der nächsten Seite und Abb. 8.3 auf Seite 221 sind die quadratisch gemittelten Unterschiede (RMSD) der Atompositionen von INSULIN zwischen den Verfeinerungen *iam* und *mul* bzw. zwischen *iam_{rb}* und *mul* mittels einer Farbcodierung dargestellt. Ob des hohen Anteils nicht verfeinerter Atome, in den genannten Abbildungen dünn gezeichnet, beträgt die über die Gesamtstruktur gemittelte Verschiebung im Modell *mul* nur 0,094 Å gegenüber *iam* und 0,005 Å beim Vergleich zu *iam_{rb}*.

Im verfeinerten Strukturteil, der in den Graphiken hervorgehoben ist, betragen die Positionsunterschiede von Nichtwasserstoffatomen im Modell *iam* 0–0,025 Å, die der Wasserstoffatome sind ob der verlängerten Abstände in *mul* höher, sie liegen zwischen 0,085 und 0,211 Å. Beim Vergleich des asphärischen Modells mit *iam_{rb}* liegen die Differenzen der Koordinaten von Nichtwasserstoffatomen und Wasserstoffatomen in einer vergleichbaren Größenordnung, da in beiden Modellen die gleichen Wasserstoffabstände Verwendung finden. Für erstere beträgt der

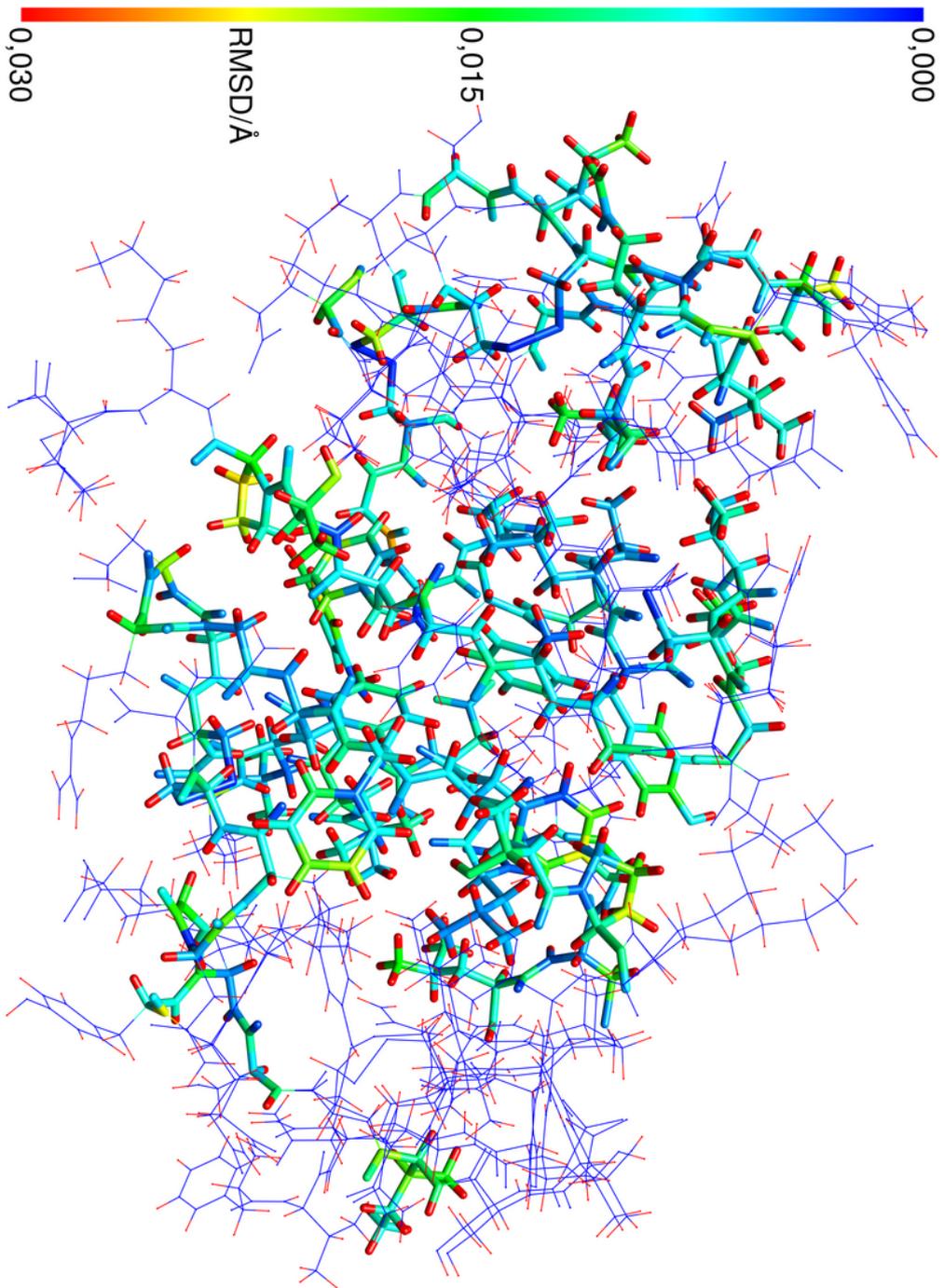


Abbildung 8.2: Quadratisch gemittelte Abweichung zwischen den Strukturen *mul* und *iam* von INSULIN. Verfeinerte Koordinaten hervorgehoben.

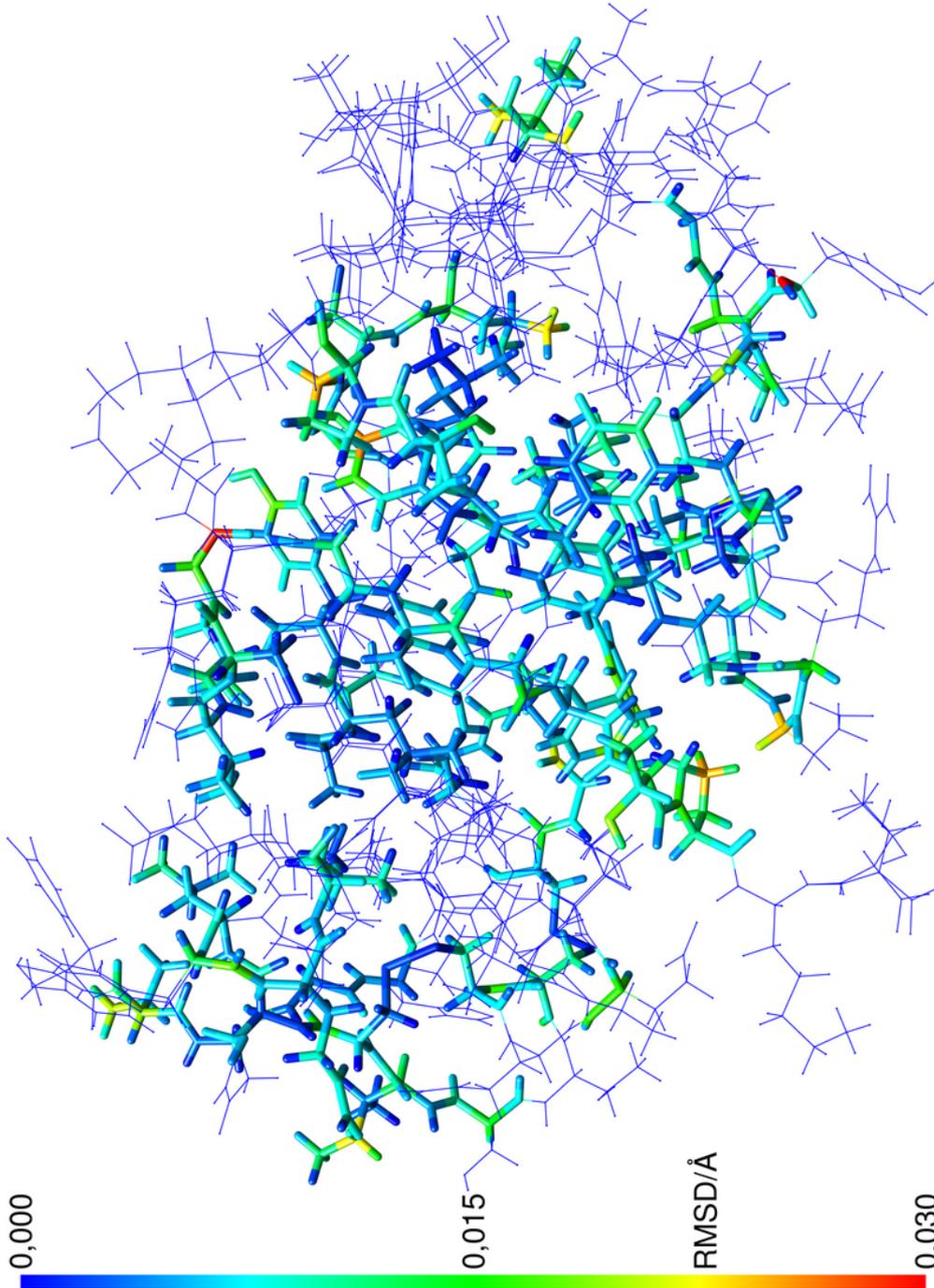


Abbildung 8.3: Quadratisch gemittelte Abweichung zwischen den Strukturen *mul* und *iam_{rb}* von INSULIN. Verfeinerte Koordinaten hervorgehoben.

Tabelle 8.3: Verfeinerte Bindungslängen der Peptidbindungen in der Hauptpopulation von INSULIN.

	$d_{E\&H}^*$	$d_{min/max}^\dagger$	$\bar{d}(\sigma_d)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–O ($n = 65$)								
<i>mul</i>		1,176/1,284	1,226(19)	1,217	1,229	1,238	0,021	0,014
<i>iam</i>	1,231(20)	1,197/1,286	1,235(18)	1,225	1,236	1,246	0,021	0,016
<i>iam_{rb}</i>		1,194/1,292	1,235(19)	1,225	1,238	1,245	0,020	0,014
C–N ($n = 74$)								
<i>mul</i>		1,265/1,394	1,337(22)	1,325	1,337	1,347	0,021	0,015
<i>iam</i>	1,329(14)	1,256/1,391	1,334(22)	1,320	1,333	1,343	0,023	0,017
<i>iam_{rb}</i>		1,260/1,386	1,336(22)	1,322	1,334	1,345	0,023	0,018
C_α–C ($n = 59$)								
<i>mul</i>		1,484/1,574	1,531(21)	1,517	1,530	1,546	0,030	0,021
<i>iam</i>	1,525(21)	1,460/1,564	1,518(23)	1,504	1,517	1,538	0,034	0,026
<i>iam_{rb}</i>		1,460/1,573	1,519(24)	1,502	1,518	1,539	0,036	0,027
C_α–N ($n = 57$)								
<i>mul</i>		1,377/1,506	1,451(23)	1,439	1,451	1,466	0,027	0,019
<i>iam</i>	1,458(19)	1,372/1,515	1,454(24)	1,440	1,451	1,468	0,027	0,019
<i>iam_{rb}</i>		1,377/1,517	1,459(24)	1,446	1,457	1,473	0,027	0,018

*Engh&Huber-Bindungslängen mit Standardabweichung; [†]Minimum, Maximum und Mittelwert mit Standardabweichung der Bindungslängen; [‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in Å.

Maximalunterschied 0,031 Å, für letztere 0,020 Å. Die am stärksten gegenüber beiden sphärischen Modellen verschobenen Nichtwasserstoffatome sind jeweils Carbonylkohlenstoffe und randständige Seitenkettenatome. Letzteres tritt eher selten auf, da solche Positionen ohnehin kaum verfeinert wurden. Aber auch die deutlichsten Positionsunterschiede liegen im Bereich der Standardabweichung der Koordinaten.

Um die Auswirkung der asphärischen Verfeinerung auf die Bindungslängen zu untersuchen, sind diese in Tab. 8.3 für eine Auswahl an Bindungen in den drei Modellen gegeben und zudem den Engh&Huber-Referenzbindungslängen gegenübergestellt. Hierbei werden nur die verfeinerten Peptidbindungen der INSULIN-Hauptpopulation betrachtet, welche weder zu randständigen Aminosäureresten noch zu Prolin oder Glycin gehören.

Die Mittelwerte der sphärischen Modelle stimmen innerhalb einer Standardabweichung mit den Vergleichslängen überein, was aber nicht für die Extremwerte eines jeden Bindungstyps zutrifft. Gleiches gilt auch für die Resultate der Verfeinerung *mul*, die sich dennoch in typischer Weise von den sphärischen Bindungslängen

Tabelle 8.4: Differenzen der verfeinerten Peptidbindungslängen der INSULIN-Hauptpopulation.

	$\Delta d_{min/max}^\dagger$	$\overline{\Delta d}(\sigma_\Delta)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–O ($n = 65$)							
<i>mul-iam</i>	–24,40/6,40	–8,70(5,32)	–11,70	–8,70	–5,50	6,20	4,74
<i>mul-iam_{rb}</i>	–23,60/4,00	–9,26(5,27)	–12,10	–8,90	–6,10	6,00	4,30
C–N ($n = 74$)							
<i>mul-iam</i>	–8,40/12,70	3,62(4,84)	–0,23	4,05	7,25	7,47	4,89
<i>mul-iam_{rb}</i>	–11,40/15,40	1,72(5,13)	–1,72	2,25	5,00	6,72	4,82
C_α–C ($n = 59$)							
<i>mul-iam</i>	–1,50/33,70	13,11(6,32)	9,25	12,30	16,85	7,60	5,34
<i>mul-iam_{rb}</i>	–3,80/29,60	12,23(6,79)	8,60	10,60	16,55	7,95	4,45
C_α–N ($n = 57$)							
<i>mul-iam</i>	–18,80/5,50	–3,62(5,55)	–6,30	–2,90	–0,40	5,90	4,45
<i>mul-iam_{rb}</i>	–24,10/3,10	–7,85(5,49)	–11,00	–7,80	–4,00	7,00	5,04

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der Bindungslängendifferenzen; [‡]0,25-Quartil, Median, 0,75-Quartil; alle Einheiten in 10^{-3}Å .

unterscheiden, was Tab. 8.4 anhand der Bindungslängendifferenzen hervorhebt.

So tritt eine Verkürzung der Carbonylbindung C–O auf, die mittlere Abnahme beträgt $-9 \cdot 10^{-3}\text{Å}$. Die Schwankungsbreite zwischen Minimal- und Maximalwerten nimmt im asphärischen Modell zu, während die restlichen Deskriptoren der Streuung eine ungefähr gleichbleibende Verteilungsbreite nahelegen. Eine geringfügige Verkürzung erfährt auch die C_α–N-Bindungslänge, insbesondere gegenüber der Verfeinerung *iam_{rb}*. Die Mediane in den Modellen *iam* und *mul* sind zwar identisch, das Mittel der Differenzen ist aber negativ. Auch bewirkt *mul* eine geringfügige Zunahme des Minimalabstands und eine Abnahme des Maximalabstands. Ansonsten sind die Streuungsindikatoren unverändert. Die zweite Stickstoff enthaltende Bindung C–N wiederum, wird in der asphärischen Beschreibung verlängert. Der Median der Bindungslängendifferenzen zu *iam* liegt bei $4 \cdot 10^{-3}\text{Å}$, zu *iam_{rb}* bei $2 \cdot 10^{-3}\text{Å}$. Hierbei nehmen Extremwerte und Streuung im Modell *mul* ab. Bei den beiden letztgenannten Bindungen ist der Unterschied zwischen den zwei sphärischen Modellen in der Differenz zu *mul* am größten, was mit den unterschiedlichen N–H-Abständen in *iam* und *iam_{rb}* begründet wird. Eine viel deutlichere Zunahme zeigt das Mittel der C_α–C-Bindungslängen, die Differenzen haben einen Median von etwa $11 \cdot 10^{-3}\text{Å}$, etwas höher bezüglich *iam*, etwas niedriger bei *iam_{rb}*. Dieser Bindungstyp zeigt auch die stärksten Veränderungen einer Bindungslänge um $0,034\text{Å}$. Nach asphärischer Verfeinerung ist die Schwankungsbreite der Längenverteilung geringer als in den sphärischen Modellen.

8.3 Verschiebungsparameter

Bei Einbeziehung aller 2238 Atomlagen, inklusive der Solvenssphäre und Wasserstoffatome, beträgt der Median der äquivalenten isotropen Verschiebungsparameter U_{eq} $0,181 \text{ \AA}^2$. Die Extremwerte liegen bei $0,081$ und $1,252 \text{ \AA}^2$, der höchste Nichtwasserstoffatomverschiebungsparameter des INSULIN-Moleküls bei $0,834 \text{ \AA}^2$. Wie in Tab. D.1 auf Seite 371 ebenfalls aufgeführt, haben die 403 verfeinerten Nichtwasserstoffatome immer noch ein mittleres U_{eq} von $125,90 \cdot 10^{-3} \text{ \AA}^2$ im Fall der Verfeinerung iam_{rb} , $125,70 \cdot 10^{-3} \text{ \AA}^2$ nach der Verfeinerung iam und $125,20 \cdot 10^{-3} \text{ \AA}^2$ im Modell mul . Somit liegt das asphärische Mittel $0,4\%$ bis $0,6\%$ unter dem der sphärischen Modelle. In dieser Untermenge verfeinerter Atome beträgt der maximale Temperaturfaktor etwa $320 \cdot 10^{-3} \text{ \AA}^2$.

In Abb. 8.4 auf der gegenüberliegenden Seite sind die U_{eq} -Werte des Modells mul über die Peptidsequenz aufgetragen. Es ist deutlich zu erkennen, wie die Werte am Anfang und Ende jeder der vier Ketten stark ansteigen, solche Atome wurden nicht verfeinert. Dort sind ebenfalls die Veränderungen zum sphärischen Modell iam als Balken aufgetragen. Der Median der Differenz zwischen den Verfeinerungen mul und iam liegt bei $-0,88 \cdot 10^{-3} \text{ \AA}^2$, der der nicht abgebildeten Differenzen zu iam_{rb} bei $-1,25 \cdot 10^{-3} \text{ \AA}^2$. Hiermit entspricht die mittlere Verbesserung $0,7\%$ bzw. 1% der sphärischen U_{eq} -Zentralwerte. Allerdings sind die Extremwerte der Modellunterschiede wesentlich höher, sie betragen -10 und $12 \cdot 10^{-3} \text{ \AA}^2$ für den Vergleich von mul mit iam , gegenüber iam_{rb} im Maximum sogar $18 \cdot 10^{-3} \text{ \AA}^2$. Wie Abb. 8.5 auf Seite 226 anhand der Auftragung der Modellunterschiede mit steigendem Verschiebungsparameter belegt, treten die drastischsten Änderungen erst zu höherem U_{eq} -Wert auf, ebenso wie mit steigendem sphärischen Temperaturfaktor häufiger eine Verschlechterung im Modell mul zu beobachten ist.

Von den 403 Veränderungen beim Übergang vom Modell iam zu mul , sind 322 negativ, der U_{eq} -Wert hat sich also verringert. Damit verbessert die asphärische Verfeinerung die Beschreibung von 80% der verfeinerten Atome. Gegenüber dem Modell iam_{rb} sind es 337 Atome, deren U_{eq} -Wert in mul geringer ist. Die Verbesserung wurde somit für 84% der Atome erzielt.

Eine Darstellung der anisotropen Verschiebungsellipsoide des verfeinerten Strukturteils bietet Abb. 8.6 auf Seite 228. Für eine Darstellung der Gesamtstruktur kann Abb. D.1 auf Seite 373 konsultiert werden. Die beiden Abbildungen machen die Auswahl von Atomen mit relativ niedrigem Temperaturfaktor ebenso deutlich wie die geringe Qualität der Kette C und etlicher randständiger Seitenketten.

Eine nach Elementen unterschiedene Statistik der Temperaturfaktoren verfeinerter Atome ist in Tab. D.2 auf Seite 372 aufgeführt, die noch weiter differenzierte Betrachtung für nicht fixierte 252 Peptidbindungsatome in Tab. 8.5 auf Seite 227. Letztere haben zumeist kleinere Verschiebungsparameter als die Gesamtheit der jeweiligen Elemente. Die höchste Verbesserung von sphärischem zu asphärischem

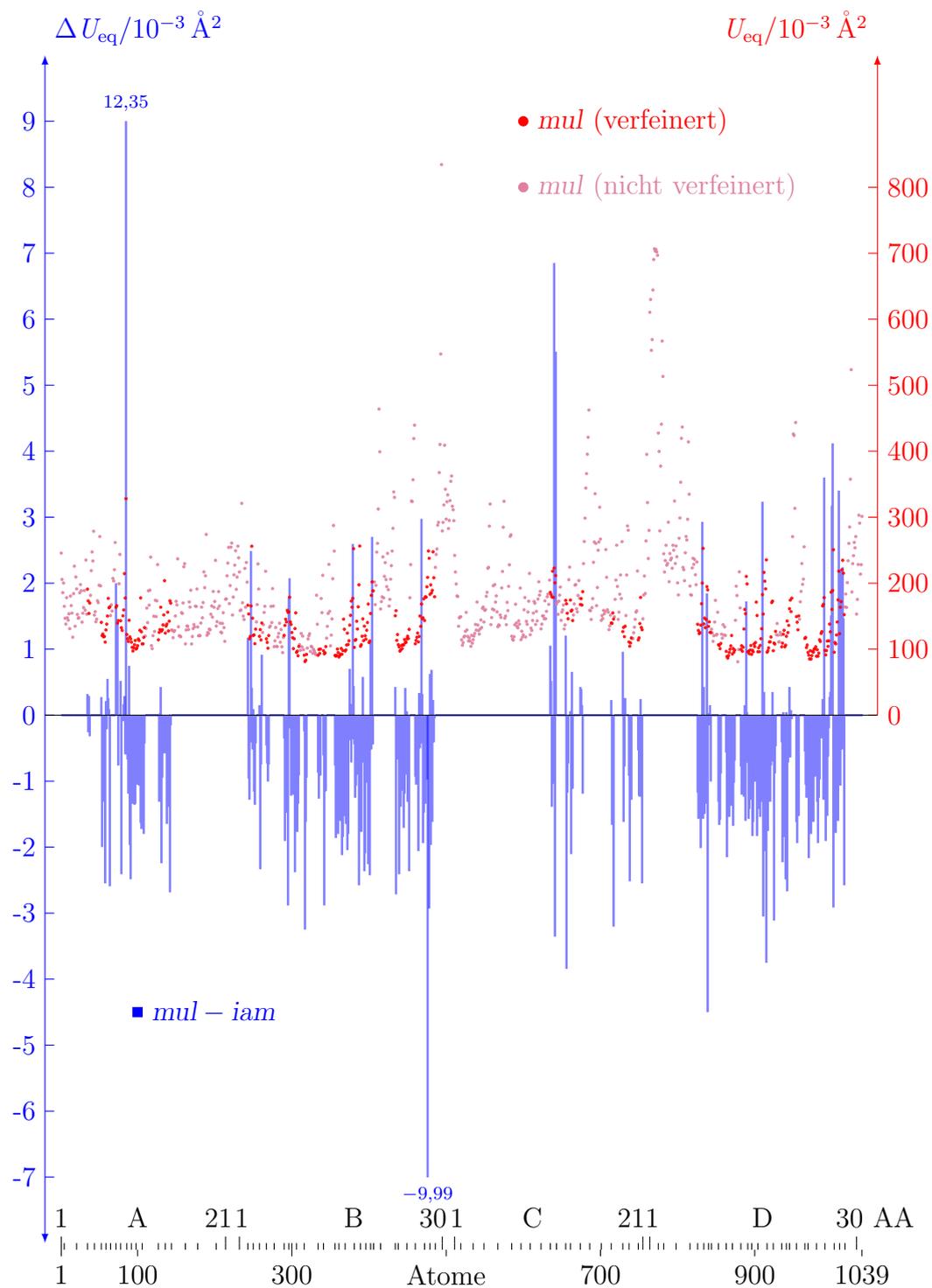


Abbildung 8.4: Veränderung der U_{eq} -Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die Nichtwasserstoffatome von INSULIN.

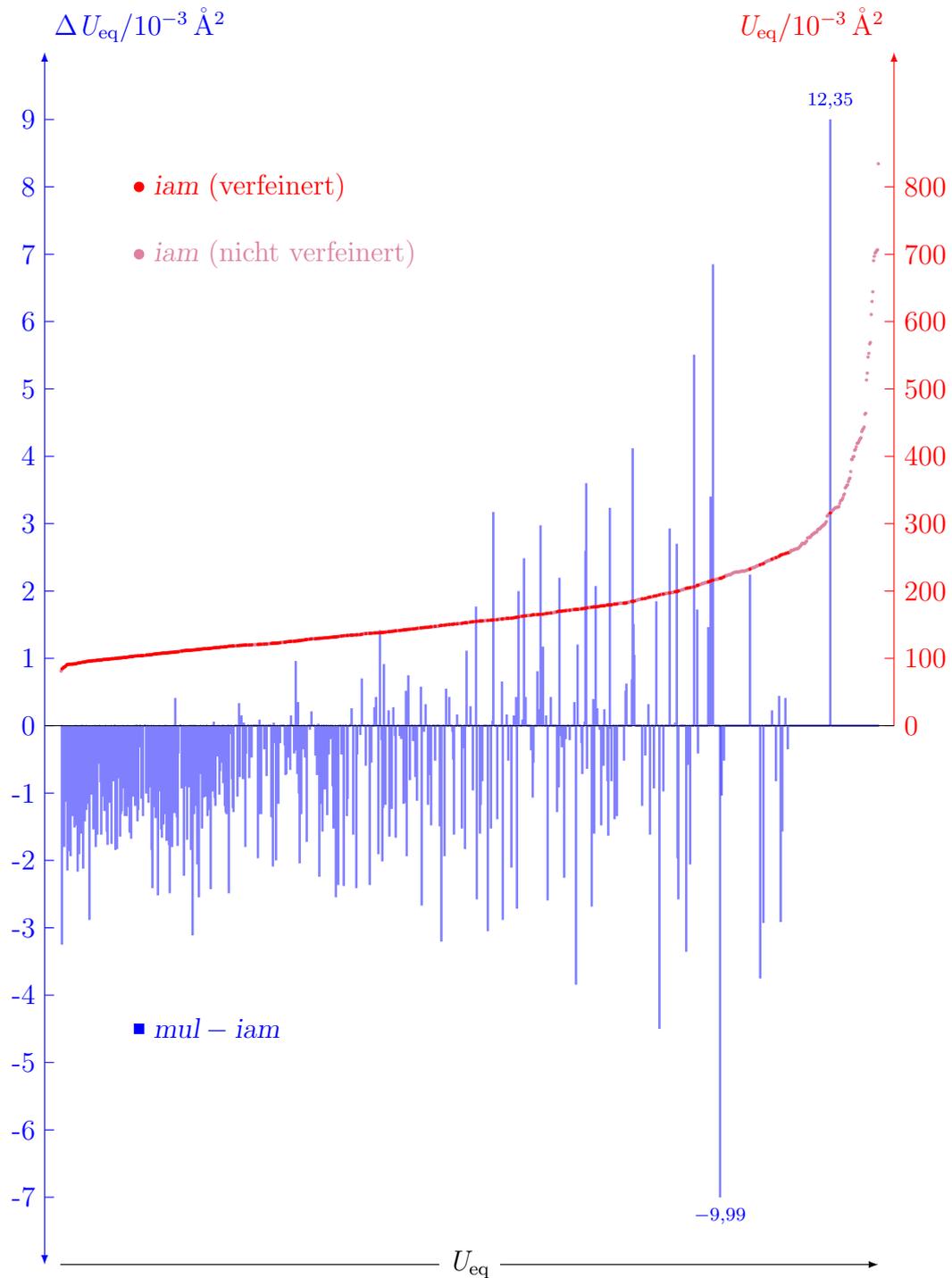


Abbildung 8.5: Veränderung der U_{eq} -Werte im asphärischen Modell mul gegenüber dem sphärischen Modell iam für die Nichtwasserstoffatome von INSULIN. Auftragung nach U_{eq} (iam).

Tabelle 8.5: Statistik verfeinerter U_{eq} von INSULIN- Atome der Peptidbindung.

	$x_{\min/\max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C$_{\alpha}$ ($n = 59$)							
<i>mul</i>	81,19/197,40	122,80(28,55)	101,30	113,30	136,90	35,64	23,26
<i>iam</i>	84,44/199,40	124,30(28,26)	102,70	114,60	138,00	35,33	22,85
<i>iam_{rb}</i>	84,81/199,30	124,70(28,07)	103,20	114,70	138,80	35,57	22,79
<i>mul-iam</i>	-3,85/0,81	-1,51(0,89)	-1,95	-1,57	-1,06	0,89	0,63
<i>mul-iam_{rb}</i>	-4,26/1,64	-1,91(1,08)	-2,46	-2,09	-1,49	0,96	0,85
C ($n = 65$)							
<i>mul</i>	83,39/201,90	121,10(27,47)	97,93	115,70	138,80	40,88	27,74
<i>iam</i>	85,00/199,20	121,80(26,89)	99,07	116,90	139,60	40,54	27,78
<i>iam_{rb}</i>	85,57/198,50	122,20(26,83)	99,44	116,70	140,30	40,85	26,80
<i>mul-iam</i>	-1,94/2,71	-0,74(0,91)	-1,35	-0,96	-0,20	1,15	0,71
<i>mul-iam_{rb}</i>	-2,68/3,65	-1,10(1,08)	-1,70	-1,41	-0,71	0,99	0,67
N ($n = 63$)							
<i>mul</i>	88,02/190,90	121,10(27,02)	97,69	115,20	138,90	41,24	28,40
<i>iam</i>	88,85/190,60	121,30(26,53)	98,34	115,60	139,20	40,81	28,02
<i>iam_{rb}</i>	89,14/190,80	121,40(26,41)	98,35	115,70	139,80	41,44	27,62
<i>mul-iam</i>	-1,61/2,00	-0,22(0,73)	-0,74	-0,37	0,19	0,93	0,64
<i>mul-iam_{rb}</i>	-2,22/2,13	-0,35(0,88)	-0,99	-0,61	0,03	1,02	0,70
O ($n = 65$)							
<i>mul</i>	90,17/220,40	133,50(31,15)	107,70	127,10	146,00	38,34	28,70
<i>iam</i>	91,44/220,90	134,30(31,02)	108,50	127,50	146,80	38,28	28,53
<i>iam_{rb}</i>	91,81/221,30	134,80(31,09)	108,90	127,70	147,00	38,08	28,58
<i>mul-iam</i>	-2,36/0,91	-0,81(0,64)	-1,19	-0,88	-0,44	0,74	0,53
<i>mul-iam_{rb}</i>	-2,60/0,63	-1,27(0,64)	-1,60	-1,26	-0,96	0,64	0,49

† Minimum, Maximum und Mittelwert mit Standardabweichung der U_{eq} und ΔU_{eq} ;

‡ 0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-3}\AA^2 .

Modell zeigen die C $_{\alpha}$ -Atome, gefolgt von den Carbonylkohlenstoffatomen. Nur halb so stark ist die mittlere Abnahme der U_{eq} -Werte des Sauerstoffs. Bei Stickstoffatomen findet sich das geringste Mittel der Verringerung der Temperaturfaktoren, hier liegt die obere Quartilgrenze bereits im Positiven.

Keine der Verteilungsbreiten dieser Atome nimmt im sphärischen Modell ab. Die Ergebnisse aus Tab. 8.5 sind auch repräsentativ für die Ergebnisse der nur nach Elementen unterschiedenen Statistik, nur Schwefelatome kommen dort zusätzlich vor. Deren Verbesserung von ca. 125 auf $122,5 \cdot 10^{-3}\text{\AA}^2$ ist ausgeprägter als die hier gezeigten Resultate, auch nimmt bei ihnen die Streuung im Modell *mul* ab.

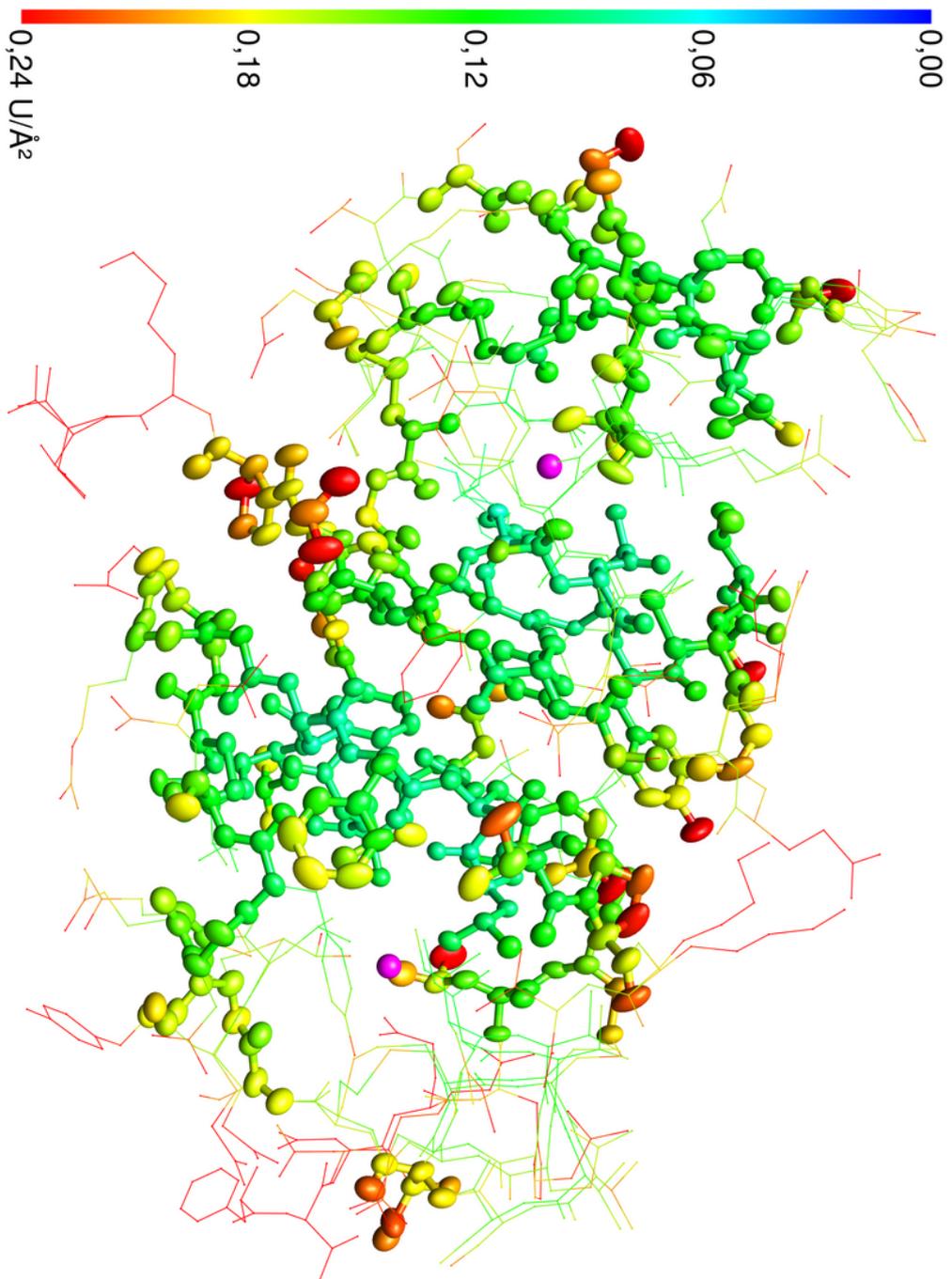


Abbildung 8.6: Anisotrope Verschiebungsparameter des verfeinerten Teils von INSULIN im Modell *ml*.
Farbcode: U_{eq} . Darstellung der massengewichteten Zentren als magentafarbene Kugeln.

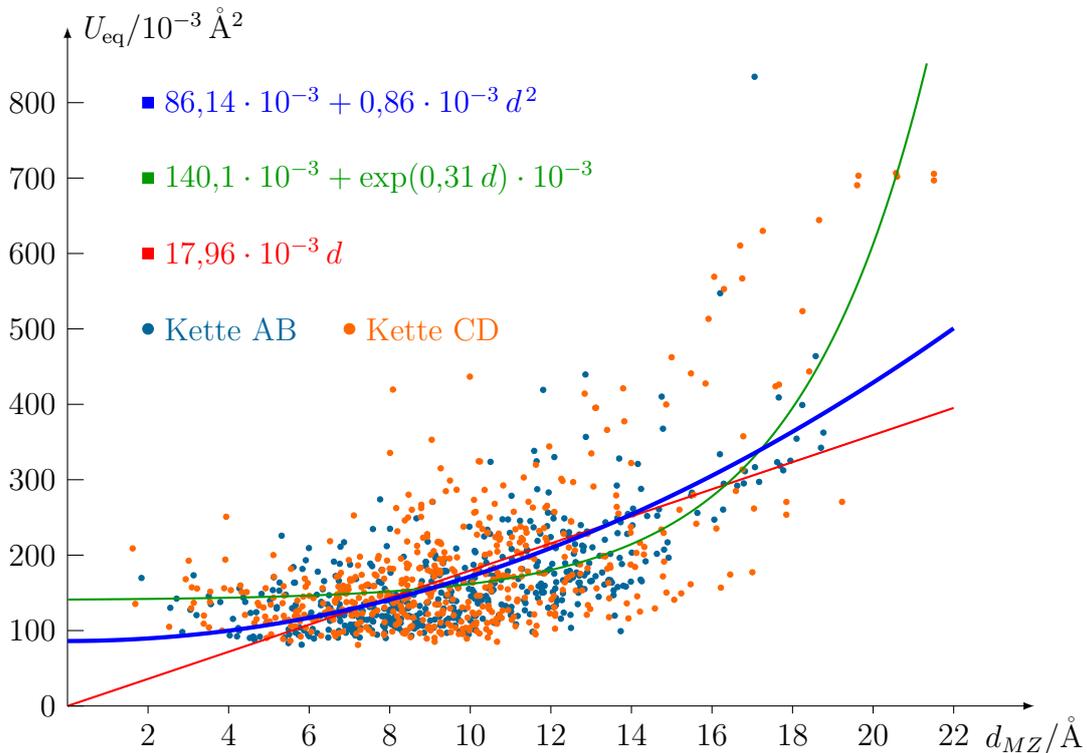


Abbildung 8.7: U_{eq} (mul) der Nichtwasserstoffatome von INSULIN in Abhängigkeit vom Abstand zu den Massenzentren.

Mittels Gl. 6.1 auf Seite 159 berechnen sich die massengewichteten Schwerpunktskoordinaten der beiden Ketten A und B zu $(55,67; 27,11; 4,52) \text{ \AA}$, für die Ketten C und D zu $(51,04; 13,12; -7,47) \text{ \AA}$, sie sind in Abb. 8.6 gekennzeichnet. Der Abstand der 1039 Nichtwasserstoffatome zum jeweiligen Schwerpunkt variiert zwischen $1,62 \text{ \AA}$ und einem Maximalabstand von $21,51 \text{ \AA}$. Über diese Abstände sind in Abb. 8.7 die äquivalenten isotropen Verschiebungsparameter aufgetragen.

Die Punktwolke scheint durch die quadratische Funktion am besten beschrieben, ausgehend von einem U_{eq} von $86 \cdot 10^{-3} \text{ \AA}^2$ steigt sie mit $1,71 \cdot 10^{-3} d \text{ \AA}$ mit der Entfernung vom Schwerpunkt an. Mit ca. 4% Parameterfehler sind die Unsicherheiten allerdings größer als bei Grad- und Exponentialfunktion, deren Koeffizienten einen Fehler von ca. 1% aufweisen. Letztere Funktionen unter- bzw. überschätzen allerdings den extrapolierten U_{eq} -Wert im Massenzentrum. Zwar wurde die Anpassung einer linearen Funktion $a + bx$ versucht, was $15,03(6,89) \cdot 10^{-3} + 16,57(0,67) \cdot 10^{-3} d$ ergab, ob der hohen Parameterunsicherheiten und gleichwertigen Beschreibung aber hier nicht gezeigt wird.

Der gemeinsame Schwerpunkt aller 4 Ketten liegt bei $(53,30; 19,96; -1,61) \text{ \AA}$, Minimal- und Maximalabstände bei $2,08$ und $29,11 \text{ \AA}$. An diese Verteilung angepasste Ausgleichsfunktionen weisen einen wesentlich geringeren Anstieg auf.

8.4 DMSDA-Werte

In den Abb. 8.8 auf der nächsten Seite und 8.9 auf Seite 232 sind die DMSDA-Beträge aller 1084 Bindungen von INSULIN in den Modellen *mul* und *iam* aufgetragen, einmal in der Reihenfolge des Auftretens in der Peptidsequenz, ein andermal nach steigenden Beträgen geordnet. Dazu werden die jeweiligen Unterschiede der Modelle *iam* und *mul* gezeigt, die in den 473 Bindungen auftreten, welche zumindest ein verfeinertes Atom beinhalten. Hierin sind 23 Bindungen zwischen Haupt- und minderpopuliertem Strukturteil enthalten, sie zeigen eher moderate Werte.

Offensichtlich variieren die DMSDA-Beträge sehr stark, ebenso die Differenzen zwischen sphärischen und asphärischen Werten. In den Modellen *iam_{rb}*, *iam* und *mul* betragen die Maximalwerte jeweils $823 \cdot 10^{-4} \text{ \AA}^2$, $833 \cdot 10^{-4} \text{ \AA}^2$ und $712 \cdot 10^{-4} \text{ \AA}^2$, die asphärische Verfeinerung verringert also den Extremwert der Verteilung. Die Unterschiede von *iam_{rb}* und *mul* liegen zwischen -181 und $163 \cdot 10^{-4} \text{ \AA}^2$, der Vergleich von *iam* mit *mul* zeitigt Differenzen von -143 und $204 \cdot 10^{-4} \text{ \AA}^2$. Für die Bindungen des verfeinerten Hauptteils betragen die Mediane in allen Modellen $\approx 72 \cdot 10^{-4} \text{ \AA}^2$. Entsprechend gering sind die Zentralwerte der Modelldifferenzen, gegenüber *iam* und *iam_{rb}* lauten sie -4 und $-3 \cdot 10^{-4} \text{ \AA}^2$. Knapp über die Hälfte der DMSDA-Beträge wird in der Verfeinerung *mul* geringer. Hauptsächlich hohe sphärische Werte erfahren eine Verringerung, und zwar umso stärker, je höher der DMSDA-Betrag ist. Allerdings liegen in diesen Bereichen auch die merklichsten Zunahmen zum asphärischen Modell.

Verschiedene Bindungstypen haben unterschiedliche DMSDA-Werte und -Differenzen. Tab. D.3 auf Seite 374 zeigt dies für die fünf nach Elementen geordneten verfeinerten Bindungen der Hauptpopulation. Die 217 Kohlenstoff-Kohlenstoff-Bindungen zeigen die stärkste Abnahme, im Median beträgt diese $-5 \cdot 10^{-4} \text{ \AA}^2$. Dieser Bindungstyp weist die höchsten Maximalwerte auf, die Mediane liegen im Modell *mul* bei 80, im Fall *iam* und *iam_{rb}* bei 82 und $77 \cdot 10^{-4} \text{ \AA}^2$. Auch die 144 Kohlenstoff-Stickstoff-Bindungen zeigen negative Differenzen von -4 und $-5 \cdot 10^{-4} \text{ \AA}^2$, die sphärischen Mediane liegen bei 58 bzw. $62,5 \cdot 10^{-4} \text{ \AA}^2$, nach der Verfeinerung *mul* bei $57,5 \cdot 10^{-4} \text{ \AA}^2$. Eine Verschlechterung erfährt hingegen der Durchschnitt der 76 Kohlenstoff-Sauerstoff-Bindungen, während in den Modellen *iam* und *iam_{rb}* der Median noch 73 bzw. $80 \cdot 10^{-4} \text{ \AA}^2$ betrug, liegt er in *mul* bei $83,5 \cdot 10^{-4} \text{ \AA}^2$. Das Mittel der Differenz zu *iam* ist mit $3 \cdot 10^{-4} \text{ \AA}^2$ ebenfalls positiv, zu *iam_{rb}* beträgt es $-0,5 \cdot 10^{-4} \text{ \AA}^2$. Besonders stark ist die DMSDA-Zunahme auf den 8 verfeinerten Kohlenstoff-Schwefel-Bindungen, in der Reihenfolge *iam*, *iam_{rb}*, *mul* lauten die Mediane 43, 33 und $46 \cdot 10^{-4} \text{ \AA}^2$. Für die Veränderungen von *iam* und *iam_{rb}* zu *mul* gar 12 und $31,5 \cdot 10^{-4} \text{ \AA}^2$. Auf den 5 Disulfidbrücken sinkt der Durchschnitt wiederum, von 50 oder $51 \cdot 10^{-4} \text{ \AA}^2$ in *iam* bzw. *iam_{rb}* auf $47 \cdot 10^{-4} \text{ \AA}^2$ im Modell *mul*, bedingt durch zwei bis drei verbesserte Bindungen.

Ganz ähnliche Entwicklungen sind für die Auswahl an Peptidbindungen in

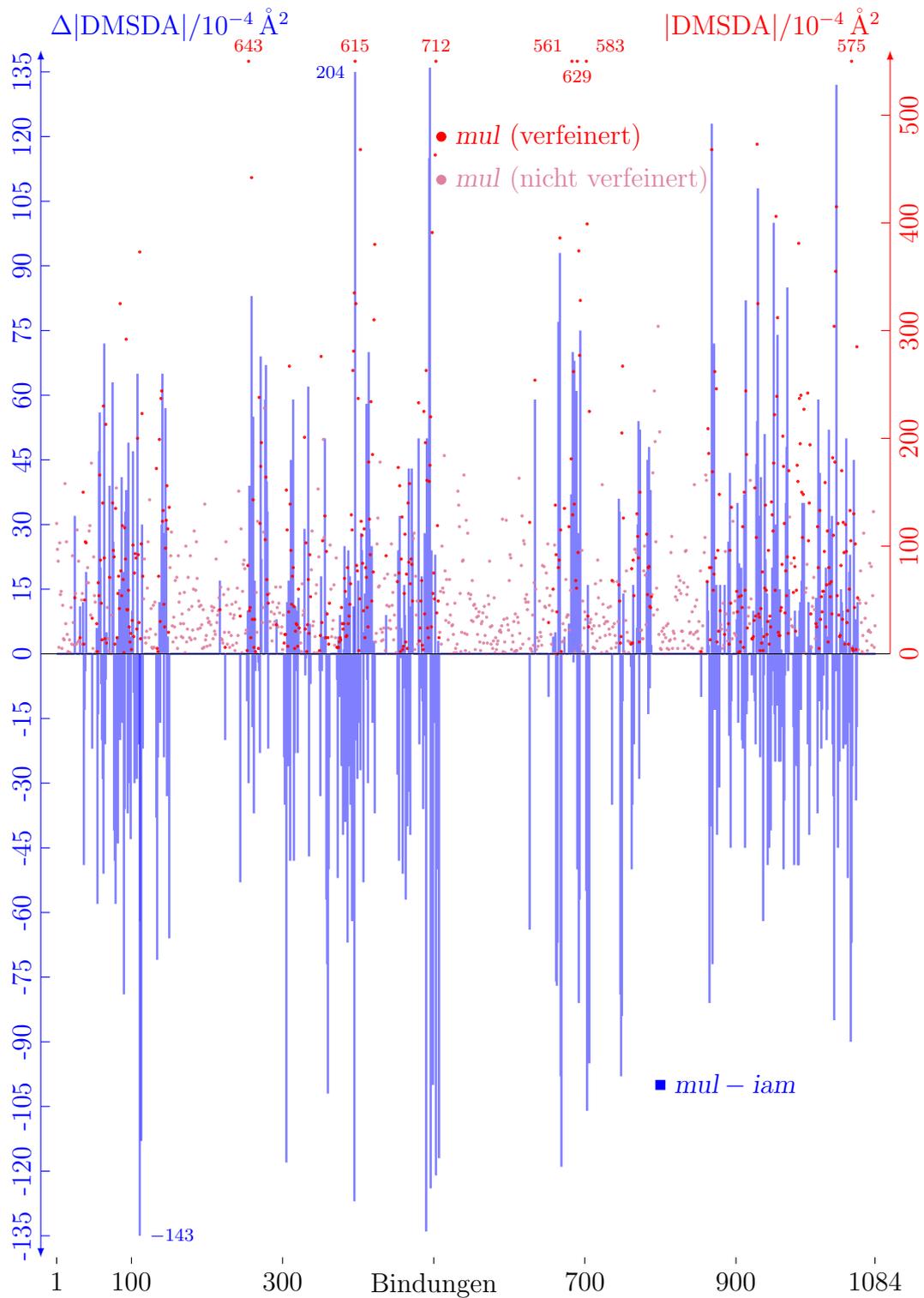


Abbildung 8.8: Veränderung der DMSDA-Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die Nichtwasserstoffatome von INSULIN.

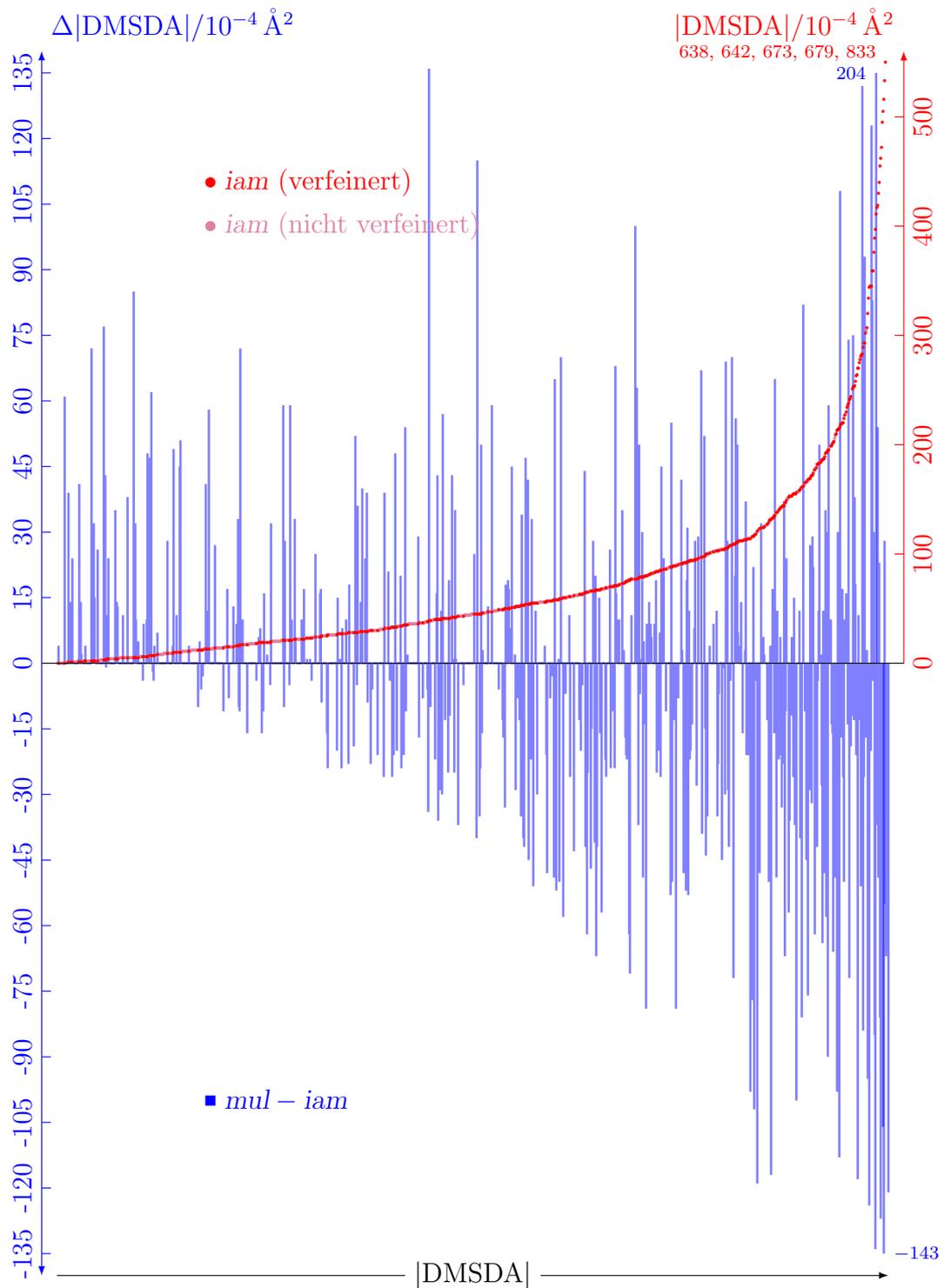


Abbildung 8.9: Veränderung der DMSDA-Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die Nichtwasserstoffatome von INSULIN. Auftragung nach steigenden $|DMSDA|$ -Werten des Modells *iam*.

Tabelle 8.6: DMSDA-Beträge der verfeinerten Peptidbindungen der INSULIN-Hauptpopulation und deren Differenzen.

	$x_{min/max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–O ($n = 65$)							
<i>mul</i>	2/380	105,10(90,96)	39,00	80,00	154,00	115,00	83,03
<i>iam</i>	1/417	103,00(90,97)	35,00	71,00	155,00	120,00	65,23
<i>iam_{rb}</i>	0/337	102,10(85,08)	38,00	79,00	155,00	117,00	74,13
<i>mul-iam</i>	-95/75	2,11(30,32)	-20,00	1,00	19,00	39,00	28,17
<i>mul-iam_{rb}</i>	-87/108	3,08(32,44)	-15,00	0,00	19,00	34,00	25,20
C–N ($n = 74$)							
<i>mul</i>	0/310	77,03(69,73)	29,00	53,50	92,00	63,00	50,41
<i>iam</i>	1/334	85,00(77,86)	32,25	63,00	104,50	72,25	48,18
<i>iam_{rb}</i>	2/315	85,43(79,23)	28,25	62,00	100,00	71,75	54,86
<i>mul-iam</i>	-118/72	-7,97(37,52)	-34,00	-7,50	15,75	49,75	34,84
<i>mul-iam_{rb}</i>	-154/81	-8,40(43,37)	-36,00	-16,50	22,75	58,75	48,18
C_α–C ($n = 59$)							
<i>mul</i>	1/629	108,10(131,50)	22,50	77,00	131,50	109,00	83,03
<i>iam</i>	2/679	118,20(139,18)	51,50	80,00	121,00	69,50	48,93
<i>iam_{rb}</i>	0/724	122,10(145,14)	48,50	83,00	128,00	79,50	51,89
<i>mul-iam</i>	-143/115	-10,12(50,67)	-49,50	-19,00	31,00	80,50	53,37
<i>mul-iam_{rb}</i>	-181/114	-14,02(59,28)	-56,00	-29,00	33,00	89,00	60,79
C_α–N ($n = 57$)							
<i>mul</i>	2/583	82,40(94,16)	17,00	58,00	121,00	104,00	66,72
<i>iam</i>	1/638	80,02(99,59)	19,00	50,00	112,00	93,00	56,34
<i>iam_{rb}</i>	1/566	77,98(90,88)	20,00	60,00	107,00	87,00	63,75
<i>mul-iam</i>	-55/63	2,39(25,08)	-13,00	-1,00	14,00	27,00	22,24
<i>mul-iam_{rb}</i>	-50/58	4,42(24,92)	-7,00	3,00	23,00	30,00	20,76

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der DMSDA-Beträge;
[‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-4}Å^2 .

Tab. 8.6 zu beobachten, bloß bei meist niedrigeren DMSDA-Beträgen und ausgeprägteren Differenzen. Anhand der Lage unterer und oberer Quartilsgrenzen ist gut zu erkennen, daß jeder Bindungstyp sowohl Verbesserungen als auch Verschlechterungen aufweist. Zudem unterscheiden sich nach Elementen gleiche, in „Nächster-Nachbar-Näherung“ aber verschiedene, Bindungstypen in ihrem Verhalten, wie der Vergleich von C–N- und C_α–N-Bindungen nahelegt. Zwischen durchschnittlicher Höhe der sphärischen DMSDA-Beträge und der Änderung mit sphärischer Verfeinerung kann keine Korrelation festgestellt werden.

Tabelle 8.7: Statistik der Restdichten INSULIN.

	$x_{min/max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
<i>mul</i>	-690,3/499,2	0,2(85,0)	-55,7	-1,5	54,8	110,5	81,8
<i>iam</i>	-792,9/500,9	0,2(86,4)	-56,7	-1,3	55,8	112,5	83,3
<i>iam_{rb}</i>	-788,9/495,9	0,2(86,2)	-56,8	-1,3	55,7	112,5	83,3

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der Restdichte;

[‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in $10^{-3} \text{ e}/\text{\AA}^3$.

8.5 Restdichten

Für die drei Modelle von INSULIN wurden die Restdichten auf einem dem Kristallsystem angepaßten Gitter berechnet. Die entsprechend der Auflösung gewählte Feinheit des Gitters führt zu $270 \times 270 \times 114$ Punkten entlang der Achsen a, b, c . In der asymmetrischen Einheit umfaßt das Gitter somit 1244918 Voxel, jedes mit einem Einzelvolumen von $0,0235 \text{ \AA}^3$. Die integrierte normierte Restelektronendichte $\sum |\Delta\rho_{res}|V_{vox}$ beträgt 1487 e nach der sphärischen Verfeinerung *iam* und 1485 e für das Modell *iam_{rb}*. Mit der asphärischen Verfeinerung *mul* nimmt sie auf 1464 e leicht ab. Dementsprechend beläuft sich die prozentuale Verbesserung auf 1,4 % bis 1,6 %. Auf das Volumen der asymmetrischen Einheit bezogen, ergibt sich eine zwischen den Strukturen vergleichbare Größe. Diese beträgt im Modell *mul* $0,0676$, in den Modellen *iam* und *iam_{rb}* $0,0686$ sowie $0,0685 \text{ e}/\text{\AA}^3$.

Anhand der Statistik der Restdichten in Tab. 8.7 und den Histogrammen der Abb. 8.10 auf der gegenüberliegenden Seite wird die Betrachtung vertieft und um den Aspekt der Vorzeichen der Dichten erweitert. In allen Modellen liegen die positiven Maximalreste bei etwa $0,5 \text{ e}/\text{\AA}^3$, während die negativen Extrema der sphärischen Verfeinerungen mit fast $-0,8 \text{ e}/\text{\AA}^3$ etwas ausgeprägter als im asphärischen Fall mit $-0,7 \text{ e}/\text{\AA}^3$ sind. Allerdings zeigen die Histogramme, daß es sich hierbei nur um eine einstellige Anzahl von Gitterpunkten handelt. Zwar zeigen die arithmetischen Mittelwerte keine Unterschiede zwischen den Verteilungen, aber anhand der Mediane erkennt man einen leicht negativeren Durchschnitt der Restdichten nach asphärischer Verfeinerung, wobei alle Streuungsindikatoren dieses Modells eine Abnahme der Verteilungsbreite um das Mittel erkennen lassen. So ist in allen Modellen das Restdichteintervall $(-0,01; 0,00] \text{ e}/\text{\AA}^3$ am stärksten populiert, mit 4,92 % in *mul* allerdings höher als in den sphärischen Modellen, die dort 4,82 % der Werte aus *iam* und 4,81 % im Fall *iam_{rb}* aufweisen. Dies bestätigt die vorherige Statistik bezüglich der allgemein eher negativen Restdichten und deren Zunahme im Modell *mul*. Auch die leichte Verbesserung durch die asphärische Verfeinerung spiegelt sich wieder, so liegen 9,8 % der Werte aus *mul* im vorteilhaftesten Intervall von $(-0,01; 0,01] \text{ e}/\text{\AA}^3$, aber nur ca. 9,6 % der sphärischen.

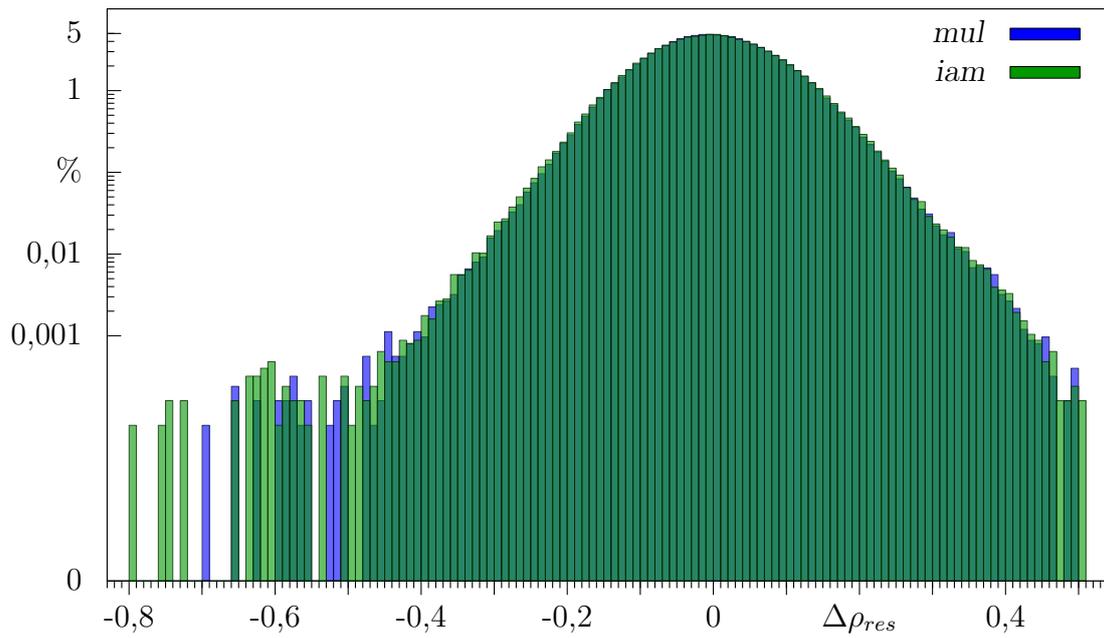
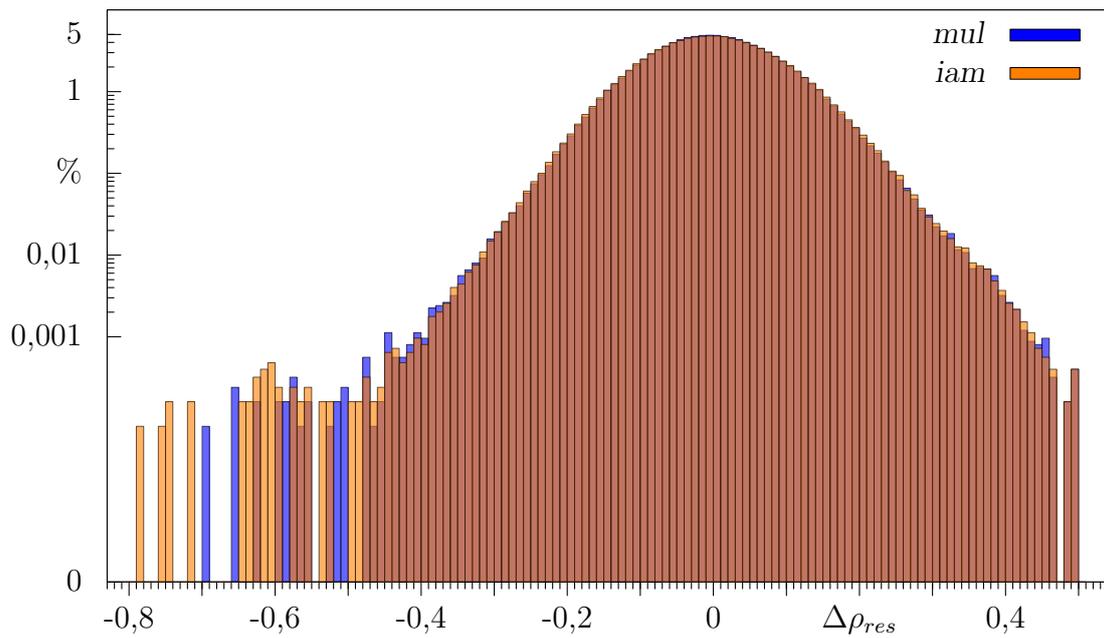
(a) Restdichten INSULIN, *mul* vs *iam*.(b) Restdichten INSULIN, *mul* vs *iam_{rb}*.

Abbildung 8.10: Histogramm der Restdichtegitter $\sum(F_o - F_c) \exp(i\phi_c)$ der sphärischen (grün/orange) und asphärischen (blau) Modelle. Intervalle $\Delta(i; i+0,01]$, Einheiten $e\text{\AA}^{-3}$.

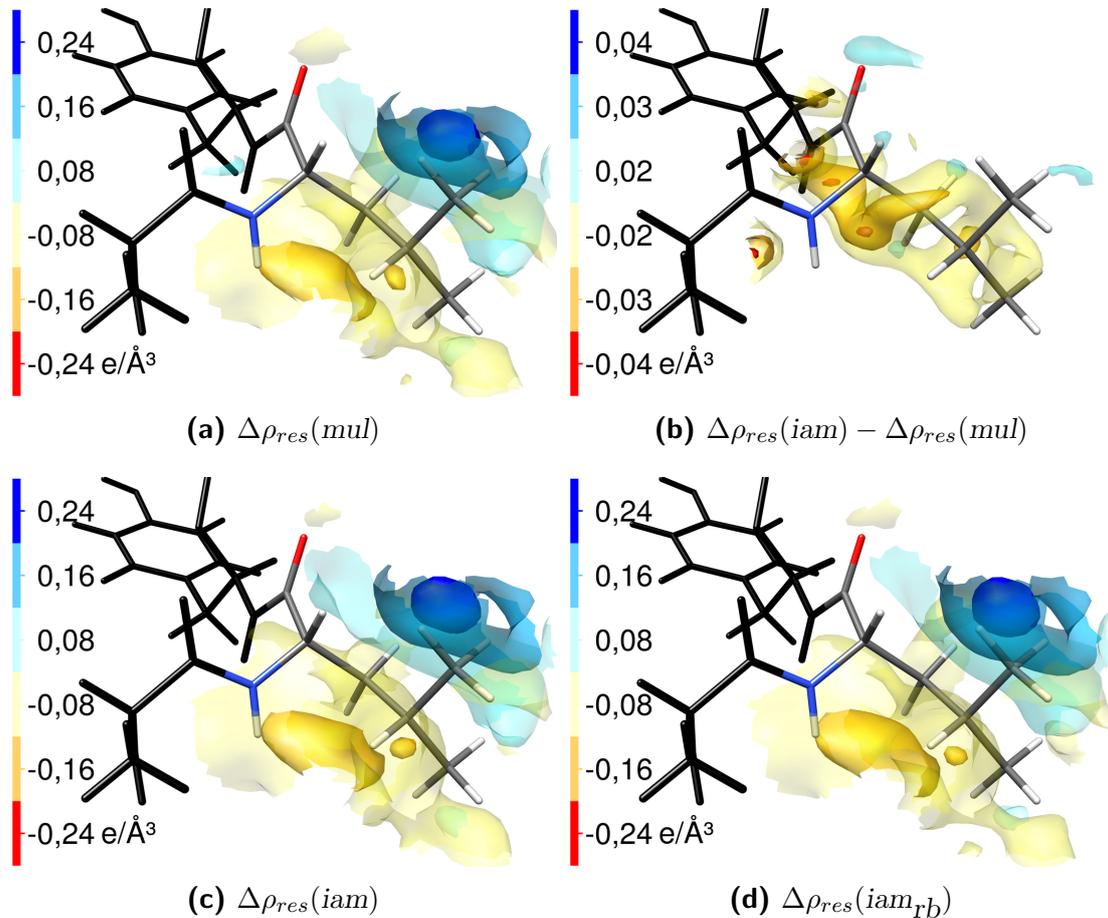


Abbildung 8.11: Restdichten und Restdichtedifferenzen bis $1,6 \text{ \AA}$ um den INSULIN-Leucinrest 215 für sphärische und asphärische Verfeinerungen.

Der Blick auf die Restdichten mittels Isooberflächendarstellungen läßt eher deren hohe Ähnlichkeit in allen Modellen erkennen. Zudem ist die ausgeprägt homogene Verteilung negativer und positiver Reste erkennbar – Dichteansammlungen beider Vorzeichen, vornehmlich in der Größenordnung $0,1\text{--}0,2 e/\text{\AA}^3$, sind willkürlich in der Struktur verteilt. So zumindest der Eindruck anhand der beispielhaft gegebenen Aminosäurereste Leucin 215 und Tyrosin 226 in den Abb. 8.11 und 8.12 auf der gegenüberliegenden Seite, der sich in den Illustrationen des gesamten Moleküls noch verstärkt. Letztere sind in Abb. 8.13 auf Seite 238 für die asphärische Verfeinerung und in Abb. 8.14 auf Seite 239 für das Modell *iam* gegeben. Eine dritte Darstellung zeigt Abb. D.2 auf Seite 375 für *iam_{rb}*, die von den ersten beiden ebensowenig zu unterscheiden ist wie diese voneinander.

Ob der geringen Modellunterschiede sind die Isooberflächenwerte der Restdichtedifferenzkarten bei ungefähr einem Fünftel der voneinander abgezogenen Dichtewerte gewählt.

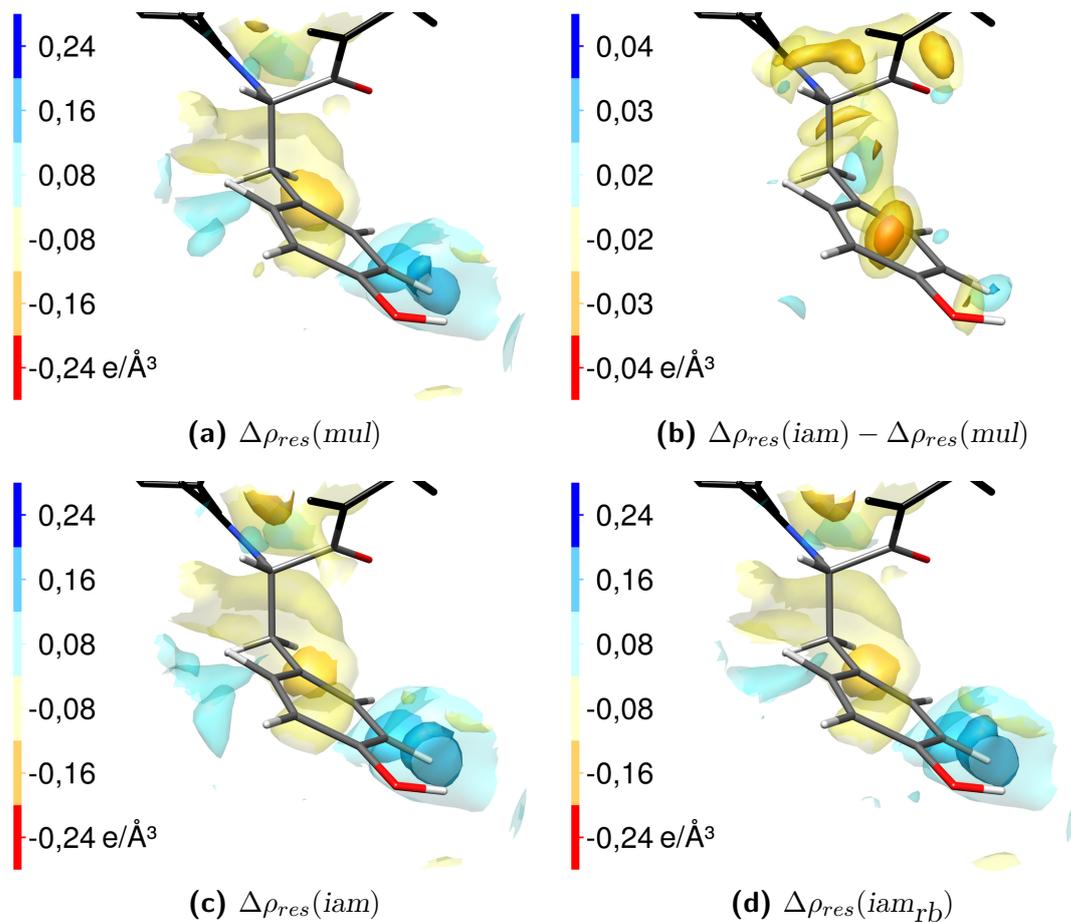


Abbildung 8.12: Restdichten und Restdichtedifferenzen bis $1,6 \text{ \AA}$ um den INSULIN-Tyrosinrest 226 für sphärische und asphärische Verfeinerungen.

In keinem der Modelle zeigen sich nennenswerte Restdichten direkt auf der Struktur, insbesondere keine auf Bindungsdichten hinweisenden positiven. Vielmehr sind die Hauptmerkmale in einiger Entfernung zum Molekül lokalisiert, wie am Leucinrest in Abb. 8.11 besonders gut zu sehen ist. Deshalb liegen die Isooberflächen der Differenz $\Delta\rho_{res}(iam) - \Delta\rho_{res}(mul)$ primär neben Bindungen und Atomen. Die durch die asphärische Verfeinerung hervorgerufenen Änderungen scheinen sich hauptsächlich auf die Reduktion negativer Restdichten zu konzentrieren, wie z. B. die Ringmitte des Tyrosins zeigt. Andererseits erhöhen sich auf den Bindungen die negativen Restdichten, da dies in der Größenordnung von $0,01 \text{ e/\AA}^3$ stattfindet, ist es hier nicht zu erkennen. Eine Zunahme negativer Restdichte offenbart auch die Region am Carbonylsauerstoff des Leucinrestes, sie wird in der Differenzkarte als positive Oberfläche deutlich. Sofern im sphärischen Modell positive Restdichten an Atomen oder Bindungen lokalisiert sind, wie es am H_c des Tyrosinringes der Fall ist, werden diese in der asphärischen Verfeinerung reduziert.

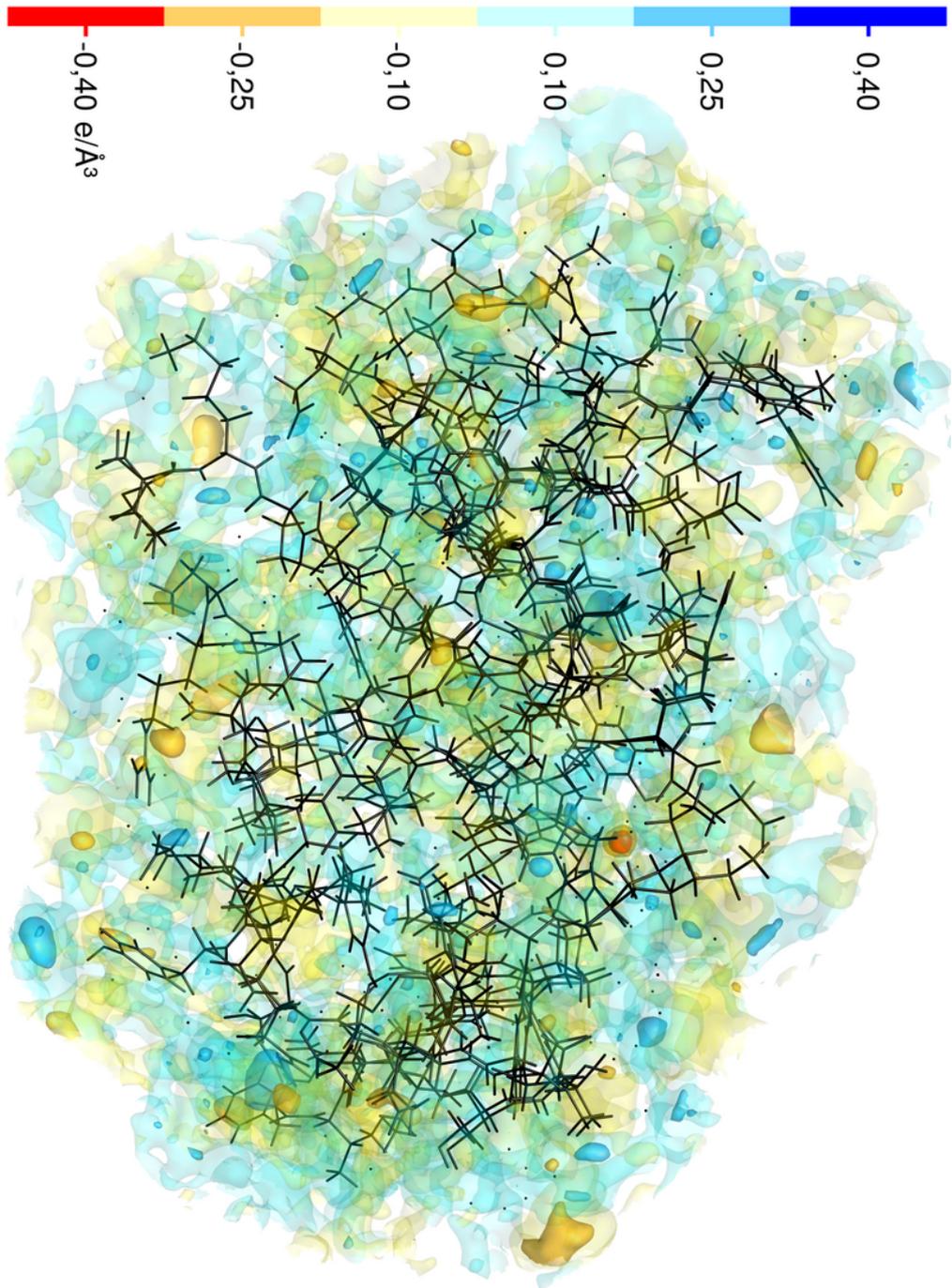


Abbildung 8.13: Restdichten von INSULIN im Modell *ml*, Darstellung bei sechs Isooberflächenwerten bis zu 4 Å um die Struktur.

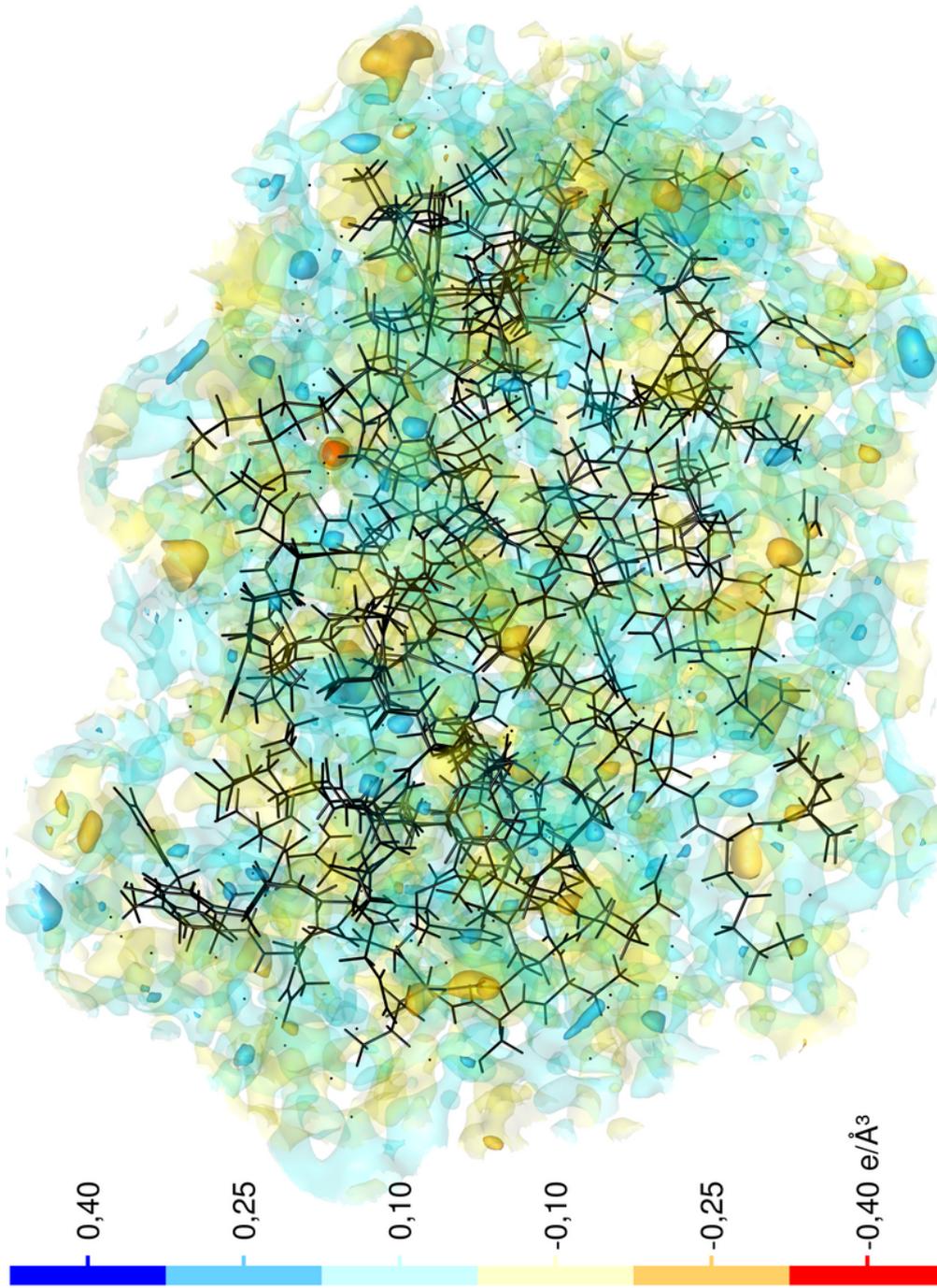


Abbildung 8.14: Restdichten von INSULIN im Modell iam, Darstellung bei sechs Isooberflächenwerten bis zu 4 Å um die Struktur.

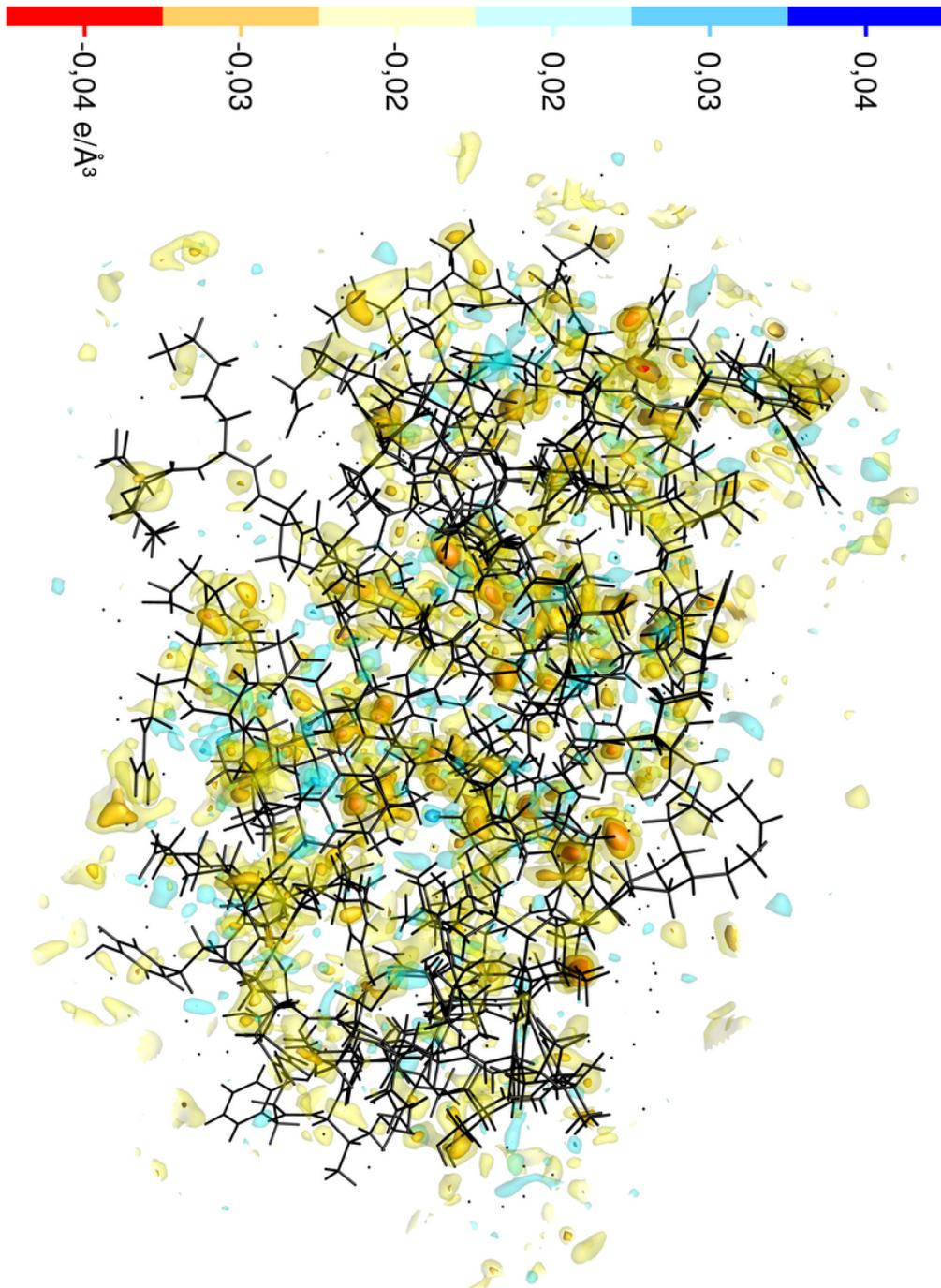


Abbildung 8.15: Restdichtedifferenzen *iam-mul* 4 Å um INSULIN.

8.6 Deformationsdichten

Die den Unterschied des asphärischen Modells zum sphärischen Modell konstituierende Deformationselektronendichte $\Delta\rho_{def}$ bildet die Graphik 8.16 für die im vorherigen Abschnitt eingeführten Aminosäurereste Leucin 215 und Tyrosin 226 ab, eine analoge Darstellung der gesamten INSULIN-Struktur zeigt Abb. 8.17 auf der folgenden Seite.

Mit der Abbildung der Molekülausschnitte sind die Eigenschaften der Deformationsdichte besser kenntlich und ebenso, daß die Ausrichtung der asphärischen Dichten ordnungsgemäß erfolgte, da die Valenzdichteverteilung den Erwartungen einer Konzentration auf den Bindungen und an den freien Elektronenpaaren entspricht. Anhand der Gesamtabbildung wird wiederum nochmal gezeigt, daß der Solvensbereich sphärisch behandelt wurde – nur die eigentlichen Insulinketten, wurden mit Invariomen belegt. Weiterhin sind Anhäufung positiver wie negativer Deformationsdichte zu sehen, die als Effekt dicht beieinanderliegender Fehlordnung gedeutet werden.

Die Summation aller Beträge der Deformationsdichte auf einem die Struktur umschließenden Gitter resultiert in insgesamt 752 e. Somit entfallen auf jedes der 1572 asphärisch behandelten Atome, verteilt auf 2055 Lagen, durchschnittlich 0,48 e Valenzdichte, die im Modell *mul* anders als in den sphärischen Modellen verteilt ist, um ihrer Deformation durch Bindungseffekte gerecht zu werden.

Zwischen der Deformationsdichte und der Restdichte, besteht ein Verhältnis von $\Delta\rho_{def}/\Delta\rho_{res} = 752\text{ e}/1464\text{ e} = 0,514$. Somit zeichnet sich das Modell *mul* durch doppelt so viel unbeschriebene Dichte wie asphärisch beschriebene aus.

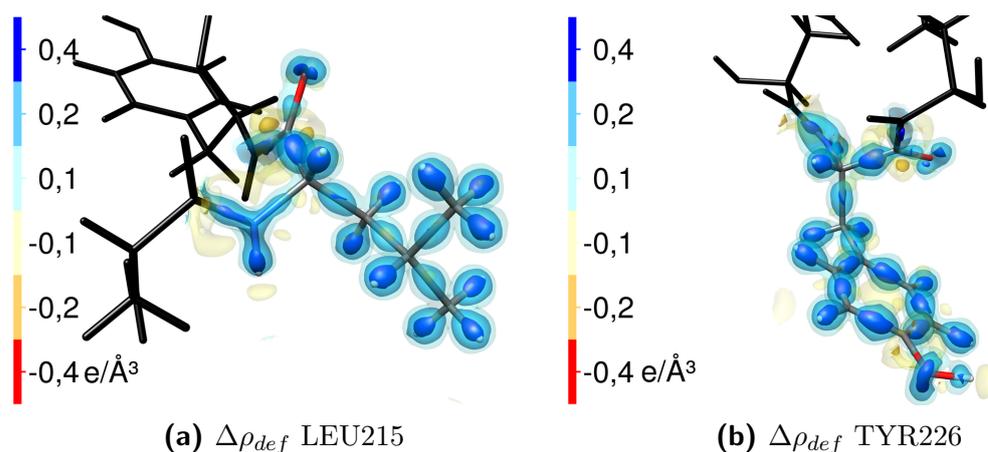


Abbildung 8.16: Deformationsdichten der INSULIN-Reste Leucin 215 und Tyrosin 226.

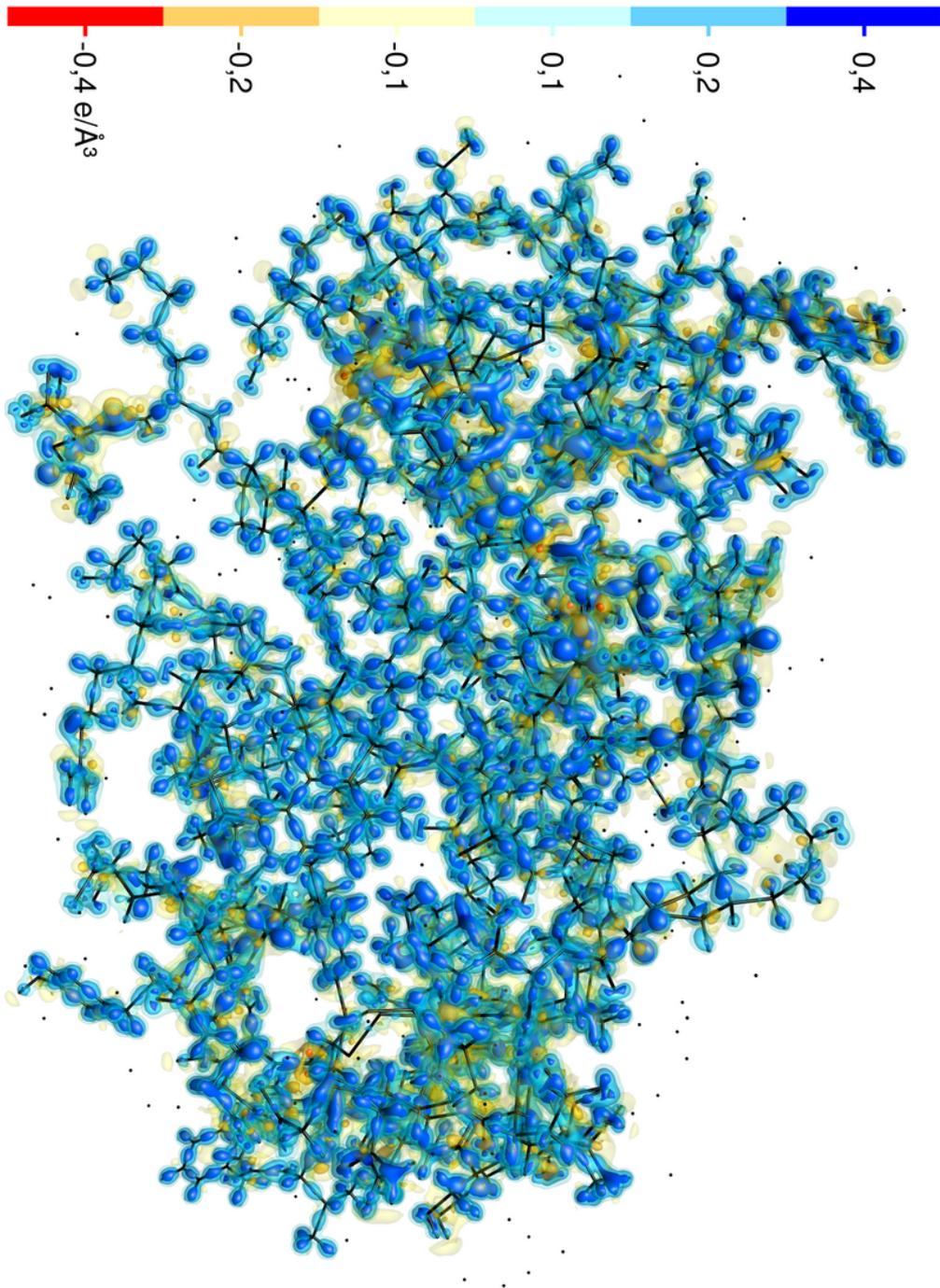


Abbildung 8.17: Deformationsdichte $\Delta\rho_{def}$ von INSULIN.

8.7 Abgeleitete Eigenschaften

In Abb. 8.18 auf der nächsten Seite ist das elektrostatische Potential von INSULIN auf der $0,001 e/a_0^3$ -Elektronendichteoberfläche gezeigt, wie jeweils aus den Multipolen der Hauptpopulation des Modells *mul* abgeleitet. Dabei ist es in zwei Orientierungen dargestellt, um die recht unterschiedlichen Werte der „Außen-“ und „Innenseite“ zu illustrieren. Erstere weist in der Hexamer-Quartärstruktur nach außen (obere Abb.), letztere zum Inneren des Hexamers und zur dreizähligen Achse, auf der auch das Zink liegt (untere Abb.).

Der Gesamtwertebereich erstreckt sich zwischen $-0,2128$ und $0,7073 e/a_0$. Auf der „Außenseite“ sind negative und positive Potentiale der Größenordnung $\pm 0,2 e/a_0$ recht ausgeglichen verteilt, die „Innenseite“ ist aber von positiven ESP-Werten dominiert, deren höchste am Zn^{2+} vorliegen. Noch stärker wird dies von Abb. 8.19(a) auf Seite 245 und 8.20(a) auf Seite 246 betont, welche die vorherigen Darstellungen in einem eingeschränkten Wertebereich von $\pm 0,18 e/a_0$ (± 191 kT/e) wiederholen.

Ihnen wird jeweils ein Punktladungs-Potential ϕ_{PC} gegenübergestellt, welches mittels *delphi* aus der Poisson-Boltzmann-Gleichung berechnet wurde, wozu jedem Atom eine der *amber03*-Parametrisierung entsprechende Ladung zugewiesen wurde. Die „Außenseite“ wird in Abb. 8.19 gezeigt, die „Innenseite“ in 8.20. Hierbei sind die Abb. 8.19/8.20(b) jeweils direkt mit den Multipolergebnissen ϕ_{mul} zu vergleichen, da die Potentiale auf der gleichen $0,001 e/a_0^3$ -Isooberfläche der Multipolelektronendichte abgebildet sind. Alternativ können Abb. 8.19(c) und (d) sowie Abb. 8.20(c) und (d) einem Vergleich unterzogen werden, bei ihnen sind die Potentiale auf der van-der-Waals-Oberfläche A_{vdW} wiedergegeben. Jeweils der volle Wertebereich des Punktladungs-ESPs auf der $0,001 e/a_0^3$ -Oberfläche wird in Abb. 8.19/8.20(e) gezeigt, der auf der vdW-Hülle in Abb. 8.19/8.20(f).

Das aus Punktladungen bestimmte elektrostatische Potential weist vermehrt elektronegative Bereiche auf, bei gleichzeitig etwas stärker negativen Werten. Im Gegenzug sind weniger elektropositive Flächen zu finden, zugleich verringert sich der Maximalwert des ESPs. So lautet der Potentialumfang von ϕ_{PC} auf der $0,001 e/a_0^3$ -Elektronendichteoberfläche $-0,2133$ – $0,5766 e/a_0$ (-226 – 611 kT/e), wogegen das Multipol-ESP ϕ_{mul} Werte von $-0,2128$ – $0,7073 e/a_0$ (-226 – 750 kT/e) ergibt. Auf der van-der-Waals-Oberfläche sind die Unterschiede ausgeprägter, hier zeigt ϕ_{PC} maximale Potential von $-0,2605 e/a_0$ (-276 kT/e) und $0,6360 e/a_0$ (674 kT/e), während die Extremwerte von ϕ_{mul} $-0,2274 e/a_0$ (-241 kT/e) und $0,8151 e/a_0$ (864 kT/e) betragen. Die höhere Polarisierung auf der vdW-Oberfläche kann auf deren geringeres Volumen von 10950 \AA^3 zurückgeführt werden, weshalb sie näher am Molekül liegt. Mit 13460 \AA^3 umschließt ρ_{mul} ein größeres Volumen und ist etwas weiter von den Atomen entfernt.

Trotz der Differenzen in den Maximalwerten ist die Übereinstimmung beider Potentiale gut, insbesondere die räumliche Verteilung elektropositiver und -negativer Bereiche ist beinahe identisch.

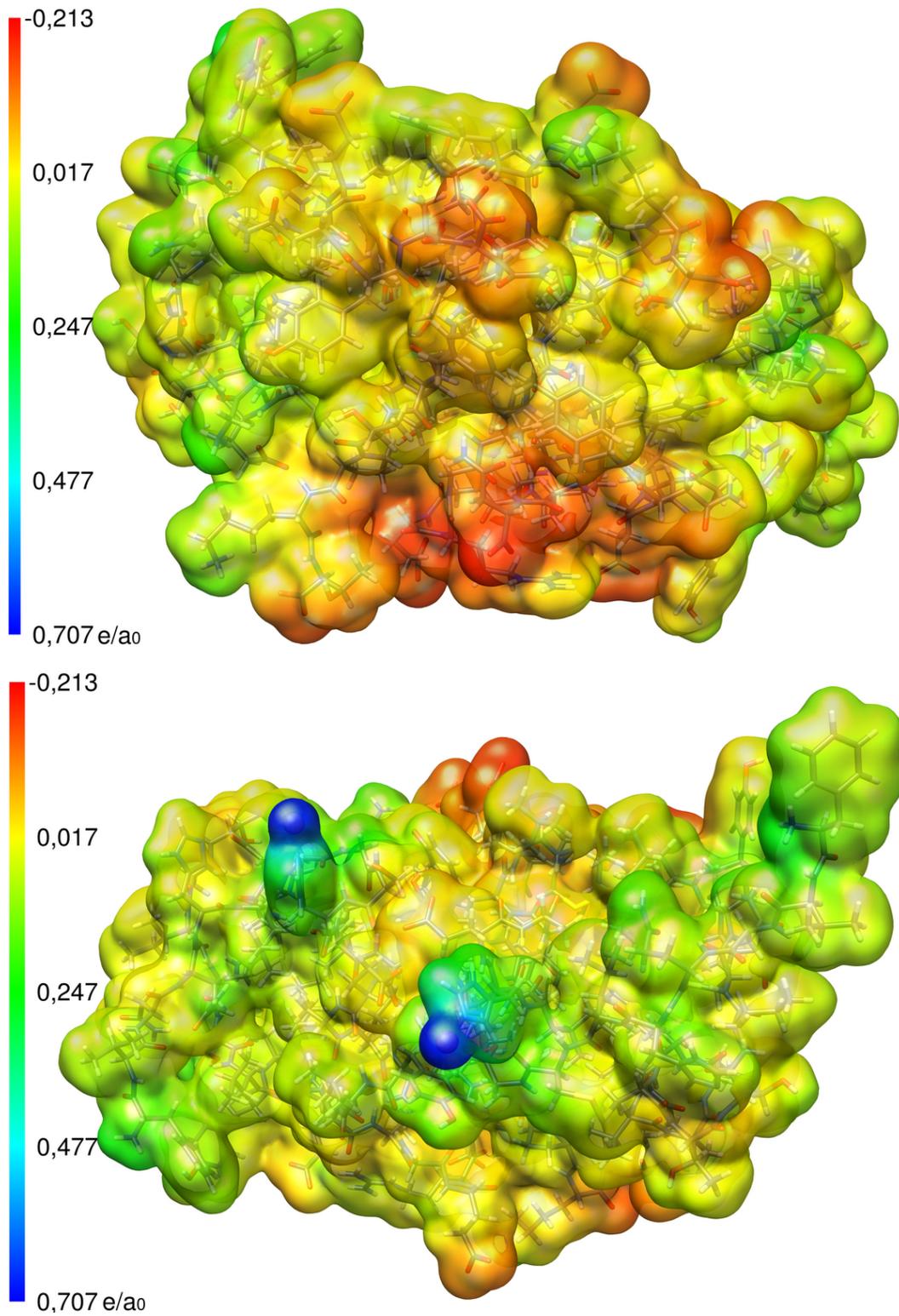


Abbildung 8.18: Elektrostatisches Potential des Multipolmodells von INSULIN auf der $0,001 e/a_0^3$ -Isooberfläche in zwei Orientierungen.

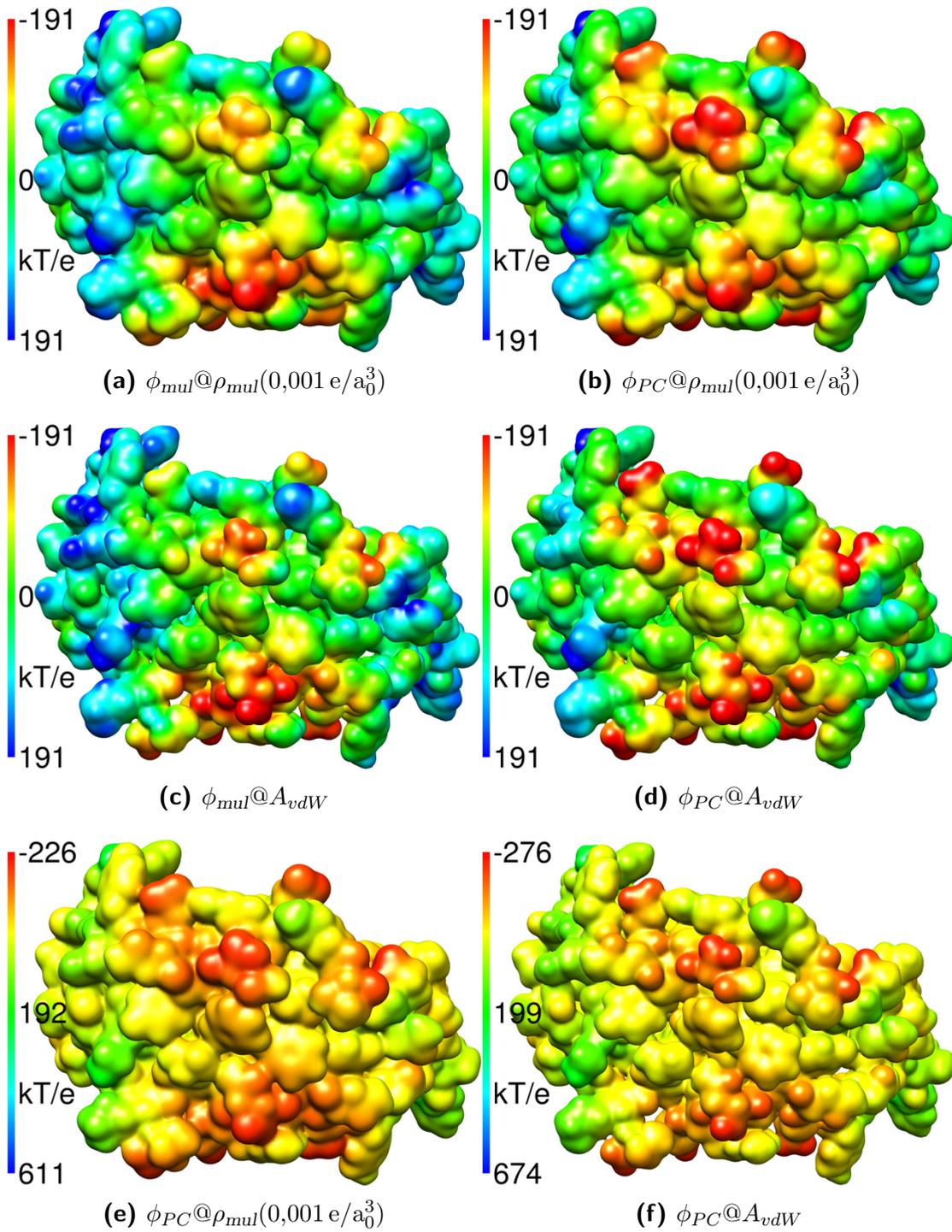


Abbildung 8.19: Multipol-ESP (ϕ_{mul}) und Punktladungs-ESP (ϕ_{PC}) auf $0,001 e/a_0^3$ -Isooberfläche der Multipolelektronendichte (ρ_{mul}) und van-der-Waals-Oberfläche (A_{vdW}). (a)-(d): Fester Wertebereich $\pm 191 \text{ kT/e} = \pm 0,18 e/a_0$. (e)-(f): Voller Wertebereich.

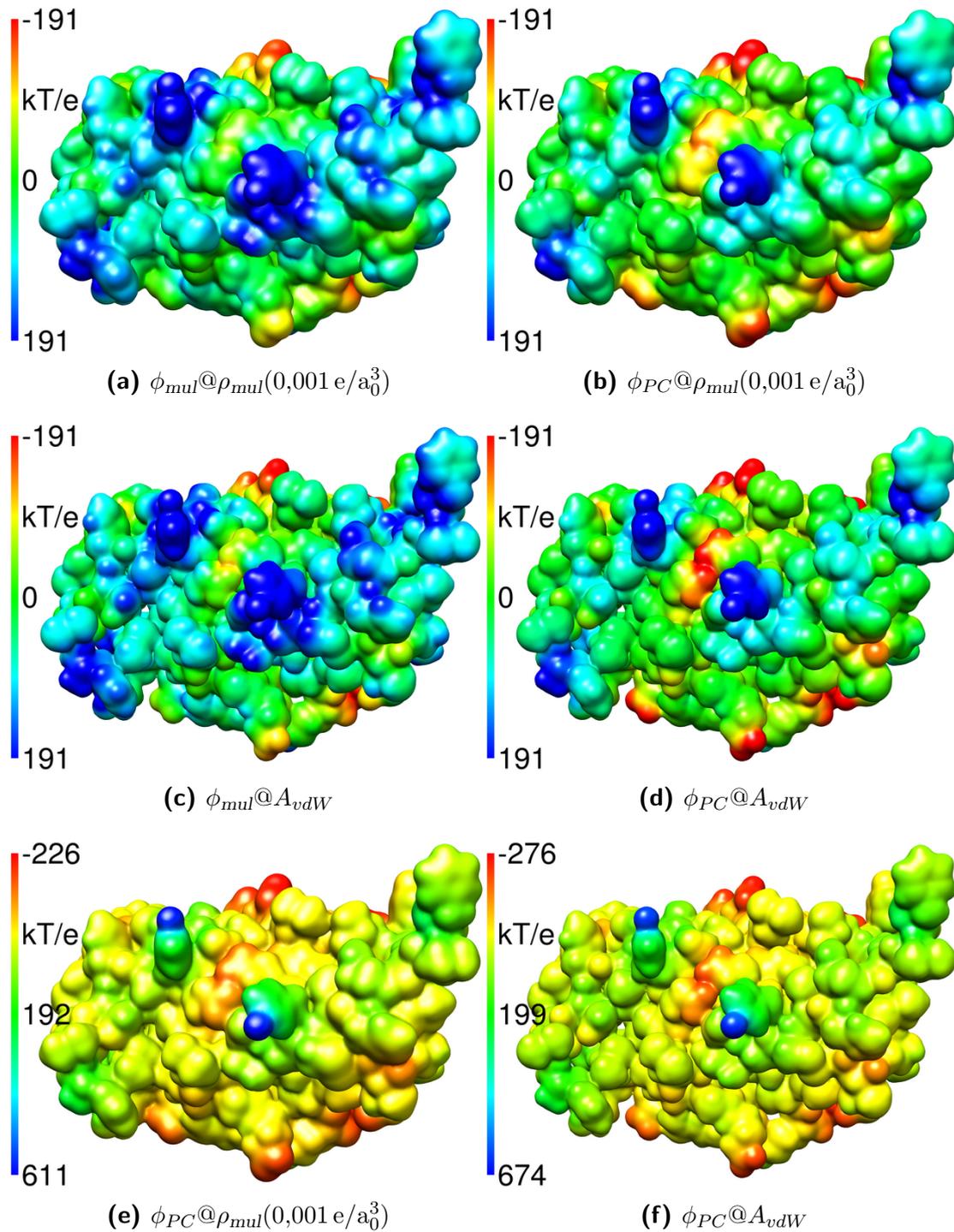


Abbildung 8.20: Multipol-ESP (ϕ_{mul}) und Punktladungs-ESP (ϕ_{PC}) auf $0,001 e/a_0^3$ -Isooberfläche der Multipolelektronendichte (ρ_{mul}) und van-der-Waals-Oberfläche (A_{vdW}). (a)-(d): Fester Wertebereich ± 191 kT/e = $\pm 0,18 e/a_0$. (e)-(f): Voller Wertebereich.

9 Trichotoxin

9.1 Gütefaktoren

Die in Absch. 4.6 auf Seite 105 eingeführten Modelle führen zu den in Tab. 9.1 gezeigten Ergebnissen. Gegenüber den sphärischen Verfeinerungen *iam* und *iam_{rb}* ergeben sich in der asphärischen Modellierung *mul* etwa 4,3 % und 7,2 % niedrigere $R(F)$ - und $R_w(F^2)$ -Werte. Im Vergleich zu *iam* sinken diese in *mul* um 0,21 und 0,45 Prozentpunkte, in Relation zu *iam_{rb}* um fast doppelt so hohe 0,36 und 0,78 Punkte. Dabei zeigen alle Verfeinerungen einen GoF_w nahe Eins und hohe Konvergenz.

Im sphärischen Modell *iam* hat die Gesamtheit aller 14153 Strukturfaktordifferenzen $\Delta F = |F_o - F_c| \sigma(F_o)^{-1}$ ein arithmetisches Mittel von 0,2003 (Median 0,1549), während das asphärische Modell *mul* eine ca. 0,01 geringere Abweichung von 0,1938 (0,1475) zeigt. Ohne Gewichtung steigen die mittleren Differenzen auf das Sechs- bis Siebenfache, die Modellunterschiede auf etwa 0,04.

Diese geringen Unterschiede zugunsten des asphärischen Modells werden für statistisch signifikant befunden: Zu den Testhypothesen $H_0 : \Delta F_w(iam) = \Delta F_w(mul)$ und $H_1 : \Delta F_w(iam) > \Delta F_w(mul)$ ergeben sich im *t*- und Wilcoxon-Test *p*-Werte von 10^{-6} , was ebenfalls für die ungewichteten Differenzen gilt.

Um mit den bisherigen Strukturen verglichen werden zu können, zeigt Abb. 9.1 auf der folgenden Seite den Quotienten F_o^2/F_c^2 des TRICHOTOXIN-Modells *iam*. Hier haben 1497 oder 10,6 % der Reflexe ein Verhältnis von ≥ 2 oder $\leq 1/2$.

Tabelle 9.1: Gütefaktoren der sphärischen und asphärischen Verfeinerung von TRICHOTOXIN.

	<i>iam</i>	<i>iam_{rb}</i>	<i>mul</i>
Reflexe $F_o^2 > 2\sigma(F_o^2)$	10659	10659	10659
Parameter	2098	2098	2098
N_{ref}/N_{par}	5,081	5,081	5,081
$R(F)$	0,0475	0,0490	0,0454
$R_{all}(F)$	0,0683	0,0698	0,0661
$R_w(F^2)$	0,1075	0,1108	0,1030
GoF_w	1,0017	1,0329	0,9599
Mittl.(Versch./esd)	$2,78 \cdot 10^{-6}$	$4,01 \cdot 10^{-6}$	$4,01 \cdot 10^{-6}$

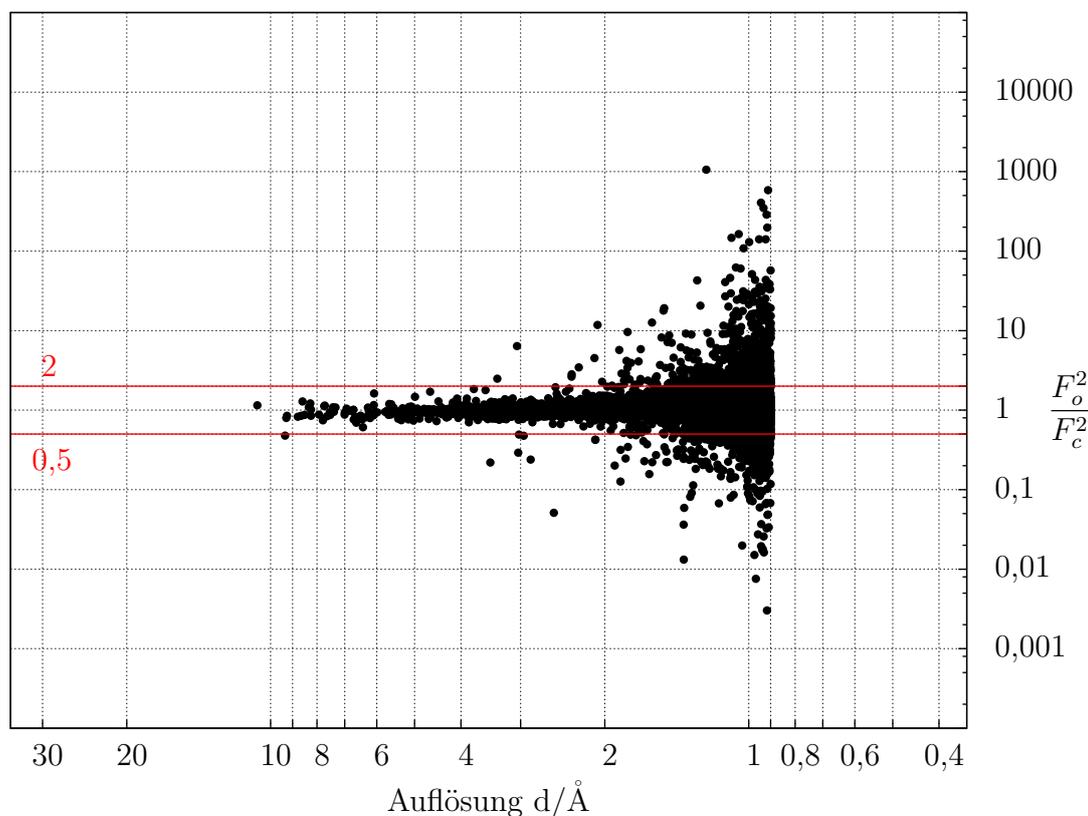


Abbildung 9.1: Quotient beobachteter und berechneter Intensitäten und ihre Auflösungsabhängigkeit im Modell *iam* von TRICHOTOXIN.

9.2 Geometrie

Der durchschnittliche Unterschied der Nichtwasserstoffatomkoordinaten der Verfeinerungen *mul* und *iam* beträgt $0,008 \text{ \AA}$, die Maximaländerung betrifft das N_{e2} des Glutaminrestes 106 mit $0,019 \text{ \AA}$. Unter Einbeziehung der Wasserstoffatompositionen verzehnfacht sich die mittlere Verschiebung auf $0,08 \text{ \AA}$, die stärksten Änderungen betreffen nun die Wasserstoffatome, deren quadratisch gemittelte Abweichung zwischen *mul* und *iam* von $0,11 \text{ \AA}$ bis $0,16 \text{ \AA}$ reicht. Das ist auch gut in Abb. 9.2(a) auf der nächsten Seite zu erkennen: Alle Wasserstoffatome tragen den roten Farbcode für die höchste Abweichung, während die Schweratome durch einen maximal grünen gekennzeichnet sind.

Beim Vergleich mit der Verfeinerung *iam_{rb}* ist die durchschnittliche Verschiebung der Nichtwasserstoffatome etwas geringer, sie beträgt nur noch $0,007 \text{ \AA}$. Die Wasserstoffpositionen sind nun zwischen den Modellen kaum zu unterscheiden, was Abb. 9.2(b) deutlich hervorhebt. Den größten Koordinatenunterschied von $0,022 \text{ \AA}$ weist das $C_{\eta 2}$ des terminalen L-Valinols 118 auf.

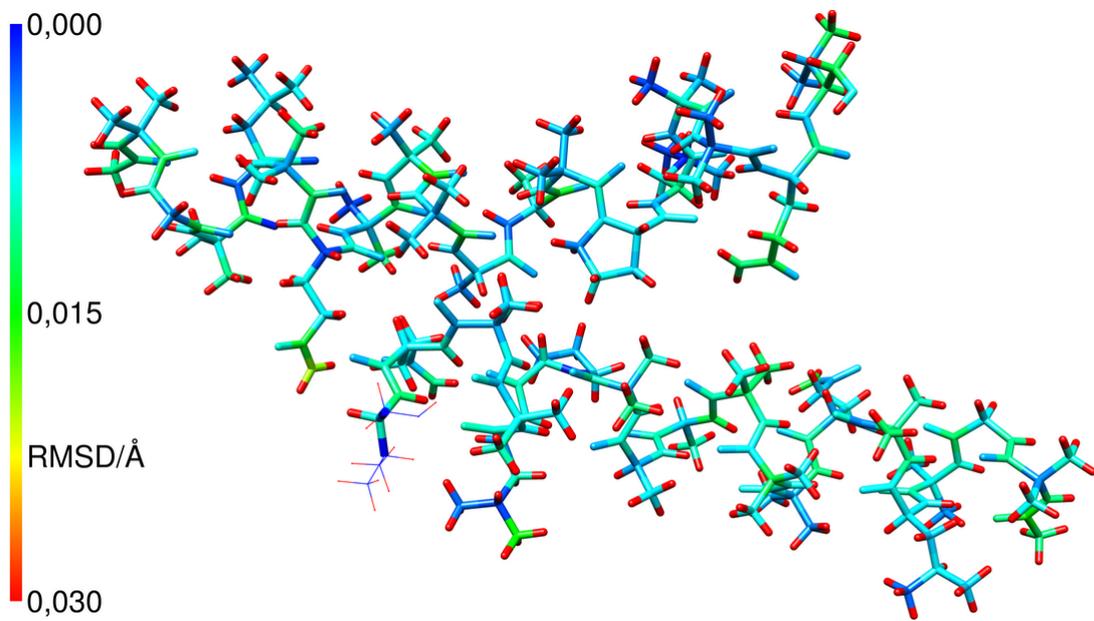
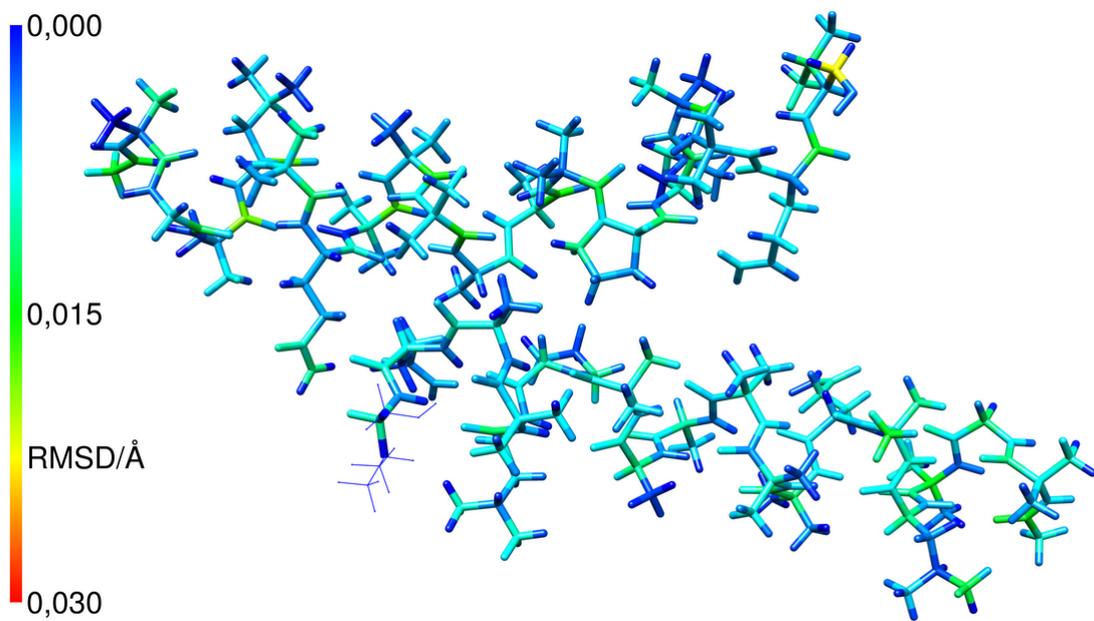
(a) *mul vs iam*(b) *mul vs iam_{rb}*

Abbildung 9.2: Quadratisch gemittelte Abweichungen zwischen asphärischer und spärlicher Verfeinerungen von TRICHOTOXIN. Fixierte Koordinaten als Strichmodell.

Tabelle 9.2: Verfeinerte Bindungslängen der Peptidbindungen von TRICHOTOXIN.

	$d_{E\&H}^*$	$d_{min/max}^\dagger$	$\bar{d}(\sigma_d)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–O ($n = 36$)								
<i>mul</i>		1,204/1,250	1,227(10)	1,219	1,227	1,233	0,014	0,011
<i>iam</i>	1,231(20)	1,210/1,260	1,238(10)	1,230	1,237	1,244	0,014	0,011
<i>iam_{rb}</i>		1,208/1,258	1,238(10)	1,231	1,239	1,244	0,014	0,011
C–N ($n = 34$)								
<i>mul</i>		1,244/1,358	1,330(19)	1,327	1,331	1,341	0,014	0,010
<i>iam</i>	1,329(14)	1,240/1,352	1,326(19)	1,322	1,325	1,337	0,015	0,009
<i>iam_{rb}</i>		1,245/1,356	1,329(19)	1,325	1,328	1,339	0,014	0,012
C_α–C ($n = 14$)								
<i>mul</i>		1,514/1,549	1,532(09)	1,528	1,530	1,535	0,007	0,004
<i>iam</i>	1,525(21)	1,501/1,539	1,520(10)	1,516	1,518	1,521	0,005	0,004
<i>iam_{rb}</i>		1,504/1,541	1,522(10)	1,517	1,521	1,524	0,007	0,006
C_α–N ($n = 14$)								
<i>mul</i>		1,444/1,472	1,454(08)	1,450	1,453	1,458	0,008	0,006
<i>iam</i>	1,458(19)	1,448/1,480	1,458(08)	1,454	1,455	1,460	0,007	0,006
<i>iam_{rb}</i>		1,454/1,480	1,462(07)	1,457	1,462	1,464	0,008	0,007

*Engh&Huber-Bindungslängen mit Standardabweichung; [†]Minimum, Maximum und Mittelwert mit Standardabweichung der Bindungslängen; [‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in Å.

Bezüglich der Abstände von Peptidbindungsatomen, die in „Nächster Nachbar Näherung“ mit der Auswahl für vorherige Strukturen übereinstimmen, können die zuvor erhaltenen Ergebnisse reproduziert werden. Tab. 9.2 zeigt die Bindungslängenstatistik, Tab. 9.3 auf der nächsten Seite die der Differenzen.

Innerhalb einer Standardabweichung stimmen die Abstände aller Modelle mit den Engh&Huber-Mittelwerten überein, mal mit positiven, mal mit negativen Diskrepanzen. Die Carbonylbindung verkürzt sich mit asphärischer Behandlung um 0,01 Å, wobei die Differenzen zu *iam* und *iam_{rb}* sehr ähnlich sind. Im Gegensatz dazu verlängert sich die C_α–C-Bindung um etwa den gleichen Betrag, stärker im Vergleich zum Modell *iam* als zu *iam_{rb}*. Bei Stickstoff enthaltenden Bindungen treten sowohl Ab- wie Zunahmen der Abstände auf. Die C–N-Bindung verlängert sich im Mittel um $5 \cdot 10^{-3}$ Å gegenüber *iam* und um nur $2 \cdot 10^{-3}$ Å verglichen mit *iam_{rb}*. Umgekehrt tritt eine durchschnittliche Verkürzung der C_α–N-Bindung um ca. -4 und $-8 \cdot 10^{-3}$ Å auf, diesmal ausgeprägter für das Modell *iam_{rb}*. Die Verteilungsbreiten der verschiedenen Bindungstypen in den drei Verfeinerungen unterscheiden sich kaum. Wie schon die quadratisch gemittelten Koordinatenunterschiede, liegen auch die Längenänderungen in der Größenordnung des Fehlers.

Tabelle 9.3: Differenzen der verfeinerten Peptidbindungsängen von TRICHOTOXIN.

	$\Delta d_{min/max}^\dagger$	$\overline{\Delta d}(\sigma_\Delta)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–O ($n = 36$)							
<i>mul-iam</i>	–22,3/–5,8	–11,19(3,05)	–12,42	–10,60	–9,50	2,93	2,22
<i>mul-iam_{rb}</i>	–24,0/–4,2	–11,63(3,31)	–14,02	–10,85	–9,38	4,65	2,82
C–N ($n = 34$)							
<i>mul-iam</i>	–2,1/13,0	4,79(3,15)	3,00	5,10	6,83	3,83	2,97
<i>mul-iam_{rb}</i>	–3,4/6,6	1,70(2,62)	–0,15	1,90	3,20	3,35	2,82
C_α–C ($n = 14$)							
<i>mul-iam</i>	7,9/18,2	11,86(2,66)	10,32	11,85	13,25	2,93	2,30
<i>mul-iam_{rb}</i>	6,6/20,1	10,07(3,56)	8,05	8,75	11,30	3,25	2,60
C_α–N ($n = 14$)							
<i>mul-iam</i>	–8,1/4,0	–3,90(2,88)	–5,85	–3,90	–3,18	2,68	1,48
<i>mul-iam_{rb}</i>	–11,0/–3,4	–7,84(2,43)	–9,45	–8,45	–6,33	3,13	2,74

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der Bindungslängendifferenzen; [‡]0,25-Quartil, Median, 0,75-Quartil; alle Einheiten in 10^{-3}Å .

9.3 Verschiebungsparameter

Gemittelt über die gesamte Struktur von TRICHOTOXIN, inklusive der Atome der Solvenssphäre und der Wasserstoffatome, ergibt sich ein Median des äquivalenten isotropen Verschiebungsparameters U_{eq} von $58 \cdot 10^{-3}\text{Å}^2$ in den sphärischen Modellen und von $56 \cdot 10^{-3}\text{Å}^2$ im Modell *mul*. Die Werte variieren zwischen den Extremen 21 und $399 \cdot 10^{-3}\text{Å}^2$.

Neben weiteren Untermengen der Struktur, zeigt Tab. E.1 auf Seite 389 auch die Statistik für die verfeinerten 233 Nichtwasserstoffatome: Hier beläuft sich der Zentralwert der U_{eq} -Verteilung auf 44,56, 43,26 und $41,11 \cdot 10^{-3}\text{Å}^2$ für die Verfeinerungen *iam_{rb}*, *iam* und *mul*. Die Streuung der Temperaturfaktoren ist im Modell *mul* etwas niedriger als in *iam*, aber nicht geringer als in *iam_{rb}*. Die Abnahme der Temperaturfaktoren durch die asphärische Verfeinerung zeigt sich ebenso im Mittelwert der Differenz zwischen den U_{eq} des Modells *mul* und *iam*, welcher bei $-2,50 \cdot 10^{-3}\text{Å}^2$ liegt. Noch etwas ausgeprägter ist diese Größe beim Vergleich mit *iam_{rb}*, sie beträgt $-3,68 \cdot 10^{-3}\text{Å}^2$. In beiden Differenzen liegen die Quartilsgrenzen im negativen Bereich, mit einem geringen Abstand von $\approx 1 \cdot 10^{-3}\text{Å}^2$, was auf eine recht gleichmäßige Verringerung der Temperaturfaktoren hindeutet.

Auch die Minimal- und Maximalwerte der Verschiebungsparameter nehmen mit der asphärischen Verfeinerung ab. Betragen sie im Modell *iam_{rb}* noch 25 und $255 \cdot 10^{-3}\text{Å}^2$ bzw. 24 und $256 \cdot 10^{-3}\text{Å}^2$ im Modell *iam*, so lauten sie im Modell *mul* auf 21 und $252 \cdot 10^{-3}\text{Å}^2$. Die Extremwerte der Differenzen liegen bei $-5,11$ und

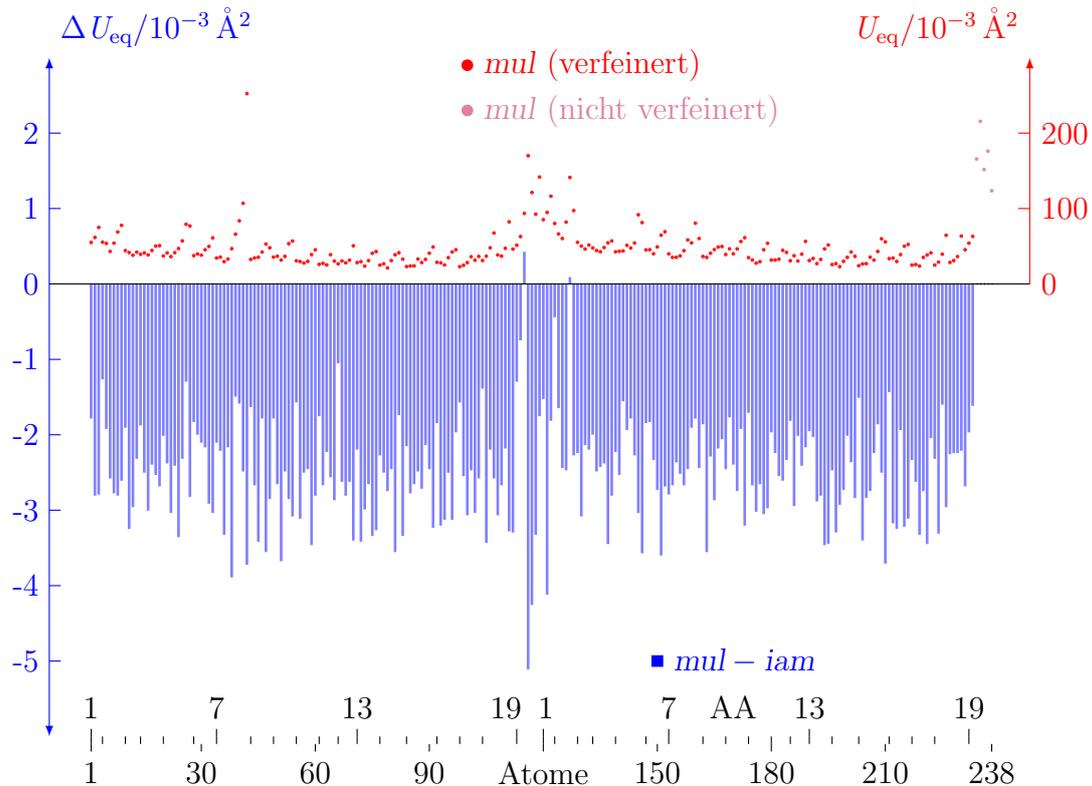


Abbildung 9.3: Veränderung der U_{eq} -Werte im Modell *mul* gegenüber dem sphärischen Modell *iam* für die Nichtwasserstoffatome von TRICHOTOXIN.

$0,42 \cdot 10^{-3} \text{Å}^2$ zu *iam* sowie $-5,59$ und $3,68 \cdot 10^{-3} \text{Å}^2$ zu iam_{rb} .

Eine Darstellung der anisotropen Temperaturellipsoide des verfeinerten Strukturteils gibt Abb. 9.5 auf Seite 255, jeweils mit einem Farbcode für den äquivalenten isotropen Temperaturfaktor versehen. Die U_{eq} -Werte nach der Verfeinerung *mul* sind in Abb. 9.3 über die Sequenz aufgetragen, was, neben dem periodischen Anstieg in den Seitenketten, die deutlich höheren Temperaturfaktoren der Atome der Kettentermini aufzeigt. Abb. 9.4 auf der nächsten Seite beinhaltet eine Auftragung der U_{eq} -Werte im Modell *iam*, und zwar nach zunehmender Größe geordnet. In beiden Graphiken sind die Differenzen zwischen sphärischem und asphärischem Modell zum jeweiligen Temperaturfaktor eingezeichnet.

Die obigen Aussagen zur Statistik der U_{eq} -Werte sieht man hier bestätigt. Nur zwei Temperaturfaktoren des asphärischen Modells liegen höher als im Modell *iam*, alle anderen 231 Werte sinken. Zumeist belaufen sich die Verbesserungen auf -2 bis $-3 \cdot 10^{-3} \text{Å}^2$, die, wie Abb. 9.4 zeigt, eine nur geringe Abhängigkeit vom jeweiligen sphärischen Verschiebungsparameter aufweisen: Positive und negative Maximaldifferenzen korrespondieren zwar mit den höchsten U_{eq} -Werten, ansonsten ist deren Abnahme aber weitestgehend konstant.

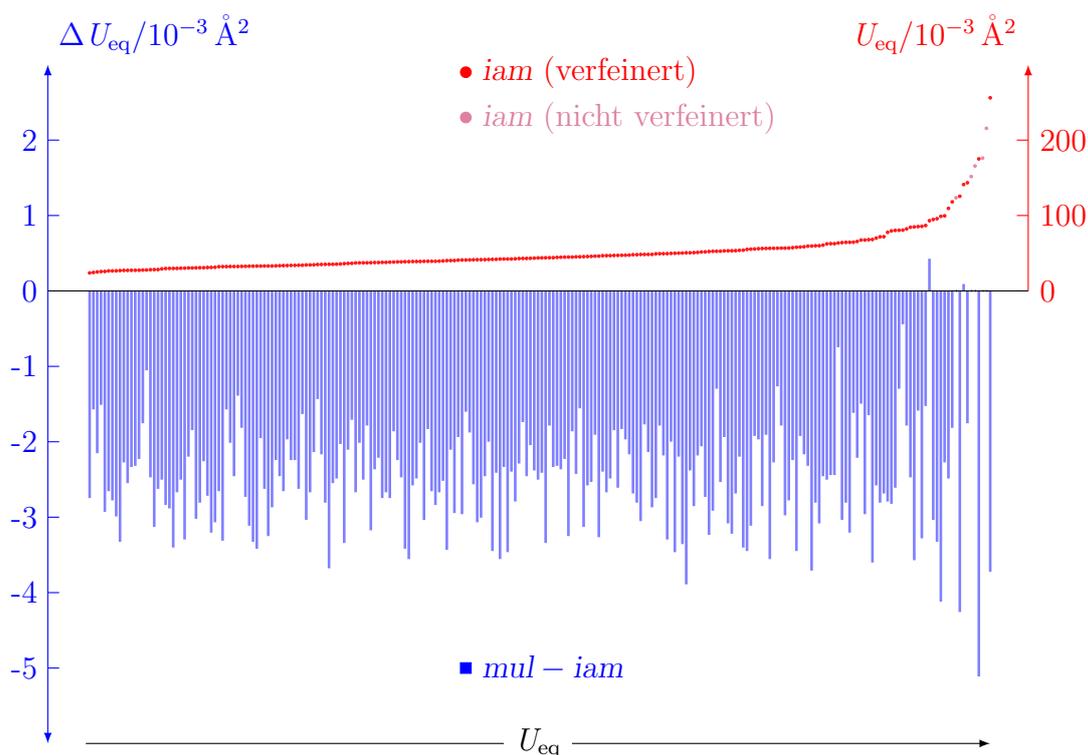


Abbildung 9.4: Veränderung der U_{eq} -Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die Nichtwasserstoffatome von TRICHOTOXIN. Auftragung nach U_{eq} (*iam*).

Neben den hohen Verschiebungsparametern der fünf nicht verfeinerten Atome, ist ein weiterer Ausreißer an 42ster Stelle in der Auftragung 9.3 zu erkennen. Es handelt sich um das $\text{N}_{\epsilon 2}$ des Glutaminsäurerestes 6, welches dennoch asphärisch verfeinert wurde und hierbei eine Verbesserung von 256 auf $252 \cdot 10^{-3} \text{Å}^2$ gegenüber dem Modell *iam* erfährt. Dieses Atom ist in Abb. 9.5 ebenfalls auffällig. Ohne Verbesserung des Temperaturfaktors blieb der Versuch eines Tausches der Sauerstoff- und Stickstoff-Position dieser Amidgruppe.

Dabei stellt dieses Stickstoffatom eine Ausnahme dar, denn im Durchschnitt aller 40 Vertreter dieses Elements, zeigen sie den kleinsten Temperaturfaktor von $35 \cdot 10^{-3} \text{Å}^2$ im asphärischen Modell bzw. 37 und $38 \cdot 10^{-3} \text{Å}^2$ in den Verfeinerungen *iam* und *iam_{rb}* (siehe Tab. E.2 auf Seite 390). Gleichfalls am niedrigsten ist der Median der Differenzen, welcher bei $-1,93$ und $-2,62 \cdot 10^{-3} \text{Å}^2$ für den Vergleich von *mul* mit *iam* und *iam_{rb}* liegt.

An zweiter Stelle bezüglich der U_{eq} -Höhe liegen die 152 verfeinerten Kohlenstoffatome. Der Median der Verschiebungsparameter beläuft sich in den Modellen *iam_{rb}*, *iam* und *mul* auf etwa $45,68$, $44,39$ und $41,72 \cdot 10^{-3} \text{Å}^2$. Gegenüber den Verfeinerungen *iam* und *iam_{rb}* ist die durchschnittliche Verbesserung im Modell *mul*

Tabelle 9.4: Statistik verfeinerte U_{eq} TRICHOTOXIN- Atome der Peptidbindung.

	$x_{\min/\max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C_α (n = 16)							
<i>mul</i>	23,83/63,01	36,09(11,88)	27,16	33,46	38,36	11,20	9,46
<i>iam</i>	26,60/64,63	38,67(11,35)	30,15	35,81	41,16	11,01	8,52
<i>iam_{rb}</i>	28,40/66,71	40,13(11,33)	31,51	37,61	42,50	10,99	9,16
<i>mul-iam</i>	-3,43/-0,75	-2,58(0,69)	-3,03	-2,65	-2,34	0,69	0,52
<i>mul-iam_{rb}</i>	-4,93/-1,64	-4,04(0,80)	-4,43	-4,30	-3,76	0,67	0,67
C (n = 36)							
<i>mul</i>	21,15/85,24	35,87(12,54)	27,44	34,04	40,17	12,73	9,70
<i>iam</i>	23,90/86,77	38,48(12,22)	30,07	36,50	42,52	12,46	9,52
<i>iam_{rb}</i>	24,78/89,20	39,91(12,46)	31,42	38,36	43,96	12,54	9,95
<i>mul-iam</i>	-3,68/-1,52	-2,61(0,54)	-3,01	-2,61	-2,27	0,74	0,58
<i>mul-iam_{rb}</i>	-5,29/-2,84	-4,04(0,51)	-4,33	-4,05	-3,77	0,57	0,43
N (n = 34)							
<i>mul</i>	22,94/80,16	37,09(12,00)	29,16	34,64	41,89	12,73	9,15
<i>iam</i>	24,51/80,60	38,92(11,86)	30,94	36,55	44,26	13,32	8,99
<i>iam_{rb}</i>	25,91/80,99	39,69(11,76)	31,64	37,33	44,43	12,79	9,35
<i>mul-iam</i>	-2,79/-0,44	-1,83(0,44)	-2,09	-1,86	-1,58	0,50	0,40
<i>mul-iam_{rb}</i>	-3,24/-0,82	-2,61(0,46)	-2,97	-2,60	-2,47	0,50	0,38
O (n = 36)							
<i>mul</i>	29,52/94,80	44,19(14,67)	34,36	40,47	48,47	14,11	11,09
<i>iam</i>	32,18/98,91	46,73(14,85)	36,85	42,88	50,93	14,08	11,10
<i>iam_{rb}</i>	33,64/100,40	48,26(14,94)	37,94	44,52	52,63	14,69	11,33
<i>mul-iam</i>	-4,11/-1,78	-2,55(0,38)	-2,74	-2,51	-2,40	0,33	0,26
<i>mul-iam_{rb}</i>	-5,59/-3,08	-4,07(0,50)	-4,32	-4,04	-3,71	0,62	0,47

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der U_{eq} und ΔU_{eq} ;

[‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-3}\AA^2 .

mit $-2,67$ und $-3,77 \cdot 10^{-3}\text{\AA}^2$ anzugeben. Zwar ist dies die stärkste Verbesserung unter den drei Elementen, allerdings sind es zwei Kohlenstoffatome, die die beiden einzigen U_{eq} -Anstiege aufweisen.

Nochmals etwas höhere äquivalente isotrope Temperaturfaktoren sind für den Median der 41 Sauerstoffwerte zu reportieren. In gleicher Reihenfolge wie zuvor betragen diese $46,11$, $44,49$ und $42,30 \cdot 10^{-3}\text{\AA}^2$. Die mittlere Differenz zum Modell *mul* beträgt $-2,49$ und $-4,01 \cdot 10^{-3}\text{\AA}^2$. Wie schon bei den beiden anderen Elementtypen, nimmt auch hier, wenn auch ebenfalls in sehr geringem Maße, die Streuung der U_{eq} -Verteilung mit der asphärischen Verfeinerung ab.

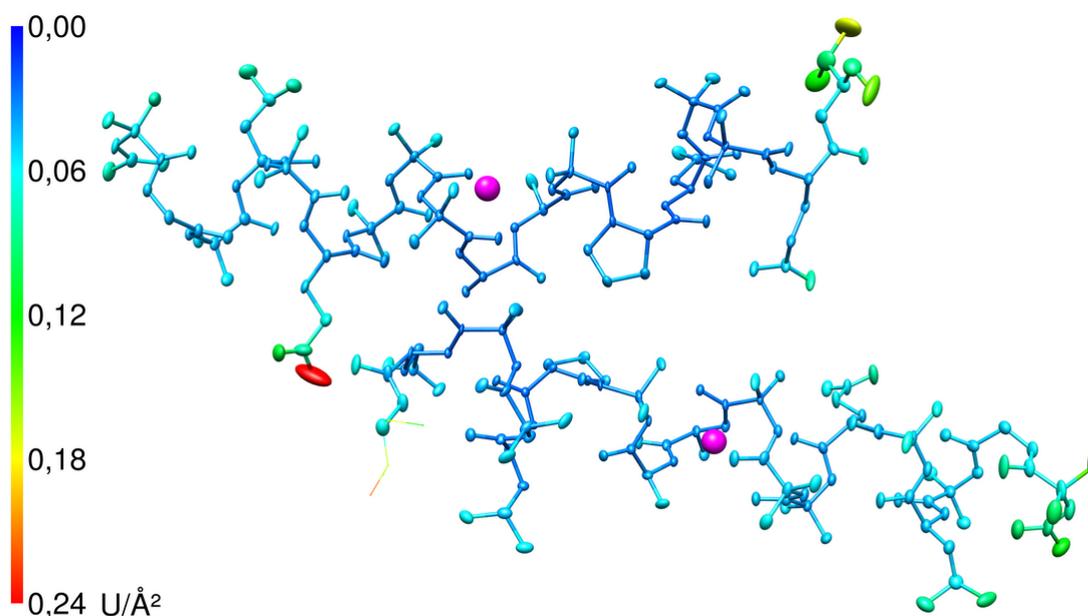


Abbildung 9.5: Anisotrope Verschiebungsparameter des verfeinerten Teils von TRICHOTOXIN nach asphärischer Verfeinerung. Farbcode: U_{eq} . Darstellung der massengewichteten Zentren beider Ketten als magentafarbene Kugeln.

Eine gleichartige Statistik über die vier Atomtypen der Peptidbindung zeigt Tab. 9.4 auf der vorherigen Seite. Diese sind offenbar besser als der Durchschnitt definiert, zumindest zeigen sie durchweg geringere U_{eq} -Werte verglichen mit den entsprechenden Elementen. Insbesondere die Kohlenstoffatome C_α und C zeigen viel niedrigere Werte, während die Unterschiede zur vorherigen Statistik bei N und O nicht so hoch sind. Dies erklärt sich schlicht aus der Tatsache, daß fast alle vorkommenden Heteroatome zur Peptidgruppe zählen, aber nur etwa ein Drittel der Kohlenstoffatome in der aktuellen Statistik berücksichtigt wird.

Weil die Temperaturfaktoren in der Peptidbindung geringer sind, die durchschnittliche Abnahme selbiger im Modell *mul* aber ungefähr gleich bleibt, ist die prozentuale Verbesserung höher. Betrug diese für die Gesamtheit der verfeinerten Atome noch ca. 5 %, sind es nun etwa 6 %. Mit 6,6 % ist die Verringerung der Mediane beider Kohlenstoffatomtypen ausgeprägter als bei Sauerstoff- (5,6 %) oder Stickstoffatomen (5,2 %) der Peptidbindung. Dies spiegelt sich auch in der Abstufung der mittleren Differenzen wieder.

Auch nehmen hier in allen vier Fällen die Maximalwerte ab, ebenso wie sich die kleinsten U_{eq} jedes Atomtyps nochmals verringern. Andererseits nimmt sowohl die Streuung um das arithmetische Mittel mit asphärischer Verfeinerung zu, wie die steigenden Standardabweichungen belegen, als auch die Schwankung um den jeweiligen U_{eq} -Median, was durch höhere Interquartilabstände und den Median der absoluten Abweichungen belegt wird.

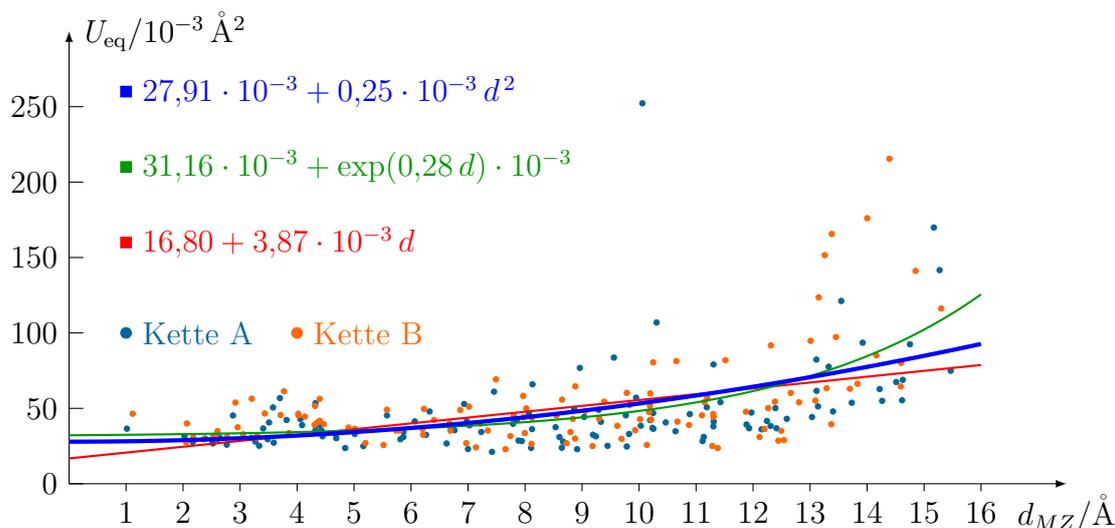


Abbildung 9.6: U_{eq} (mul) der Nichtwasserstoffatome von TRICHOTOXIN in Abhängigkeit vom Abstand zum Massenzentrum.

Diagramm 9.6 zeigt die Abhängigkeit der U_{eq} -Werte der Atome des asphärischen Modells vom Abstand zum Massenzentrum der Ketten A und B. Dabei liegt das Zentrum der Kette A bei $(5,40; 5,67; 15,31)$ Å, das der Kette B bei $(0,15; -0,63; 4,82)$ Å, die Koordinaten sind in Abb. 9.5 auf der vorherigen Seite durch magentafarbene Kugeln angezeigt. Die Anpassung einer linearen, quadratischen und exponentiellen Funktion, verlief in allen Fällen akzeptabel. Bezüglich des Ordinatenabschnittes liegen die Exponential- und die quadratische Funktion etwas hoch, während die Geradengleichung etwas zu niedrige Werte angibt. Die Entwicklung mit zunehmendem Abstand ist in allen drei Fällen in guter Übereinstimmung mit der Anmutung der Daten. Unterschiedlich sind allerdings die mit den jeweiligen Parametern verbundenen Unsicherheiten: Die e-Funktion hat Fehlermargen von ca. 7% und 2%, die quadratische Funktion von 10% und 11%, und die lineare Funktion von 26% und 12%. Mit weniger zu verfeinernden Parametern sinken die Fehler, z.B. hat die Anpassung einer Ursprungsgerade mit der Steigung $5,54 \cdot 10^{-3}$ Å nur noch einen Fehler von 4%, aber den Nachteil einer inadäquaten Beschreibung des Verhaltens bei null Abstand zum massengewichteten Schwerpunkt. Die abstandsabhängige Temperaturfaktorzunahme ist offenbar nur moderat ausgeprägt, bspw. liegt der Anstieg nach der quadratischen Funktion bei $0,5 \cdot 10^{-3} d$ Å, was aber auch an der gestreckten Molekülform liegen mag.

Betrachtet man einen gemeinsamen Schwerpunkt $(2,78; 2,52; 10,06)$ Å beider Ketten, so erstreckt sich der Abstands d_{MZ} von 2,06 bis 21,24 Å. Daran angepaßte Ausgleichsfunktionen ergeben noch deutlich geringere Steigungen.

9.4 DMSDA-Werte

Von den insgesamt 235 Bindungen in TRICHOTOXIN, die wenigstens ein frei verfeinertes Atom beinhalten, haben 135 oder 58 % nach der asphärischen Verfeinerung einen niedrigeren DMSDA-Betrag als nach der Beschreibung im Modell *iam*. Gegenüber der Verfeinerung *iam_{rb}* verbessert eine etwas geringere Zahl von 132 (56 %) Bindungen ihren DMSDA-Wert im Modell *mul*.

In den beiden sphärischen Modellen liegt der Median über alle verfeinerten Bindungen bei 54 und $50 \cdot 10^{-4} \text{ \AA}^2$. Dahingegen liegt das Mittel des asphärischen Modells bei $45 \cdot 10^{-4} \text{ \AA}^2$, was in Tab. E.3 auf Seite 390 detailliert ist. Die Differenz zwischen *mul* und *iam* bzw. *iam_{rb}* ist durchschnittlich $-6 \cdot 10^{-4} \text{ \AA}^2$ hoch und umfaßt so extreme Werte wie -90 und $76 \cdot 10^{-4} \text{ \AA}^2$ im Fall des Vergleichs mit *iam* oder gar -139 bis $120 \cdot 10^{-4} \text{ \AA}^2$ gegenüber *iam_{rb}*. Dadurch nimmt der Maximalwert des asphärischen Modells auf $899 \cdot 10^{-4} \text{ \AA}^2$ zu, nach vorherigen 878 und $867 \cdot 10^{-4} \text{ \AA}^2$.

Die stark schwankenden DMSDA-Beträge des Modells *mul* sind in Abb. 9.7 auf der folgenden Seite zusammen mit der ebenfalls ausgesprochen volatilen Differenz zum Modell *iam* aufgetragen. Die Relation der Skalierung der Achsen der Beträge und Differenzen beträgt 3 : 1, hierbei mußten einige besonders hohe DMSDA-Werte nicht maßstabsgetreu angegeben und mit kleinen Zahlen gekennzeichnet werden. Abb. 9.8 auf Seite 259 zeigt die nach zunehmender Größe geordneten DMSDA-Beträge des Modells *iam*, dazu die gleichen Differenzen wie zuvor. Aus dieser Darstellung geht hervor, daß die negativen DMSDA-Differenzen mit steigendem sphärischen DMSDA-Betrag zunehmen, sowohl in ihrer Anzahl wie Höhe. Für niedrige ($< 16 \cdot 10^{-4} \text{ \AA}^2$) Beträge des Modells *iam* sind fast nur Verschlechterungen festzustellen, für hohe Beträge zumeist Verbesserungen.

Anhand der in Tab. E.3 auf Seite 390 nach Elementen differenzierten Statistik, sowie analoger Betrachtungen zu den Bindungen zwischen den Atomen der Peptidbindung in Tab. 9.5 auf Seite 260, wird ersichtlich, daß die Beiträge verschiedener Bindungstypen unterschiedlich sind.

So verbessert sich die Beschreibung von 116 Kohlenstoff-Kohlenstoff-Bindungen von 50 und $43 \cdot 10^{-4} \text{ \AA}^2$ in den Modellen *iam* und *iam_{rb}* auf $36 \cdot 10^{-4} \text{ \AA}^2$ im asphärischen Fall. Gegenüber beiden sphärischen Verfeinerungen werden durchschnittliche Differenzen von $-9 \cdot 10^{-4} \text{ \AA}^2$ erzielt, bei gleichzeitiger Abnahme der Streubreite der DMSDA-Beträge. Die 14 C_α-C-Bindungen der Tab. 9.5 weisen jedoch im Modell *mul* eine starke Zunahme des DMSDA-Medians verglichen mit *iam* auf, statt 33 beträgt dieser nun $55 \cdot 10^{-4} \text{ \AA}^2$.

Bezüglich der Verschlechterung der C-O- und der Verbesserung der C-N-DMSDA-Werte besteht wiederum Übereinstimmung zwischen den Gesamt- und Peptid-Statistiken, die allerdings hierbei auch wesentlich höhere Schnittmengen untereinander haben. Auf den C-O-Bindungen betragen die Veränderung durchschnittlich $10 \cdot 10^{-4} \text{ \AA}^2$. Im Modell *mul* liegen die DMSDA-Beträge dann bei $74 \cdot 10^{-4} \text{ \AA}^2$, wäh-

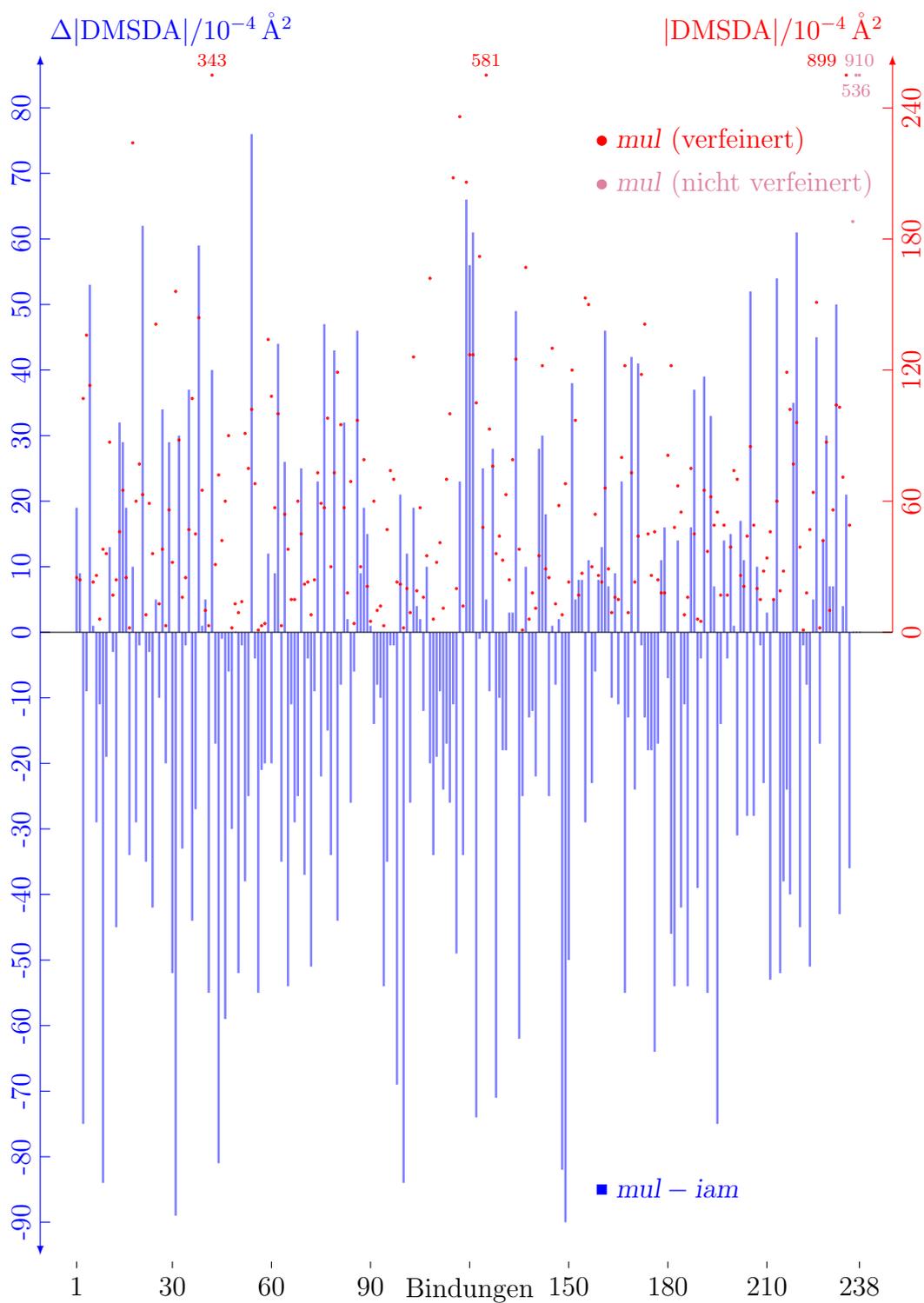


Abbildung 9.7: Veränderung der DMSDA-Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die Nichtwasserstoffatome von TRICHOTOXIN.

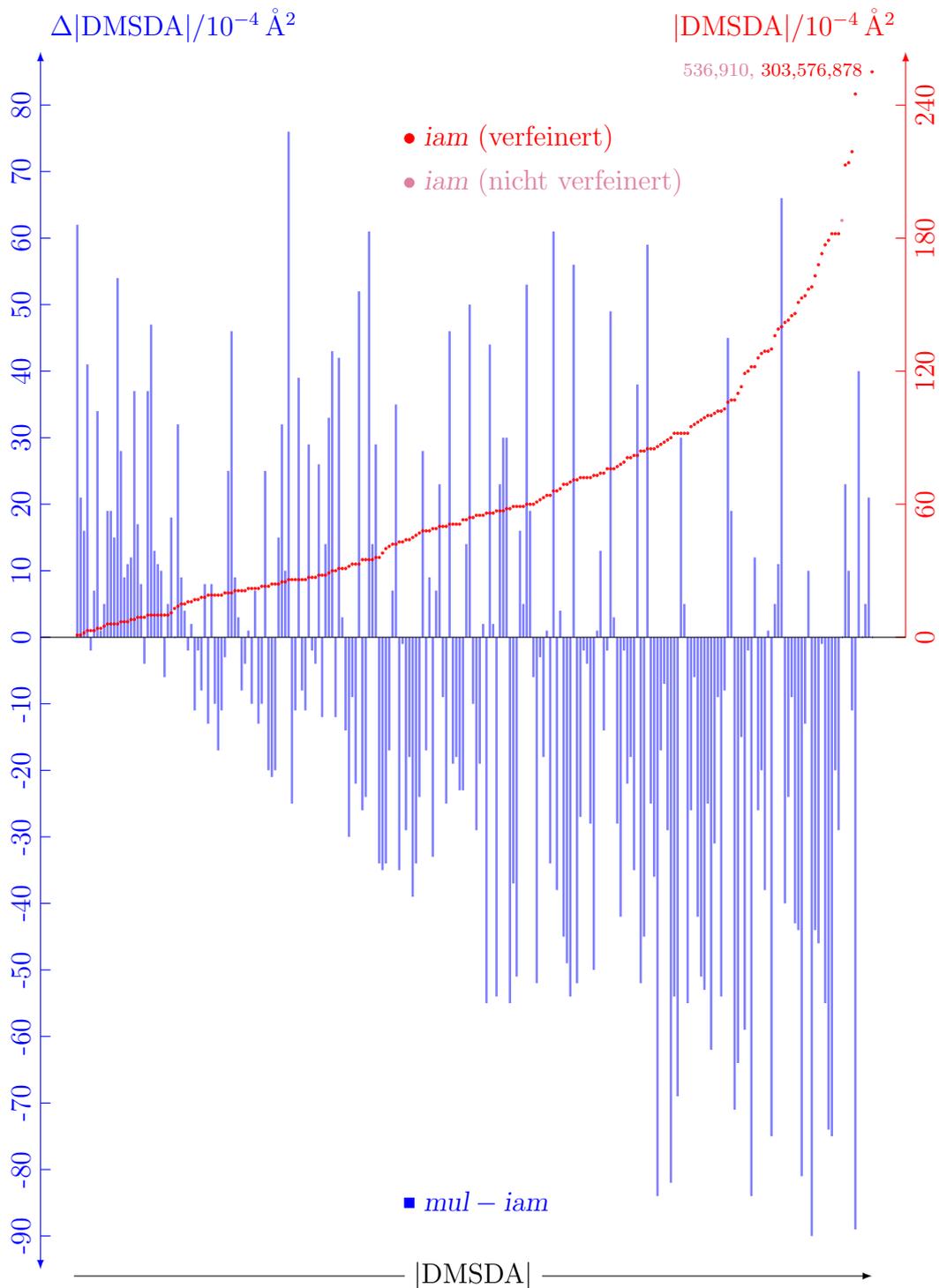


Abbildung 9.8: Veränderung der DMSDA-Werte im asphärischen Modell *mul* gegenüber dem sphärischen Modell *iam* für die Nichtwasserstoffatome von TRICHOTOXIN. Auftragung nach steigenden $|DMSDA|$ -Werten des Modells *iam*.

Tabelle 9.5: DMSDA-Beträge der verfeinerten Peptidbindungen des TRICHOTOXIN-Hauptmoleküls und deren Differenzen.

	$x_{min/max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–O ($n = 36$)							
<i>mul</i>	6/224	81,33(49,53)	45,00	74,50	119,20	74,25	45,96
<i>iam</i>	6/214	73,81(51,09)	28,00	71,50	100,80	72,75	49,67
<i>iam_{rb}</i>	6/194	69,33(49,79)	27,50	64,50	97,75	70,25	55,60
<i>mul-iam</i>	-27/56	7,53(19,99)	-6,50	10,00	20,00	26,50	20,76
<i>mul-iam_{rb}</i>	-36/88	12,00(26,82)	-8,50	13,00	31,25	39,75	31,88
C–N ($n = 34$)							
<i>mul</i>	2/144	53,32(42,23)	21,25	37,00	94,50	73,25	31,88
<i>iam</i>	2/182	75,32(52,27)	27,50	73,50	100,20	72,75	55,60
<i>iam_{rb}</i>	0/244	92,59(63,87)	50,50	82,00	121,00	70,50	56,34
<i>mul-iam</i>	-84/76	-22,00(38,27)	-44,25	-30,50	3,75	48,00	35,58
<i>mul-iam_{rb}</i>	-139/89	-39,26(52,74)	-71,25	-51,00	-3,25	68,00	36,32
C_α–C ($n = 14$)							
<i>mul</i>	11/151	60,00(45,83)	16,25	55,00	94,00	77,75	57,82
<i>iam</i>	1/168	47,86(43,62)	22,75	33,00	54,75	32,00	23,72
<i>iam_{rb}</i>	0/184	56,50(43,22)	37,75	47,50	60,50	22,75	19,27
<i>mul-iam</i>	-54/62	12,14(39,57)	-19,50	22,50	44,75	64,25	41,51
<i>mul-iam_{rb}</i>	-62/63	3,50(47,97)	-39,25	-2,00	49,50	88,75	70,42
C_α–N ($n = 14$)							
<i>mul</i>	4/100	45,57(31,05)	17,25	50,50	69,25	52,00	42,25
<i>iam</i>	9/126	51,07(33,60)	26,25	51,50	70,75	44,50	37,06
<i>iam_{rb}</i>	3/92	49,86(30,83)	21,00	56,50	74,00	53,00	33,36
<i>mul-iam</i>	-28/9	-5,50(11,45)	-11,00	-4,00	3,50	14,50	11,12
<i>mul-iam_{rb}</i>	-40/14	-4,29(13,19)	-7,25	-0,50	2,00	9,25	5,93

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der DMSDA-Beträge;

[‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-4}Å^2 .

rend die sphärischen Mediane aller 41 C–O-Bindungen nur $63 \cdot 10^{-4}\text{Å}^2$ betragen. Im Fall der auf die Carbonylbindung beschränkten Statistik liegt das Mittel von *iam* bei $72 \cdot 10^{-4}\text{Å}^2$, das des Modells *iam_{rb}* bei $65 \cdot 10^{-4}\text{Å}^2$.

Mit einer durchschnittlichen Differenz von $-12 \cdot 10^{-4}\text{Å}^2$ verringern sich die DMSDA-Werte aller 78 C–N-Bindungen, und zwar von 55 und $67 \cdot 10^{-4}\text{Å}^2$ in den sphärischen Modellen auf $44 \cdot 10^{-4}\text{Å}^2$ im asphärischen Modell. Die Differenzierung in C–N-Bindung und C_α–N- zeigt besonders starke Verbesserungen im ersten Fall und wesentlich geringere Abnahmen im letzten.

Tabelle 9.6: Statistik der Restdichten TRICHOTOXIN.

	$x_{min/max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
<i>mul</i>	-371,3/688,6	-0,0(60,1)	-39,7	-0,5	38,8	78,5	58,3
<i>iam</i>	-394,9/716,6	-0,0(62,4)	-40,8	0,4	41,1	81,9	60,8
<i>iam_{rb}</i>	-387,7/677,7	0,0(63,6)	-41,4	0,4	41,9	83,3	61,8

[†]Minimum, Maximum und Mittelwert mit Standardabweichung der Restdichte;

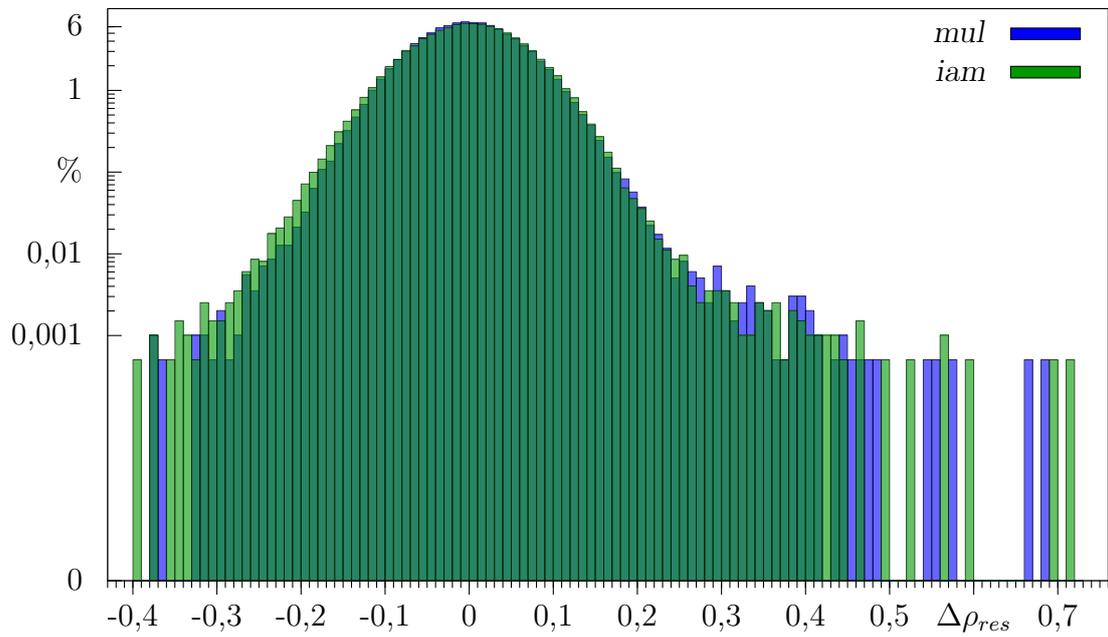
[‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in $10^{-3} \text{ e}/\text{\AA}^3$.

9.5 Restdichten

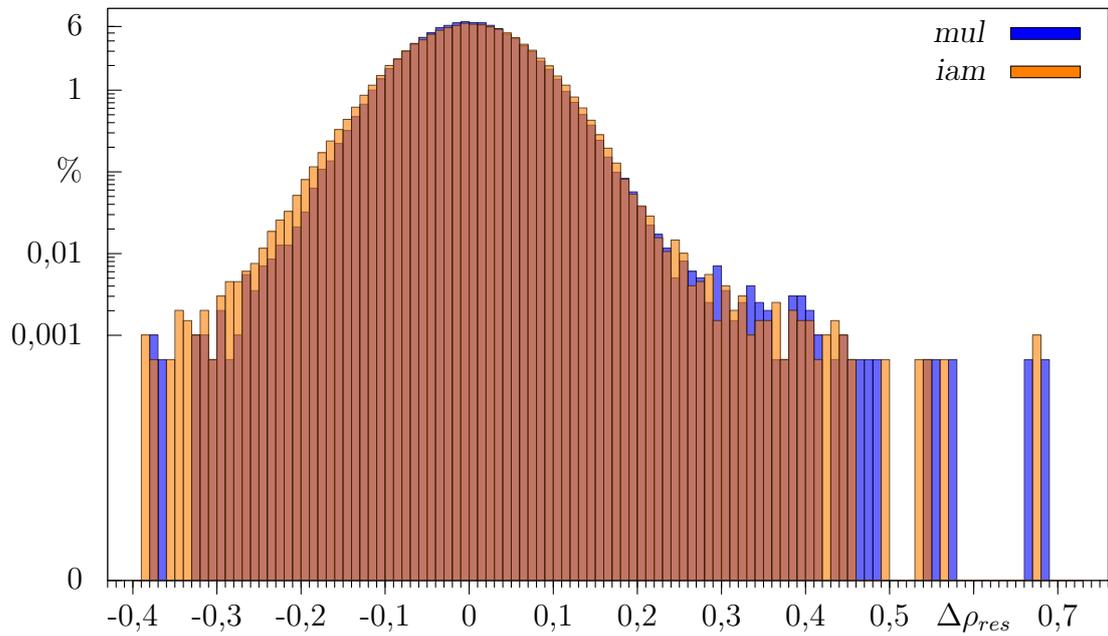
Die Restelextronendichte in der asymmetrischen Einheit wurde auf einem $34 \times 56 \times 104 = 198016$ Punkte umfassenden Gitter berechnet. Das Volumen eines Voxels betragt $0,0248 \text{ \AA}^3$, womit die normierte Gesamtrestelextronendichte $\sum |\Delta\rho_{res}|V_{vox}$ in den Modellen *mul*, *iam* und *iam_{rb}* zu jeweils 231,896 e, 241,014 e und 245.682 e bestimmt wird. Auf das Volumen der Zelle bezogen, ergeben sich in der gleichen Reihenfolge Werte von 0,0473, 0,0491 und $0,0501 \text{ e}\text{\AA}^{-3}$, die mit den restlichen Strukturen direkt vergleichbar sind. Das aspharische Modell enthalt 3,8 % weniger Restdichte als das Modell *iam* und 5,6 % weniger als *iam_{rb}*.

Ebenfalls geringer als in den spharischen Modellen sind die negativen und positiven Extremwerte der aspharischen Restdichteverteilung, wie man sowohl anhand der Statistik in Tab. 9.6, als auch in den Histogrammen 9.9(a) und (b) auf der nachsten Seite erkennen kann. Auch sinkt die Breite der Verteilung der Restdichten um den nahe Null liegenden Mittelwert. In Tab. 9.6 auert sich das durch geringere Standardabweichung, schmalere Interquartilabstand und kleineren Median der absoluten Abweichungen vom Mittel. Die Histogramme zeigen dies durch hohere Haufigkeiten nahe Null und abnehmende Haufigkeiten in steigender Entfernung zum Nullpunkt. So liegt zwar in allen drei Modellen das Maximum der Restdichten im Intervall $(-0,01; 0,00] \text{ e}/\text{\AA}^3$, in *mul* mit 6,86 % aber ein groerer Anteil der Werte als in *iam* und *iam_{rb}* mit jeweils 6,60 % und 6,57 %. Auch die Population des besonders gunstigen Intervalls $(-0,01; 0,01]$ ist mit 13,58 %, im Gegensatz zu 13,16 % und 12,97 %, im aspharischen Fall angehoben.

Der Zentralwert der Restdichteverteilung des Modells *mul* ist zwar zu einem leicht negativen statt positiven Wert verschoben, da die Restdichte insgesamt abnimmt erfolgt die Reduktion positiver Dichten aber nicht auf Kosten zunehmender negativer Restdichte. In den Histogrammen ist dieser Sachverhalt optisch zu erfassen. Zudem enthalt das aspharische Modell 99 % der Restdichten in einem Bereich von $(-0,1555; 0,1621] \text{ e}/\text{\AA}^3$, wahrend fur *iam* und *iam_{rb}* dieses Intervall auf $(-0,1681; 0,1623]$ und $(-0,1723; 0,1661]$ lautet. Obere und untere Grenze nehmen in *mul* ab, die negative sogar starker als die positive.



(a) Restdichten TRICHOTOXIN, *mul* vs *iam*.



(b) Restdichten TRICHOTOXIN, *mul* vs *iam_{rb}*.

Abbildung 9.9: Histogramm der Restdichtegitter $\sum(F_o - F_c) \exp(i\phi_c)$ der sphärischen (grün/orange) und asphärischen (blau) Modelle. Intervalle $\Delta(i; i + 0,01]$, Einheiten $e\text{Å}^{-3}$.

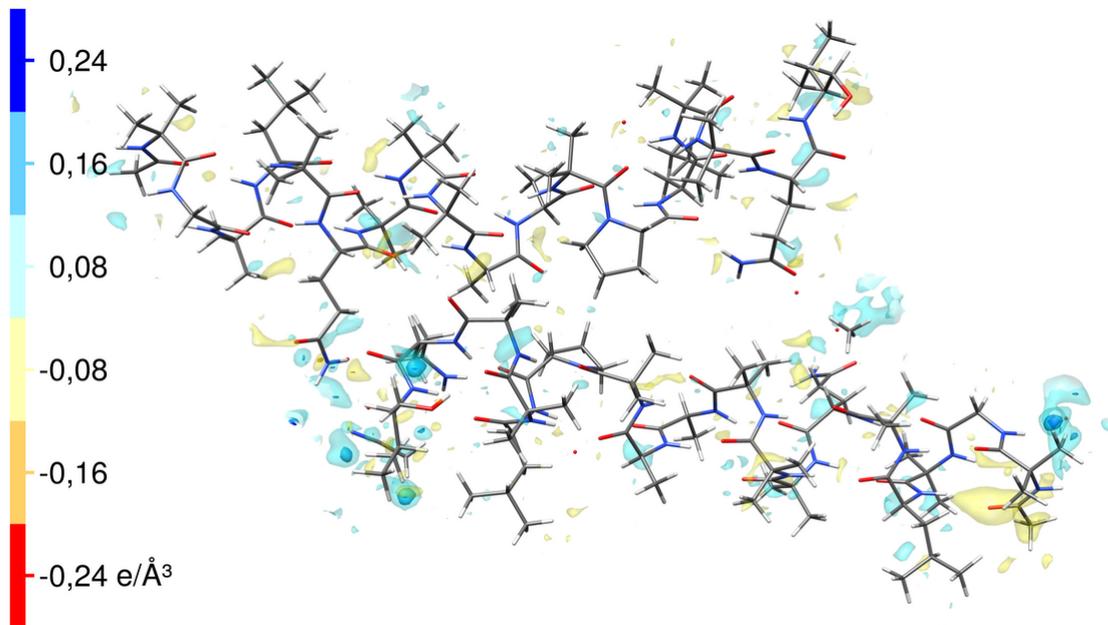
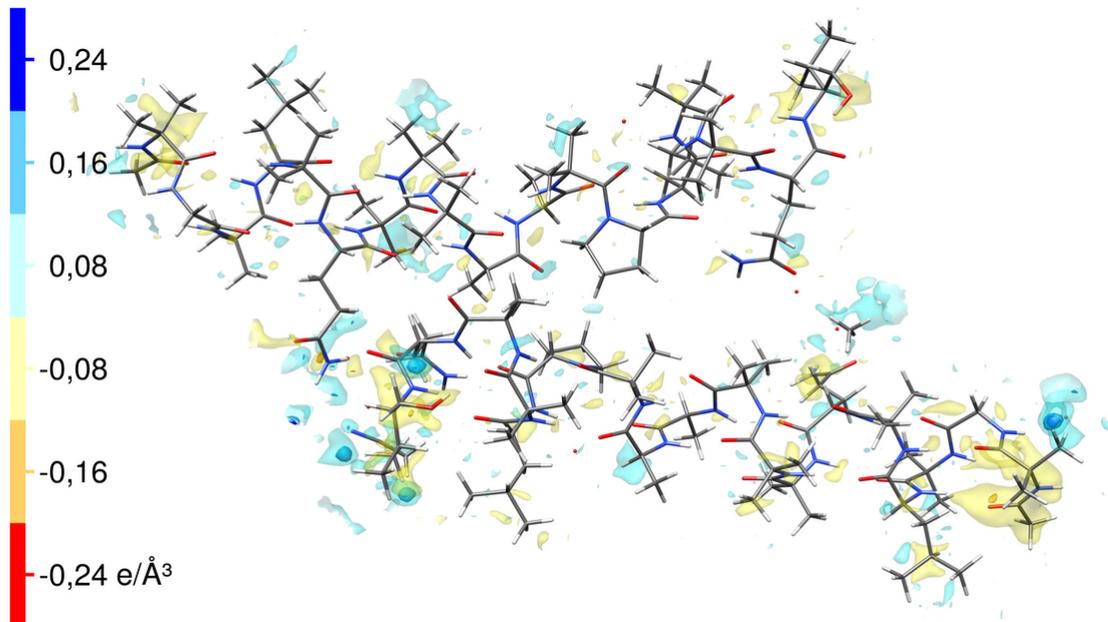
(a) $\Delta\rho_{res}(mul)$ (b) $\Delta\rho_{res}(iam)$

Abbildung 9.10: Restdichten der Modelle *mul* und *iam* von TRICHOTOXIN und deren Restdichtedifferenz. Darstellung der Isooberflächen im Umkreis von 2 Å.

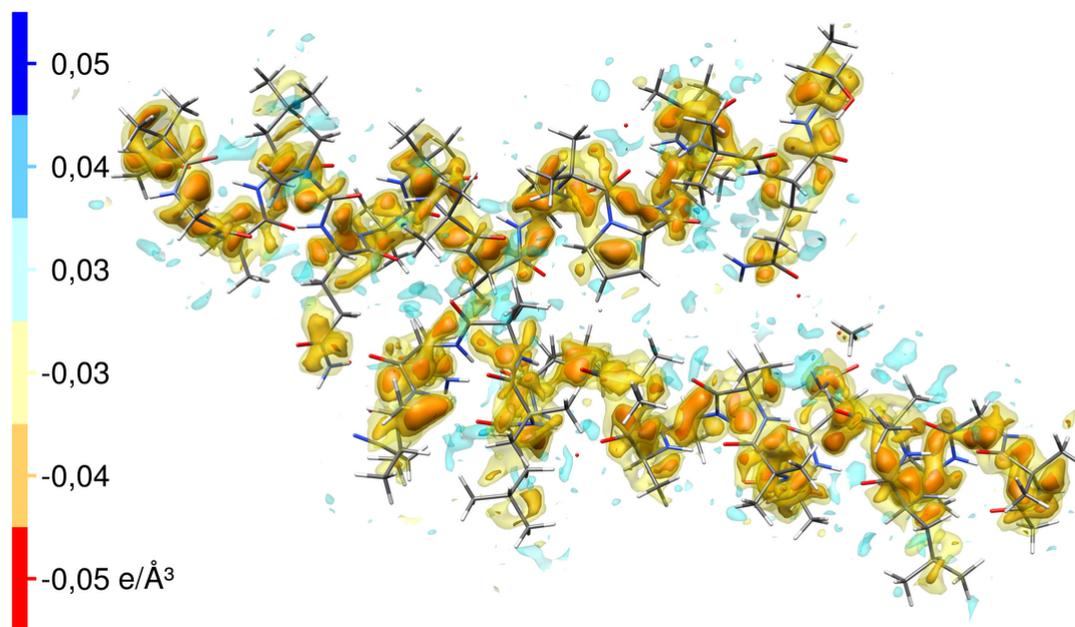
(c) $\Delta\rho_{res}(iam) - \Delta\rho_{res}(mul)$

Abbildung 9.10: Restdichten der Modelle *mul* und *iam* von TRICHOTOXIN und deren Restdichtedifferenz. Darstellung der Isooberflächen im Umkreis von 2 Å (Forts.).

Abb. 9.10(a) und (b) zeigen die Restdichten der Modelle *mul* und *iam* als Isooberflächen bis 2 Å um die beiden Ketten. Abb. E.1 auf Seite 391 enthält die Illustration des Modells *iam_{rb}*. Die Isowerte wurden übereinstimmend mit den bisherigen Strukturen gewählt. Alle Modelle zeigen unveränderte Akkumulationen der Größenordnung $0,15 \text{ e}\text{\AA}^{-3}$ an den Resten VOL 218 und AIB 201 (jeweils links und rechts der zuunterst abgebildeten Kette), was auf nicht behandelte Fehlordnung oder sonstige Modellfehler hinweist. Darüber hinaus sind in diesen Darstellungen nur geringe Restdichten erkennbar. Aber auch bei kleineren Isooberflächenwerten als den abgebildeten, sind keine Bindungsrestdichten in den sphärischen Modellen zu finden. Diese zeigen auf den endständigen Atomen und Bindungen vielmehr negative Restdichten der Größenordnung $-0,05 \text{ e}\text{\AA}^{-3}$. In gleicher Höhe findet sich negative Restdichte in unmittelbarer Umgebung der restlichen Kette, aber nicht auf dieser. Sehr geringe positive Restdichten verteilt sich diffus um das Molekülgerüst.

Die Differenz $\rho_{res}(iam) - \rho_{res}(mul)$ ist gering: Für Abb. 9.10(c) wurde ein deutlich kleinerer Wertebereich als für die Restdichtekarten selbst gewählt. Die ausgeprägt negativen Bereiche, die um die Atome und Bindungen in dieser Darstellung zu erkennen sind, waren bereits im Modell *iam* durch negative Restdichten gekennzeichnet, welche im Modell *mul* etwas gemindert wurde, so daß die Differenz negativ bleibt. Positive Differenzrestdichten entsprechen Bereichen, in denen das Multi-

polmodell entweder mehr negative Restdichte als das sphärische aufweist oder in denen es geringere positive zeigt. Ersteres betrifft die weiter entfernten Bereiche, letzteres unmittelbar das Molekül.

Abschließend ergibt sich der Eindruck, daß auf den Bindungen und Atomen fast keine Änderung durch die asphärische Behandlung stattfindet, aber in deren unmittelbarer Nähe die negativen Restdichten des sphärischen Modells reduziert werden, in geringerem Umfang auch die positiven. Umgekehrt aber, in weiterer Entfernung vom Molekülgerüst, etwas mehr negative Restdichte im Modell *mul* als im Modell *iam* vorhanden ist. Die Abnahme der Restdichten überwiegt hierbei die Zunahme andernorts. Alle Veränderungen betreffen Bereiche geringer und diffuser Restdichteverteilung, die höheren Restdichteakkumulationen, welche in allen drei Modellen sehr ähnlich gestaltet sind, werden kaum verändert.

9.6 Deformationsdichten

Auf 510 Atome der Struktur verteilen sich insgesamt 190,339 e Deformationsdichte. Im Mittel bedeutet dies pro Atom 0,37 e Valenzelektronendichte, die gegenüber dem sphärischen Modell verändert wurde. Der Deformations- zu Restdichtequotient liegt bei $\Delta\rho_{def}/\Delta\rho_{res} = 190,339\text{ e}/231,896\text{ e} = 0,821$ – die vom Modell nicht berücksichtigte Restelektronendichte übersteigt die deformierte Valenzdichte um ca. ein Fünftel.

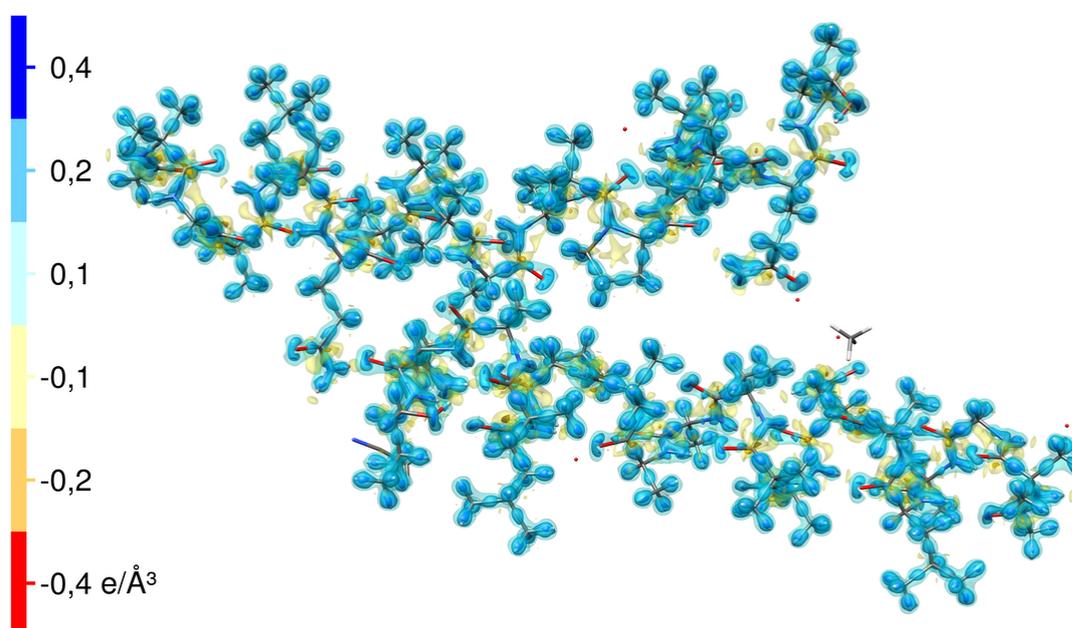


Abbildung 9.11: Deformationsdichte $\Delta\rho_{def}$ von TRICHOTOXIN.

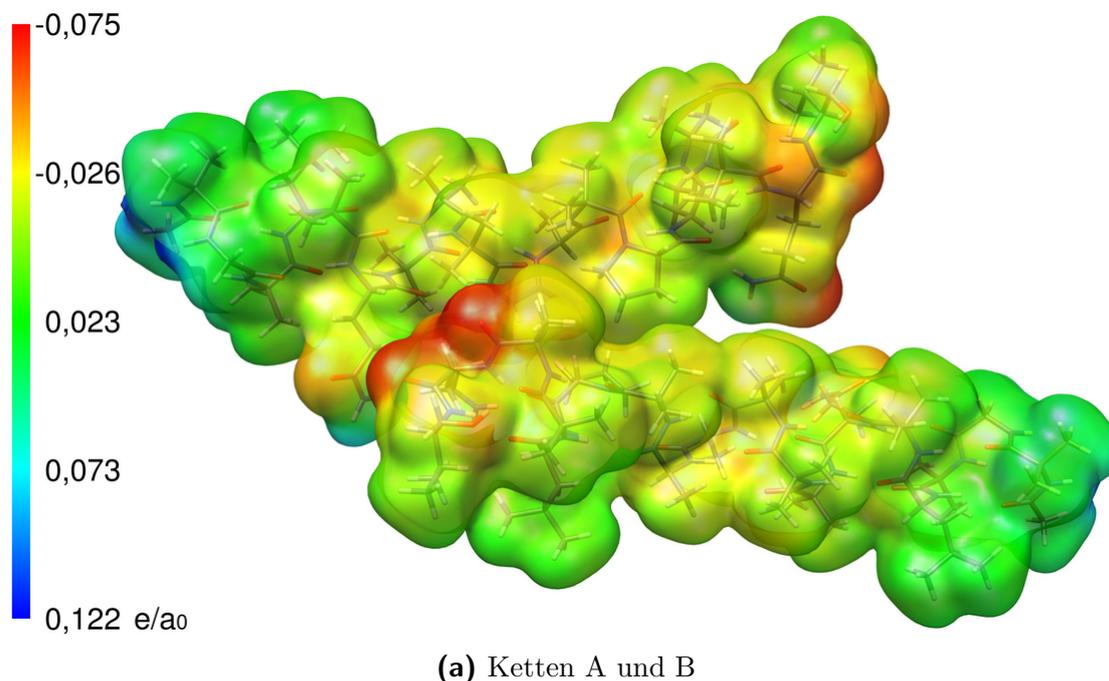


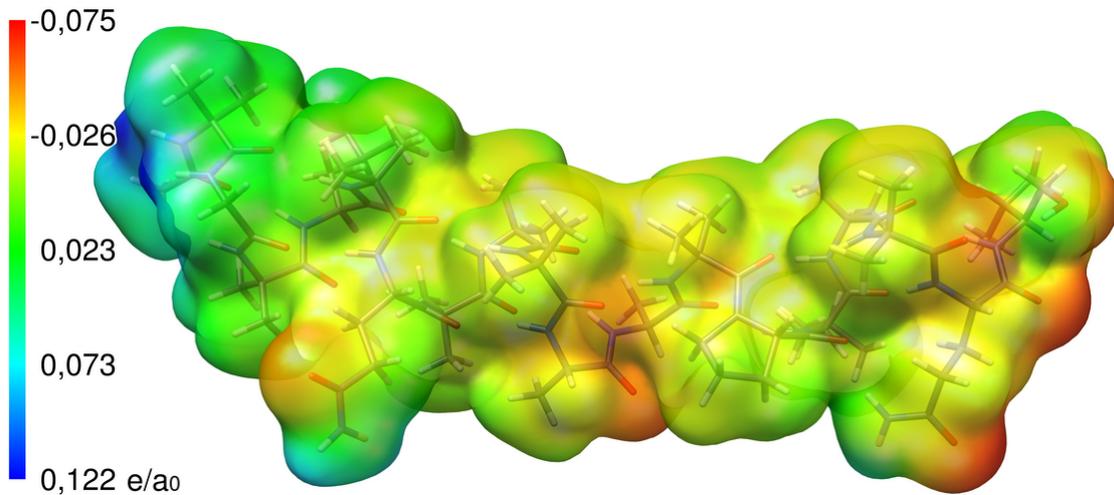
Abbildung 9.12: Elektrostatiches Potential des Multipolmodells von TRICHOTOXIN auf der $0,001 \text{ e/a}_0^3$ -Isooberfläche.

9.7 Abgeleitete Eigenschaften

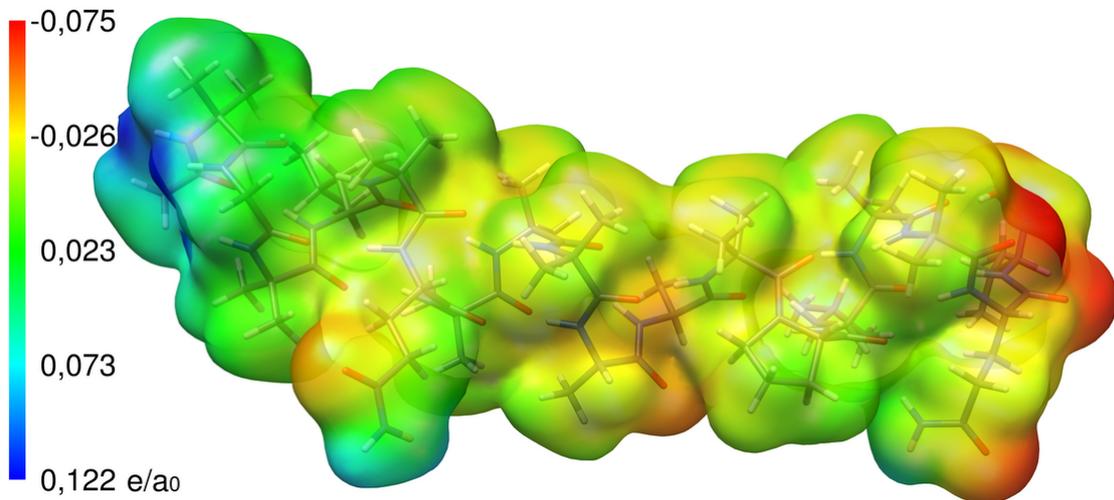
Das für TRICHOTOXIN aus den Multipolen abgeleitete elektrostatiches Potential zeigt Abb. 9.12. In diesen Darstellungen ist es als Farbcode auf die Isooberfläche der Multipolelektronendichte bei $0,001 \text{ e/a}_0^3$ aufgetragen. Zum einen beinhaltet Abb. 9.12(a) die Oberfläche und das ESP aus der gemeinsamen Bestimmung beider Ketten und in der gewohnten Orientierung. Zum anderen bietet Abb. 9.12(b) und (c) eine Darstellung der Isooberflächen der einzelnen Ketten A und B, jeweils ähnlich orientiert.

Auf der gewählten Oberfläche nimmt das Potential Werte zwischen $-0,075$ und $0,122 \text{ e/a}_0$ (-80 – 130 kT/e) an. Jede Kette zeichnet sich durch einen elektropositiven Anfang und ein elektronegatives Ende aus. Bedingt durch geringe Konformationsunterschiede, gibt es kleine Abweichungen zwischen den elektrostaticheren Potentialen beider Ketten: Die B-Kette weist sowohl ein etwas stärker negatives als auch ein stärker positives Potential auf.

Ganz ähnliche Ergebnisse werden für das aus `amber03`-Ladungen bestimmte ESP ϕ_{PC} dieser Struktur erhalten, wiewohl es tendentiell elektronegativere Werte als das Multipolpotential ϕ_{mul} hervorbringt. Abb. 9.13(a) auf Seite 268 stellt eine Wiederholung von Abb. 9.12(a) dar, allerdings mit einer auf $\pm 0,18 \text{ e/a}_0$ ($\pm 191 \text{ kT/e}$) festgelegten Farbskala. Durch den Vergleich zu Abb. 9.13(b), dem Punktladungs-



(b) Kette A



(c) Kette B

Abbildung 9.12: Elektrostatisches Potential des Multipolmodells von TRICHOTOXIN auf der $0,001 \text{ e/a}_0^3$ -Isooberfläche (Forts.)

potential auf der gleichen $0,001 \text{ e/a}_0^3$ -Multipolelektronendichteoberfläche, wird die Ähnlichkeit beider ESPs deutlich. Jeweils auf die van-der-Waals-Oberfläche A_{vdW} aufgetragen, zeigt ϕ_{PC} (9.13(d)) seinen gegenüber ϕ_{mul} (9.13(c)) ins Elektronegative verschobenen Charakter deutlicher. Auf ρ_{mul} nimmt es Werte zwischen $-0,111 \text{ e/a}_0$ (-116 kT/e) und $0,094 \text{ e/a}_0$ (100 kT/e) an, auf A_{vdW} von $-0,134 \text{ e/a}_0$ (-142 kT/e) und $0,100 \text{ e/a}_0$ (106 kT/e), wie Abb. 9.13(e) und (f) zeigen. Wiederum hat die Kette B einen etwas weiteren Potentialbereich als Kette A und auch die Potentialverteilung ist den Multipolergebnissen sehr ähnlich.

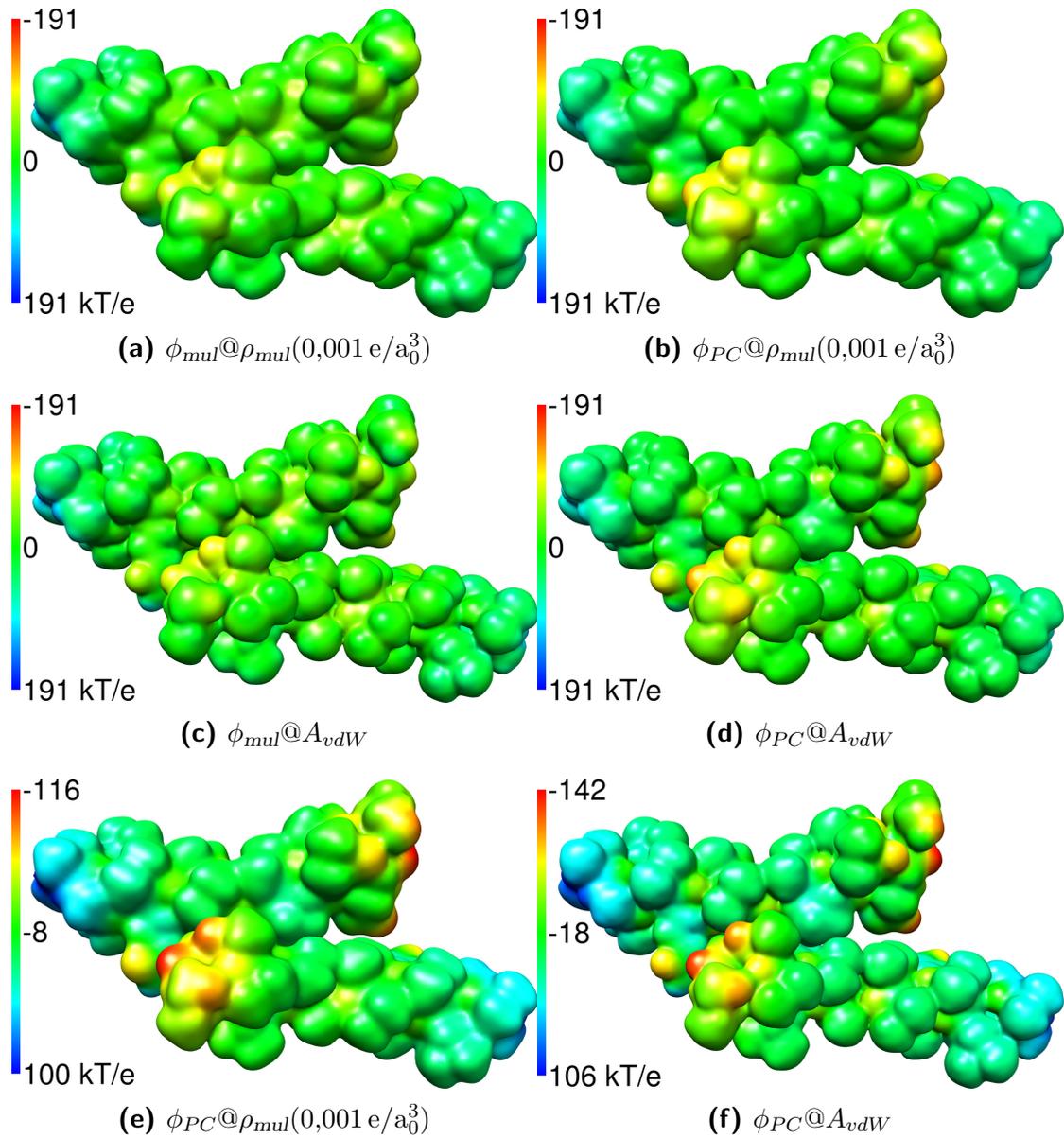


Abbildung 9.13: Multipol-ESP (ϕ_{mul}) und Punktladungs-ESP (ϕ_{PC}) auf $0,001 e/a_0^3$ -Isooberfläche der Multipolelektronendichte (ρ_{mul}) und van-der-Waals-Oberfläche (A_{vdW}). (a)-(d): Fester Wertebereich $\pm 191 \text{ kT/e} = \pm 0,18 e/a_0$. (e)-(f): Voller Wertebereich.

10 Vergleich der Strukturen und Modellverbesserungen

Die fünf behandelten Strukturen unterscheiden sich erheblich voneinander: Zwischen den Molekülgrößen, der Auflösung der Daten, den Gütefaktoren und Restdichten, sowie den Temperaturfaktoren können mehrere Größenordnungen liegen. Ebenso variiert der Einfluß des asphärischen Verfeinerungsmodells beträchtlich. Auf den erzielten Verbesserungen und ihrer Abhängigkeit von verschiedenen Parametern liegt das Augenmerk im aktuellen Kapitel.

Tab. 10.1 stellt für alle behandelten Strukturen mehrere ausgewählte Parameter einander gegenüber, die der Beurteilung der jeweiligen Strukturgüte und des Effekts der verschiedenen Verfeinerungsmodelle dienen. Zur Einschätzung des Erfolges der Einführung der asphärischen Streufaktoren kann z. B. das Verhältnis der R -Werte $R1(iam)/R1(mul)$ herangezogen werden. Je höher dieser Quotient ausfällt, desto stärker ist die Verbesserung. Nach diesem Verhältnis sind die Strukturen in Tab. 10.1 von links nach rechts geordnet. Zwar ist die Verringerung der Gütefaktoren ein wichtiges Merkmal für eine Modellverbesserung, kann aber nicht das alleinige Kriterium dafür sein. Ebenso bedeutsam ist die Abnahme der Temperaturfaktoren, der DMSDA-Beträge oder der Restdichte. Allerdings entwickeln sich diese Größen weitestgehend parallel zu den R -Werten, so daß das R -Wertverhältnis ein guter Indikator der Modellverbesserung zu sein scheint. In den nachfolgenden Abschnitten wird der Versuch unternommen, diese Kennzahl des Verfeinerungserfolges in einen funktionalen Zusammenhang mit den restlichen Merkmalen jeder Verfeinerung zu bringen.

Tabelle 10.1: Vergleich ausgewählter Parameter der verfeinerten Strukturen.

	APA0.83	APA	TRICHO- TOXIN	P1LYSO	TETRALY- SO	INSULIN
Gütefaktoren						
$R1(iam)/\%$	2,43	3,52	4,75	8,30	12,48	11,23
$R1(mul)/\%$	1,09	2,71	4,54	8,18	12,37	11,23
iam/mul	2,2294	1,2989	1,0463	1,0147	1,0089	1,0000
N_{ref}/N_{par}	5,6	47,6	5,1	30,4	15,9	14,9
Datenqualität						
$d/\text{Å}$	0,83	0,37	0,90	0,65	0,80	0,90

10 Vergleich der Strukturen und Modellverbesserungen

Vergleich ausgewählter Parameter der verfeinerten Strukturen (Forts.).

	APA0.83	APA	TRICHO- TOXIN	PILYSO	TETRALY- SO	INSULIN
Meßtemp./K	100	100	150	100	100	100
$R_{mrgd}^\dagger/\%$	0,7 (1,6)	4,9 (38,2)	5,1 (26,1)	4,5 (18,4)	3,4 (50,6)	4,1 (68,0)
$I/\sigma(I)^\dagger$	128 (68,2)	30,1 (3,8)	15,0 (2,4)	36,2 (4,2)	47,5 (3,7)	26,3 (2,6)
Redund. [†]	45,6 (23,6)	12,0 (2,6)	2,8 (2,2)	7,1 (2,7)	35,5 (13,2)	9,5 (3,5)
$> 2\sigma^\dagger/\%$	99,1	87,1	75,3	93,5	87,9	87,1
† Durchschnitt und Angabe in Klammern für höchste Auflösungsschale.						
Strukturunsicherheit und -normalität						
$\bar{\sigma}(xyz)/\text{Å}$	0,00218	0,00058	0,01181	0,02006	0,01606	0,04134
$\bar{\sigma}(U_{eq})/\text{Å}^2$	0,00028	0,00004	0,00162	0,00280	0,00285	0,00634
G	-0,12	-0,04	-0,02	-0,01	-0,09	-0,16
Median U_{eq} verf. Nicht-H-Atome des Hauptmol./ 10^{-3}Å^2						
<i>iam</i>	16,72	14,52	43,26	51,98	97,23	125,70
<i>mul</i>	15,13	14,40	41,11	51,46	96,18	125,20
<i>iam/mul</i>	1,1051	1,0083	1,0523	1,0101	1,0109	1,0040
Median DMSDA verf. Bindungen zw. Nicht-H-Atomen des Hauptmol./ 10^{-4}Å^2						
<i>iam</i>	7,0	3,5	54,0	15,0	42,0	71,5
<i>mul</i>	5,0	2,5	45,0	13,0	40,0	71,5
<i>iam/mul</i>	1,4000	1,4000	1,2000	1,1539	1,0500	1,0000
Restdichten $\sum \Delta\rho_{res} V_{asym}^{-1}/e\text{Å}^{-3}$						
<i>iam</i>	0,0278	0,0508	0,0491	0,1054	0,1312	0,0687
<i>mul</i>	0,0118	0,0375	0,0473	0,1038	0,1284	0,0676
<i>iam/mul</i>	2,3451	1,3540	1,0388	1,0150	1,0217	1,0157
Anteil Bindungen zw. Nicht-H-Atomen des Hauptmol. mit $\rho_{res} > 0/\%$						
<i>iam</i>	94,44	100	59,00	64,97	46,24	37,05
Verhältnis zu Deformationsdichte $\Delta\rho_{def}/\Delta\rho_{res}$						
<i>mul</i>	3,810	1,203	0,821	0,384	0,225	0,514
Asymmetrische Einheit, Solvensanteil via vdW-Vol. und Matthewscoeff.						
$V_{ASU}/\text{Å}^3$	335,2	335,2	4908	25057	28729	21658
$V_{Solv}(vdW)/\%$	26	26	30	42	51	46
$V_{Solv}(Mk.)/\%$	6	6	13	30	38	33
Nichtwasserstoffatome des Hauptmoleküls						
<i>N</i>	18	18	238	1001	1001	810
Fehlord./%	0	0	0	32,2	6,9	28,3
Verf./%	100	100	97,9	63,8	75,0	49,8

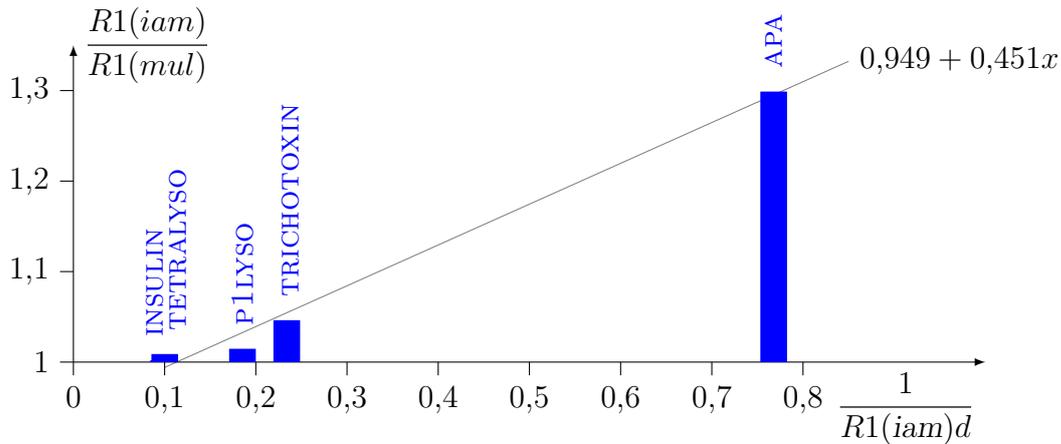


Abbildung 10.1: *R*-Wertverbesserung durch die asphärische Verfeinerung: Abhängigkeit vom sphärischen *R*-Wert (in %) und der Auflösung.

10.1 Gütefaktoren

Die drastischste Änderung der Gütefaktoren erfährt der APA0.83-Datensatz: Der $R1(iam0.83) = 2,43\%$ sinkt um 55 % auf $R1(mul0.83) = 1,09\%$. In den Verfeinerungen von APA liegen $R1(iam)$ und $R1(mul)$ bei 3,52 % und 2,71 % – es wird ein Quotient von 1,3 erzielt, der sphärische *R*-Wert liegt um 30 % über dem asphärischen. Wesentlich geringer stellt sich die Verbesserung im Fall TRICHOTOXIN dar. Hier nimmt der sphärische Gütefaktor von 4,75 % auf 4,54 % im Modell *mul* ab. Damit beträgt der relative *R*-Wertunterschied 4,6 %. Nochmals geringer ist das Verhältnis der Struktur P1LYSO, die *R*-Wertdifferenz, ausgehend von $R1(iam) = 8,30\%$, beträgt nur 0,12 Prozentpunkte zu $R1(mul) = 8,18\%$, entsprechend einer relativen Verbesserung um 1,5 %. Im Fall der Verbindung TETRALYSO beträgt der Absolutunterschied zwischen sphärischem und asphärischem *R*-Wert zwar ganz ähnliche 0,11 Prozentpunkte, was aber relativ nur noch 0,9 % Verbesserung entspricht, da der $R1(iam)$ 12,48 % beträgt und der $R1(mul)$ 12,37 %. Die Verfeinerung des INSULINS schließlich, erbrachte keine Verbesserung des *R*-Wertes. Entsprechend ist der obengenannte Quotient 1,0 – für beide Modelle *iam* und *mul* ist der $R1 = 11,23\%$.

Damit weisen die *R*-Wertverbesserungen fast die gleiche Reihenfolge wie die sphärischen *R*-Werte selbst auf. Mit Ausnahme von TETRALYSO geht mit steigendem $R(iam)$ -Wert eine geringere Abnahme durch die asphärische Verfeinerung einher. Normiert man den prozentualen $R1(iam)$ mit der Auflösung, wie in Abb. 10.1 geschehen, ergibt sich ein linearer Zusammenhang von $R1(iam)d$ und der relativen Änderung $R1(iam)/R1(mul)$, solange man die außergewöhnlich starke Verbesserung im Fall APA0.83 außer acht läßt.

10.2 Strukturgüte

Die Güte einer Struktur soll hier anhand der erreichten Auflösung und einigen Maßzahlen der Reflexqualität untersucht werden. Zusätzlich zu diesen auf den jeweiligen Datensatz bezogenen Parametern, werden auch die unterschiedlich hohen Standardabweichungen der Verfeinerung von Koordinaten und Verschiebungsparameter diskutiert.

Die unangefochten höchste Auflösung hat der APA-Datensatz mit $d = 0,37 \text{ \AA}$. Bemerkenswerter ist aber die bei P1LYSO erreichte zweithöchste Auflösung von $d = 0,65 \text{ \AA}$, was es von den ähnlich großen Strukturen TETRALYSO und INSULIN abhebt, für die nur Daten bis $0,8$ und $0,9 \text{ \AA}$ gewonnen werden konnten. Gleichfalls auffällig ist die geringe Auflösung von TRICHOTOXIN, welche nur $0,9 \text{ \AA}$ beträgt. Da diese Struktur eine merkliche Verbesserung erfährt, gibt es offenbar keinen direkten Zusammenhang von Auflösung und erzielter Modellverbesserung.

Auch rangiert TRICHOTOXIN an letzter Stelle, den Anteil starker Reflexe ($F_o^2 > 2\sigma(F_o^2)$) betreffend. Hierbei liegt P1LYSO an erster Stelle vor den restlichen Strukturen. Wenn die jeweilig erzielte Auflösung berücksichtigt wird, da die Reflexintensität mit dieser abfällt, relativiert sich das etwas und APA zeigt den höchsten Anteil starker Reflexe, gefolgt von P1LYSO. Erst mit einigem Abstand folgen die beiden anderen Makromoleküle und wieder zuletzt TRICHOTOXIN. Bis auf TRICHOTOXIN folgt das R -Wertverhältnis diesem Parameter.

Alle Messungen sind fast 100 % vollständig. Ob der hohen Auflösung, zeigen APA und P1LYSO etwas geringere Werte, insbesondere zu höchster Auflösung. Die durchschnittliche Redundanz von TETRALYSO ist am höchsten, die von TRICHOTOXIN am niedrigsten, dazwischen ordnen sich die restlichen Strukturen ein, zu maximalem d bleibt diese Relation erhalten. Fast genauso verhalten sich die Signal-zu-Untergrund-Verhältnisse $I/\sigma(I)$, wobei hier P1LYSO und APA die besten Werte in der letzten Auflösungsschale erreichen.

Auch die durchschnittlichen Werte des R_{mrgd} zeigen keine Korrelation mit den R -Werten, selbst bei Hinzunahme der Auflösung nicht. Gleiches gilt für die R_{mrgd} der jeweils äußersten Auflösungsschale oder deren Verhältnis zum Durchschnittswert. Allerdings ist diese relative Zunahme für TETRALYSO und INSULIN besonders ausgeprägt, die Datenqualität läßt hier zu höchster Auflösung ausgesprochen stark nach. Die geringste Zunahme ist für P1LYSO zu beobachten. Unter Einbeziehung der Auflösung, gilt dies auch für APA. Im Fall TRICHOTOXIN ist dieser Parameter nur wenig schlechter.

Für alle verglichenen Deskriptoren der Datenqualität, bis auf die Auflösung, zeigt APA0.83 drastisch bessere Werte. Damit verstärkt sich der Eindruck, daß bei guten Werten der Reflexstatistik eher niedrige Gütefaktoren und starke Modellverbesserungen mit den asphärischen Streufaktoren resultieren. Insbesondere sollten die Beobachtungen über den gesamten Auflösungsbereich von akzeptabler Qualität bleiben. Allerdings kann kein monotoner Zusammenhang einer dieser

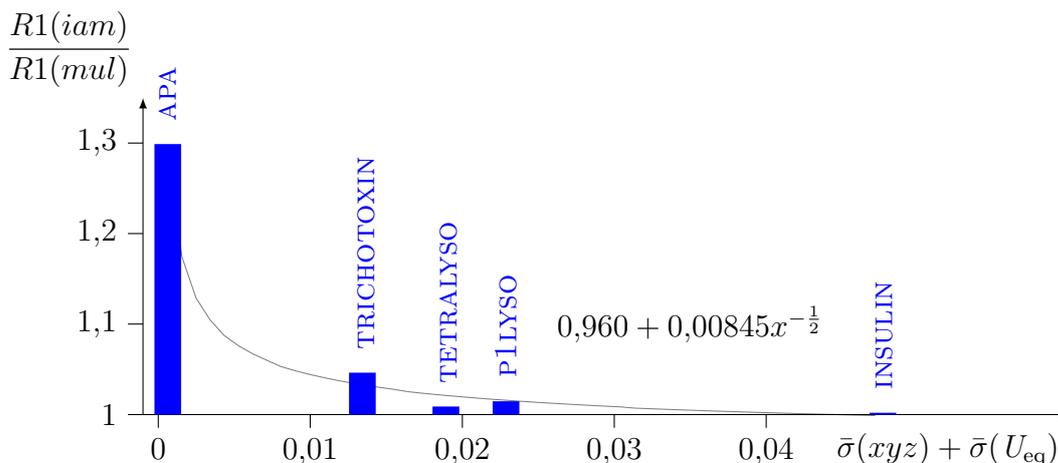


Abbildung 10.2: R -Wertverbesserung durch die asphärische Verfeinerung: Abhängigkeit von mittleren Fehlern der Koordinaten und U_{eq} .

Parameter mit dem Verfeinerungserfolg gefunden werden. Auch die Kombination mehrerer Parameter ergibt nicht die gesuchte Korrelation mit den R -Werten und deren Verbesserung.

Als Maß für eine unverzerrt und sicher bestimmte Struktur, listet Tab. 10.1 die mittleren Fehler der Koordinaten xyz und der Temperaturfaktoren U der Nichtwasserstoffatome des Hauptmoleküls jeder Struktur, wie sie als Standardabweichung der Parameter aus der Kleinsten-Quadrate-Verfeinerung erhalten werden, sowie die Gesamt-G-Faktoren (siehe hierzu Absch. 3.2 auf Seite 40) auf.

Da der G-Faktor aus dem Vergleich mit gemittelten stereochemischen Eigenschaften eines Satzes von guten Proteinstrukturen gewonnen wird, ergeben sich bei andersartigen Strukturen oder auch schon bei ungewöhnlichen Auflösungsbereichen, eventuell irreführende Einschätzungen, was aber bereits in den Abschnitten zu den einzelnen Molekülen angesprochen wurde. Alle Strukturen zeigen akzeptable bis gute Werte des Normalitätsfaktors, wobei INSULIN den schlechtesten G-Faktor aufweist. Den am wenigsten negativen und damit besten Wert erreichen P1LYSO und TRICHOTOXIN, dazwischen ordnen sich TETRALYSO und APA ein.

Von den Beschränkungen des G-Faktors unbeeinflusst sind die Parameterfehler, deren kleinster bei der hochaufgelösten APA-Struktur vorliegt. Bei Beschränkung des Datenumfangs, steigen die Unsicherheiten von APA0.83 auf etwa das Vierfache. Alle anderen aufgeführten Verfeinerungen weisen um den Faktor 20–70 höhere Unsicherheiten auf, wobei auf TRICHOTOXIN mit weiter steigendem Fehler TETRALYSO, P1LYSO und INSULIN folgen. Insbesondere letztgenannte Verbindung weist eine auffallend hohe mittlere Standardabweichung auf. Anhand Abb. 10.2 ist ein starker Zusammenhang zwischen relativer R -Wertverbesserung und der Höhe der mittleren Fehler der Koordinaten und Temperaturfaktoren zu erkennen.

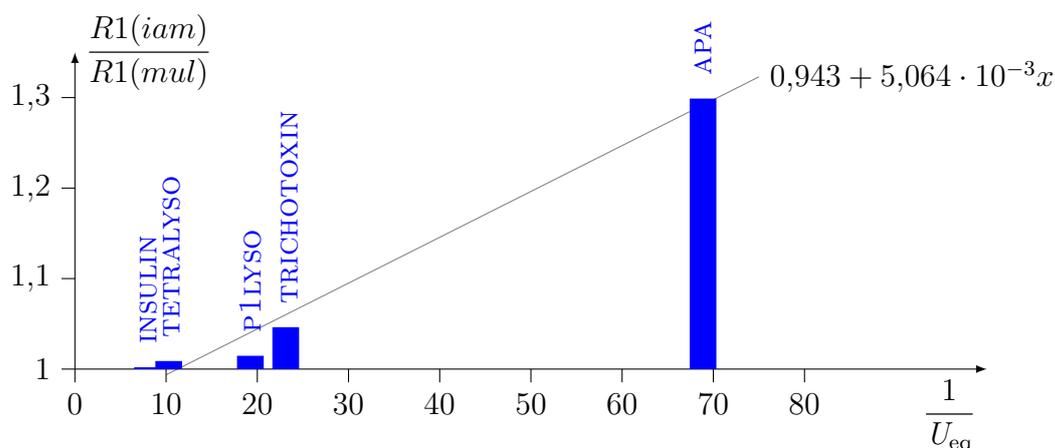


Abbildung 10.3: R -Wertverbesserung durch die asphärische Verfeinerung: Abhängigkeit vom sphärischen U_{eq} -Wert.

10.3 Verschiebungsparameter

Aufgrund der postulierten besseren Trennung von Beiträgen temperaturbedingter Schwingungen und Bindungseffekten zur Verschmierung der Elektronendichte, sollte die Größe der Temperaturfaktoren mit der asphärischen Verfeinerung abnehmen, da sie sich nicht mehr an Elektronendichte entlang von Bindungen oder in freien Elektronenpaaren anzupassen brauchen, um diese teilweise zu absorbieren. Tab. 10.1 zeigt die Entwicklung anhand des Medians aller verfeinerten Nichtwasserstoffatome des jeweiligen Moleküls, bzw. anhand des Quotientens dieser Größe im sphärischen und asphärischen Fall.

Wieder zeigt der in der Auflösung beschränkte Datensatz von APA die stärkste Veränderung, hier liegt der Median nach der Verfeinerung *iam0.83* um 10,5 % über dem der Verfeinerung *mul0.83*. Nun folgt in der Abstufung aber nicht der vollaufgelöste APA-Datensatz, der nur eine Änderung von 0,8 % aufweist. Auch liegt das Mittel bei hoher Auflösung niedriger als bei den beschränkten Daten, die höhere Auflösung bedingt also kleinere Temperaturfaktoren, die bereits im sphärischen Fall so gut sind, daß mit der Verwendung der Invariome kaum eine zusätzliche Verbesserung möglich ist. Die mittelgroße TRICHOTOXIN-Struktur weist mit $43 \cdot 10^{-3} \text{ \AA}^2$ einen durchschnittlich dreimal so hohen äquivalenten isotropen Temperaturfaktor wie die Kleinmolekülstruktur APA auf, der Einfluß des Modells *mul* beträgt ungefähr 5 %. Die beiden Lysozymstrukturen zeigen eine Verringerung des Medians der U_{eq} -Werte von nur noch 1 %, allerdings auf deutlich unterschiedlichem Niveau: Während P1LYSO mit $52 \cdot 10^{-3} \text{ \AA}^2$ Verschiebungsparameter in ähnlicher Höhe wie TRICHOTOXIN aufweist, liegen diese bei TETRALYSO mit $97 \cdot 10^{-3} \text{ \AA}^2$ doppelt so hoch. Einen nochmals höheren Median der Verschiebungsparameter findet man für INSULIN mit $126 \cdot 10^{-3} \text{ \AA}^2$, was das 8,7-fache der APA-Struktur und

immerhin das 2,4-fache von P1LYSO ist. Mit einem Quotienten von 1,004 liegt hier auch die geringste Verbesserung des U_{eq} -Mittelwertes vor.

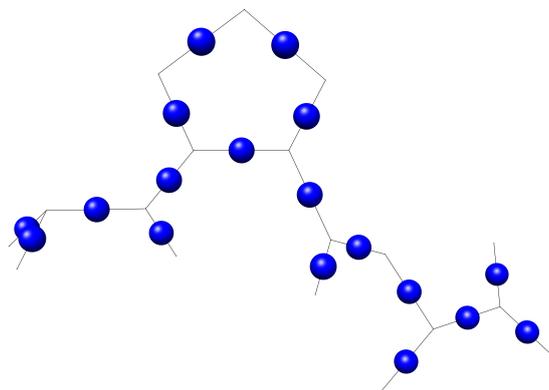
Während man zwischen sphärischem U_{eq} und der Verbesserung desselben nur einen ungefähren Zusammenhang erkennen kann, ist das relative R -Wertverhältnis umgekehrt proportional zum gemittelten Verschiebungsparameter der Strukturen, wie Abb. 10.3 zeigt. Diese Relation – je höher die Temperaturfaktoren, desto geringer die Verbesserung der R -Werte – wird bei Verwendung des Medians der U_{eq} aller Atome, statt nur der verfeinerten, noch deutlicher. Allerdings wäre wieder die besonders starke Verbesserung von APA0.83, bei gleichzeitig etwas höherem U_{eq} -Median als APA, eine Abweichung vom linearen Zusammenhang.

Noch uneinheitlicher ist das Verhältnis zwischen dem Median der DMSDA-Beträge nach sphärischer Verfeinerung und der im Modell *mul* erzielten Verbesserung, ebenso wie der Zusammenhang zum $R1$ -Verhältnis. Man kann nur statuieren, daß APA und APA0.83 mit den niedrigsten $|DMSDA|$ -Mittelwerten die höchste relative Verbesserung von 40 % ebendieser erfahren, was mit den stärksten relativen Abnahmen der Gütefaktoren zusammenfällt und daß genau Gegenteiliges für INSULIN gilt – zeh- bis zwanzigfach höherer $|DMSDA|$ -Median, der keinerlei Abnahme im Modell *mul* erfährt, in Kombination mit unveränderten R -Werten. Die Werte von TRICHOTOXIN und TETRALYSO liegen zwischen diesen Extremen und sind von ähnlicher Höhe, dennoch zeigt sich einmal eine Verbesserung von 20 %, im anderen Fall von nur 5 %. Einen verhältnismäßig guten Median der DMSDA-Beträge findet man bei P1LYSO, aber nur eine etwa 15 %ige Verbesserung.

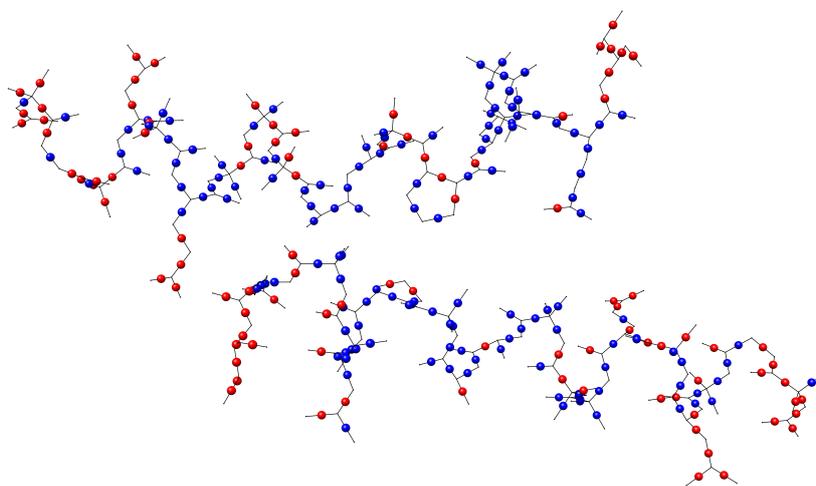
10.4 Restdichten

Die fünf Strukturen weisen auch recht unterschiedliche auf das jeweilige Zellvolumen normierte Restdichten auf, die, bei Berücksichtigung der jeweiligen Auflösung, mit dem Gütefaktor korrelieren. Hierbei fällt besonders INSULIN auf, das für den erreichten R -Wert erstaunlich niedrige Restdichten aufweist. Die mit der Invariomverfeinerung erzielte Senkung der Restdichte in den Strukturen verhält sich ganz ähnlich dem R -Wertverhältnis: Während im Fall APA und APA0.83 eine deutliche Reduktion der Restdichten um 30-60 % erzielt werden konnte, liegen die prozentualen Abnahmen bei den größeren Strukturen im niedrigen einstelligen Bereich. Letzteres erklärt sich zum einen aus der Tatsache, daß der Großteil an Restdichte außerhalb des Moleküls in der Solvenssphäre diffus verteilt ist, wo das veränderte Verfeinerungsmodell keinen direkten Einfluß nimmt und deshalb auch nichts an den Restdichten ändert. Zum anderen ist die sinnvolle und erfolgreiche Verwendung der asphärischen Modellierung nur möglich, wenn die Restdichten im Molekül der Erwartung entsprechen, daß sie an Orten von Bindungen und freien Elektronenpaaren lokalisiert sind.

Um dies zu testen, wurden die Restdichtewerte des Modells *iam* an den Koordina-



(a) APA: $\Delta\rho_{res}(iam)@Bindungen$



(b) TRICHOTOXIN: $\Delta\rho_{res}(iam)@Bindungen$

Abbildung 10.4: Restdichte des Modells *iam* auf den Bindungsmitten: blau – positiv, grün – null, rot – negativ.

ten der Bindungsmitten, wie in Absch. 3.4 auf Seite 44 beschrieben, ermittelt. Mit dieser Näherung ist zumindest eine Abschätzung der Frage nach den benötigten Restdichtelokalisierung möglich. Allerdings werden keine Restdichten berücksichtigt, die durch freie Elektronenpaare verursacht werden, obwohl auch diese mit der asphärischen Verfeinerung modelliert werden können und so zur Modellverbesserung beitragen. Durch die Spline-Interpolation beeinflusst zudem auch die unmittelbare Umgebung der Bindungen das Ergebnis. Mit den Abbildungen 10.4(a) bis 10.4(e) wird das Vorzeichen der Restdichte nach sphärischer Verfeinerung auf den Bindungsmitten der Nichtwasserstoffatome dargestellt: Blaue, grüne und rote Platzhalter symbolisieren positive, keine und negative Restdichten auf den Bindungen der fünf Moleküle.

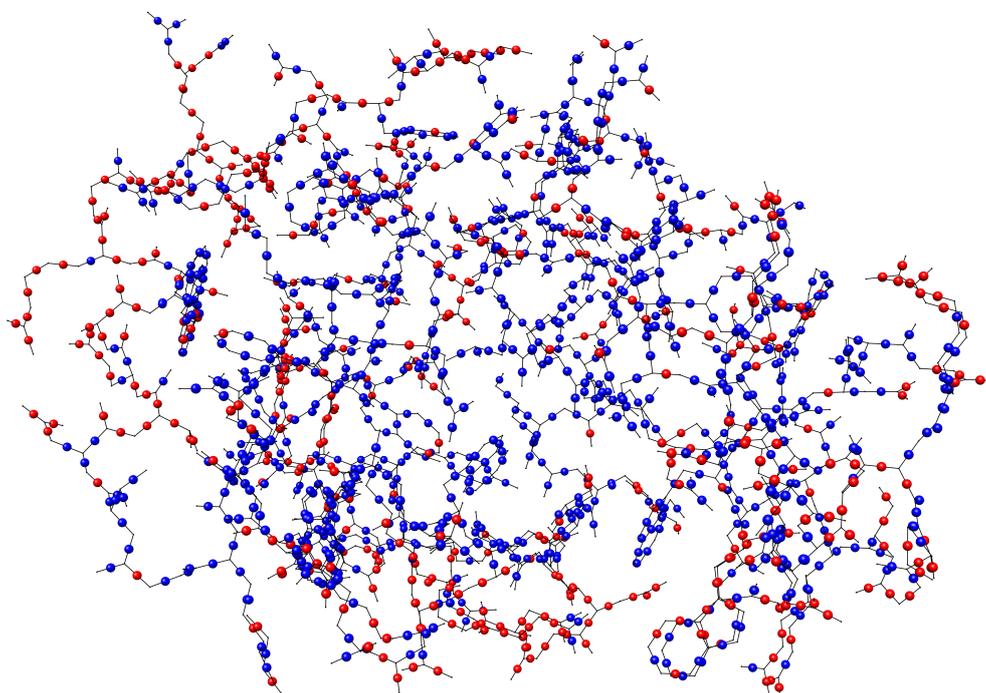
(c) P1LYSO: $\Delta\rho_{res}(iam)@$ Bindungen

Abbildung 10.4: Restdichte des Modells *iam* auf den Bindungsmitten: blau – positiv, grün – null, rot – negativ. (Forts.)

In der Kleinmolekülstruktur APA verbleiben bei sphärischer Verfeinerung deutliche Restdichten auf den Bindungen (siehe auch Absch. 5.5 auf Seite 123), alle 18 Platzhalter sind in Abb. 10.4(a) blau gefärbt. Durch die Verfeinerung *mul* sinkt der Mittelwert aller so ermittelten Bindungsrestdichten um 95 %, die Hälfte der Bindungen zeigt nun eine schwach negative Restdichte.

Bei der mittelgroßen TRICHOTOXIN-Struktur tragen 98 oder 41 % der 238 Bindungen eine Restdichte kleiner als Null (Abb. 10.4(b)), das arithmetische Mittel über alle Bindungen ist positiv. Insbesondere an den randständigen, schlechter bestimmten, Bereichen ist das Problem lokalisiert. Wie in Absch. 9 auf Seite 247 gezeigt wurde, gibt es dabei eine Übereinstimmung mit Strukturteilen mit höherem U_{eq} und der Nachbarschaft zu den höchsten Restdichtekonzentrationen. Mit der asphärischen Verfeinerung nimmt der Anteil der Bindungen mit negativer Restdichte auf 58 % zu und der Mittelwert verläßt den positiven Bereich.

Die Makromolekülstruktur P1LYSO weist 483 Bindungen mit negativen Restdichten auf, was 35 % aller 1378 Bindungen entspricht. Auch hier ist der Durchschnitt der Restdichten an den Bindungsmitten positiv, blau illustrierte Bindungen überwiegen in Abb. 10.4(c). Der Zusammenhang mit Bereichen höherer mittlerer Verschiebungsparameter ist nicht so deutlich wie im vorhergehenden Fall, aber vor-

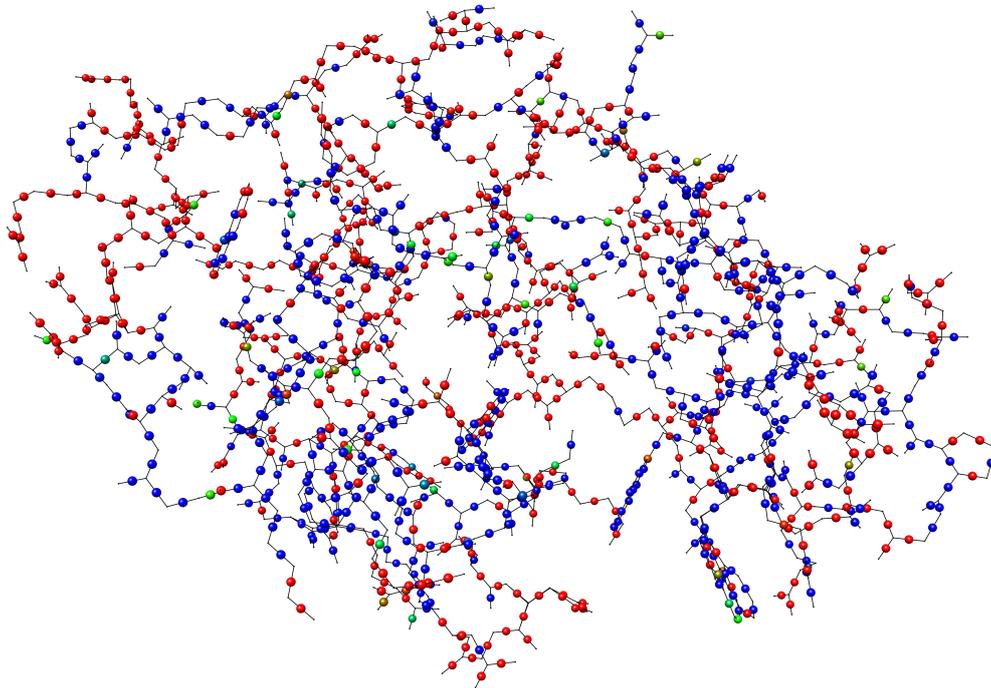
(d) TETRALYSO: $\Delta\rho_{res}(iam)$ @Bindungen

Abbildung 10.4: Restdichte des Modells *iam* auf den Bindungsmitten: blau – positiv, grün – null, rot – negativ. (Forts.)

handen (siehe hierzu auch Absch. 6.3 auf Seite 156). Ebenso fallen die betroffenen Bereiche häufig mit den nicht zur freien Verfeinerung freigegebenen Strukturteilen zusammen. Auch sind die 353 Bindungen der Minderpopulation mit 42 % häufiger betroffen als die der Hauptpopulation (33 %). Das gilt aber nicht generell für die fehlgeordnet beschriebenen Bereiche, da diese 628 Bindungen mit 37 % nur geringfügig häufiger eine negative Restdichte aufweisen als die 750 in nur einer Konformation beschriebenen Bindungen, welche zu 33 % dieses Phänomen ausprägen. Durch Anwendung der Invariome kommt es zu einem negativen Mittelwert der Restdichten auf den Bindungsmitten, 66 % aller Bindungen weisen nun eine Restdichte unter Null aus.

Die TETRALYSO-Struktur zeigt auf 593 oder 54 % aller 1102 Bindungen eine negative Restdichte (Abb. 10.4(d)). Nun ist hierbei auch der Mittelwert aller Bindungsrestdichten negativ. Ein Vergleich mit der U_{eq} -Verteilung in Absch. 7.3 auf Seite 194 ergibt keinen eindeutigen Zusammenhang zum Ausbleiben positiver Bindungsrestdichten, da sich auch Bereiche mit niedrigem Temperaturfaktor durch negative Restdichten auszeichnen, wiewohl Strukturteile mit hohem U_{eq} fast durchgängig dieses Problem aufweisen. Auch besteht kein Zusammenhang zur Einteilung der Struktur in fehlgeordnet und geordnet modellierte Bereiche. In

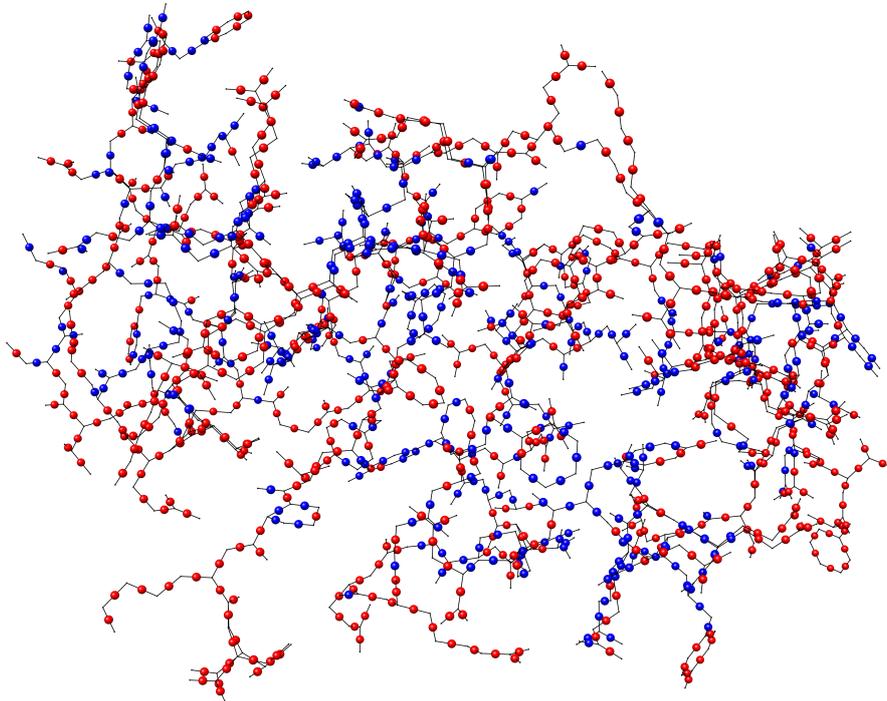
(e) INSULIN: $\Delta\rho_{res}(iam)@$ Bindungen

Abbildung 10.4: Restdichte des Modells *iam* auf den Bindungsmitten: blau – positiv, grün – null, rot – negativ. (Forts.)

den fehlgeordnet beschriebenen Bereichen tragen 48 % der 137 Bindungen eine negative Restdichte, während es bei den geordnet modellierten 965 Bindungen sogar 55 % sind. Gleiches gilt für den Vergleich der Bindungen von Haupt- und Minderpopulation. Mit dem Modell *mul* verstärkt sich der negative Mittelwert und es sind nun insgesamt 61 % aller Bindungen, deren Restdichte negativ ist.

Im Fall von INSULIN ergibt sich wieder ein negativer Durchschnitt der Restdichten in Bindungsmitte, es sind 683 der insgesamt 1084 Bindungen betroffen, was 63 % entspricht. Deutlich wird Abb. 10.4(e) durch rotgefärbte Bindungstypen dominiert. Wieder gilt dies für beinahe alle Bindungen zwischen Nichtwasserstoffatomen mit hohem Verschiebungsparameter, allerdings auch für sehr viele Bindungen von Atomen mit geringerem U_{eq} (siehe auch Absch. 8.3 auf Seite 224). Die 640 geordnet und 444 fehlgeordnet beschriebenen Bindungen sind jeweils zu 61 % und 67 % mit negativer Restdichte versehen. Noch ausgewogener ist das Verhältnis in Bezug auf Haupt- und Nebenpopulation, hier zeigen 63 % der 833 Bindungen der stärker besetzten Konformation und 64 % der 252 Bindungen im niederpopulierten Strukturteil eine Restdichte kleiner Null. Nach der asphärischen Verfeinerung tragen 68 % der Bindungen eine negative Restdichte, der Mittelwert sinkt ebenfalls auf einen negativeren Wert.

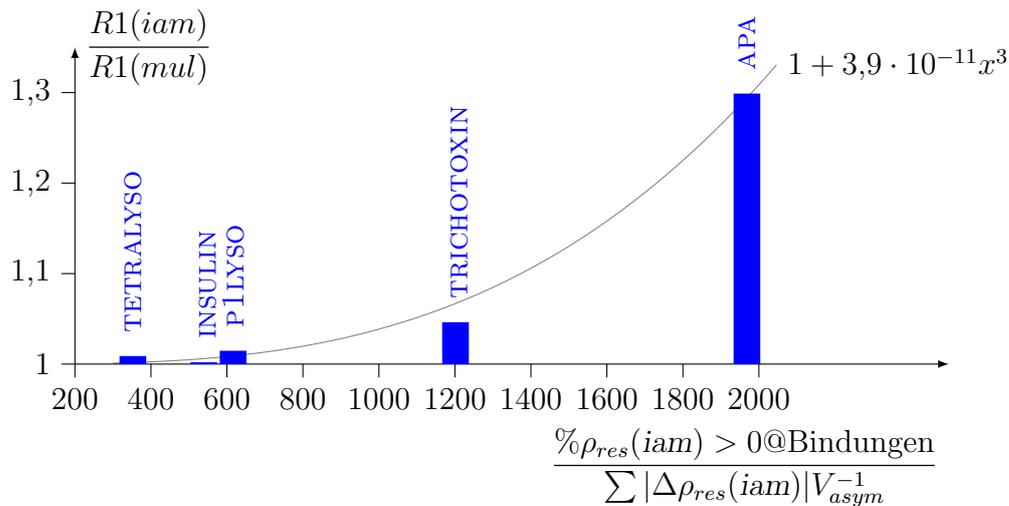


Abbildung 10.5: R -Wertverbesserung durch die asphärische Verfeinerung: Abhängigkeit von der Gesamtdichte und dem Anteil von Bindungen mit positiver Restdichte.

Der Unterschied zwischen der kleinen APA-Struktur und den größeren Molekülen ist augenfällig. Es gibt allerdings keinen direkten Zusammenhang zu den Verfeinerungs- R -Werten. Eine Kombination aus dem Anteil von Bindungen mit positiver Restdichte und der Gesamtdichte pro Volumeneinheit der Zelle zeigt nun eine Korrelation mit dem zur Abschätzung des Verfeinerungserfolges benutzten $R1(iam)/R1(mul)$ -Quotienten: Wie Abb. 10.5 nahelegt, ist dieser umso größer, je kleiner die sphärische Gesamtdichte ist und je höher dabei der Anteil der Bindungen an der positiven Restdichte ist. Allerdings ist der Zusammenhang nicht linear, was allein anhand der nicht passenden Anordnung von TETRALYSO und INSULIN in Abb. 10.5 zu erkennen ist.

Weitestgehend kongruent zu den Verhältnissen des Gütefaktors sind auch die jeweiligen $\Delta\rho_{def}/\Delta\rho_{res}$ -Quotienten, die das Verhältnis von Deformationsdichte und der Restdichte der jeweiligen Struktur im Modell mul beziffern. Je stärker der Umfang der Deformationsdichte die vom Modell unbeschriebene Restdichte überschreitet, desto höher der R -Wertquotient und damit die erzielte Verbesserung im Modell mul (siehe Abb. 10.6 auf der nächsten Seite). Alle größeren Moleküle haben dabei ein Verhältnis kleiner eins, die verbliebene Restdichte übersteigt also die Deformationsdichte, was die geringen Auswirkungen der asphärischen Verfeinerung auf z. B. die Gütefaktoren plausibel erscheinen läßt. Ein weiteres Mal steht INSULIN einem monotonen Zusammenhang im Wege, was auf die außergewöhnlich niedrige Restdichte zurückzuführen ist. Allerdings handelt es sich hierbei auch um die Struktur mit besonders vielen Bindungen ohne positive Restdichte, die von der Deformationsdichte vorteilhaft beeinflusst werden könnte.

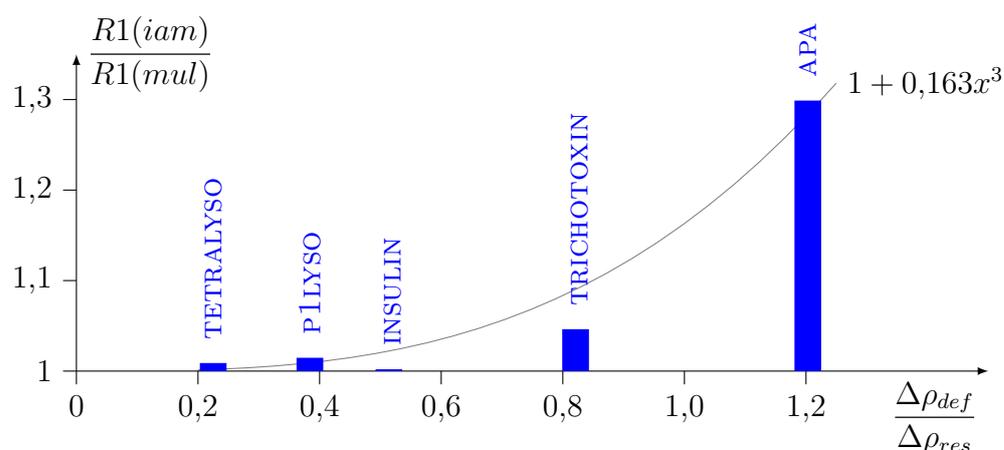


Abbildung 10.6: *R*-Wertverbesserung durch die asphärische Verfeinerung: Abhängigkeit vom Verhältnis von Deformations- zu Restdichte.

10.5 Geometrieänderung

Für alle Strukturen sind die durch das asphärische Modell hervorgerufenen Geometrieänderungen sehr gering, insbesondere im Hinblick auf die mittleren Fehler der Koordinaten, wenn die Nichtwasserstoffatome und deren Bindungen untereinander betrachtet werden. So beläuft sich der quadratisch gemittelte Positionsunterschied der vier größeren Strukturen in den Modellen *iam_{rb}* und *mul*, auf 0,003 – 0,006 Å, bei mittleren Positionsunsicherheiten von 0,012 – 0,041 Å. In den Fällen APA und APA0.83, liegen die mittleren Verschiebungen der Nichtwasserstoffatome bei 0,0008 und 0,003 Å, denen Fehler von nur 0,0006 und 0,002 Å gegenüberstehen. Nur das kleinste Molekül zeigt höhere Mittelwerte der Koordinatendifferenz als der Koordinatenfehler. Interessant ist die wesentlich stärkere Verschiebung bei APA0.83, die einen ähnlich hohen Wert wie die Makromoleküle hat – die verringerte Auflösung ergibt höhere Verschiebungen und Unsicherheiten. Mit Maximalabweichungen von 0,019–0,045 Å der Nichtwasserstoffatomkoordinaten der vier größeren Strukturen, liegen diese etwa zehnmal höher als das Mittel. Wesentlich geringer ist diese Relation für APA und APA0.83, deren Maximalabweichungen bei jeweils 0,002 und 0,006 Å liegen, somit ungefähr dem Doppelten des Mittels. Der RMSD-Mittelwert der Gesamtstrukturen in den Modellen *iam* und *mul* beträgt für alle Strukturen etwa 0,095 Å, die Abweichung durch die verschiedenen Wasserstoffatomabstände ist das dominante Merkmal.

Ganz ähnliche Zahlenwerte ergibt die Betrachtung der Bindungslängenunterschiede in der Peptidgruppe, wobei nur die in Absch. 6.2 auf Seite 151 erläuterten und in „Nächster Nachbar Näherung“ differenzierten Bindungstypen berücksichtigt wurden. Der Durchschnitt der Mediane der Längendifferenzen liegt zwischen 0,0015 und 0,0072 Å, wobei P1LYSO die geringsten Werte aufweist, gefolgt von

TRICHOTOXIN und TETRALYSO mit vergleichbaren 0,004 Å und schließlich INSULIN, dessen Bindungslängen in den verschiedenen Modellen am stärksten differieren. Für die entsprechenden Bindungen in APA und APA0.83, ist die durchschnittliche Bindungslängendifferenz mit 0,0005 und 0,001 Å wieder geringer, bei verminderter Auflösung nähert sie sich den Werten der Makromoleküle an. Die Maximaländerungen liegen für APA bei 0,0013 Å und etwa doppelt so hoch für APA0.83. Bei den vier anderen Strukturen können diese aber wesentlich stärker ausfallen: 0,023 Å im Fall von P1LYSO, 0,024 Å für TRICHOTOXIN, 0,033 Å bei INSULIN und 0,040 Å zwischen den Modellen von TETRALYSO.

Hieraus folgt, daß mit abnehmender Auflösung und Strukturqualität die Modellunterschiede bezüglich Atompositionen und Bindungslängen zunehmen, wie es auch die Standardabweichungen der Koordinaten tun. Zumindest für die Makromoleküle sind die mittleren Änderungen geringer als die Unsicherheiten, während die ca. zehnfach höheren Maximalunterschiede unplausibel sind.

Trotzdem scheinen die Geometrieänderungen nicht völlig zufällig zu sein, da in allen Strukturen ähnliche Entwicklungen der Peptidbindungslängen vom sphärischen zum asphärischen Modell beobachtet werden konnten: Am deutlichsten ist die Verkürzung der C–O-Carbonylbindung, die, jeweils für APA, P1LYSO, TETRALYSO, INSULIN und TRICHOTOXIN, Differenzmediane von $-0,001$, $-0,003$, $-0,005$, $-0,009$ und $-0,011$ Å aufweist, um im Modell *mul* einen Median der Bindungsabstände von 1,240, 1,230, 1,223, 1,229 und 1,227 Å zu zeigen. Im Fall der C_α–N-Bindung betragen die mittleren Änderungen und Endwerte, jeweils in obiger Reihenfolge, 0,000, $-0,001$, $-0,002$, $-0,003$ und $-0,004$ Å, sowie 1,462, 1,452, 1,456, 1,451 und 1,453 Å. Hier ist die Verkürzung geringer ausgeprägt, die einzige Bindung dieses Typs in APA bleibt sogar unverändert. Die C_α–C-Bindungen zeigen fast durchgehend eine Verlängerung, nur bei APA tritt eine kleine Verkürzung auf. Die Werte nach obigem Schema lauten $-0,0004$, 0,001, 0,008, 0,012 und 0,012 Å für die mittleren Unterschiede und 1,538, 1,529, 1,536, 1,530 und 1,530 Å. Im Fall der C–N-Bindung der fünf Strukturen, kommt es dreimal zu einer Zunahme des Abstandes und zweimal zu einer geringen Abnahme. Die Differenzen betragen nun im Median $-0,0002$, $-0,0006$, 0,002, 0,004 und 0,005 Å, die Bindungslängen im asphärischen Modell 1,346, 1,338, 1,339, 1,337 und 1,331 Å.

Bei APA0.83 verlängert sich die C–O-Bindung um durchschnittlich 0,008 Å auf 1,236 Å, die C_α–N-Bindung nimmt um $-0,002$ Å auf 1,463 Å ab, ebenso wie die C_α–C-Bindung um $-0,008$ Å auf 1,540 Å und die C–N-Bindung um $-0,0001$ Å auf 1,344 Å. Es handelt sich bei APA0.83 und APA insofern um Sonderfälle, als in deren „Statistik“ jeweils nur ein bis zwei Werte eingehen, was die hohen Abweichungen erklären könnte.

Ansonsten ist der übereinstimmende Trend gut zu erkennen, aber auch, daß die Mittelwerte der jeweiligen Bindungstypen in den verschiedenen Strukturen recht unterschiedlich sind. Zudem sind die Differenzen geringer als die korrespondierenden Streuungsmaße, die sich kaum zwischen den Modellen unterscheiden.

10.6 Strukturgröße

Zwar ist die Größe einer Verbindung und die Menge an Solvens nicht per se mit Strukturqualität und Verbesserung im asphärischen Modell zu assoziieren, aber zumindest fällt auf, daß die kleinsten R -Werte aller fünf Strukturen für APA und TRICHOTOXIN erreicht werden und dies mit den kleinsten Volumina der asymmetrischen Einheit korrespondiert. Die wesentlich größeren Proteinstrukturen P1LYSO, TETRALYSO und INSULIN haben 75-mal so große asymmetrische Einheiten wie APA und immerhin noch fünfmal so große wie TRICHOTOXIN. Dabei weisen sie auch deutlich höhere R -Werte auf.

Weiterhin haben APA und TRICHOTOXIN den kleinsten Solvensanteil. Dies gilt sowohl für die Betrachtung nach der Anzahl der Atome als auch für die Volumenanteile, die aus dem Matthews-Koeffizienten oder aus dem van-der-Waals-Volumen berechnet werden. Letzterer prozentualer Volumenanteil liegt prinzipiell höher als der nach Matthews und zwar umso stärker, je kleiner die Struktur ist. Das könnte auf die abnehmende Gültigkeit der Gl. 2.2 auf Seite 32 zugrunde liegenden Annahme zur Proteindichte von $\rho_{prot} = 1,35 \text{ g/cm}^3$ zurückgeführt werden, da TRICHOTOXIN und APA keine „richtigen“ Proteine sind. Mit der Bestimmung der Volumenanteile über das vdW-Volumen ist man von dieser Annahme unabhängig. Zwischen den beiden kleineren Strukturen und den drei Makromolekülen gibt es einen starken Unterschied im Solvensanteil, gleich nach welchem Maß, der Anstieg ist sprunghaft beim Übergang zu P1LYSO, TETRALYSO und INSULIN.

Zudem sind es die beiden kleineren Moleküle, die fehlordnungsfrei sind. Wobei einzuschränken ist, daß die Modellierung oder Negierung von Fehlordnung keinen strengen Regeln gehorcht, sondern je nach Deutlichkeit und Nützlichkeit vorgenommen wird. Wenn ein schlecht bestimmter Strukturbereich zwar Fehlordnung nahelegt, Dichten und Restdichten aber keine Lokalisierung zulassen und zudem dadurch weder R -Werte noch Verschiebungsparameter verbessert werden, braucht dort auch keine weitere Konformation modelliert zu werden. Ebenso kann es angebracht erscheinen, Atome fehlgeordnet zu verfeinern, obwohl es keine starken Anzeichen für die Zweitkonformation gibt, um beispielsweise ein zusammenhängenderes und besser erklärbares Modell zu erhalten. So könnte man im Fall TRICHOTOXIN durchaus annehmen, daß ein oder zwei Randgruppen eine weitere Konformation aufweisen, die allerdings schlecht zu erkennen und zu modellieren ist. Ähnliches gilt für TETRALYSO, welches den geringsten Fehlordnungsanteil der drei Makromoleküle aufweist.

Ebenso ist die Auswahl der in der asphärischen Verfeinerung freigegebenen Atome etwas willkürlich, auch wenn dabei immer beachtet wurde, möglichst viele Parameter zu verfeinern, ohne zu hohe Verschiebungen nach sich zu ziehen. Beim Tri- und Oktadekapeptid wurden alle, bzw. fast alle Koordinaten und Verschiebungsparameter verfeinert. Die entsprechenden Anteile der Makromoleküle betragen aber nur 50–75 %, der geringeren Strukturqualität Rechnung tragend.

10.7 Schlußfolgerungen zu den Voraussetzungen für den Erfolg einer asphärischen Verfeinerung

Die letzten Abschnitte verdeutlichten, daß es keinen einzelnen Parameter der Strukturen gibt, der den Erfolg einer asphärischen Verfeinerung allein erklärt, da es immer mindestens eine Struktur gibt, die den monotonen Zusammenhang zum Erfolgsindikator $R1(iam)/R1(mul)$ verhindert, gleiches gilt auch für die relativen Verbesserungen der Verschiebungsparameter, DMSDA-Werte oder Restdichten. Zudem ist die erzielte Güte der Anpassung von Ausgleichsfunktionen an die vermuteten Zusammenhänge nur befriedigend. Dennoch soll anhand der untersuchten Abhängigkeiten versucht werden, die Voraussetzungen für eine erfolgreiche asphärische Verfeinerung zu quantifizieren.

Letztere bemißt sich in der erzielten Modellverbesserung. Man kann sowohl eine Verkleinerung der Temperaturfaktoren und DMSDA-Werte erwarten als auch eine Verbesserung der Geometrie im Sinne von Bindungslängen und -winkeln. Ebenso sollte die Restelextronendichte abnehmen, was natürlich nur im Einflußbereich der asphärischen Verfeinerung auf den Bindungen und den freien Elektronenpaaren geschehen kann. Insgesamt sollte dies zu besseren Gütefaktoren führen, weshalb deren relativer Unterschied als Indikator gewählt wurde, der auch die sonstigen Verbesserungen weitestgehend synchron begleitet.

Ob die Verwendung eines asphärischen Streumodells die Verfeinerung verbessert, ist von der Güte der sphärisch beschriebenen Struktur abhängig, da die asphärische Beschreibung eine verhältnismäßig kleine Korrektur zum sphärischen Modell darstellt. Dabei ist die „Strukturgüte“ schwierig zu definieren und subsummiert letztlich alle zuvor detailliert besprochenen Merkmale.

Auflösung und initiale Gütefaktoren

Sowohl eine hohe Auflösung als auch niedrige Gütefaktoren sprechen für eine gute Startstruktur. Zwar ist ein mit der Auflösung moderat steigender R -Wert akzeptabel, allerdings nicht unbegrenzt: Setzt man als Ziel der asphärischen Verfeinerung eine wenigstens zehnpromtente Verbesserung an, legt Abb. 10.1 auf Seite 271 einen maximalen Wert von $R1(iam) \cdot d \approx 3$ nahe. Daraus folgt, daß bei einer Auflösung von $0,6 \text{ \AA}$ ein $R1$ von 5% hinnehmbar ist, dieser bei z. B. $d = 0,8$ aber nur noch 3,75% betragen dürfte, um eine relevante Verbesserung zu erzielen.

Verschiebungsparameter

Mit einer guten Struktur sind üblicherweise niedrige Verschiebungsparameter verbunden, die auf wohldefinierte Atome und Bindungen hinweisen, wodurch Bindungseffekte beobachtbar werden. Für eine zehnpromtente relative Verbesserung der Gütefaktoren kann anhand Abb. 10.3 auf Seite 274 ein maximaler Zentralwert der sphärischen äquivalenten isotropen Temperaturfaktoren von $\approx 32 \cdot 10^{-3} \text{ \AA}^2$ festgelegt werden. Ein solcher Grenzwert entspricht etwa dem Doppelten der

Struktur APA und etwas mehr als der Hälfte von P1LYSO.

Die Verbesserung der Verschiebungsparameter selbst wird ebenfalls mit niedrigeren Ausgangswerten steigen. Hier gilt die Einschränkung, daß der Effekt mit steigender Auflösung kleiner wird, da die Hochwinkeldaten bereits im sphärischen Fall für eine ausreichend gute Bestimmung sorgen und die Absorption von Bindungsdichte durch die Temperaturfaktoren verhindern.

Präzision der Lage- und Verschiebungsparameter

Eine hohe Genauigkeit der Bestimmung von Koordinaten und Verschiebungsparametern ist laut Abb. 10.2 auf Seite 273 nötig, um einen 10 % besseren $R1(mul)$ als $R1(iam)$ zu erreichen: Der Term $\bar{\sigma}(xyz) + \bar{\sigma}(U_{eq})$ dürfte einen Wert von 0,0036 nicht überschreiten. Da die mittlere Standardabweichung der Koordinaten etwa zehnmal höher ist als die der Temperaturfaktoren (siehe Tab. 10.1), kann das Ergebnis hauptsächlich mit der Koordinatenunsicherheit assoziiert werden. Es entspricht etwa dem Sechsfachen des Fehlers von APA und umgekehrt ungefähr einem Sechstel der Unsicherheiten der vier größeren Strukturen.

Restdichte

Besonders eng mit den Gütefaktoren ist die vom jeweiligen Modell unbeschriebene Restelektronendichte verbunden. Für die erzielte Verbesserung durch das asphärische Modell gilt dies auch, aber hier spielt die Lokalisierung der Restdichten eine große Rolle. Sie sollte auf Bindungen und an den freien Elektronenpaaren lokalisiert sein, um vom asphärischen Modell beeinflußt werden zu können.

Der Quotient aus prozentualem Anteil der Bindungen mit positiver Restdichte und der Gesamtrestdichte pro Volumeneinheit der Zelle (Abb. 10.5 auf Seite 280), sollte über ≈ 1400 liegen, um die angestrebte relative R -Wertabnahme von 10 % zu erreichen. Tragen 100 % der Bindungen positive Restdichte, würde dies bis zu $0,073 e/\text{\AA}^3$ Gesamtrestdichte erlauben. Sind nur noch 50 % der Bindungen mit positiven Beiträgen der Restdichte versehen, so sinkt die tolerable Gesamtrestdichte auf $0,036 e/\text{\AA}^3$.

Liegen die Gesamtrestdichtewerte wesentlich höher, weil z. B. der Solvensbereich viel unbeschriebene Dichte beinhaltet, so kann das asphärische Modell diese nicht inkorporieren, weshalb die Modellverbesserung entsprechend gering bleibt. Vermutlich stellt die Population der Hälfte der Bindungen mit positiver Restdichte die untere Grenze einer sinnvollen asphärischen Behandlung dar. Im Vergleich zu obigen Grenzwerten, zeigt APA eine etwa 30 % bessere Restdichte, während z. B. TRICHOTOXIN eine ca. 14 % zu hohe Restdichte aufweist. Die drei großen Proteinstrukturen verfehlen die Limits um mehrere Vielfache.

Das Verhältnis aus Deformationsdichte, welche den Unterschied zwischen sphärischer und asphärischer Modellierung kennzeichnet, und der unbeschriebenen Restelektronendichte, stellt eine andere Sichtweise auf das gleiche Problem dar. Aus Abb. 10.6 auf Seite 281 ergibt sich ein Quotient von mindestens 0,85, um die

genannte Verbesserung zu erreichen. Somit sollte die Restdichte die Deformationsdichte um nicht mehr als 18 % überschreiten, als sehr grobe Näherung kann man von 0,4 e pro Atom der Struktur ausgehen, um die Menge an Deformationsdichte abzuschätzen. Sowohl für Rest- wie Deformationsdichte sind immer Betragswerte gemeint.

Datenqualität

Die Datensatzstatistiken können nicht in einen gleichförmigen Zusammenhang mit den R -Werten oder der Modellverbesserung gebracht werden, was auf den Einfluß durch ein evtl. unvollständiges Strukturmodell hindeutet. Für die gewählten Größen interner R -Wert R_{mrgd} , Signal-zu-Rausch-Verhältnis $I/\sigma(I)$, Redundanz und Anteil starker Reflexe $I > 2\sigma(I)$, scheint die Strukturgröße am ehesten mit der Relation zwischen dem jeweiligen Mittelwert und dem entsprechenden Wert der äußersten Auflösungsschale zu korrespondieren.

Strukturgröße

Die zuvor genannten Voraussetzungen für eine sinnvolle und erfolgreiche asphärische Verfeinerung sind für kleine Strukturen wie APA leicht zu erfüllen. Kleine und starre Moleküle mit starken intermolekularen Wechselwirkungen ergeben eher eine hochwertige Startstruktur, die in einer kleinen asymmetrischen Einheit mit hoher Packungsdichte vorliegt. Sie kristallisieren leichter in einem Energieminimum, werden kaum Fehlordnung aufweisen und eine geordnete Solvenssphäre ausprägen. Deshalb sind Daten zu höherer Auflösung zugänglich und alle Koordinaten diskret verfeinerbar, weshalb die Restdichten nach sphärischer Verfeinerung lokalisiert auf den Bindungen, statt diffus verteilt, vorliegen.

Für Makromoleküle sind die Anforderungen schwer zu erfüllen. Diese großen Strukturen tendieren, ob der vielen konformativen Freiheitsgrade, zu beweglichen Bereichen, deren Fehlordnung nur bedingt diskret beschreibbar ist, weshalb häufig hohe, ungenau bestimmte Verschiebungsparameter resultieren. Diese tragen zu verringerten Bindungsrestdichten bei, da sie nicht nur die recht große Flexibilität der Proteinkette beschreiben, sondern auch die Restdichten auf den Bindungen und sonstige Modellungenauigkeiten absorbieren. Bei proteintypischer Daten- und Strukturqualität ist die Fähigkeit des asphärischen Modells zur Dekonvolution von Bindungsdichte und Verschiebungsparametern eingeschränkt. Nur in besonders sicher bestimmten Bereichen, zeigen deshalb kleine Verschiebungsparameter auch die stärkste Verbesserung, da sie hauptsächlich durch nicht beschriebene Bindungsdichte verfälscht waren. Proteine sind zudem verhältnismäßig lose gepackt, der Freiraum zwischen den Molekülen ist mit sehr beweglichen und fast durchgängig fehlgeordneten Solvensmolekülen, zumeist Wasser, besetzt. Das Strukturmodell hat einen mehr oder weniger starken Näherungscharakter, insbesondere der Solvensbereich wird nie vollständig beschreibbar sein und Bindungseffekte überlagernde Restdichten aufweisen.

Dabei ist zu bedenken, daß die hier untersuchten Proteinstrukturen bezüglich Auflösung und Gütefaktoren überdurchschnittlich gut sind. So zählt die triklin Lysozymmodifikation P1LYSO von Wang u. a. (2007) mit einer Auflösung von $d = 0,65 \text{ \AA}$ und einem konventionellen $R_{all}(F)$ von 8,4% zu den besten literaturbekannten Messungen (siehe auch Absch. 2.1 auf Seite 25). Trotzdem sind die Verbesserungen ausgesprochen gering. Ebenso übertrifft die im Rahmen dieser Arbeit gemessene tetragonale Lysozymmodifikation TETRALYSO mit $d = 0,80 \text{ \AA}$ und $R_{all}(F) = 13,1\%$ alle bisherigen Datensätze analoger Strukturen. Hier werden noch geringere Fortschritte als bei P1LYSO erzielt. Fast keine Verbesserung erbringt die asphärische Verfeinerung von INSULIN, obwohl auch dieser Datensatz mit $d = 0,90 \text{ \AA}$ mit $R_{all}(F) = 11,5\%$ weit höhere Auflösung und niedrigere Gütefaktoren als sonstige Insulinstrukturen in der PDB aufweist.

Auch bei deutlicher Verbesserung des aktuellen Prozederes – z. B. durch ein elaborierteres Modell zur Behandlung der diffusen Lösungsmittelbereichs oder die restrainte Verfeinerung aller Atompositionen und Temperaturfaktoren – erscheinen höchstens Verbesserungen auf dem Niveau von TRICHOTOXIN realistisch. Mit einer dem Gros an Proteinstrukturen entsprechenden Strukturqualität jenseits von TETRALYSO oder INSULIN, werden keine nennenswerten Verbesserungen möglich sein. Ein Vorteil durch die asphärische Modellierung wie bei kleinen Molekülen ist ausgeschlossen.

10.8 Vergleich zu literaturbekannten Studien

Die erzielten Ergebnisse sind insofern in Übereinstimmung mit den in technischer Hinsicht von dieser Arbeit recht verschiedenen Arbeiten von Jelsch u. a. (2000a) zu Crambin und Guillot u. a. (2008) zu Aldosereduktase, als auch für diese Strukturen nur sehr geringe relative R -Wertverbesserungen von 2,2% respektive 4,7% berichtet werden. Nur für Aldosereduktase sind auch die Änderungen der Verschiebungsparameter und DMSDA-Werte angegeben, sie verbessern sich mit 2,4% und 20% etwas stärker als bei den Proteinstrukturen dieser Arbeit. Die beiden Proteine gehören zusammen mit P1LYSO (Wang u. a. 2007) zu den zehn besten PDB-Einträgen mit einem Gütefaktor-Auflösungs-Produkt von $R \cdot d < 6$, unter denen Aldosereduktase und P1LYSO die zwei größten Strukturen sind.

Zudem kann das R -Wertverhältnis der sphärischen zur asphärischen Verfeinerung mit den hier gefundenen funktionalen Zusammenhängen in guter Näherung prognostiziert werden: Für Crambin wird anhand des $R \cdot d$ -Indikators ein Verhältnis von 1,04 bestimmt, tatsächlich liegt die Veränderung bei nur 1,02. Im Fall der Aldosereduktase ergibt sich anhand dieses Indikators die Prognose 1,03, anhand der Relation zum mittleren Temperaturfaktor die Vorhersage 1,04 – das tatsächliche Verhältnis liegt bei 1,05.

Zusammenfassung

In dieser Arbeit wurden erstmalig drei Proteine nach Transfer theoretisch abgeleiteter Invariom-Multipolpopulationen asphärisch verfeinert: Zwei im Rahmen dieser Studie bestimmte Strukturen rhomboedrischen Insulins und tetragonalen Lysozyms sowie triklines Lysozym (2VB1; Wang u. a. 2007). Dies wurde durch analoge Untersuchungen an dem Oktadekapeptid Trichotoxin (1M24; Chugh u. a. 2002) und dem Tripeptid L-Alanyl-L-Prolyl-L-Alanin ergänzt. Die Veränderungen gegenüber einem sphärischen Streumodell wurden ausführlich anhand struktureller und elektronischer Eigenschaften analysiert.

Obwohl bezüglich Auflösung und Gütefaktoren überdurchschnittlich gute Proteinstrukturen untersucht wurden, resultierten nur geringe Verbesserungen durch die asphärische Verfeinerung: Die relative Abnahme der Gütefaktoren lag bei höchstens 1,5 %, das Mittel der Verschiebungsparameter sank um etwa 1 %, das der DMSDA-Werte um maximal 15 %. Aus der asphärischen Verfeinerung des Oktadekapeptids folgten eine deutlichere relative Abnahme des Gütefaktors um 4,4 % und um 5 % bzw. 20 % verringerte mittlere Verschiebungsparameter und DMSDA-Werte. Damit kontrastieren die wesentlich stärkeren Verbesserungen des Tripeptids bei hoher und niedriger Auflösung, hier nahmen Gütefaktoren und Restdichten um 25 % bis 50 % ab, der mittlere Verschiebungsparameter um bis zu 10 % und das DMSDA-Mittel um 30 %.

Allen Proteinen sind recht hohe, diffus über das Molekül und insbesondere den Solvensbereich verteilte Restdichten gemein. Dem stehen nur geringe Akkumulationen auf 37–65 % der Bindungen gegenüber, weshalb der Einfluß des atomzentrierten asphärischen Modells beschränkt blieb.

Die Geometrieänderungen bspw. der Peptidgruppe um 0,001–0,010 Å folgten zwar einem allgemeinen Trend verkürzter C–O- und C_α–N-Bindungslängen sowie verlängerter C_α–C- und C–N-Bindungen, lagen für die Proteine aber innerhalb der Koordinatenunsicherheit.

Für jede Struktur wurde das elektrostatische Potential aus den Multipolen abgeleitet. Der Vergleich mit einem Punktladungspotential zeigte nur für die beiden Lysozymstrukturen deutliche Unterschiede. Ob einer der beiden Beschreibungen der Vorzug zu geben ist, bedarf weiterer Untersuchungen. Ebenso bleibt zu klären, inwieweit eine eventuelle Anpassung der Invariome an die generell niedrigen Bindungsrestdichten und die Geometrieverzerrungen von Proteinen eine verbesserte Modellierung ermöglicht.

Es wurden funktionale Zusammenhänge zwischen dem Effekt des asphärischen

Modells und sphärischen Ausgangsparametern untersucht. Um relevante Verbesserungen zu erzielen, sollte nach sphärischer Verfeinerung unter anderem der Temperaturfaktormedian $U_{\text{eq}} \approx 32 \cdot 10^{-3} \text{ \AA}^2$ ($B_{\text{eq}} \approx 2,5 \text{ \AA}^2$) nicht überschreiten sowie das Produkt aus prozentualem $R1$ -Wert und Auflösung unter 3 liegen.

Mittels dieser Abhängigkeiten lassen sich auch die Ergebnisse der Studien zu Crambin (Jelsch u. a. 2000a) und Aldosereduktase (Guillot u. a. 2008) prognostizieren, welche dieser Arbeit ähnliche Resultate ergaben.

Im Lichte dieser Studie und der Literaturbeispiele kann die aufwendigere asphärische Verfeinerung von Proteinstrukturen nicht empfohlen werden, solange keine wesentlich höherwertigen Daten gewonnen werden können. Davon unberührt bleibt der Ansatz, transferierbare asphärische Dichten zur Ableitung elektrostatischer Eigenschaften einzusetzen.

Summary

In this study three proteins were aspherically refined after transfer of theoretically predicted Invariom multipole populations for the first time. Refinement was performed on structures of rhombohedral insulin, tetragonal lysozyme and triclinic lysozyme (2VB1; Wang *et al.* 2007), the first two were measured within this study. In addition the octadecapeptide Trichotoxin (1M24; Chugh *et al.* 2002) and the tripeptide L-Alanyl-L-Prolyl-L-Alanin were also studied. The changes in relation to a spherical scattering model were analyzed in detail on the basis of structural and electronic characteristics.

Although the proteins under investigation were of above average quality concerning resolution and residual factors, only small improvements resulted from the aspherical refinement: The highest relative reduction of residual factors was about 1.5 %, the mean of the displacement parameters decreased by approximately 1 %, the DMSDA mean values by at most 15 %. The aspherical refinement of the octadecapeptide resulted in the more distinct relative decrease of the residual factor of 4.4 %, the mean displacement parameter and DMSDA value reduced by 5 % and 20 %, respectively. Much stronger improvements were found for the tripeptide at high and low resolution. The residuals decreased by about 25 % and 50 %, respectively, the mean displacement parameter by about 10 % and the mean DMSDA value by 30 %.

All proteins showed a diffuse distribution of substantial residual density over the whole molecule, in particular in the solvent region. Conversely only small accumulations are found on 37–65 % of the bonds. Therefore, the influence of the atom-centered aspherical model was rather limited.

The refinements resulted in geometry changes mostly between 0.001 and 0.010 Å. For instance the trend in the peptide groups was a shortening of the C–O- and C_α–N-bonds and an elongation of the C_α–C- und C–N-bonds. However, these are within the uncertainty of the protein coordinates.

For each structure, the electrostatic potential was derived from the multipole density. In comparison to point-charge potentials, there were clear differences only in case of the two lysozyme modifications. Whether one or the other description should be preferred, requires further investigations. Likewise, it remains to be clarified, to what extent a possible adjustment of the Invarioms, for the small bonding density and the distorted geometry in proteins. would make an improved modelling possible.

Functional relations between the effect of the aspherical model and initial

spherical parameters were examined. In order to obtain relevant improvements, among other things, the median of the temperature factors should not exceed $U_{\text{eq}} \approx 32 \cdot 10^{-3} \text{ \AA}^2$ ($B_{\text{eq}} \approx 2.5 \text{ \AA}^2$) after spherical refinement. Furthermore, the product of $R1$ (value in percent) and the resolution should be less than 3.

By means of these functions, predictions of the results from studies on crambin (Jelsch *et al.* 2000a) and aldose reductase (Guillot *et al.* 2008) are possible, where improvements similar to this work were reported.

In the light of this study and of examples from the literature, the complex aspherical refinement of protein structures can not be recommended as long as no data of substantially higher quality is available. Nevertheless, the possibility remains to use transferable aspherical densities to derive electrostatic characteristics.

Danksagung

Ich bedanke mich bei Herrn Professor Dr. Peter Luger für die Aufnahme in die Arbeitsgruppe und die Unterstützung während der Arbeit, besonders für die Ermöglichung etlicher Besuche nationaler und internationaler Tagungen.

Bei den Arbeitsgruppenmitgliedern Herrn Dr. Wolfgang Dreissig, Frau Dr. Diana Förster, Herrn Dipl.-Chem. Simon Grabowsky, Herrn Dr. Christian Hübschle, Herrn Dr. Stefan Mebs, Herrn Dr. Stephan Scheins, Frau Dr. Marianna Strümpel, Frau Manuela Weber und Herrn Dr. Dieter Zobel bedanke ich mich für ihre stete Hilfsbereitschaft, die angenehme Arbeitsatmosphäre und anregende Diskussionen. Insbesondere danke ich Stefan Mebs für das Korrekturlesen dieser Arbeit.

Die Zusammenarbeit mit Herrn Dr. Birger Dittrich war sehr wichtig, ich bedanke mich für seine freundliche fachliche und moralische Unterstützung und die Bereitstellung einer von Prof. Dr. Edgar Weckert parallelisierten XD-Version.

Entscheidend für das Gelingen der Arbeit waren auch die Kooperationen mit Herrn Dr. Armin Wagner an der DLS und Herrn Dr. Alke Meents von der SLS, deren Kompetenz und Hilfsbereitschaft ich die Datensätze des Insulins und des tetragonalen Lysozyms verdanke.

Ferner danke ich Herrn Dr. Carsten Paulmann und Herrn Dr. Wolfgang Morgenroth für die Betreuung der Messungen der Fullerene am HASYLAB.

Mein Dank gilt auch Herrn Prof. Dr. Dieter Lentz für seine Hilfs- und Diskussionsbereitschaft sowie die Bereitstellung seiner Laborkapazitäten.

Für die finanzielle Förderung danke ich der DFG (Projekte SPP1178 und Lu222/27-1,2) und dem Graduiertenkolleg Nr. 788.

Literaturverzeichnis

- Afonine u. a. 2007** AFONINE, P. V. ; GROSSE-KUNSTLEVE, R. W. ; ADAMS, P. D. ; LUNIN, V. Y. ; URZHUMTSEV, A. : On macromolecular refinement at subatomic resolution with interatomic scatterers. In: *Acta Crystallographica Section D* 63 (2007), Nov, Nr. 11, S. 1194–1197
- Afonine u. a. 2004** AFONINE, P. V. ; LUNIN, V. Y. ; MUZET, N. ; URZHUMTSEV, A. : On the possibility of the observation of valence electron density for individual bonds in proteins in conventional difference maps. In: *Acta Cryst.* D60 (2004), S. 260–274
- Alkorta u. Picazo 2005** ALKORTA, I. ; PICAZO, O. : Influence of protonation on the electron density derived properties. In: *ARKIVOC* ix (2005), S. 305–320
- Allen u. a. 1992** *Kapitel 9.5.* In: ALLEN, F. H. ; KENNARD, O. ; WATSON, D. G. ; BRAMMER, L. ; ORPEN, A. G. ; TAYLOR, R. : *International Tables for X-ray Crystallography*. Bd. C. Amsterdam : Kluwer Academic Publishers, 1992, S. 685–706
- Allen 2002a** ALLEN, F. H.: The Cambridge Structural Database: A Quarter of a Million Crystal Structures and Rising. In: *Acta Crystallogr. Sect. B* B58 (2002), S. 380–388
- Allen 2002b** ALLEN, F. H.: The Cambridge Structural Database: a quarter of a million crystal structures and rising. In: *Acta Cryst.* B58 (2002), S. 380–388
- Ambrogelly u. a. 2007** AMBROGELLY, A. ; PALIOURA, S. ; SÖLL, D. : Natural expansion of the genetic code. In: *Nature Chemical Biology* 3 (2007), S. 29–35
- Andersson u. Hovmöller 1998** ANDERSSON, K. M. ; HOVMÖLLER, S. : The average atomic volume and density of proteins. In: *Z. Krist.* 213 (1998), S. 369–373
- Andersson u. Hovmöller 2000** ANDERSSON, K. M. ; HOVMÖLLER, S. : The protein content in crystals and packing coefficients in different space groups. In: *Acta Cryst.* D56 (2000), S. 789–790
- Atkins 1996** ATKINS, P. W. ; HÖPFNER, A. (Hrsg.): *Physikalische Chemie*. 2. Aufl. Weinheim; New York; Basel; Cambridge; Tokyo : VCH, 1996

- Bader 1990** BADER, R. F. W.: *Atoms in Molecules - A Quantum Theory*. Oxford : University of Oxford Press, 1990
- Bader 1991** BADER, R. F. W.: A Quantum Theory of Molecular Structure and Its Applications. In: *Chem. Rev.* 91 (1991), S. 893–928
- Bader 2008** BADER, R. F. W.: Nearsightedness of Electronic Matter As Seen by a Physicist and a Chemist. In: *J. Phys. Chem.* 112 (2008), S. 13717–13728
- Bader 2009** BADER, R. F. W.: Bond Paths Are Not Chemical Bonds. In: *J. Phys. Chem.* 113 (2009), S. 10391–10396
- Bader u. Gatti 1998** BADER, R. F. W. ; GATTI, C. : A Green's Function for the Density. In: *Chem. Phys. Lett.* 287 (1998), May, S. 233–238
- Bak u. a. 2009** BAK, J. M. ; DOMINIAK, P. M. ; WILSON, C. C. ; WOŹNIAK, K. : Experimental charge-density study of paracetamol – multipole refinement in the presence of a disordered methyl group. In: *Acta Crystallographica Section A* 65 (2009), Nr. 6, S. 490–500
- Berman 2008** BERMAN, H. M.: The Protein Data Bank: a historical perspective. In: *Acta Crystallographica Section A* 64 (2008), Jan, Nr. 1, 88–95. <http://dx.doi.org/10.1107/S0108767307035623>. – DOI 10.1107/S0108767307035623
- Berman u. a. 2000** BERMAN, H. ; WESTBROOK, J. ; FENG, Z. ; GILLILAND, G. ; BHAT, T. ; WEISSIG, H. ; SHINDYALOV, I. ; BOURNE, P. : The Protein Data Bank. In: *Nucleic Acids Research* 28 (2000), S. 235–242
- Blake u. a. 1965** BLAKE, C. C. F. ; KOENIG, D. F. ; MAIR, G. A. ; NORTH, A. C. T. ; PHILLIPS, D. C. ; SARMA, V. R.: Structure of Hen Egg-White Lysozyme: A Three-dimensional Fourier Synthesis at 2 Å Resolution. In: *Nature* 206 (1965), S. 757–761
- Brock u. a. 1991** BROCK, C. P. ; DUNITZ, J. D. ; HIRSHFELD, F. L.: Transferability of Deformation Densities among Related Molecules : Atomic Multipole Parameters from Perylene for Improved Estimation of Molecular Vibrations in Naphthalene and Anthracene. In: *Acta Crystallogr. Sect. B* B47 (1991), S. 789–797
- Bunkoczi u. a. 2003** BUNKOCZI, G. ; SCHIELL, M. ; VERTESY, L. ; SHELDRIK, G. : Crystal Structures of Cephaibols. In: *J. Pept. Sci.* 9 (2003), S. 745
- Chandler u. Spackman 1978** CHANDLER, G. S. ; SPACKMAN, M. A.: Fourier Transform of Gaussian orbital products. In: *Acta Crystallogr. Sect. A* A34 (1978), S. 341–343

- Chang u. Bader 1992** CHANG, C. ; BADER, R. F. W.: Theoretical Construction of a Polypeptide. In: *J. Phys. Chem.* 96 (1992), Nr. 4, S. 1654–1662
- Chugh u. a. 2002** CHUGH, J. K. ; BRÜCKNER, H. ; WALLACE, B. : Model for a helical bundle channel based on the high-resolution crystal structure of trichotoxin A50E. In: *Biochemistry* 41 (2002), S. 12934–12941
- Chugh u. Wallace 2001** CHUGH, J. K. ; WALLACE, B. A.: Peptaibols: Models for Ion Channels. In: *Biochem. Soc. Trans.* 29 (2001), S. 565–570
- Clementi u. Raimondi 1963** CLEMENTI, E. ; RAIMONDI, D. L.: Atomic Screening Constants from SCF Functions. In: *J. Chem. Phys.* 38 (1963), June, Nr. 11, S. 2686–2689
- Clementi u. Roetti 1974** CLEMENTI, E. ; ROETTI, C. : Roothaan-Hartree-Fock Atomic Wavefunctions. In: *Atomic and Nuclear Data Tables* 14 (1974), S. 177–478
- Collaborative Computational Project 1994** COLLABORATIVE COMPUTATIONAL PROJECT, N. ...: The CCP4 suite: programs for protein crystallography. In: *Acta Cryst.* D50 (1994), S. 760–763
- Connolly 1996** CONNOLLY, M. L.: Molecular Surfaces: A Review. (1996). <http://www.netsci.org/Science/Compchem/feature14.html>, Abruf: 27.01.2010
- Coppens 1967** COPPENS, P. : Comparative X-Ray and Neutron Diffraction Study of Bonding Effects in s-Triazine. In: *Science* 158 (1967), S. 1577–1579
- Coppens 1997** COPPENS, P. : *X-Ray Charge Densities and Chemical Bonding*. 1. Oxford : Oxford University Press, 1997 (IUCr Texts on Crystallography 4). – ISBN 0–19–509823–4
- Coppens 2005** COPPENS, P. : Charge Densities Come of Age. In: *Angew. Chem. Int. Ed.* 44 (2005), S. 6810–6811
- Cornell u. a. 1995** CORNELL, W. D. ; CIEPLAK, P. ; BAYLY, C. I. ; GOULD, I. R. ; MERZ, K. M. ; FERGUSON, D. M. ; SPELLMEYER, D. C. ; FOX, T. ; CALDWELL, J. W. ; KOLLMAN, P. A.: A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. In: *J. Am. Chem. Soc.* 117 (1995), S. 5179–5197
- Cortes-Guzman u. Bader 2004** CORTES-GUZMAN, F. ; BADER, R. : Role of functional groups in linear regression analysis of molecular properties. In: *J. Phys. Org. Chem.* 17 (2004), S. 95–99

- Cremer u. Pople 1975** CREMER, D. ; POPLE, J. A.: A general definition of ring puckering coordinates. In: *J. Am. Chem. Soc.* 97 (1975), S. 1354–1358
- Dauter u. a. 1995** DAUTER, Z. ; LAMZIN, V. S. ; WILSON, K. S.: Proteins at atomic resolution. In: *Current Opinion in Structural Biology* 5 (1995), Nr. 6, S. 784 – 790
- Dauter u. a. 1997** DAUTER, Z. ; LAMZIN, V. S. ; WILSON, K. S.: The benefits of atomic resolution. In: *Current Opinion in Structural Biology* 7 (1997), Nr. 5, S. 681 – 688
- Dawson 1967** DAWSON, B. : A General Structure Factor Formalism for Interpreting Accurate X-ray and Neutron Diffraction Data. In: *Proc. R. Soc. Lond. A* 298 (1967), S. 255–263
- Debye 1915** DEBYE, P. : In: *Ann. Phys.* 46 (1915), 809–823 S.
- Dewar u. a. 1985** DEWAR, M. J. S. ; ZOEBISCH, E. G. ; HEALY, E. F. ; STEWART, J. J. P.: AM1: A New General Purpose Quantum Mechanical Model. In: *J. Am. Chem. Soc.* 107 (1985), S. 3902–3909
- Diederichs u. a. 2003** DIEDERICHS, K. ; MCSWEENEY, S. ; RAVELLI, R. B. G.: Zero-dose extrapolation as part of macromolecular synchrotron data reduction. In: *Acta Cryst.* D59 (2003), S. 903–909
- Dittrich u. a. 2009a** DITTRICH, B. ; HÜBSCHLE, C. B. ; HOLSTEIN, J. J. ; FABBIANI, F. P. A.: Towards extracting the charge density from normal-resolution data. In: *Journal of Applied Crystallography* 42 (2009), Dec, Nr. 6, S. 1110–1121
- Dittrich u. a. 2006a** DITTRICH, B. ; HÜBSCHLE, C. B. ; LUGER, P. ; SPACKMAN, M. A.: Introduction and Validation of an Invariom Database for Amino Acid, Peptide and Protein Molecules. In: *Acta Crystallogr. Sect. D* D62 (2006), S. 1325–1335
- Dittrich u. a. 2005** DITTRICH, B. ; HÜBSCHLE, C. B. ; MESSERSCHMIDT, M. ; KALINOWSKI, R. ; GIRNT, D. ; LUGER, P. : The invariom model and its application: Refinement of D,L-serine at different temperatures and resolution. In: *Acta Crystallogr. Sect. A* A61 (2005), S. 314–320
- Dittrich u. a. 2004** DITTRICH, B. ; KORITSÁNSZKY, T. ; LUGER, P. : A Simple Approach to Nonspherical Electron Densities by Using Invarioms. In: *Angew. Chem. Int. Ed.* 43 (2004), S. 2718–2721

- Dittrich u. a. 2008** DITTRICH, B. ; MCKINNON, J. J. ; WARREN, J. E.: Improvement of anisotropic displacement parameters from invariom-model refinements for three L-hydroxylysine structures. In: *Acta Crystallographica Section B* 64 (2008), Dec, Nr. 6, S. 750–759
- Dittrich u. a. 2006b** DITTRICH, B. ; MUNSHI, P. ; SPACKMAN, M. A.: Structure determination and invariom refinement of L-Valinol. In: *Acta Crystallogr. Sect. C* (2006), S. to be published
- Dittrich u. a. 2007** DITTRICH, B. ; MUNSHI, P. ; SPACKMAN, M. A.: Redetermination, invariom-model and multipole refinement of L-ornithine hydrochloride. In: *Acta Crystallographica Section B* 63 (2007), Jun, Nr. 3, S. 505–509
- Dittrich u. a. 2003** DITTRICH, B. ; SCHEINS, S. ; PAULMANN, C. ; LUGER, P. : Transferability of Atomic Volumes and Charges in the Peptide Bond Region in the Solid State. In: *J. Phys. Chem. A* 107 (2003), S. 7471–7474
- Dittrich u. Spackman 2007** DITTRICH, B. ; SPACKMAN, M. A.: Can the interaction density be measured? The example of the non-standard amino acid sarcosine. In: *Acta Crystallographica Section A* 63 (2007), Sep, Nr. 5, S. 426–436
- Dittrich u. a. 2006c** DITTRICH, B. ; STRÜMPPEL, M. ; KORITSÁNSZKY, T. ; SCHÄFER, M. ; SPACKMAN, M. A.: Invarioms for improved absolute structure determination of light-atom molecular structures. In: *Acta Crystallogr. Sect. A* (2006), S. submitted.
- Dittrich u. a. 2000** DITTRICH, B. ; FLAIG, R. ; KORITSÁNSZKY, T. ; KRANE, H.-G. ; MORGENROTH, W. ; LUGER, P. : Topological Properties of the Peptid Bond in Glycyl-L-threonine Dihydrate Based on a Fast Synchrotron/CCD-Diffraction Experiment. In: *Chem. Eur. J.* 6 (2000), S. 2582–2589
- Dittrich u. a. 2002** DITTRICH, B. ; KORITSÁNSZKY, T. ; GROSCHE, M. ; SCHERER, W. ; FLAIG, R. ; WAGNER, A. ; KRANE, H.-G. ; KESSLER, H. ; RIEMER, C. ; SCHREURS, A. M. M. ; LUGER, P. : Reproducibility and Transferability of Topological Properties; Experimental Charge Density of the Hexapeptide cyclo-(D,L-Pro₂)-(L-Ala)₄ Monohydrate. In: *Acta Crystallogr. Sect. B* 58 (2002), S. 721–727
- Dittrich u. a. 2009b** DITTRICH, B. ; WARREN, J. E. ; FABBIANI, F. P. A. ; MORGENROTH, W. ; CORRY, B. : Temperature dependence of rotational disorder in a non-standard amino acid from X-ray crystallography and molecular dynamics simulation. In: *Phys. Chem. Chem. Phys.* 11 (2009), Nr. 15, S. 2601–2609

- Dittrich u. a. 2009c** DITTRICH, B. ; WEBER, M. ; KALINOWSKI, R. ; GRABOWSKY, S. ; HUEBSCHLE, C. B. ; LUGER, P. : How to easily replace the independent atom model - the example of bergenin, a potential anti-HIV agent of traditional Asian medicine. In: *Acta Cryst.* B65 (2009), Nr. 6, S. 749–756
- Dominiak u. a. 2009** DOMINIAK, P. M. ; VOLKOV, A. ; DOMINIAK, A. P. ; JARZEMBSKA, K. N. ; COPPENS, P. : Combining crystallographic information and an aspherical-atom data bank in the evaluation of the electrostatic interaction energy in an enzyme-substrate complex: influenza neuraminidase inhibition. In: *Acta Cryst.* D65 (2009), S. 485–499
- Dominiak u. a. 2007** DOMINIAK, P. M. ; VOLKOV, A. ; LI, X. ; MESSERSCHMIDT, M. ; COPPENS, P. : A Theoretical Databank of Transferable Aspherical Atoms and its Application to Electrostatic Interaction Energy Calculations of Macromolecules. In: *J Chem. Theory. Comput.* 3 (2007), Nr. 1, S. 232–247
- Duan u. a. 2003** DUAN, Y. ; WU, C. ; CHOWDHURY, S. ; LEE, M. C. ; XIONG, G. ; ZHANG, W. ; YANG, R. ; CIEPLAK, P. ; LUO, R. ; LEE, T. ; CALDWELL, J. ; WANG, J. ; KOLLMAN, P. : A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. In: *J. Comp. Chem.* 24 (2003), S. 1999–2012
- Dupradeau u. a. 2008** DUPRADEAU, F.-Y. ; CEZARD, C. ; LELONG, R. ; STANISLAWIAK, E. ; PECHER, J. ; DELEPINE, J. C. ; CIEPLAK, P. : R.E.D.D.B.: A database for RESP and ESP atomic charges, and force field libraries. In: *Nucl. Acids Res.* 36 (2008), S. D360–D367
- Emsley u. Cowtan 2004** EMSLEY, P. ; COWTAN, K. : Coot: Model-Building Tools for Molecular Graphics. In: *Acta Crystallographica Section D - Biological Crystallography* 60 (2004), S. 2126–2132
- Engh u. Huber 1991** ENGH, R. ; HUBER, R. : Accurate Bond and Angle Parameters for X-ray Protein Structure Refinement. In: *Acta Cryst.* A47 (1991), S. 392–400
- Fischer u. Tillmanns 1988** FISCHER, R. X. ; TILLMANNS, E. : The equivalent isotropic displacement factor. In: *Acta Cryst.* C44 (1988), S. 775–776
- Flack 1983** FLACK, H. D.: On Enantiomorph-Polarity Estimation. In: *Acta Crystallogr. Sect. A* A39 (1983), S. 876–881
- Fleming 1922** FLEMING, A. : On a remarkable bacteriolytic element found in tissues and secretions. In: *Proc. Roy. Soc.* B93 (1922), S. 306–317

- Fogolari u. a. 2002** FOGOLARI, F. ; BRIGO, A. ; MOLINARI, H. : The Poisson-Boltzmann equation for biomolecular electrostatics: a tool for structural biology. In: *J. Mol. Recognit.* 15 (2002), S. 377–392
- Fox Jr. u. Richards 1982** FOX JR., R. ; RICHARDS, F. : A voltage-gated ion channel model inferred from the crystal structure of alamethicin at 1.5-Å resolution. In: *Nature* 300 (1982), S. 325–330
- Fradera u. a. 2000** FRADERA, X. ; DURAN, M. ; MESTRES, J. : Atomic transferability within the exchange-correlation density. In: *J. Comp. Chem.* 21 (2000), S. 1361–1374
- Friedrich u. a. 1913** FRIEDRICH, W. ; KNIPPING, P. ; LAUE, M. : Interferenzenerscheinungen bei Röntgenstrahlen. In: *Ann. Phys.* 41 (1913), S. 971
- Frisch u. a. 2004** FRISCH, M. J. ; TRUCKS, G. W. ; SCHLEGEL, H. B. ; SCUSERIA, G. E. ; ROBB, M. A. ; CHEESEMAN, J. R. ; MONTGOMERY JR., J. A. ; VREVEN, T. ; KUDIN, K. N. ; BURANT, J. C. ; MILLAM, J. M. ; IYENGAR, S. S. ; TOMASI, J. ; BARONE, V. ; MENNUCCI, B. ; COSSI, M. ; SCALMANI, G. ; REGA, N. ; PETERSSON, G. A. ; NAKATSUJI, H. ; HADA, M. ; EHARA, M. ; TOYOTA, K. ; FUKUDA, R. ; HASEGAWA, J. ; ISHIDA, M. ; NAKAJIMA, T. ; HONDA, Y. ; KITAO, O. ; NAKAI, H. ; KLENE, M. ; LI, X. ; KNOX, J. E. ; HRATCHIAN, H. P. ; CROSS, J. B. ; BAKKEN, V. ; ADAMO, C. ; JARAMILLO, J. ; GOMPERTS, R. ; STRATMANN, R. E. ; YAZYEV, O. ; AUSTIN, A. J. ; CAMMI, R. ; POMELLI, C. ; OCHTERSKI, J. W. ; AYALA, P. Y. ; MOROKUMA, K. ; VOTH, G. A. ; SALVADOR, P. ; DANNENBERG, J. J. ; ZAKRZEWSKI, V. G. ; DAPPRICH, S. ; DANIELS, A. D. ; STRAIN, M. C. ; FARKAS, O. ; MALICK, D. K. ; RABUCK, A. D. ; RAGHAVACHARI, K. ; FORESMAN, J. B. ; ORTIZ, J. V. ; CUI, Q. ; BABOUL, A. G. ; CLIFFORD, S. ; CIOSLOWSKI, J. ; STEFANOV, B. B. ; LIU, G. ; LIASHENKO, A. ; PISKORZ, P. ; KOMAROMI, I. ; MARTIN, R. L. ; FOX, D. J. ; KEITH, T. ; AL-LAHAM, M. A. ; PENG, C. Y. ; NANAYAKKARA, A. ; CHALLACOMBE, M. ; GILL, P. M. W. ; JOHNSON, B. ; CHEN, W. ; WONG, M. W. ; GONZALEZ, C. ; ; POPLE, J. A.: *Gaussian 03, Revision D.01*. Wallingford CT, 2004
- Frisch u. a. 2002** FRISCH, M. J. ; TRUCKS, G. W. ; SCHLEGEL, H. B. ; SCUSERIA, G. E. ; ROBB, M. A. ; CHEESEMAN, J. R. ; ZAKRZEWSKI, V. G. ; MONTGOMERY JR., J. A. ; STRATMANN, R. E. ; BURANT, J. C. ; DAPPRICH, S. ; MILLAM, J. M. ; DANIELS, A. D. ; KUDIN, K. N. ; STRAIN, M. C. ; FARKAS, O. ; TOMASI, J. ; BARONE, V. ; COSSI, M. ; CAMMI, R. ; MENNUCCI, B. ; POMELLI, C. ; ADAMO, C. ; CLIFFORD, S. ; OCHTERSKI, J. ; PETERSSON, G. A. ; AYALA, P. Y. ; CUI, Q. ; MOROKUMA, K. ; MALICK, D. K. ; RABUCK, A. D. ; RAGHAVACHARI, K. ; FORESMAN, J. B. ; CIOSLOWSKI, J. ; ORTIZ, J. V.

; BABOUL, A. G. ; STEFANOV, B. B. ; LIU, G. ; LIASHENKO, A. ; PISKORZ, P. ; KOMAROMI, I. ; GOMPERTS, R. ; MARTIN, R. L. ; FOX, D. J. ; KEITH, T. ; AL-LAHAM, M. A. ; PENG, C. Y. ; NANAYAKKARA, A. ; CHALLACOMBE, M. ; GILL, P. M. W. ; JOHNSON, B. ; CHEN, W. ; WONG, M. W. ; ANDRES, J. L. ; GONZALEZ, C. ; HEAD-GORDON, M. ; REPLOGLE, E. S. ; ; POPLE, J. A.: *Gaussian 98, Revision A.11.3*. Pittsburgh PA, 2002

Gatti u. Bertini 2004 GATTI, C. ; BERTINI, L. : The local form of the source function as a fingerprint of strong and weak intra- and intermolecular interactions. In: *Acta Cryst.* A60 (2004), S. 438–449

Gatti u. a. 2003 GATTI, C. ; CARGNONI, F. ; BERTINI, L. : Chemical information from the source function. In: *J. Comp. Chem.* 24 (2003), S. 422–436

Grabowsky u. a. 2009 GRABOWSKY, S. ; KALINOWSKI, R. ; WEBER, M. ; FOESTER, D. ; PAULMANN, C. ; LUGER, P. : Transferability and reproducibility in electron-density studies - bond-topological and atomic properties of tripeptides of the type L-alanyl-X-L-alanine. In: *Acta Cryst.* B65 (2009), Nr. 4, S. 488–501

Graña u. Mosquera 1999 GRAÑA, A. M. ; MOSQUERA, R. A.: Atomic and bond properties in functionalized esters and amides. In: *J. Comput. Chem.* 20 (1999), S. 1444–1454

Guillot u. a. 2008 GUILLOT, B. ; JELSCH, C. ; PODJARNY, A. ; LECOMTE, C. : Charge-density analysis of a protein structure at subatomic resolution: the human aldose reductase case. In: *Acta Crystallographica Section D* 64 (2008), May, Nr. 5, S. 567–588

Guillot u. a. 2001 GUILLOT, B. ; VIRY, L. ; GUILLOT, R. ; LECOMTE, C. ; JELSCH, C. : Refinement of proteins at subatomic resolution with *MOPRO*. In: *Journal of Applied Crystallography* 34 (2001), Apr, Nr. 2, S. 214–223

Hamilton 1965 HAMILTON, W. C.: Significance tests on the crystallographic R factor. In: *Acta Cryst.* 18 (1965), S. 502–510

Handl 2006 HANDL, A. : *Skript zur Einführung in die Statistik mit R*. Version: 2006. <http://www.wiwi.uni-bielefeld.de/~frohn/Mitarbeiter/Handl/stagrund.html>, Abruf: 02.10.2009

Hansen u. Coppens 1978 HANSEN, N. K. ; COPPENS, P. : Testing Aspherical Atom Refinements on Small-Molecule Data Sets. In: *Acta Cryst.* A34 (1978), S. 909–921

- Hirshfeld 1971** HIRSHFELD, F. L.: Difference Densities by Least-Squares Refinement: Fumaramic Acid. In: *Acta Crystallogr. Sect. B* B27 (1971), S. 769–781
- Hirshfeld 1976** HIRSHFELD, F. L.: Can X-ray Data Distinguish Bonding Effects from Vibrational Smearing? In: *Acta Crystallogr. Sect. A* A32 (1976), S. 239–244
- Hodgkin 1972** HODGKIN, D. C.: *Nobel Lectures*. Bd. Chemistry 1963–1970. Amsterdam : Elsevier Publishing Company, 1972
- Hohenberg u. Kohn 1964** HOHENBERG, P. ; KOHN, W. : Inhomogeneous Electron Gas. In: *Phys. Rev.* 136 (1964), Nr. 3B, S. B864–B871
- Hooft u. a. 1994** HOOFT, R. W. W. ; SANDER, C. ; VRIEND, G. : Reconstruction of symmetry related molecules from protein data bank (PDB) files. In: *J. Appl. Cryst.* 27 (1994), S. 1006–1009
- Hooft u. a. 1996a** HOOFT, R. W. W. ; SANDER, C. ; VRIEND, G. : Positioning hydrogen atoms by optimizing hydrogen bond networks in protein structures. In: *PROTEINS* 26 (1996), S. 363–376
- Hooft u. a. 1996b** HOOFT, R. W. W. ; VRIEND, G. ; SANDER, C. ; ABOLA, E. E.: Errors in protein structures. In: *Nature* 381 (1996), S. 272–272
- Housset u. a. 2000** HOUSSET, D. ; BENABICHA, F. ; PICHON-PESME, V. ; JELSCH, C. ; MAIERHOFER, A. ; DAVID, S. ; FONTECILLA-CAMPS, J. C. ; LECOMTE, C. : Towards the Charge-Density Study of Proteins: A Room-Temperature Scorpion-Toxin Structure at 0.96 Resolution as a First Test Case. In: *Acta Crystallogr. Sect. D* D56 (2000), S. 151–160
- Howard u. a. 2004** HOWARD, E. ; SANISHVILI, R. ; CACHAU, R. ; MITSCHLER, A. ; CHEVRIER, B. ; BARTH, P. ; LAMOUR, V. ; VAN ZANDT, M. ; SIBLEY, E. ; BON, C. ; MORAS, D. ; SCHNEIDER, T. ; JOACHIMIAK, A. ; PODJARNY, A. : Ultrahigh Resolution Drug Design I: Details of Interactions in Human Aldose Reductase-Inhibitor Complex at 0.66 Å. In: *Proteins: Struct., Funct., Genet.* 55 (2004), S. 792
- Huber u. Schneider 1985** HUBER, R. ; SCHNEIDER, M. : A group refinement procedure in protein crystallography using Fourier transforms. In: *J. Appl. Cryst.* 18 (1985), S. 165–169
- Hübschle u. a. 2007** HÜBSCHLE, C. B. ; LUGER, P. ; DITTRICH, B. : *Automation of invariom and of experimental charge density modelling of organic molecules with the preprocessor program InvariomTool*. Jun 2007

- Hübschle u. a. 2008** HÜBSCHLE, C. B. ; DITTRICH, B. ; GRABOWSKY, S. ; MESSERSCHMIDT, M. ; LUGER, P. : Comparative experimental electron density and electron localization function study of thymidine based on 20K X-ray diffraction data. In: *Acta Crystallographica Section B* 64 (2008), Jun, Nr. 3, S. 363–374
- Jaworski 2000** JAWORSKI, A. : *Sequenzierung von Peptaibol-Antibiotika mittels Flüssigkeitschromatographie-Elektrosprayionisierungs-Massenspektrometrie*, Justus-Liebig-Universität Gießen, Diss., 2000. <http://bibd.uni-giessen.de/ghtm/2000/uni/d000010.htm>
- Jayatilaka u. Grimwood 2003** JAYATILAKA, D. ; GRIMWOOD, D. J.: Ton-to: A Fortran Based Object-Oriented System for Quantum Chemistry and Crystallography. In: *Computational Science - ICCS 2003* 2660 (2003), S. 142–151
- Jelsch u. a. 2000a** JELSCH, C. ; TEETER, M. ; LAMZIN, V. ; PICHON-PESME, V. ; BLESSING, R. ; LECOMTE, C. : Accurate protein crystallography at ultra-high resolution: valence electron distribution in crambin. In: *Proc. Natl. Acad. Sci. USA* 97 (2000), S. 3171–3176
- Jelsch u. a. 2005** JELSCH, C. ; GUILLOT, B. ; LAGOUTTE, A. ; LECOMTE, C. : Advances in protein and small-molecule charge-density refinement methods using *MoPro*. In: *Journal of Applied Crystallography* 38 (2005), Feb, Nr. 1, S. 38–54
- Jelsch u. a. 1998** JELSCH, C. ; PICHON-PESME, V. ; LECOMTE, C. ; AUBRY, A. : Transferability of Multipole Charge-Density Parameters: Application to Very High Resolution Oligopeptide and Protein Structures. In: *Acta Crystallogr. Sect. D* D54 (1998), S. 1306–1318
- Jelsch u. a. 2000b** JELSCH, C. ; TEETER, M. M. ; LAMZIN, V. ; PICHON-PESME, V. ; BLESSING, R. H. ; LECOMTE, C. : Accurate Protein Crystallography at Ultra-High Resolution: Valence Electron Distribution in Crambin. In: *Proc. Natl. Acad. Sci. USA* 97 (2000), Nr. 7, S. 3171–3176
- Johnas u. a. 2009** JOHNAS, S. K. J. ; DITTRICH, B. ; MEENTS, A. ; MESSERSCHMIDT, M. ; WECKERT, E. F.: Charge-density study on cyclosporine A. In: *Acta Crystallographica Section D* 65 (2009), Nr. 3, S. 284–293
- Jones 1992** *Kapitel A, yaap, asap, A set of averaging programs.* In: JONES, T. A.: *Molecular Replacement*. Warrington : SERC Daresbury Laboratory, 1992, S. 91–105

- Jones u. Thirup 1986** JONES, T. A. ; THIRUP, S. : Using known substructures in protein model building and crystallography. In: *EMBO J.* 5 (1986), S. 819–822
- Kaarsholm u. a. 1989** KAARSHOLM, N. C. ; KO, H. C. ; DUNN, M. F.: Comparison of solution structural flexibility and zinc binding domains for insulin, proinsulin, and miniproinsulin. In: *Biochemistry* 26 (1989), Nr. 10, S. 4427–4435
- Kabsch 1993** KABSCH, W. : Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. In: *J. Appl. Cryst.* 26 (1993), S. 795–800
- Kabsch u. Sander 1983** KABSCH, W. ; SANDER, C. : Dictionary of protein secondary structure: pattern recognition of hydrogen bond and geometrical features. In: *Biopolymers* 22 (1983), S. 2577–2637
- Kalinowski u. a. 2007** KALINOWSKI, R. ; DITTRICH, B. ; HUEBSCHLE, C. B. ; PAULMANN, C. ; LUGER, P. : Experimental charge density of L-alanyl-L-prolyl-L-alanine hydrate: classical multipole and invariom approach, analysis of intra- and intermolecular topological properties. In: *Acta Cryst.* B63 (2007), Nr. 5, S. 753–767
- Kdzierski u. Sokalski 2001** KDZIERSKI, P. ; SOKALSKI, W. A.: Analysis of the transferability of atomic multipoles for amino acids in modeling macromolecular charge distribution from fragments. In: *J. Comp. Chem.* 22 (2001), S. 1082–1097
- Kingsford-Adaboh u. a. 2006** KINGSFORD-ADABOH, R. ; DITTRICH, B. ; HÜBSCHLE, C. B. ; GBEWONYO, W. S. K. ; OKAMOTO, H. ; KIMURA, M. ; ISHIDA, H. : Invariom structure refinement, electrostatic potential and toxicity of 4-O-methylalpinumisoflavone, O,O-dimethylalpinumisoflavone and 5-O-methyl-4-O-(3-methylbut-2-en-1-yl)alpinum-isoflavone. In: *Acta Crystallogr. Sect. B* 62 (2006), S. 843–849
- Kissinger u. a. 1999** KISSINGER, C. R. ; GEHLHAAR, D. K. ; FOGEL, D. B.: Rapid automated molecular replacement by evolutionary search. In: *Acta Cryst.* D55 (1999), S. 484–491
- Kissinger u. a. 2001** KISSINGER, C. R. ; SMITH, B. A. ; GEHLHAAR, D. K. ; BOUZIDA, D. : Molecular replacement by evolutionary search. In: *Acta Cryst.* D57 (2001), S. 1474–1479
- Kissinger 2008** KISSINGER, C. R.: *EPMR: Molecular replacement by evolutionary search*. <http://www.epmr.info/>. Version: 8.03, 2008. – Kissinger u. a. (1999, 2001); Huber u. Schneider (1985)

- Kleywegt 1996** KLEYWEGT, G. J.: *Making the most of your search model*. CCP4/ESF-EACBM Newsletter on Protein Crystallography 32, 1996
- Kleywegt 1999** KLEYWEGT, G. J.: Experimental assessment of differences between related protein crystal structures. In: *Acta Cryst.* D55 (1999), S. 1878–1857
- Kleywegt u. a. 2004** KLEYWEGT, G. J. ; HARRIS, M. R. ; ZOU, J. Y. ; TAYLOR, T. C. ; WAHLBY, A. ; JONES, T. A.: The Uppsala Electron Density Server. In: *Acta Cryst.* D60 (2004), S. 2240–2249
- Kleywegt u. Jones 1996** KLEYWEGT, G. J. ; JONES, T. A.: xdlMAPMAN and xdlDATAMAN - programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. In: *Acta Cryst.* D52 (1996), S. 826–828
- Kleywegt u. Jones 1999** KLEYWEGT, G. J. ; JONES, T. A.: Software for handling macromolecular envelopes. In: *Acta Cryst.* D55 (1999), S. 941–944
- Kleywegt u. Jones 1994** *Kapitel* Halloween ... Masks and Bones. In: KLEYWEGT, G. J. ; JONES, T. : *From First Map to Final Model*. Warrington : SERC Daresbury Laboratory, 1994, S. 59–66
- Kleywegt u. Read 1997** KLEYWEGT, G. J. ; READ, R. J.: Not your average density. In: *Structure* 5 (1997), S. 1557–1569
- Kleywegt u. a. 2001** *Kapitel* Around O. In: KLEYWEGT, G. J. ; ZOU, J. Y. ; KJELDGAARD, M. ; JONES, T. A.: *Crystallography of Biological Macromolecules*. Bd. F: *International Tables for Crystallography*. Dordrecht : Kluwer Academic Publishers, 2001, S. Chapter 17.1, pp. 353–356, 366–367
- Kleywegt u. Jones 2008** KLEYWEGT, G. J. ; JONES, T. A.: *MAPMAN - manipulation and analysis of electron-density maps*. Uppsala Software Factory. <http://xray.bmc.uu.se/usf/>. Version: 080625/7.8.5, 2008. – Jones u. Thirup (1986); Jones (1992); Kleywegt u. Jones (1994, 1996); Kleywegt (1996); Kleywegt u. Read (1997); Kleywegt u. Jones (1999); Kleywegt (1999); Read u. Kleywegt (2001); Kleywegt u. a. (2004, 2001)
- Koch u. a. 1995** KOCH, U. ; POPELIER, P. L. A. ; STONE, A. J.: Conformational Dependence of Atomic Multipole Moments. In: *Chem. Phys. Lett.* 238 (1995), S. 253–260
- Koch u. Stone 1996** KOCH, U. ; STONE, A. J.: Conformational dependence of the molecular charge distribution and its influence on intermolecular interactions. In: *J. Chem. Soc., Faraday Trans.*, 92 (1996), S. 1701 – 1708

- Koritsanszky u. a. 2003** KORITSANSZKY, T. ; RICHTER, T. ; MACCHI, P. ; VOLKOV, A. ; GATTI, C. ; HOWARD, S. ; MALLINSON, P. R. ; FARRUGIA, L. ; SU, Z. W. ; HANSEN, N. K.: *XD, A Computer Program Package for Multipole Refinement and Analysis of Electron Densities from Diffraction Data*. Freie Universität Berlin, 2003
- Koritsánszky u. a. 2002** KORITSÁNSZKY, T. ; VOLKOV, A. ; COPPENS, P. : Aspherical-atom scattering factors from molecular wave functions. 1. Transferability and conformation dependence of atomic electron densities of peptides within the multipole formalism. In: *Acta Crystallogr. Sect. A* A58 (2002), S. 464–472
- Koritsanszky u. Coppens 2000** KORITSANSZKY, T. S. ; COPPENS, P. : Chemical Applications of X-ray Charge-Density Analysis. In: *Chemical Reviews* (2000)
- Kurki-Suonio 1968** KURKI-SUONIO, K. : On the Information about Deformations of the Atoms in X-ray Diffraction Data. In: *Acta Cryst.* A24 (1968), S. 379–390
- Kurki-Suonio 1977** KURKI-SUONIO, K. : Symmetry and its Implications. In: *Isr. J. Chem.* 16 (1977), S. 115–123
- Lamzin u. a. 1999** LAMZIN, V. S. ; MORRIS, R. J. ; DAUTER, Z. ; WILSON, K. S. ; TEETER, M. M.: Experimental Observation of Bonding Electrons in Proteins. In: *J. Bio. Chem.* 274 (1999), S. 20753–20755
- Laskowski u. a. 1993** LASKOWSKI, R. A. ; MACARTHUR, M. W. ; MOSS, D. S. ; THORNTON, J. M.: PROCHECK: a program to check the stereochemical quality of protein structures. In: *J. Appl. Cryst.* 26 (1993), S. 283–291
- Laskowski u. a.** LASKOWSKI, R. A. ; MACARTHUR, M. W. ; SMITH, D. K. ; JONES, D. T. ; HUTCHINSON, E. G. ; MORRIS, A. L. ; MOSS, D. S. ; THORNTON, J. M.: *PROCHECK*. Biochemistry Biocomputing Group, University College London. <http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html>. Version: 3.5.4. – Laskowski u. a. (1993); Morris u. a. (1992)
- Lecomte u. a. 2004** LECOMTE, C. ; GUILLOT, B. ; MUZET, N. ; PICHON-PESME, V. ; JELSCH, C. : Ultra-high-resolution X-ray structure of proteins. In: *Cellular and Molecular Life Sciences* 61 (2004), S. 774–782
- Liu u. a. 2003** LIU, Q. ; HUANG, Q. ; TENG, M. ; WEEKS, C. M. ; JELSCH, C. ; ZHANG, R. ; NIU, L. : The Crystal Structure of a Novel, Inactive, Lysine 49 PLA2 from *Agkistrodon acutus* Venom. In: *J. Bio. Chem.* 278 (2003), S. 41400–41408

- Luger 2007** LUGER, P. : Fast electron density methods in the life sciences – a routine application in the future? In: *Org. Biomol. Chem.* 5 (2007), S. 2529–2540
- Luger 2000** LUGER, P. : *Vorlesungsskript "Praxis der Röntgenstrukturanalyse"*. Freie Universität Berlin, 2000
- Luzzati 1952** LUZZATI, V. : Traitement statistique des erreurs dans la détermination des structures cristallines. In: *Acta Cryst.* 5 (1952), S. 802–810
- Madsen u. a. 1998** MADSEN, D. ; FLENSBURG, C. ; LARSEN, S. : Properties of the Experimental Crystal Charge Density of Methylammonium Hydrogen Maleate. A Salt with a Very Short Intramolecular O–H–O Hydrogen Bond. In: *J. Phys. Chem. A* A102 (1998), Nr. 12, S. 2177–2188
- Martin 2009** MARTIN, A. : *ProFit*. <http://www.bioinf.org.uk/software/profit/>. Version:3.1, 2009. – Fitting nach dem McLachlan-Algorithmus McLachlan (1982)
- Maslen u. a. 1992** *Kapitel 6.1.1*. In: MASLEN, E. N. ; FOX, A. G. ; O'KEEFLE, M. A.: *International Tables for X-ray Crystallography*. Bd. C. Amsterdam : Kluwer Academic Publishers, 1992, S. 476–512
- Massa 2002** MASSA, W. : *Kristallstrukturbestimmung*. 3., überarb. und erw. Aufl. Stuttgart; Leipzig; Wiesbaden : B. G. Teubner, 2002
- Matta u. Bader 2002** MATTA, C. F. ; BADER, R. F. W.: Atoms-in-Molecules Study of Genically Encoded Amino Acids. II Computational Study of Molecular Geometries. In: *Proteins:Structure, Function , and Genetics* 48 (2002), S. 519–538
- Matta u. Bader 2003** MATTA, C. F. ; BADER, R. F. W.: Atoms-in-Molecules Study of Genically Encoded Amino Acids. III Bond and Atomic Properties and their Correlation with Experiment including Mutation-Induced Changes in Protein Stability and Genetic Coding. In: *Proteins: Structure, Function and Genetics* 52 (2003), S. 360–399
- Matta 2001** MATTA, C. : Theoretical reconstruction of the electron density of large molecules from fragments determined as proper open quantum systems: The properties of the oripavine PEO, enkephalins, and morphine. In: *J. Phys. Chem.* 105 (2001), S. 11088–11101
- Matta 2003** MATTA, C. : Application of the quantum theory of atoms in molecules to selected physico-chemical and biophysical problems: Focus on correlation with experiment. In: *J. Comp. Chem.* 24 (2003), S. 453–462

- Matta u. Bader 2000** MATTA, C. ; BADER, R. F. W.: An Atoms-In-Molecules Study of the Genetically-Encoded Amino Acids : I. Effects of Conformation and of Tautomerization on Geometric, Atomic, and Bond Properties. In: *Proteins: Structure, Function, and Genetics* 40 (2000), S. 310–329
- Matthews 1968** MATTHEWS, D. W.: Solvent Content of Protein Crystals. In: *J. Mol. Biol.* 33 (1968), S. 491–497
- McLachlan 1982** MCLACHLAN, A. : Rapid Comparison of Protein Structures. In: *Acta Cryst.* A38 (1982), S. 871–873
- Mebs u. a. 2006** MEBS, S. ; MESSERSCHMIDT, M. ; LUGER, P. : Experimental charge density of an L-phenylalanine formic acid complex with a short hydrogen bond determined at 25 K. In: *Z. Kristallogr.* 221 (2006), S. 656–664
- M.J. u. a. 1969** M.J., A. ; T.L., B. ; E.J., D. ; G.G., D. ; M., V. ; E.N., B. ; M.M., H. ; D.C., H. ; B., R. ; S., S. : Structure of Rhombohedral 2 Zinc Insulin Crystals. In: *Nature* 224 (1969), S. 491–495
- Morris u. a. 1992** MORRIS, A. L. ; MACARTHUR, M. W. ; HUTCHINSON, E. G. ; THORNTON, J. M.: Stereochemical quality of protein structure coordinates. In: *Proteins* 12 (1992), S. 345–364
- Müller 1996** MÜLLER, U. : *Anorganische Strukturchemie*. 3., überarb. und erw. Aufl. Stuttgart : B. G. Teubner, 1996
- Muzet u. a. 2003** MUZET, N. ; GUILLOT, B. ; JELSCH, C. ; HOWARD, E. ; LECOMTE, C. : Electrostatic complementarity in an aldose reductase complex from ultra-high-resolution crystallography and first principle calculations. In: *Proc. Natl. Acad. Sci. USA* 100 (2003), S. 8742–8747
- Nanao u. a. 2005** NANAHO, M. H. ; SHELDRIK, G. M. ; RAVELLI, R. B. G.: Improving radiation-damage substructures for RIP. In: *Acta Crystallographica Section D* 61 (2005), Sep, Nr. 9, S. 1227–1237
- Nishikawa u. Ooi 1986** NISHIKAWA, K. ; OOI, T. : Radial locations of amino-acid residues in a globular protein - correlation with the sequence. In: *J. Biochem.* 100 (1986), S. 1043–1047
- Norrman u. Schluckebier 2007** NORRMAN, M. ; SCHLUCKEBIER, G. : Crystallographic characterization of two novel crystal forms of human insulin induced by chaotropic agents and a shift in pH. In: *BMC Structural Biology* 7 (2007), Nr. 83

- Norrman u. a. 2006** NORRMAN, M. ; STÅHL, K. ; SCHLUCKEBIER, G. ; AL-KARADAGHI, S. : Characterization of insulin microcrystals using powder diffraction and multivariate data analysis. In: *Journal of Applied Crystallography* 39 (2006), Jun, Nr. 3, S. 391–400
- Parkinson u. a. 1996** PARKINSON, G. ; VOITECHOVSKY, J. ; CLOWNEY, L. ; BRÜNGER, A. T. ; BERMAN, H. : New parameters for the refinement of nucleic acid-containing structures. In: *Acta Cryst.* D52 (1996), S. 57–64
- Pauling u. a. 1951** PAULING, L. ; COREY, R. ; BRANSON, H. : The Structure of Proteins: Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain. In: *Proc. Nat. Acad. Sci.* 37 (1951), S. 205–211
- Petrova u. Podjarny 2004** PETROVA, T. ; PODJARNY, A. : Protein crystallography at subatomic resolution. In: *Rep. Prog. Phys.* 67 (2004), S. 1565–1605
- Pettersen u. a. 2004** PETTERSEN, E. F. ; GODDARD, T. D. ; HUANG, C. C. ; COUCH, G. S. ; GREENBLATT, D. M. ; MENG, E. C. ; FERRIN, T. E.: UCSF Chimera—a visualization system for exploratory research and analysis. In: *J. Comput. Chem.* 25 (2004), Nr. 13, S. 1605–1612
- Phillips 1966** PHILLIPS, D. C.: The three-dimensional structure of an enzyme molecule. In: *Scientific American* 215 (1966), S. 78–90
- Pichon-Pesme u. a. 2004** PICHON-PESME, V. ; JELSCH, C. ; GUILLOT, B. ; LECOMTE, C. : A Comparison between experimental and theoretical aspherical-atom scattering factors for charge-density refinement of large molecules. In: *Acta Crystallogr. Sect. A* 60 (2004), S. 204–208
- Pichon-Pesme u. a. 1995** PICHON-PESME, V. ; LECOMTE, C. ; LACHEKAR, H. : On Building a Data Bank of Transferable Experimental Electron Density Parameters: Application to Polypeptides. In: *J. Phys. Chem.* 99 (1995), Nr. 16, S. 6242–6250
- Pichon-Pesme u. a. 2000** PICHON-PESME, V. ; LACHEKAR, H. ; SOUHASSOU, M. ; LECOMTE, C. : Electron Density and Electrostatic Properties of Two Peptide Molecules: Tyrosyl-Glycyl-Glycine Monohydrate and Glycyl-Aspartic Acid Dihydrate. In: *Acta Crystallogr. Sect. B* 56 (2000), S. 728–737
- Pichon-Pesme u. Lecomte 1998** PICHON-PESME, V. ; LECOMTE, C. : Experimental Charge Density and Electrostatic Potential of Triglycine. In: *Acta Crystallogr. Sect. B* B54 (1998), S. 485–493
- Popelier u. Bader 1994** POPELIER, P. L. A. ; BADER, R. F. W.: Effect of Twisting a Polypeptide on Its Geometry and Electron Distribution. In: *J. Phys. Chem.* 98 (1994), Nr. 16, S. 4473–4481

- Popelier u. a. 2004** POPELIER, P. L. A. ; DEVEREUX, M. ; RAFAT, M. : The quantum topological electrostatic potential as a probe for functional group transferability. In: *Acta Cryst.* A60 (2004), S. 427–433
- Popelier u. Aicken 2003a** POPELIER, P. ; AICKEN, F. : Atomic properties of selected biomolecules: Quantum topological atom types of carbon occurring in natural amino acids and derived molecules. In: *J. Am. Chem. Soc.* 125 (2003), S. 1284–1292
- Popelier u. Aicken 2003b** POPELIER, P. ; AICKEN, F. : Atomic properties of selected biomolecules: Quantum topological atom types of hydrogen, oxygen, nitrogen and sulfur occurring in natural amino acids and their derivatives. In: *Chem. Europ. J.* 9 (2003), S. 1207–1216
- Prugovecki u. a. 2008** PRUGOVECKI, B. ; DODSON, E. ; DODSON, G. ; MATKOVIC-CALOGOVIC, D. : Structure of the T6 human insulin derivative with nickel at 1.35 Å resolution. In: *to be published* (2008)
- Quillin u. Matthews 2000** QUILLIN, M. L. ; MATTHEWS, B. W.: Accurate calculation of the density of proteins. In: *Acta Cryst.* D56 (2000), S. 791–794
- Quinonez u. a. 2003** QUINONEZ, P. ; VILA, A. ; GRANA, A. ; MOSQUERA, R. : AIM study on the influence of fluorine atoms on the alkyl chain. In: *Chem. Phys.* 287 (2003), S. 227–236
- R Development Core Team 2009** R DEVELOPMENT CORE TEAM: *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2009. <http://www.R-project.org>. – ISBN 3-900051-07-0
- Ramachandran u. a. 1963** RAMACHANDRAN, G. ; RAMAKRISHNANA, C. ; SASISEKHARAN, V. : Stereochemistry of polypeptide chain configurations. In: *J. Mol. Bio.* 7 (1963), S. 95–99
- Read u. Kleywegt 2001** *Kapitel* Density modification: theory and practice. In: READ, R. J. ; KLEYWEGT, G. : *Methods in Macromolecular Crystallography*. Amsterdam : IOS Press, 2001, S. 123–135
- Richards 1977** RICHARDS, F. : Areas, volumes, packing and protein structure. In: *Annu. Rev. Biophys. Bioeng.* 6 (1977), S. 151–176
- Richardson 1981** RICHARDSON, J. S.: The Anatomy and Taxonomy of Protein Structure. In: *Advances in Protein Chemistry* 34 (1981), S. 167–339

- Rocchia u. a. 2001** ROCCHIA, W. ; ALEXOV, E. ; HONIG, B. : Extending the Applicability of the Nonlinear Poisson-Boltzmann Equation: Multiple Dielectric Constants and Multivalent Ions. In: *J. Phys. Chem.* B105 (2001), Nr. 28, S. 6507–6514
- Rocchia u. a. 2002** ROCCHIA, W. ; ALEXOV, E. ; HONIG, B. : *DelPhi*. http://wiki.c2b2.columbia.edu/honiglab_public/index.php/ Software:DelPhi. Version: V. 4 Release 1.1, 2002. – Rocchia u. a. (2001)
- Sanner u. a. 1996** SANNER, M. ; SPEHNER, J.-C. ; OLSON, A. : Reduced surface: an efficient way to compute molecular surfaces. In: *Biopolymers* 38 (1996), S. 305–320
- Sansom 1993a** SANSOM, M. S. P.: Alamethicin and related peptaibols - model ion channels. In: *Eur. Biophys. J.* 22 (1993), S. 105–124
- Sansom 1993b** SANSOM, M. S. P.: Structure and function of channel-forming peptaibols. In: *Quarterly Reviews of Biophysics* 26 (1993), S. 365–421
- Sauter u. a. 2001** SAUTER, C. ; OTÁLORA, F. ; GAVIRA, J.-A. ; VIDAL, O. ; GIEGÉ, R. ; GARCÍA-RUIZ, J. M.: Structure of tetragonal hen egg-white lysozyme at 0.94 Å from crystals grown by the counter-diffusion method. In: *Acta Crystallogr. Sect. D* 57 (2001), S. 1119–1126
- Scheins u. a. 2005** SCHEINS, S. ; MESSERSCHMIDT, M. ; LUGER, P. : Sub-molecular partitioning of morphine hydrate based on its experimental charge density at 25 K. In: *Acta Crystallogr. Sect. B* 61 (2005), S. 443–448
- Schmidt u. a. 2003** SCHMIDT, A. ; JELSCH, C. ; ØSTERGAARD, P. ; RYPNIEWSKI, W. ; LAMZIN, V. S.: Trypsin Revisited – Crystallography at (sub) atomic resolutions and quantum chemistry revealing details of catalysis. In: *J. Bio. Chem.* 278 (2003), S. 43357–43362
- Schnieders u. a. 2009** SCHNIEDERS, M. J. ; FENN, T. D. ; PANDE, V. S. ; BRUNGER, A. T.: Polarizable atomic multipole X-ray refinement: application to peptide crystals. In: *Acta Cryst.* D65 (2009), S. 952–965
- Sevcik u. a. 1996** SEVCIK, J. ; DAUTER, Z. ; LAMZIN, V. S. ; WILSON, K. S.: Ribonuclease from *Streptomyces aureofaciens* at Atomic Resolution. In: *Acta Cryst.* D52 (1996), S. 327–344
- Sharp u. Honig 1990** SHARP, K. ; HONIG, B. : Electrostatic interactions in macromolecules: theory and applications. In: *Annu. Rev. Biophys. Biophys. Chem.* 19 (1990), S. 301–332

- Sheldrick 1997** SHELDRIK, G. M.: *SHELXH*. University of Göttingen. <http://shelx.uni-ac.gwdg.de/SHELX/>. Version: 97-2, 1997. – Huge UNIX version, adapted to OpenMP by Kay Diederichs
- Sheldrick 2008** SHELDRIK, G. M.: A short history of SHELX. In: *Acta Cryst.* A64 (2008), S. 112–122
- Smith u. a. 2003** SMITH, G. D. ; PANGBORN, W. A. ; BLESSING, R. H.: The structure of T₆ human insulin at 1.0 Å resolution. In: *Acta Crystallogr. Sect. D* 59 (2003), S. 474–482
- Smith u. Blessing 2003** SMITH, G. D. ; BLESSING, R. H.: Lessons from an aged, dried crystal of T6 human insulin. In: *Acta Cryst.* D59 (2003), S. 1384–1394
- Smith u. a. 2000** SMITH, G. D. ; CISZAK, E. ; MAGRUM, L. A. ; PANGBORN, W. A. ; BLESSING, R. H.: R6 hexameric insulin complexed with m-cresol or resorcinol. In: *Acta Cryst.* D56 (2000), S. 1541–1548
- Snook u. a. 1998** SNOOK, C. ; WOOLLEY, G. ; OLIVA, G. ; PATTABHI, V. ; WOOD, S. ; BLUNDELL, T. ; WALLACE, B. : The structure and function of antiamoebin I, a proline-rich membrane-active polypeptide. In: *Structure* 6 (1998), S. 783–792
- Sokalski u. Poirier 1983** SOKALSKI, W. ; POIRIER, R. : Cumulative atomic multipole representation of the molecular charge distribution and its basis set dependence. In: *Chem. Phys. Lett.* 98 (1983), S. 86–92
- Spackman u. a. 1999** SPACKMAN, M. A. ; BYROM, P. G. ; ALFREDSSON, M. ; HERMANSSON, K. : Influence of Intermolecular Interactions on Multipole-Refined Electron Densities. In: *Acta Crystallogr. Sect. A* A55 (1999), S. 30–47
- Stefanov u. Cioslowski 1996** STEFANOV, B. ; CIOSLOWSKI, J. : Variability of shapes and properties of atoms in molecules: A case study of the carbonyl oxygen. In: *Canad. J. Chem.* 74 (1996), S. 1263–1270
- Stewart 2008** STEWART, J. J. P.: *MOPAC2009*. Stewart Computational Chemistry. [HTTP://OpenMOPAC.net](http://OpenMOPAC.net). Version: 9.211L, 2008
- Stewart 1969** STEWART, R. F.: Generalized X-ray Scattering Factors. In: *J. Chem. Phys.* 51 (1969), Nr. 10, S. 4569–4577
- Stone 1981** STONE, A. : Distributed multipole analysis, or how to describe a molecular charge distribution. In: *Chem. Phys. Lett.* 83 (1981), S. 233–239

- Stone u. a. 1996** STONE, A. ; HATTIG, C. ; JANSEN, G. ; ANGYAN, J. : Transferability of topologically partitioned polarizabilities: The case of n-alkanes. In: *Mol. Phys.* 89 (1996), S. 595–605
- Stryer 1990** STRYER, L. : *Biochemie*. Heidelberg : Spektrum der Wissenschaft, 1990
- Strynadka u. James 1991** STRYNADKA, N. C. J. ; JAMES, M. N. G.: Lysozyme revisited: Crystallographic evidence for distortion of an N-acetylmuramic acid residue bound in site D. In: *Journal of Molecular Biology* 220 (1991), Nr. 2, S. 401 – 424. – ISSN 0022–2836
- Su u. Coppens 1992** SU, Z. ; COPPENS, P. : On the Mapping of Electrostatic Properties from the Multipole Description of the Charge Density. In: *Acta Cryst.* A48 (1992), S. 188–197
- Tantau u. Wibrow 2008** TANTAU, T. ; WIBROW, M. : *The TikZ and PGF Packages*. University Lübeck. <http://sourceforge.net/projects/pgf>. Version: 2.0, 2008
- Tsirelson u. a. 2006** TSIRELSON, V. G. ; STASH, A. I. ; POTEMKIN, V. A. ; RYKOUNOV, A. A. ; SHUTALEV, A. D. ; ZHUROVA, E. A. ; ZHUROV, V. V. ; PINKERTON, A. A. ; GURSKAYA, G. V. ; ZAVODNIK, V. E.: Molecular and crystal properties of ethyl 4,6-dimethyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxylate from experimental and theoretical electron densities. In: *Acta Cryst.* B62 (2006), S. 676–688
- UCSF Computer Graphics Laboratory 2009** UCSF COMPUTER GRAPHICS LABORATORY: *Chimera package*. Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081). <http://www.cgl.ucsf.edu/chimera/>. Version: 1.4-b28596, 2009. – Pettersen u. a. (2004)
- Uy u. Wold 1977** UY, R. ; WOLD, F. : Posttranslational covalent modification of proteins. In: *Science* 198 (1977), S. 890–896
- Vocadlo u. a. 2001** VOCADLO, D. J. ; DAVIES, G. J. ; LAINE, R. ; WITHERS, S. G.: Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. In: *Nature* 412 (2001), S. 835–838
- Volkov u. Coppens 2004** VOLKOV, A. ; COPPENS, P. : Calculation of electrostatic interaction energies in molecular dimers from atomic multipole moments obtained by different methods of electron density partitioning. In: *J. Comp. Chem.* 25 (2004), S. 921–934

- Volkov u. a. 2006** VOLKOV, A. ; KING, H. F. ; COPPENS, P. : Dependence of the Intermolecular Electrostatic Interaction Energy on the Level of Theory and the Basis Set. In: *J. Chem. Theory Comput.* 2 (2006), S. 81–89
- Volkov u. a. 2004a** VOLKOV, A. ; KORITSANSZKY, T. ; COPPENS, P. : Combination of the exact potential and multipole methods (EP/MM) for evaluation of intermolecular electrostatic interaction energies with pseudoatom representation of molecular electron densities. In: *Chem. Phys. Lett.* 391 (2004), S. 170–175
- Volkov u. a. 2004b** VOLKOV, A. ; KORITSÁNSZKY, T. ; LI, X. ; COPPENS, P. : Response to the paper A comparison between experimental and theoretical aspherical-atom scattering factors for charge-density refinement of large molecules, by Pichon-Pesme, Jelsch, Guillot & Lecomte (2004). In: *Acta Crystallogr. Sect. A* A60 (2004), S. 638–639
- Volkov u. a. 2007** VOLKOV, A. ; MESSERSCHMIDT, M. ; COPPENS, P. : Improving the scattering-factor formalism in protein refinement: application of the University at Buffalo Aspherical-Atom Databank to polypeptide structures. In: *Acta Crystallographica Section D* 63 (2007), Nr. 2, S. 160–170
- Volkov u. a. 2004c** VOLKOV, A. ; LI, X. ; KORITSÁNSZKY, T. ; COPPENS, P. : Ab initio quality electrostatic atomic and molecular properties including intermolecular energies from a transferable theoretical pseudoatom databank. In: *J. Phys. Chem. A* 108 (2004), S. 4283–4300
- Voss 2007** VOSS, N. R.: *Geometric Studies of RNA and Ribosomes, and Ribosome Crystallization*, Yale University, Diss., 2007
- Voss 2009** VOSS, N. R.: *3v*. <http://geometry.molmovdb.org/3v/>. Version: 1.2, 2009. – Voss u. a. (2006); Voss (2007)
- Voss u. a. 2006** VOSS, N. R. ; GERSTEIN, M. ; STEITZ, T. A. ; MOORE, P. B.: The geometry of the ribosomal polypeptide exit tunnel. In: *J. Mol. Biol.* 360 (2006), Nr. 4, S. 893–906
- Vrieling u. Sampson 2003** VRIELINK, A. ; SAMPSON, N. : Sub-Ångstrom resolution enzyme X-ray structures: is seeing believing? In: *Current Opinion in Structural Biology* 13 (2003), Nr. 6, S. 709 – 715
- Vriend 1990** VRIEND, G. : WHAT IF: a molecular modelling and drug design program. In: *J. Mol. Graph.* 8 (1990), S. 52–56
- Vriend 2008** VRIEND, G. : *WHAT_CHECK*. Centre for Molecular and Biomolecular Informatics, Radboud University Nijmegen. <http://swift.cmbi.>

ru.nl/gv/whatcheck/. Version: 7.0, 2008. – Hooft u. a. (1996b); Kabsch u. Sander (1983); Vriend (1990); Engh u. Huber (1991); Parkinson u. a. (1996); Matthews (1968); Cremer u. Pople (1975); Vriend u. Sander (1993); Hooft u. a. (1994, 1996a)

Vriend u. Sander 1993 VRIEND, G. ; SANDER, C. : Quality control of protein models: directional atomic contact analysis. In: *J. Appl. Cryst.* 26 (1993), S. 47–60

de Vries u. a. 2000 VRIES, R. Y. ; FEIL, D. ; TSIRELSON, V. G.: Extracting Charge Density Distributions from Diffraction Data: a Model Study on Urea. In: *Acta Crystallogr. Sect. B* 56 (2000), S. 118–123

Walsh u. a. 1998 WALSH, M. ; SCHNEIDER, T. ; SIEKER, L. ; DAUTER, Z. ; LAMZIN, V. ; WILSON, K. : Refinement of triclinic hen egg-white lysozyme at atomic resolution. In: *Acta Cryst.* D54 (1998), S. 522–546

Wang u. a. 2007 WANG, J. ; DAUTER, M. ; ALKIRE, R. ; JOACHIMIAK, A. ; DAUTER, Z. : Triclinic Lysozyme at 0.65 Å Resolution. In: *Acta Cryst.* D63 (2007), S. 1254–1268

Whitmore u. a. 2003 WHITMORE, L. ; CHUGH, J. K. ; SNOOK, C. F. ; WALLACE, B. A.: The peptaibol database: a sequence and structure resource. In: *J. Pept. Sci.* 9 (2003), S. 663–665

Williams u. a. 2009 WILLIAMS, T. ; KELLEY, C. ; LANG, R. ; KOTZ, D. ; CAMPBELL, J. ; ELBER, G. ; WOO, A. : *gnuplot - A free plotting program*. <http://gnuplot.sourceforge.net>. Version: 4.3, 2009. – LUA-Terminal von Peter Hedwig

Wilson 1992 WILSON, A. J. C. (Hrsg.): *International Tables for Crystallography*. Dordrecht; Boston; London : Kluwer Academic Publishers, 1992

Zarychta u. a. 2007 ZARYCHTA, B. ; PICHON-PESME, V. ; GUILLOT, B. ; LECOMTE, C. ; JELSCH, C. : On the application of an experimental multipolar pseudo-atom library for accurate refinement of small-molecule and protein crystal structures. In: *Acta Crystallogr. Sect. A* A63 (2007), S. 108–125

Anhang

A Alanyl-Prolyl-Alanin

Tabelle A.1: Quadratisch gemittelte Geometrieunterschiede zwischen asphärischer und sphärischer Verfeinerung von APA.

<i>mul vs iam</i>				<i>mul vs iam_{rb}</i>			
AS-Nr.	AS	Atom	RMSd	AS-Nr.	AS	Atom	RMSd
101	ALA	N	0.001	101	ALA	N	0.001
101	ALA	H0A	0.118	101	ALA	H0A	0.041
101	ALA	H0B	0.096	101	ALA	H0B	0.036
101	ALA	H0C	0.124	101	ALA	H0C	0.070
101	ALA	CA	0.001	101	ALA	CA	0.001
101	ALA	HA	0.102	101	ALA	HA	0.056
101	ALA	C	0.000	101	ALA	C	0.000
101	ALA	O	0.001	101	ALA	O	0.002
101	ALA	CB	0.001	101	ALA	CB	0.000
101	ALA	HB1	0.216	101	ALA	HB1	0.102
101	ALA	HB2	0.155	101	ALA	HB2	0.110
101	ALA	HB3	0.128	101	ALA	HB3	0.088
102	PRO	N	0.001	102	PRO	N	0.001
102	PRO	CA	0.001	102	PRO	CA	0.000
102	PRO	HA	0.148	102	PRO	HA	0.080
102	PRO	C	0.001	102	PRO	C	0.001
102	PRO	O	0.001	102	PRO	O	0.001
102	PRO	CB	0.001	102	PRO	CB	0.001
102	PRO	HB1	0.149	102	PRO	HB1	0.087
102	PRO	HB2	0.127	102	PRO	HB2	0.064
102	PRO	CG	0.001	102	PRO	CG	0.001
102	PRO	HG1	0.126	102	PRO	HG1	0.047
102	PRO	HG2	0.077	102	PRO	HG2	0.054
102	PRO	CD	0.000	102	PRO	CD	0.000
102	PRO	HD1	0.142	102	PRO	HD1	0.039
102	PRO	HD2	0.179	102	PRO	HD2	0.080
103	ALA	N	0.000	103	ALA	N	0.000
103	ALA	H0	0.129	103	ALA	H0	0.037
103	ALA	CA	0.000	103	ALA	CA	0.001
103	ALA	HA	0.126	103	ALA	HA	0.035
103	ALA	CB	0.001	103	ALA	CB	0.001
103	ALA	HB1	0.140	103	ALA	HB1	0.080

A Alanyl-Prolyl-Alanin

103	ALA	HB2	0.120	103	ALA	HB2	0.034
103	ALA	HB3	0.107	103	ALA	HB3	0.070
103	ALA	C	0.001	103	ALA	C	0.001
103	ALA	OT1	0.001	103	ALA	OT1	0.001
103	ALA	OT2	0.002	103	ALA	OT2	0.002
201	HOH	O	0.005	201	HOH	O	0.005
201	HOH	H1	0.103	201	HOH	H1	0.029
201	HOH	H2	0.078	201	HOH	H2	0.117
pro Aminosäure							
101	ALA		0.107	101	ALA		0.059
102	PRO		0.098	102	PRO		0.047
103	ALA		0.084	103	ALA		0.037
201	HOH		0.075	201	HOH		0.070
Gesamtmolekül							
							0.096
							0.051

Tabelle A.2: Quadratisch gemittelte Geometrieunterschiede zwischen asphärischer und sphärischer Verfeinerung von APA0.83 bei einer Auflösung von $d = 0,83 \text{ \AA}$.

AS-Nr.	<i>mul0.83 vs iam0.83</i>			AS-Nr.	<i>mul0.83 vs iam0.83_{rb}</i>		
	AS	Atom	RMSd		AS	Atom	RMSd
101	ALA	N	0.002	101	ALA	N	0.001
101	ALA	H0A	0.156	101	ALA	H0A	0.005
101	ALA	H0B	0.151	101	ALA	H0B	0.086
101	ALA	H0C	0.114	101	ALA	H0C	0.019
101	ALA	CA	0.004	101	ALA	CA	0.003
101	ALA	HA	0.135	101	ALA	HA	0.032
101	ALA	C	0.002	101	ALA	C	0.002
101	ALA	O	0.003	101	ALA	O	0.001
101	ALA	CB	0.004	101	ALA	CB	0.004
101	ALA	HB1	0.175	101	ALA	HB1	0.091
101	ALA	HB2	0.163	101	ALA	HB2	0.109
101	ALA	HB3	0.141	101	ALA	HB3	0.040
102	PRO	N	0.003	102	PRO	N	0.003
102	PRO	CA	0.005	102	PRO	CA	0.003
102	PRO	HA	0.155	102	PRO	HA	0.039
102	PRO	C	0.004	102	PRO	C	0.005
102	PRO	O	0.002	102	PRO	O	0.003
102	PRO	CB	0.006	102	PRO	CB	0.005
102	PRO	HB1	0.163	102	PRO	HB1	0.081
102	PRO	HB2	0.198	102	PRO	HB2	0.027
102	PRO	CG	0.004	102	PRO	CG	0.002
102	PRO	HG1	0.147	102	PRO	HG1	0.061

102	PRO	HG2	0.090	102	PRO	HG2	0.035
102	PRO	CD	0.002	102	PRO	CD	0.002
102	PRO	HD1	0.171	102	PRO	HD1	0.038
102	PRO	HD2	0.203	102	PRO	HD2	0.101
103	ALA	N	0.005	103	ALA	N	0.004
103	ALA	H0	0.139	103	ALA	H0	0.021
103	ALA	CA	0.002	103	ALA	CA	0.002
103	ALA	HA	0.124	103	ALA	HA	0.007
103	ALA	CB	0.002	103	ALA	CB	0.003
103	ALA	HB1	0.118	103	ALA	HB1	0.072
103	ALA	HB2	0.159	103	ALA	HB2	0.041
103	ALA	HB3	0.133	103	ALA	HB3	0.054
103	ALA	C	0.002	103	ALA	C	0.002
103	ALA	OT1	0.002	103	ALA	OT1	0.002
103	ALA	OT2	0.004	103	ALA	OT2	0.005
201	HOH	O	0.008	201	HOH	O	0.009
201	HOH	H1	0.087	201	HOH	H1	0.007
201	HOH	H2	0.071	201	HOH	H2	0.021
pro Aminosäure							
101	ALA		0.114	101	ALA		0.051
102	PRO		0.117	102	PRO		0.043
103	ALA		0.091	103	ALA		0.030
201	HOH		0.065	201	HOH		0.014
Gesamtmolekül							
0.106							
0.041							

Tabelle A.3: Bindungslängen asphärischer und sphärischer Verfeinerung von APA und APA0.83.

Bindung	<i>iam</i>	<i>iam_{rb}</i>	<i>mul</i>	<i>iam0.83</i>	<i>iam0.83_{rb}</i>	<i>mul0.83</i>
N(101A) – C(101AA)	1,4958	1,4961	1,4956	1,4986	1,4996	1,4973
C(101AA) – C(101A)	1,5376	1,5374	1,5372	1,5425	1,5420	1,5411
C(101AA) – C(101AB)	1,5342	1,5342	1,5337	1,5352	1,5358	1,5320
C(101A) – O(101A)	1,2402	1,2402	1,2389	1,2340	1,2343	1,2342
C(101A) – N(102A)	1,3539	1,3539	1,3534	1,3559	1,3529	1,3526
N(102A) – C(102AA)	1,4754	1,4753	1,4752	1,4798	1,4806	1,4771
N(102A) – C(102AD)	1,4840	1,4838	1,4846	1,4848	1,4846	1,4852
C(102AA) – C(102A)	1,5393	1,5390	1,5389	1,5395	1,5367	1,5393
C(102AA) – C(102AB)	1,5456	1,5456	1,5466	1,5432	1,5427	1,5471
C(102A) – O(102A)	1,2418	1,2421	1,2412	1,2360	1,2398	1,2374
C(102A) – N(103A)	1,3461	1,3462	1,3459	1,3445	1,3446	1,3444
C(102AB) – C(102AG)	1,5389	1,5394	1,5394	1,5360	1,5388	1,5382
C(102AG) – C(102AD)	1,5371	1,5370	1,5368	1,5343	1,5336	1,5364

N(103A) – C(103AA)	1,4624	1,4623	1,4624	1,4653	1,4671	1,4630
C(103AA) – C(103AB)	1,5457	1,5456	1,5456	1,5446	1,5448	1,5453
C(103AA) – C(103A)	1,5436	1,5436	1,5434	1,5471	1,5455	1,5431
C(103A) – O(103AT1)	1,2661	1,2663	1,2650	1,2628	1,2644	1,2630
C(103A) – O(103AT2)	1,2710	1,2713	1,2700	1,2716	1,2711	1,2694

Tabelle A.4: Bindungslängendifferenzen und Bindungstypen asphärischer und sphärischer Verfeinerung von APA und APA0.83. Kennzeichnung der allen Aminosäuren gemeinsamen Peptidbindungen mit *.

Bindung	Typ	$\Delta d/\text{\AA}$			
		<i>mul</i> –		<i>mul0.83</i> –	
		<i>iam</i>	<i>iam_{rb}</i>	<i>iam0.83</i>	<i>iam0.83_{rb}</i>
C(101AA) – C(101A)	Cncch–Conc *	–0,0004	–0,0002	–0,0014	–0,0009
C(101AA) – C(101AB)	Cncch–Cchhh	–0,0005	–0,0005	–0,0032	–0,0038
C(102AA) – C(102A)	Cncch–Conc *	–0,0004	–0,0001	–0,0002	0,0026
C(102AA) – C(102AB)	Cncch–Ccchh	0,0010	0,0010	0,0039	0,0044
C(102AB) – C(102AG)	Ccchh–Ccchh	0,0005	0,0000	0,0022	–0,0006
C(102AG) – C(102AD)	Ccchh–Cnchh	–0,0003	–0,0002	0,0021	0,0028
C(103AA) – C(103AB)	Cncch–Cchhh	–0,0001	0,0000	0,0007	0,0005
C(103AA) – C(103A)	Cncch–Cooc	–0,0002	–0,0002	–0,0040	–0,0024
N(101A) – C(101AA)	Nchhh–Cncch	–0,0002	–0,0005	–0,0013	–0,0023
C(101A) – N(102A)	Conc–Nccc	–0,0005	–0,0005	–0,0033	–0,0003
N(102A) – C(102AA)	Nccc–Cncch	–0,0002	–0,0001	–0,0027	–0,0035
N(102A) – C(102AD)	Nccc–Cnchh	0,0006	0,0008	0,0004	0,0006
C(102A) – N(103A)	Conc–Ncch *	–0,0002	–0,0003	–0,0001	–0,0002
N(103A) – C(103AA)	Ncch–Cncch *	0,0000	0,0001	–0,0023	–0,0041
C(101A) – O(101A)	Conc–Oc *	–0,0013	–0,0013	0,0002	–0,0001
C(102A) – O(102A)	Conc–Oc *	–0,0006	–0,0009	0,0014	–0,0024
C(103A) – O(103AT1)	Cooc–Oc	–0,0011	–0,0013	0,0002	–0,0014
C(103A) – O(103AT2)	Cooc–Oc	–0,0010	–0,0013	–0,0022	–0,0017

Tabelle A.5: Statistik U_{eq} und ΔU_{eq} APA (alle Einheiten 10^{-3}\AA^2 ; †Minimum, Maximum und Mittelwert mit Standardabweichung; ‡0,25-, 0,5-, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen).

	$x_{min/max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
alle Atome ($n = 40$)							
<i>mul</i>	11,92/48,78	25,56(11,68)	15,13	25,66	34,97	19,84	15,19
<i>iam</i>	12,03/52,98	21,68(8,94)	15,22	18,52	27,68	12,46	8,74
<i>iam_{rb}</i>	12,04/60,68	23,97(11,50)	15,22	21,60	28,56	13,34	10,18

<i>mul-iam</i>	-4,20/15,61	3,88(5,74)	-0,14	-0,09	8,88	9,01	2,10
<i>mul-iam_{rb}</i>	-17,69/12,64	1,59(5,71)	-0,14	-0,11	4,59	4,74	1,70
ohne Solvens ($n = 37$)							
<i>mul</i>	11,92/46,07	24,50(11,10)	15,08	25,20	33,59	18,51	14,92
<i>iam</i>	12,03/36,54	20,37(7,19)	15,20	17,90	25,57	10,37	7,74
<i>iam_{rb}</i>	12,04/49,20	22,43(9,61)	15,21	20,38	28,42	13,21	11,38
<i>mul-iam</i>	-2,01/15,61	4,13(5,75)	-0,12	-0,09	9,00	9,12	2,04
<i>mul-iam_{rb}</i>	-8,90/12,64	2,07(4,93)	-0,13	-0,11	4,50	4,63	1,44
ohne Solvens, ohne H ($n = 18$)							
<i>mul</i>	11,92/18,32	14,44(2,21)	12,31	14,40	16,04	3,73	3,02
<i>iam</i>	12,03/18,47	14,55(2,20)	12,45	14,52	16,18	3,73	3,00
<i>iam_{rb}</i>	12,04/18,49	14,56(2,21)	12,46	14,53	16,18	3,72	3,00
<i>mul-iam</i>	-0,15/-0,08	-0,11(0,02)	-0,14	-0,11	-0,09	0,04	0,04
<i>mul-iam_{rb}</i>	-0,17/-0,08	-0,12(0,02)	-0,13	-0,13	-0,11	0,02	0,02

Tabelle A.6: Statistik U_{eq} und ΔU_{eq} APA0.83 (alle Einheiten 10^{-3}\AA^2 ; †Minimum, Maximum und Mittelwert mit Standardabweichung; ‡0,25-, 0,5-, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen).

	$x_{\min/\max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
alle Atome ($n = 40$)							
<i>mul</i>	12,38/60,00	27,26(12,96)	15,85	27,53	37,43	21,58	17,05
<i>iam</i>	11,14/62,28	21,87(8,92)	16,08	18,81	26,18	10,09	5,75
<i>iam_{rb}</i>	14,18/59,45	26,46(11,63)	17,84	24,06	31,20	13,37	10,11
<i>mul-iam</i>	-2,60/20,90	5,39(8,67)	-2,08	-1,04	13,34	15,41	2,17
<i>mul-iam_{rb}</i>	-15,87/17,74	0,80(6,72)	-2,35	-1,74	4,36	6,71	1,78
ohne Solvens ($n = 37$)							
<i>mul</i>	12,38/48,79	25,96(11,89)	15,59	26,54	36,23	20,63	16,06
<i>iam</i>	11,14/34,30	20,34(5,87)	15,81	18,37	24,83	9,02	5,07
<i>iam_{rb}</i>	14,18/50,74	24,74(9,44)	17,41	21,67	29,50	12,08	9,63
<i>mul-iam</i>	-2,40/20,90	5,61(8,75)	-1,76	-0,85	13,57	15,32	2,29
<i>mul-iam_{rb}</i>	-15,87/17,74	1,21(6,69)	-2,34	-1,70	4,45	6,79	1,35
ohne Solvens, ohne H ($n = 18$)							
<i>mul</i>	12,38/18,69	15,02(2,25)	12,84	15,13	16,51	3,67	3,38
<i>iam</i>	14,02/20,45	16,94(2,16)	14,94	16,72	18,40	3,46	2,61
<i>iam_{rb}</i>	14,18/20,78	17,13(2,21)	15,09	16,89	18,45	3,36	2,63
<i>mul-iam</i>	-2,40/-1,23	-1,92(0,38)	-2,32	-1,91	-1,68	0,65	0,54
<i>mul-iam_{rb}</i>	-2,61/-1,60	-2,10(0,31)	-2,35	-2,09	-1,85	0,50	0,40

Tabelle A.7: Statistik U_{eq} und ΔU_{eq} APA (ohne Solvens) – nach Element (alle Einheiten 10^{-3}\AA^2 ; \dagger Minimum, Maximum und Mittelwert mit Standardabweichung; \ddagger 0,25-, 0,5-, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen).

	$x_{\min/\max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
Kohlenstoff ($n = 11$)							
<i>mul</i>	11,92/18,32	14,08(2,30)	12,19	12,53	15,67	3,48	0,90
<i>iam</i>	12,03/18,47	14,20(2,30)	12,32	12,68	15,78	3,46	0,96
<i>iam_{rb}</i>	12,04/18,49	14,21(2,30)	12,32	12,67	15,78	3,46	0,94
<i>mul-iam</i>	-0,15/-0,09	-0,12(0,02)	-0,14	-0,12	-0,11	0,03	0,03
<i>mul-iam_{rb}</i>	-0,17/-0,08	-0,13(0,03)	-0,14	-0,13	-0,11	0,03	0,03
Stickstoff ($n = 3$)							
<i>mul</i>	12,24/13,72	12,98(0,74)	12,60	12,96	13,34	0,74	1,06
<i>iam</i>	12,38/13,84	13,08(0,73)	12,71	13,04	13,44	0,73	0,97
<i>iam_{rb}</i>	12,38/13,85	13,09(0,74)	12,71	13,04	13,45	0,74	0,99
<i>mul-iam</i>	-0,14/-0,08	-0,11(0,03)	-0,12	-0,11	-0,10	0,03	0,04
<i>mul-iam_{rb}</i>	-0,13/-0,09	-0,11(0,03)	-0,13	-0,13	-0,11	0,02	0,00
Sauerstoff ($n = 4$)							
<i>mul</i>	15,55/17,82	16,51(1,14)	15,57	16,35	17,29	1,72	1,16
<i>iam</i>	15,67/17,90	16,62(1,11)	15,71	16,46	17,37	1,66	1,13
<i>iam_{rb}</i>	15,68/17,93	16,64(1,12)	15,70	16,47	17,40	1,70	1,15
<i>mul-iam</i>	-0,14/-0,08	-0,11(0,03)	-0,13	-0,10	-0,08	0,04	0,03
<i>mul-iam_{rb}</i>	-0,13/-0,11	-0,12(0,01)	-0,13	-0,12	-0,11	0,02	0,01

Tabelle A.8: Statistik U_{eq} und ΔU_{eq} APA0.83 (ohne Solvens) – nach Element (alle Einheiten 10^{-3}\AA^2 ; \dagger Minimum, Maximum und Mittelwert mit Standardabweichung; \ddagger 0,25-, 0,5-, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen).

	$x_{\min/\max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
Kohlenstoff ($n = 11$)							
<i>mul</i>	12,38/18,69	14,65(2,30)	12,72	13,43	16,30	3,58	1,55
<i>iam</i>	14,02/20,45	16,59(2,10)	14,96	15,81	17,89	2,93	2,16
<i>iam_{rb}</i>	14,18/20,78	16,76(2,21)	14,99	15,77	18,21	3,21	2,37
<i>mul-iam</i>	-2,40/-1,45	-1,94(0,38)	-2,37	-1,76	-1,68	0,69	0,45
<i>mul-iam_{rb}</i>	-2,61/-1,61	-2,11(0,31)	-2,35	-2,09	-1,93	0,42	0,39

Stickstoff ($n = 3$)							
<i>mul</i>	12,52/14,67	13,59(1,08)	13,04	13,57	14,12	1,08	1,56
<i>iam</i>	14,75/16,17	15,24(0,81)	14,78	14,80	15,49	0,71	0,07
<i>iam_{rb}</i>	14,93/16,37	15,49(0,77)	15,05	15,17	15,77	0,72	0,36
<i>mul-iam</i>	-2,23/-1,23	-1,65(0,52)	-1,87	-1,50	-1,36	0,50	0,41
<i>mul-iam_{rb}</i>	-2,41/-1,60	-1,91(0,44)	-2,06	-1,70	-1,65	0,40	0,15
Sauerstoff ($n = 4$)							
<i>mul</i>	16,00/18,56	17,13(1,30)	16,04	16,98	18,07	2,04	1,42
<i>iam</i>	18,07/20,28	19,19(1,11)	18,32	19,21	20,07	1,75	1,39
<i>iam_{rb}</i>	18,33/20,53	19,36(1,13)	18,42	19,30	20,24	1,82	1,34
<i>mul-iam</i>	-2,36/-1,71	-2,06(0,26)	-2,16	-2,08	-1,98	0,18	0,21
<i>mul-iam_{rb}</i>	-2,40/-1,97	-2,23(0,19)	-2,35	-2,28	-2,16	0,19	0,13

Tabelle A.9: DMSDA-Werte sphärischer und asphärischer Verfeinerung von APA und APA0.83.

Bindung	<i>iam</i>	<i>iam_{rb}</i>	<i>mul</i>	<i>iam0.83</i>	<i>iam0.83_{rb}</i>	<i>mul0.83</i>
N(101A) - C(101AA)	1	1	1	-6	-5	-3
C(101AA) - C(101A)	-2	-2	-5	14	12	-25
C(101AA) - C(101AB)	0	0	0	9	13	2
C(101A) - O(101A)	-7	-7	-5	-8	-16	-15
C(101A) - N(102A)	-6	-6	-4	-26	-14	-6
N(102A) - C(102AA)	5	5	5	2	-4	5
N(102A) - C(102AD)	6	6	5	3	-9	4
C(102AA) - C(102A)	3	2	-1	8	8	-17
C(102AA) - C(102AB)	0	0	1	-1	-2	-5
C(102A) - O(102A)	-6	-5	-3	-19	-20	-18
C(102A) - N(103A)	-4	-4	-3	-14	-21	-1
C(102AB) - C(102AG)	-1	-1	-1	-5	-12	-9
C(102AG) - C(102AD)	-1	-1	-1	0	2	2
N(103A) - C(103AA)	4	4	3	-9	-13	-4
C(103AA) - C(103AB)	3	4	3	6	13	8
C(103AA) - C(103A)	0	-1	-1	-2	-4	5
C(103A) - O(103AT1)	-6	-6	-2	-3	-15	0
C(103A) - O(103AT2)	-5	-5	-1	-34	-22	0
$(\sum_i^n DMSDA _i)n^{-1}$	3,333	3,333	2,500	9,389	11,389	7,167

Tabelle A.10: DMSDA-Betragsdifferenzen $|\Delta\text{DMSDA}|$ und Bindungstypen asphärischer und sphärischer Verfeinerung von APA und APA0.83. (Allen Aminosäuren gemeinsame Peptidbindungen mit * gekennzeichnet.)

Bindung	Typ	$\Delta \text{DMSDA} /10^{-4} \text{ \AA}$			
		<i>mul</i> –		<i>mul0.83</i> –	
		<i>iam</i>	<i>iam_{rb}</i>	<i>iam0.83</i>	<i>iam0.83_{rb}</i>
C(101AA) – C(101A)	Cncch–Conc *	3	3	11	13
C(101AA) – C(101AB)	Cncch–Cchhh	0	0	-7	-11
C(102AA) – C(102A)	Cncch–Conc *	-2	-1	9	9
C(102AA) – C(102AB)	Cncch–Ccchh	1	1	4	3
C(102AB) – C(102AG)	Ccchh–Ccchh	0	0	4	-3
C(102AG) – C(102AD)	Ccchh–Cnchh	0	0	2	0
C(103AA) – C(103AB)	Cncch–Cchhh	0	-1	2	-5
C(103AA) – C(103A)	Cncch–Cooc	1	0	3	1
N(101A) – C(101AA)	Nchhh–Cncch	0	0	-3	-2
C(101A) – N(102A)	Conc–Nccc	-2	-2	-20	-8
N(102A) – C(102AA)	Nccc–Cncch	0	0	3	1
N(102A) – C(102AD)	Nccc–Cnchh	-1	-1	1	-5
C(102A) – N(103A)	Conc–Nech *	-1	-1	-13	-20
N(103A) – C(103AA)	Nech–Cncch *	-1	-1	-5	-9
C(101A) – O(101A)	Conc–Oc *	-2	-2	7	-1
C(102A) – O(102A)	Conc–Oc *	-3	-2	-1	-2
C(103A) – O(103AT1)	Cooc–Oc	-4	-4	-3	-15
C(103A) – O(103AT2)	Cooc–Oc	-4	-4	-34	-22

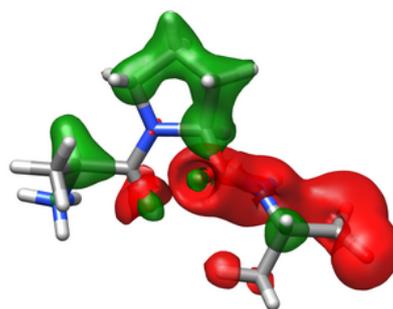
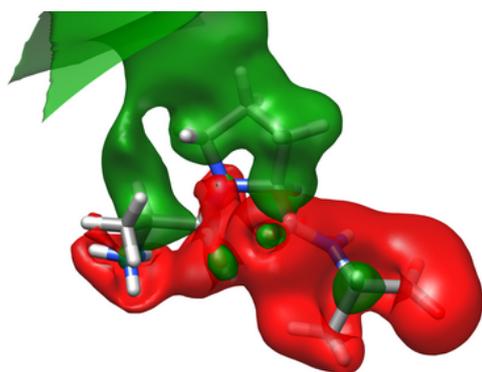
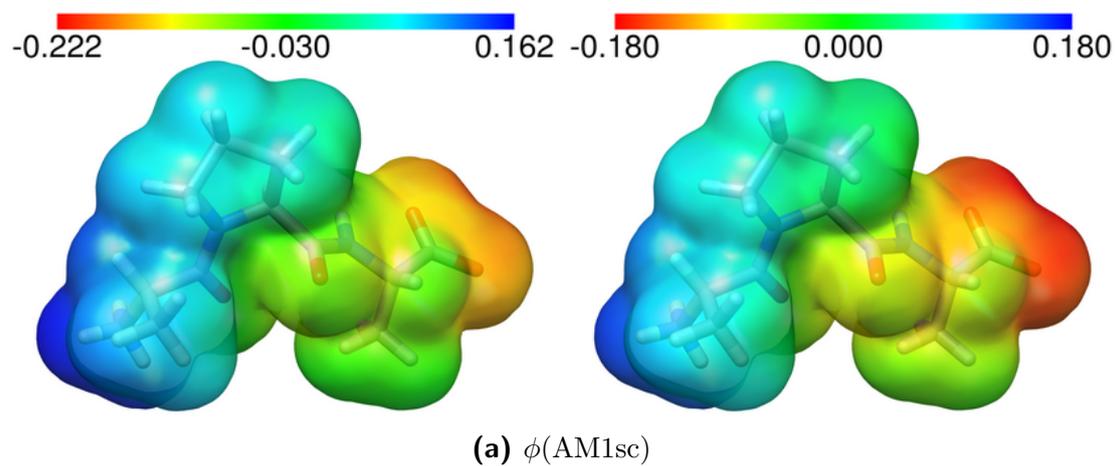


Abbildung A.1: (a): Skaliertes AM1-ESP (e/a_0) auf $0,001 e/a_0^3$ -Oberfläche, vollständiger (links) und auf $\pm 0,18 e/a_0$ beschränkter (rechts) Wertebereich. (b) und (c): Isooberflächen der Differenz AM1sc-*mul* bei $\pm 0,05\phi e/a_0$ und $\pm 0,1\phi e/a_0$ (positiv: grün, negativ: rot).

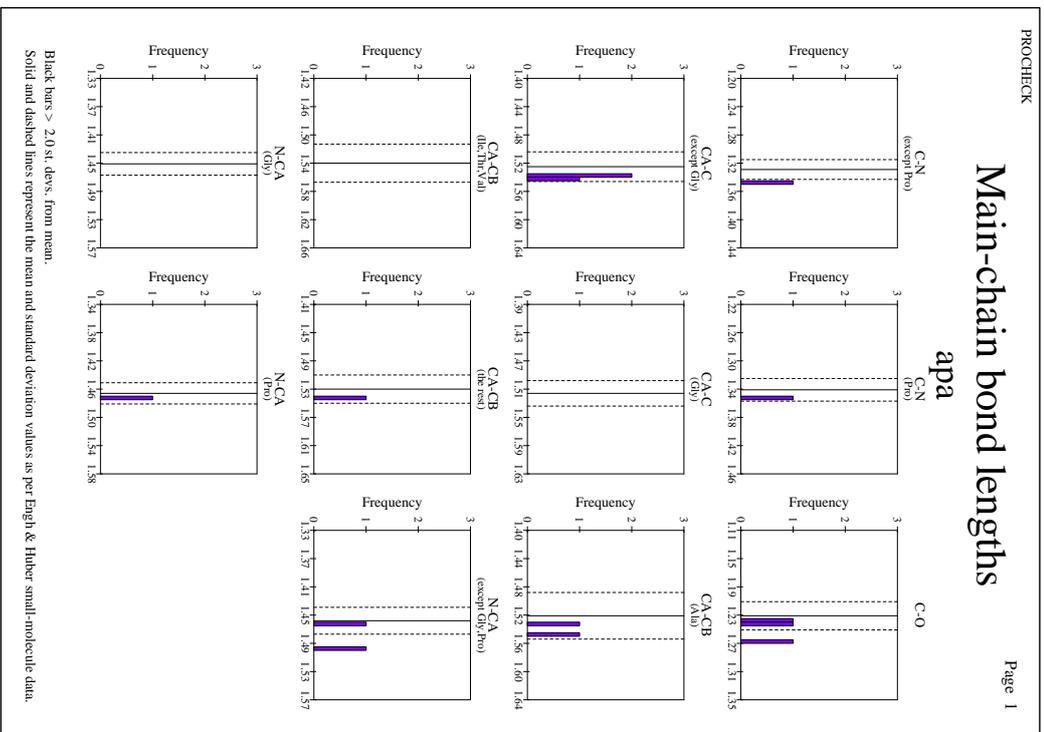
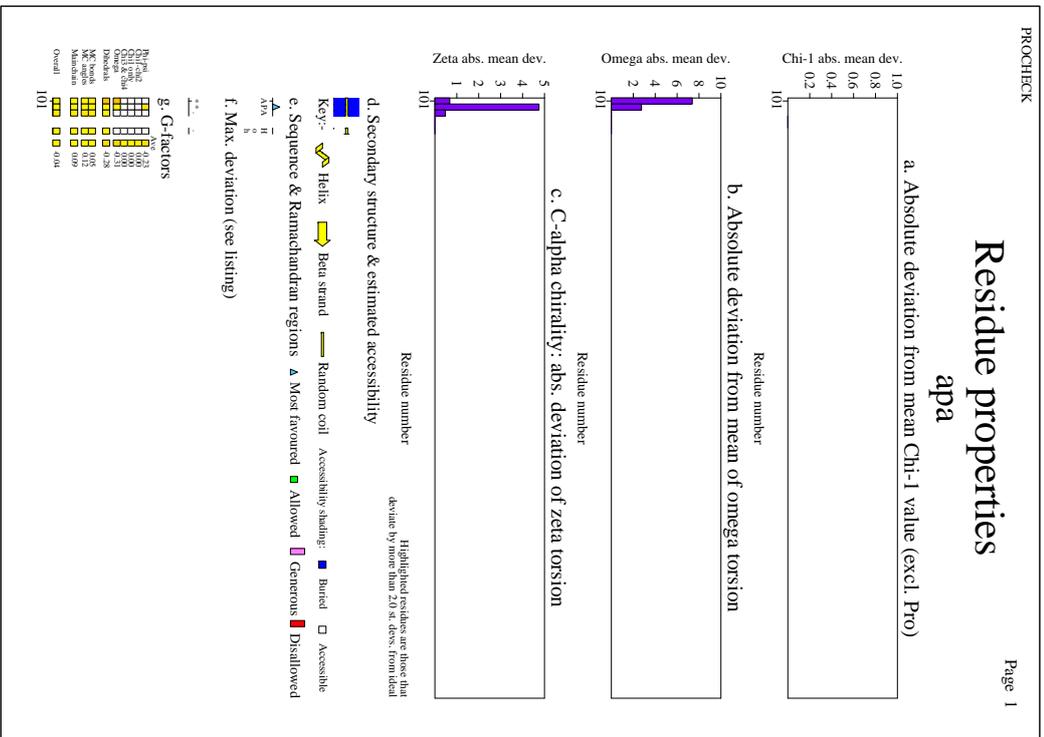


Abbildung A.2: PROCHECK-Ausgabe APA.

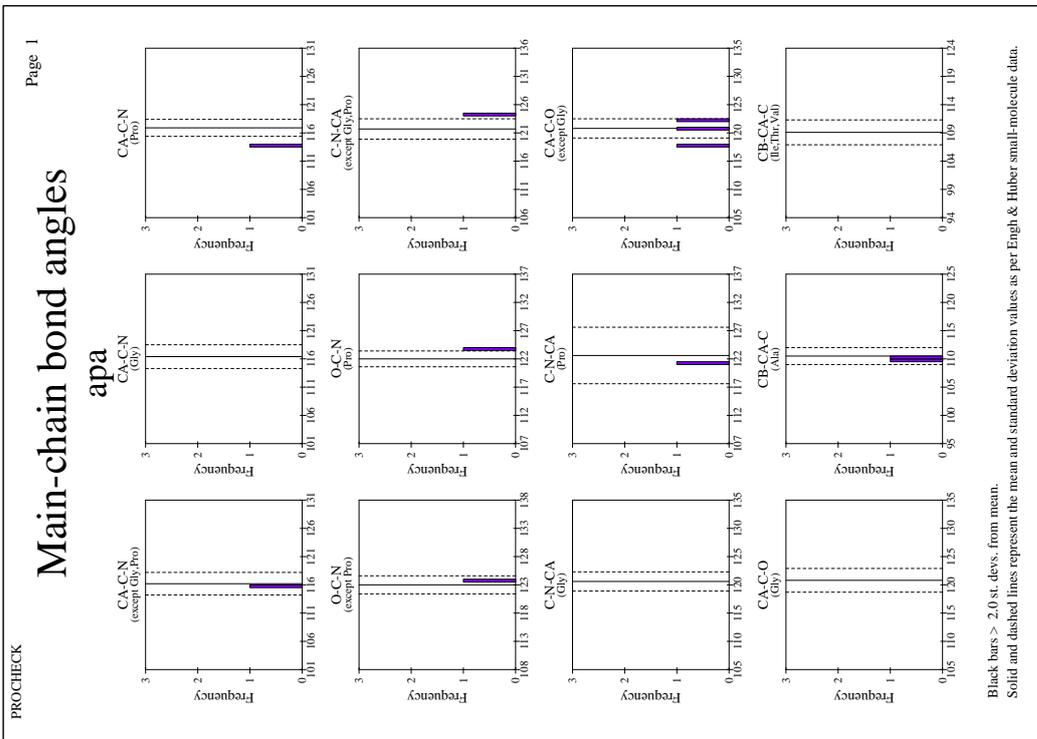
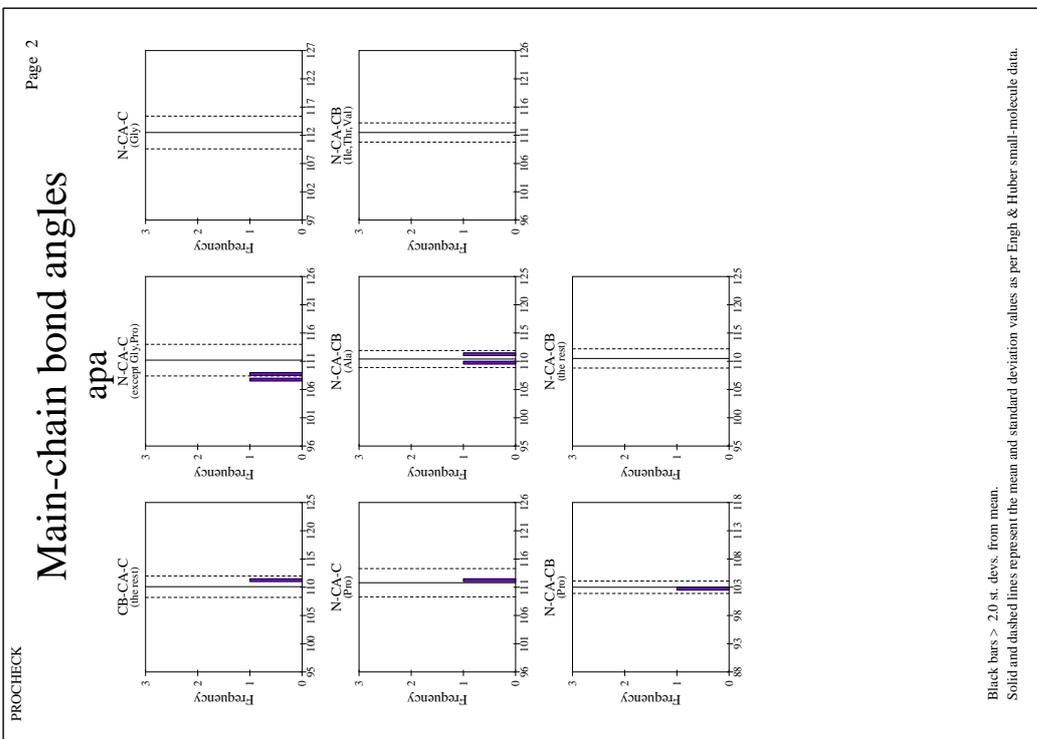


Abbildung A.2: PROCHECK-Ausgabe APA (Forts.).

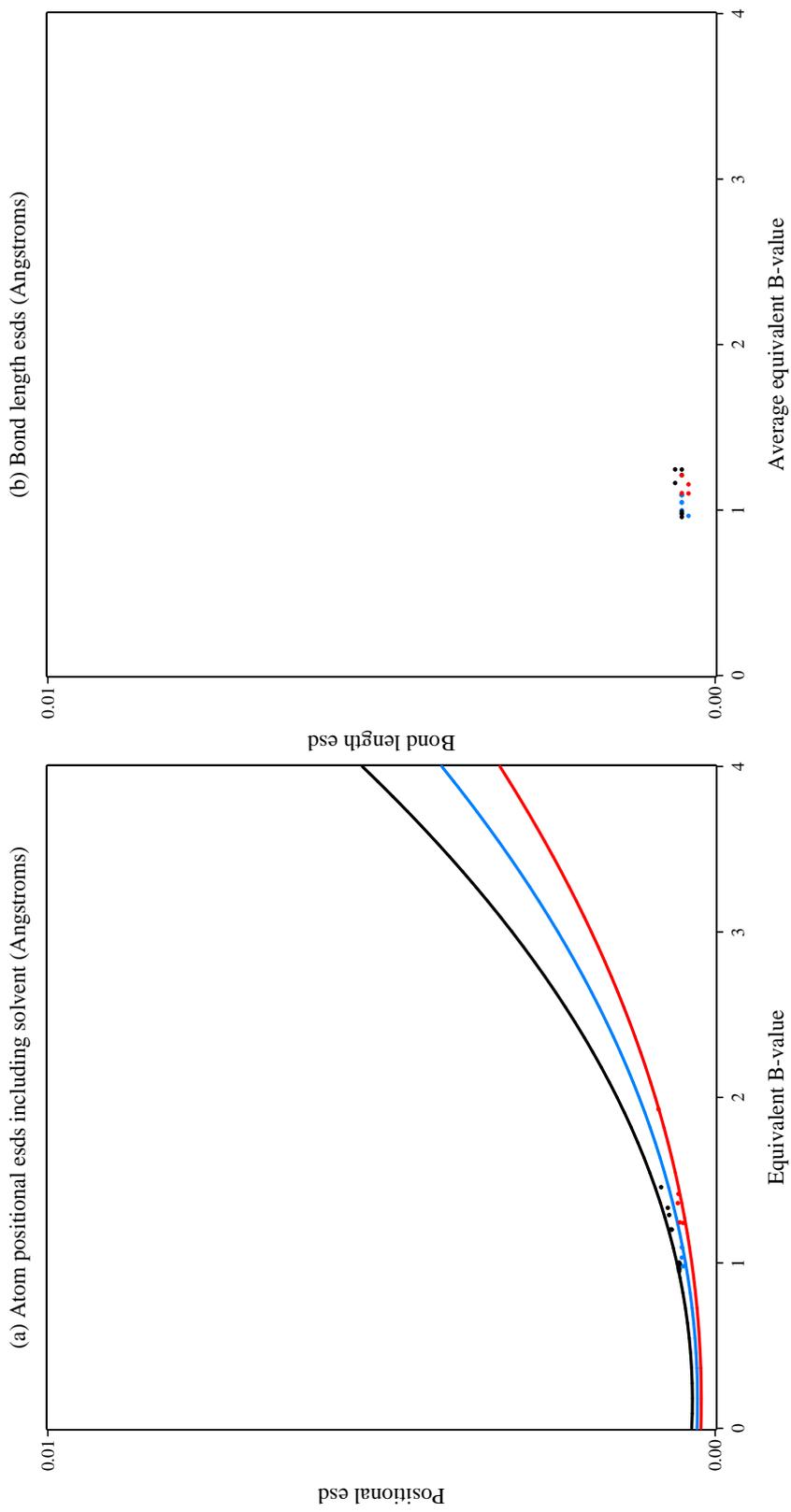


Abbildung A.4: SHELXPRO-Ausgabe APA.

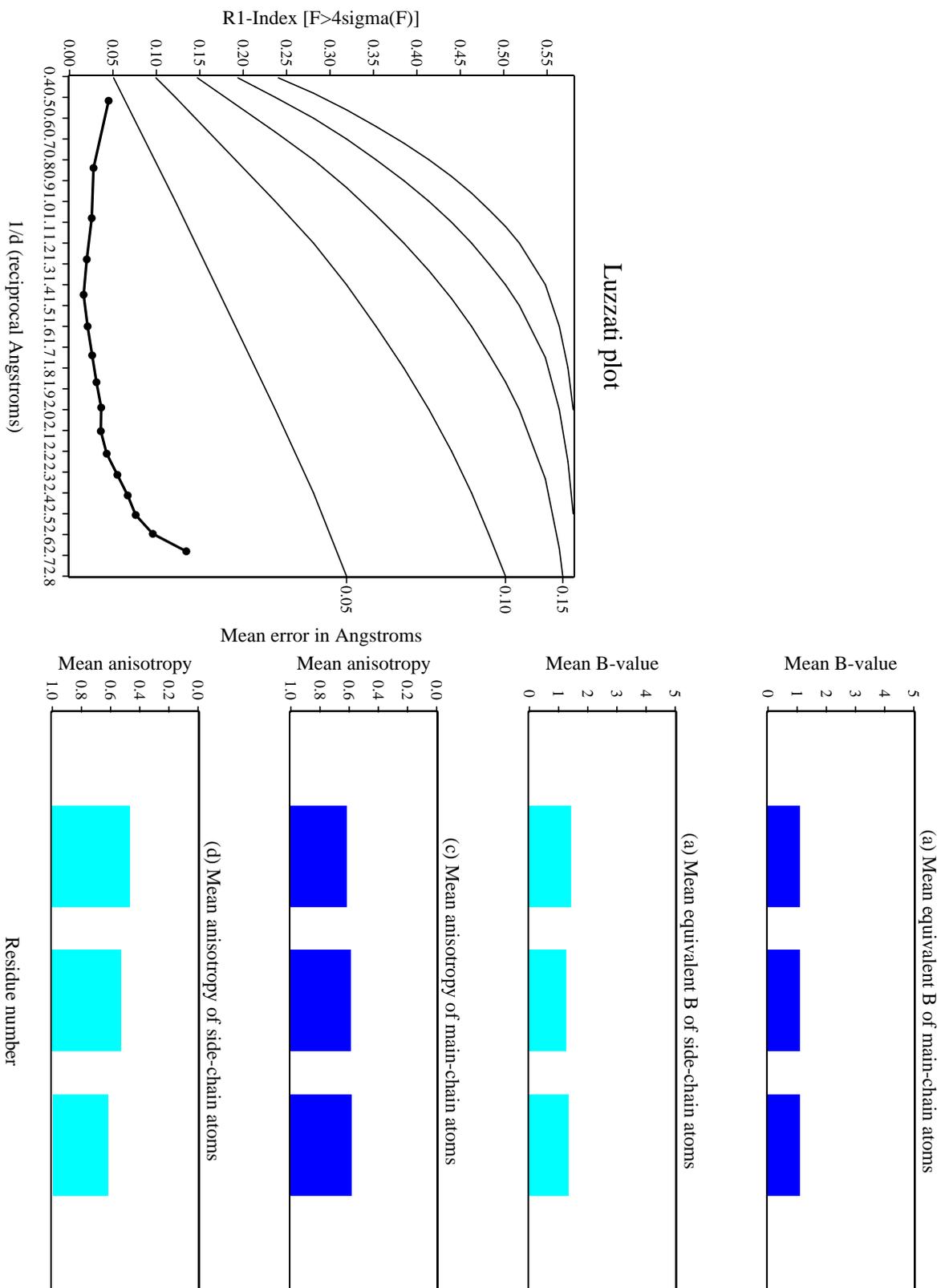


Abbildung A.4: SHELXPRO-Ansgabe APA (Forts.).

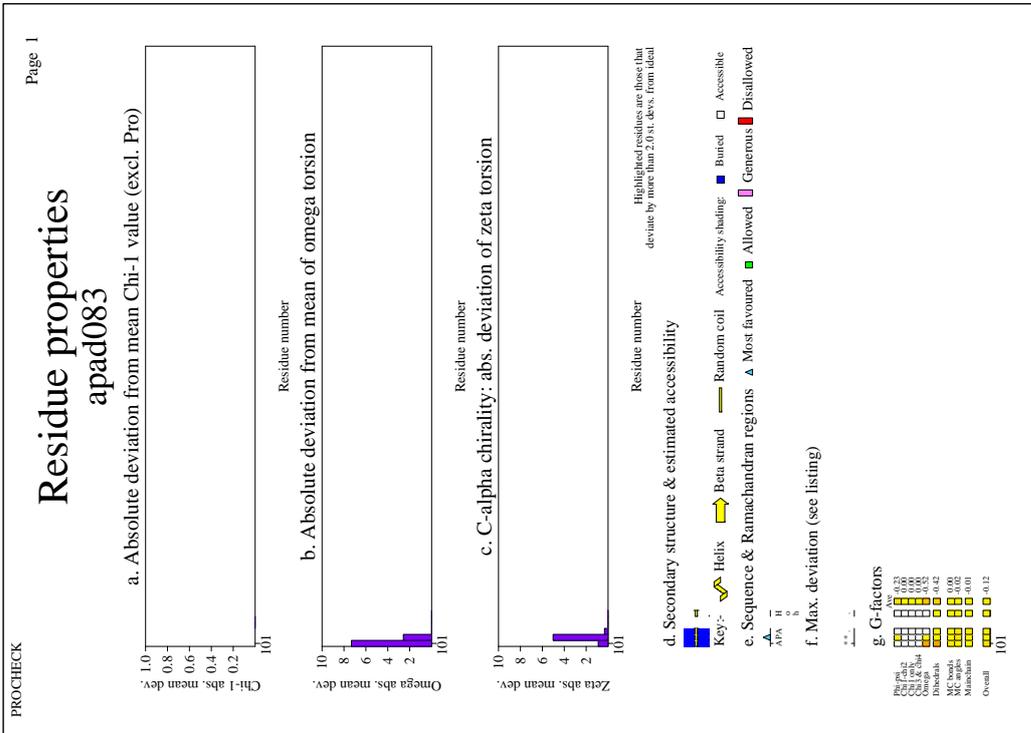
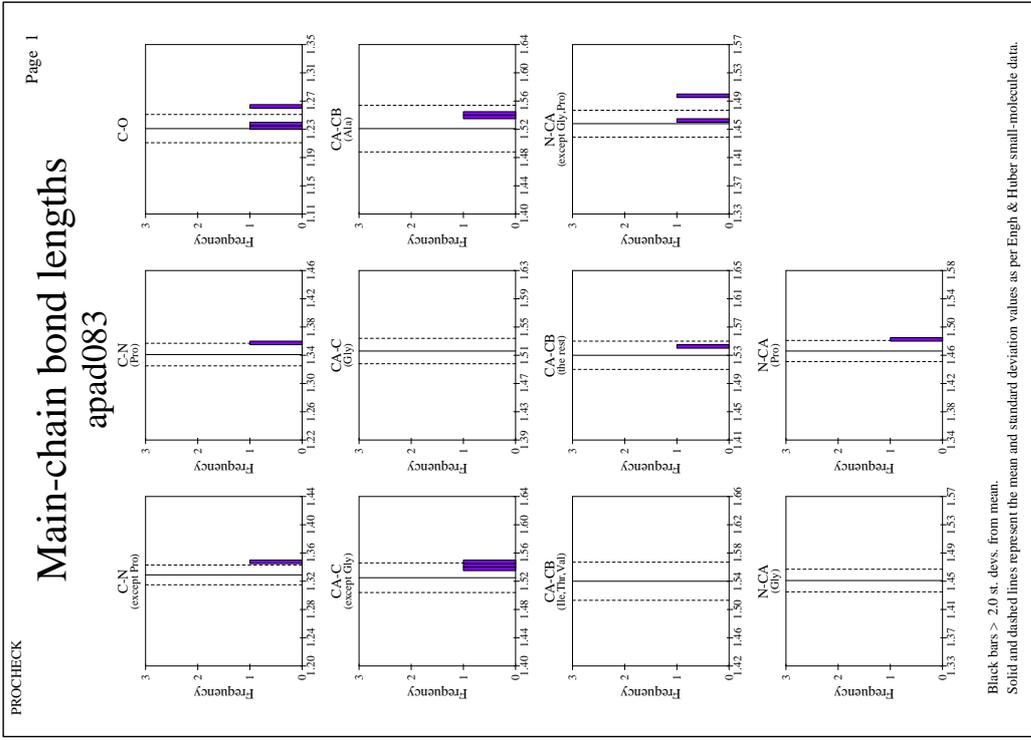


Abbildung A.5: PROCHECK-Ausgabe APA0.83.

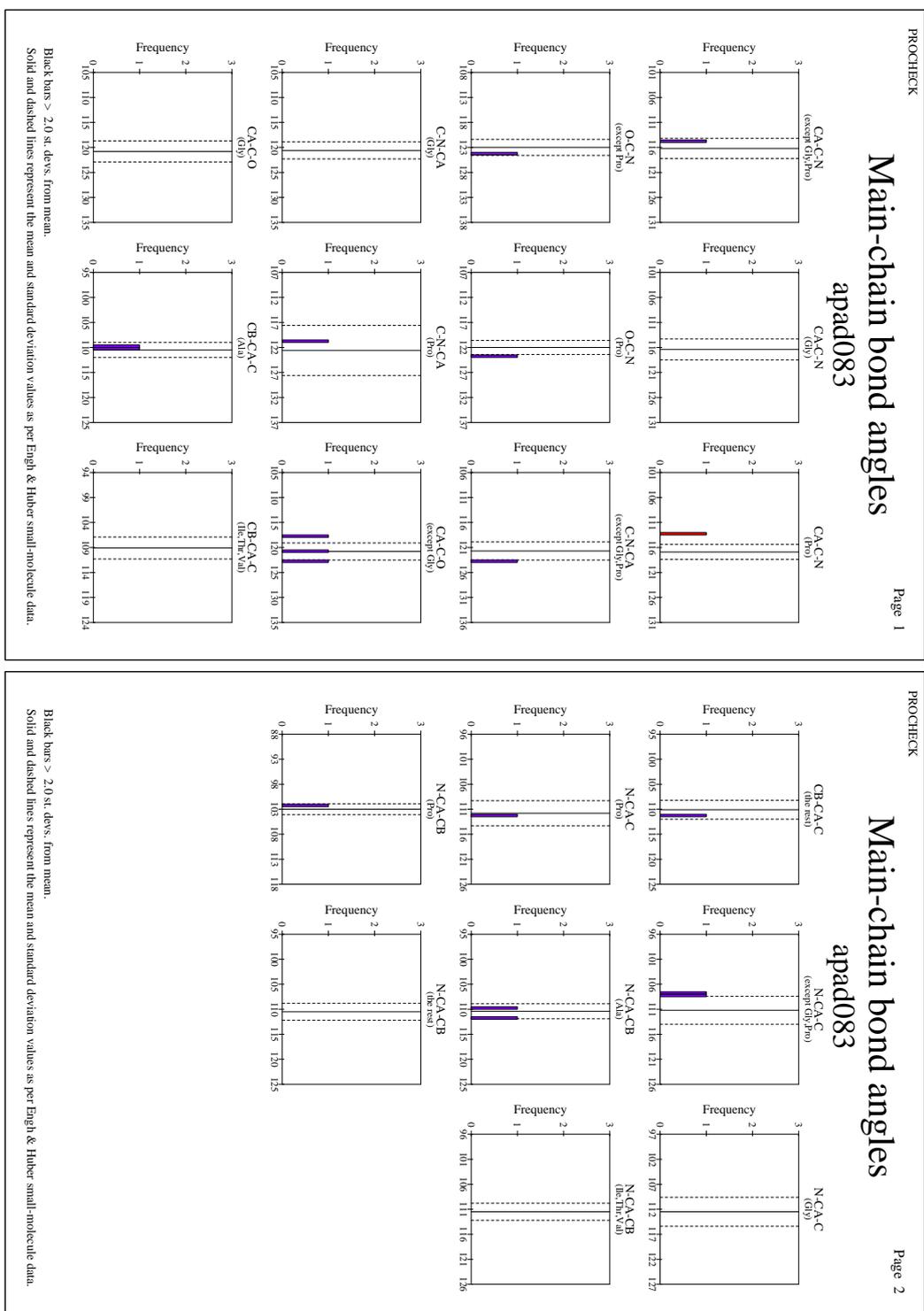


Abbildung A.5: PROCHECK-Ausgabe APA0.83 (Forts.).

1 apad083.pdb

1.1 Sect 0) Symmetry related problems

1.1.1 Error: Matthews Coefficient (V_m) too low

Molecular weight of all polymer chains: 255.274
Volume of the Unit Cell V = 1340.859
Cell multiplicity: 4
Matthews coefficient for observed atoms V_m = 1.313

1.1.2 Note: All atoms are sufficiently far away from symmetry axes

1.2 Sect 3) Coordinate problems, unexpected atoms, B-factor and occupancy checks

1.2.1 Note: Normal distribution of occupancy values

1.2.2 Note: All occupancies seem to add up to 1.0.

1.2.3 Note: B-factor distribution normal

RMS Z-score : 0.394 over 15 bonds
Average difference in B over a bond : 0.18
RMS difference in B over a bond : 0.24

1.3 Sect 5) Geometric checks

1.3.1 Note: All bond lengths OK

1.3.2 Note: Normal bond length variability

RMS Z-score for bond lengths: 0.991
RMS-deviation in bond distances: 0.023

1.3.3 Note: No bond length directionality

1.3.4 Note: All bond angles OK

1.3.5 Note: Normal bond angle variability

RMS Z-score for bond angles: 0.902
RMS-deviation in bond angles: 1.891

1.3.6 Note: Chirality OK

All protein atoms have proper chirality.

1.3.7 Note: Improper dihedral angle distribution OK

Improper dihedral RMS Z-score : 0.803

1.3.8 Note: Side chain planarity OK

1.3.9 Note: Atoms connected to aromatic rings OK

1.4 Sect 6) Torsion-related checks

1.4.1 Note: PRO puckering amplitude OK

1.4.2 Note: PRO puckering phases OK

1.4.3 Note: Torsion angles OK

1.4.4 Note: Backbone torsion angles OK

1.4.5 Note: Ramachandran Z-score OK

B-Ramachandran Z-score : 0.011

1.4.6 Note: Rotamers checked OK

1.4.7 Note: Backbone conformations OK

1.4.8 Error: Backbone conformation Z-score very low

Backbone conformation Z-score : -6456.004

1.5 Sect 7) Bump checks

1.5.1 Note: No Van der Waals overlaps

1.6 Sect 8) Packing, accessibility and threading

1.6.1 Warning: Inside/Outside residue distribution unusual

inside/outside RMS Z-score : 1.340

1.6.2 Note: Second generation packing environment OK

1.6.3 Note: No series of residues with abnormal new packing environment

1.6.4 Note: Structural average packing Z-score OK

All contacts : Average = -0.397 Z-score = -2.25
BB-BB contacts : Average = 1.024 Z-score = 6.10
BB-SC contacts : Average = -1.342 Z-score = -9.43
SC-BB contacts : Average = 1.005 Z-score = 6.40
SC-SC contacts : Average = -2.459 Z-score = -15.92

1.7 Sect 9) Water and hydrogenbond related checks

1.7.1 Note: Water contacts OK

1.7.2 Note: Water hydrogen bonds OK

1.7.3 Warning: Buried unsatisfied hydrogen bond donors

Residue
3 ALA (103) - - N

1.7.4 Note: Buried hydrogen bond acceptors OK

1.8 Final summary

1.8.1 Note: Summary report for users of a structure

This is an overall summary of the quality of the structure as compared with current reliable structures. This summary is most useful for biologists seeking a good structure to use for modelling calculations.

The second part of the table mostly gives an impression of how well the model conforms to common refinement constraint values. The first part of the table shows a number of constraint-independent quality indicators.

Structure Z-scores, positive is better than average:

1st generation packing quality : -8.267 (bad)

2nd generation packing quality : -2.252

Ramachandran plot appearance : 0.011

RMS Z-scores, should be close to 1.0:

Bond lengths : 0.991

Bond angles : 0.902

Improper dihedral distribution : 0.803

B-factor distribution : 0.394

Inside/Outside distribution : 1.340 (unusual)

Abbildung A.6: Gekürzte WHATCHECK-Ausgabe APA $d = 0,83 \text{ \AA}$.

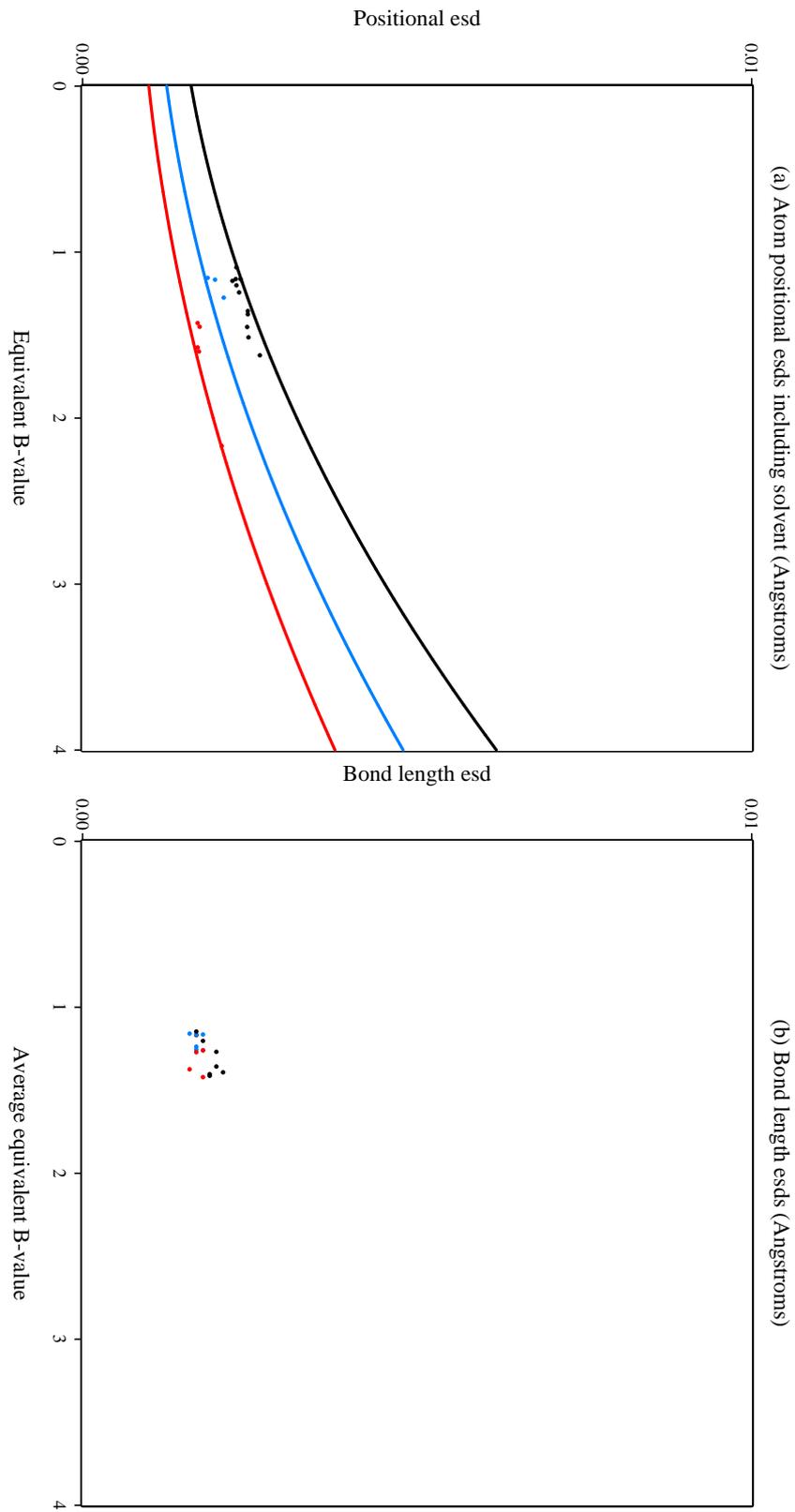


Abbildung A.7: SHELXPRO-Ausgabe APA0.83.

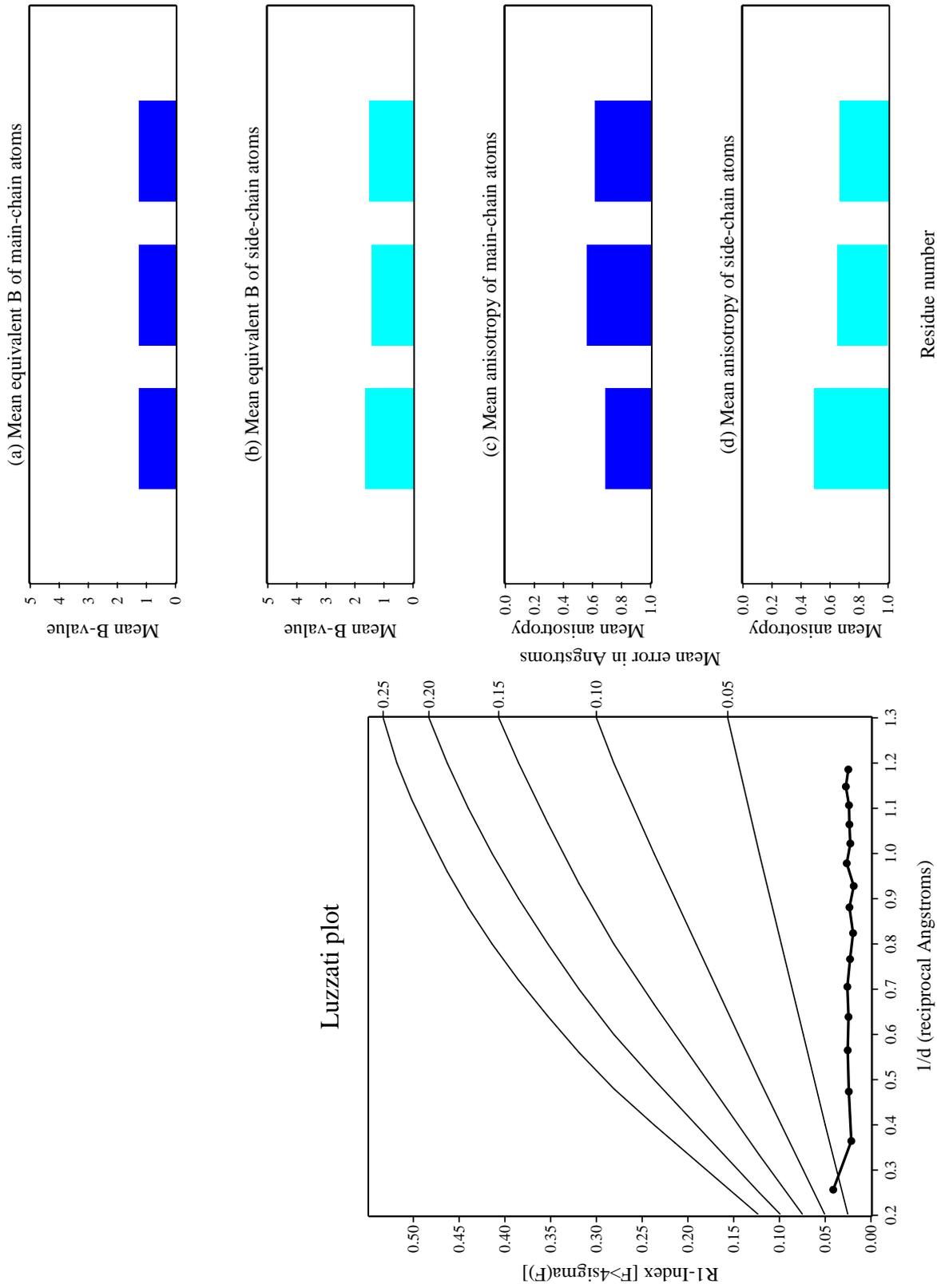


Abbildung A.7: SHELXPRO-Ausgabe APA0.83 (Forts.).

B Triklines Lysozym

Tabelle B.1: Bindungstypen P1LYSO.

Bindungstyp		gesamt			verfeinert			
		AA	AB	BB	AA	AB	BB	
C[ccc] – C[ccc]	9	6	0	3	3	3	0	0
C[ccc] – C[cch]	21	18	0	3	13	13	0	0
C[ccc] – C[ncc]	9	6	0	3	3	3	0	0
C[ccc] – C[nch]	9	6	0	3	3	3	0	0
C[ccch] – C[cchh]	21	14	0	7	6	6	0	0
C[ccch] – C[chhh]	48	34	0	14	18	18	0	0
C[cch] – C[cch]	45	36	0	9	25	25	0	0
C[cch] – C[cco]	6	6	0	0	4	4	0	0
C[cchh] – C[ccc]	16	12	0	4	8	8	0	0
C[cchh] – C[cchh]	46	30	1	15	13	13	0	0
C[cchh] – C[chhh]	10	6	0	4	2	2	0	0
C[cchh] – C[ncc]	2	1	0	1	0	0	0	0
C[cchh] – C[nchh]	34	19	2	13	4	4	0	0
C[cchh] – C[onc]	24	17	0	7	10	10	0	0
C[cchh] – C[ooc]	12	9	0	3	6	6	0	0
C[cchh] – C[schh]	2	2	0	0	2	2	0	0
C[ncc] – C[cch]	9	6	0	3	3	3	0	0
C[ncc] – C[nch]	2	1	0	1	0	0	0	0
C[ncc] – C[ccch]	18	12	0	6	6	6	0	0
C[ncc] – C[cchh]	98	68	8	22	53	45	7	1
C[ncc] – C[chhh]	13	12	0	1	11	11	0	0
C[ncc] – C[occh]	9	7	0	2	5	5	0	0
C[ncc] – C[ochh]	15	10	3	2	11	8	3	0
C[ncc] – C[onc]	150	116	12	22	100	90*	10	0
C[ncc] – C[ooc]	1	1	0	0	0	0	0	0
C[ncc] – C[schh]	8	8	0	0	8	8	0	0
C[nchh] – C[onc]	14	12	0	2	10	10	0	0
C[occh] – C[chhh]	9	7	0	2	5	5	0	0
C[cco] – O[ch]	3	3	0	0	2	2	0	0
C[occh] – O[ch]	9	7	0	2	5	5	0	0
C[ochh] – O[ch]	15	10	0	5	6	6	0	0
C[onc] – O[c]	177	145	2	30	110	110*	0	0

C[ooc] – O[c]	25	19	0	6	11	11	0	0
C[ooc] – O[ch]	1	1	0	0	1	1	0	0
C[ncc] – N[cch]	11	7	0	4	3	3	0	0
C[onc] – N[ccc]	3	2	0	1	1	1	0	0
C[onc] – N[cch]	157	126	15	16	121	107*	14	0
C[onc] – N[chh]	24	17	0	7	10	10	0	0
C[nchh] – N[cch]	35	23	0	12	11	11	0	0
C[nchh] – N[chhh]	9	6	0	3	2	2	0	0
C[nnh] – N[cc]	2	1	0	1	0	0	0	0
C[nnm] – N[chh]	42	22	0	20	2	2	0	0
C[nccch] – N[cch]	147	114	13	20	85	82*	3	0
C[nch] – N[cc]	2	1	0	1	0	0	0	0
C[nch] – N[cch]	9	6	0	3	3	3	0	0
C[nccch] – N[ccc]	3	2	0	1	0	0	0	0
C[nchh] – N[ccc]	4	2	1	1	0	0	0	0
C[nnh] – N[cch]	2	1	0	1	0	0	0	0
C[nnm] – N[cch]	21	11	0	10	1	1	0	0
C[nccch] – N[chhh]	1	1	0	0	1	1	0	0
C[schh] – S[cc]	2	2	0	0	2	2	0	0
C[schh] – S[sc]	8	8	0	0	8	8	0	0
C[shhh] – S[cc]	2	2	0	0	2	2	0	0
S[sc] – S[sc]	4	4	0	0	4	4	0	0
Σ	1378	1025	57	296	723	685	37	1

Tabelle B.2: Statistik der U_{eq} von P1LYSO (alle Einheiten in 10^{-3}\AA^2).

	$x_{min/max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
alle Atome ($n = 2901$)							
<i>mul</i>	24,50/981,00	97,27(80,69)	52,01	72,56	107,00	55,02	35,74
<i>iam</i>	24,83/981,00	97,33(80,65)	52,05	72,58	107,20	55,11	35,68
<i>iam_{rb}</i>	25,23/981,00	97,34(80,65)	52,08	72,56	107,20	55,08	35,67
ohne Solvens ($n = 2658$)							
<i>mul</i>	24,50/607,30	91,08(72,22)	50,48	70,15	98,83	48,34	33,09
<i>iam</i>	24,83/607,30	91,14(72,18)	50,54	70,35	98,83	48,28	33,13
<i>iam_{rb}</i>	25,23/607,30	91,15(72,17)	50,57	70,35	98,83	48,26	33,12
ohne Solvens, ohne H ($n = 1323$)							
<i>mul</i>	24,50/607,30	76,66(59,77)	44,44	59,04	82,01	37,57	24,72
<i>iam</i>	24,83/607,30	76,79(59,69)	44,54	59,37	82,19	37,65	24,92
<i>iam_{rb}</i>	25,23/607,30	76,81(59,68)	44,54	59,51	82,23	37,68	24,85
ohne Solvens, nur Hauptpopulation ($n = 1961$)							
<i>mul</i>	28,15/607,30	82,50(68,04)	47,75	63,41	88,67	40,92	27,78
<i>iam</i>	28,68/607,30	82,59(67,98)	47,89	63,58	88,67	40,78	27,79

<i>iam_{rb}</i>	28,66/607,30	82,60(67,98)	47,91	63,60	88,59	40,68	27,82
ohne Solvens, Hauptpopulation, exkl. H ($n = 1001$)							
<i>mul</i>	28,15/607,30	69,20(54,64)	42,38	54,31	72,60	30,22	20,21
<i>iam</i>	28,68/607,30	69,37(54,56)	42,56	54,53	72,67	30,11	20,16
<i>iam_{rb}</i>	28,66/607,30	69,40(54,56)	42,63	54,64	72,74	30,11	20,03
nur verfeinerte Atome ($n = 639$)							
<i>mul</i>	24,50/199,00	57,55(24,08)	41,20	51,46	66,04	24,85	17,89
<i>iam</i>	24,83/190,00	57,81(23,91)	41,48	51,98	66,52	25,04	17,80
<i>iam_{rb}</i>	25,23/195,10	57,86(23,91)	41,49	52,01	66,57	25,07	17,78
Differenzen der verfeinerten Atome							
<i>mul-iam</i>	-6,39/11,34	-0,27(0,79)	-0,49	-0,34	-0,16	0,33	0,24
<i>mul-iam_{rb}</i>	-6,15/9,44	-0,31(0,79)	-0,53	-0,36	-0,19	0,34	0,25

Tabelle B.3: Statistik der verfeinerten U_{eq} von P1LYSO- nach Element (alle Einheiten in 10^{-3}\AA^2).

	$x_{\min/\max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
verfeinerte C-Atome ($n = 385$)							
<i>mul</i>	24,50/165,50	55,53(20,81)	40,53	50,85	64,29	23,76	17,22
<i>iam</i>	24,83/171,90	55,91(20,89)	40,93	51,38	64,70	23,77	17,30
<i>iam_{rb}</i>	25,23/171,70	55,95(20,91)	40,92	51,51	64,80	23,88	17,55
<i>mul-iam</i>	-6,39/3,45	-0,38(0,66)	-0,51	-0,39	-0,24	0,27	0,20
<i>mul-iam_{rb}</i>	-6,15/3,21	-0,42(0,66)	-0,57	-0,42	-0,27	0,31	0,23
verfeinerte N-Atome ($n = 109$)							
<i>mul</i>	29,60/191,20	56,09(26,14)	38,45	48,38	64,29	25,84	18,62
<i>iam</i>	29,83/190,00	56,17(25,91)	38,71	48,68	64,71	25,99	18,89
<i>iam_{rb}</i>	29,86/181,70	56,12(25,46)	38,80	48,81	64,55	25,75	19,09
<i>mul-iam</i>	-1,97/3,38	-0,08(0,54)	-0,31	-0,16	0,01	0,31	0,22
<i>mul-iam_{rb}</i>	-1,92/9,44	-0,03(1,14)	-0,34	-0,23	-0,07	0,27	0,20
verfeinerte O-Atome ($n = 135$)							
<i>mul</i>	32,57/199,00	65,15(29,68)	44,19	55,26	75,58	31,39	18,64
<i>iam</i>	32,86/187,60	65,24(29,07)	44,44	55,47	76,49	32,05	18,53
<i>iam_{rb}</i>	32,95/195,10	65,36(29,34)	44,44	55,78	76,49	32,05	18,91
<i>mul-iam</i>	-2,86/11,34	-0,09(1,18)	-0,43	-0,27	-0,07	0,35	0,26
<i>mul-iam_{rb}</i>	-2,81/4,04	-0,20(0,77)	-0,49	-0,33	-0,12	0,37	0,26
verfeinerte S-Atome ($n = 10$)							
<i>mul</i>	35,17/68,94	48,42(12,18)	37,97	47,10	53,61	15,64	12,66
<i>iam</i>	35,70/69,63	48,99(12,24)	38,52	47,61	54,20	15,68	12,61
<i>iam_{rb}</i>	35,71/69,70	49,02(12,27)	38,53	47,65	54,22	15,69	12,66
<i>mul-iam</i>	-0,70/-0,51	-0,57(0,07)	-0,59	-0,54	-0,52	0,07	0,04
<i>mul-iam_{rb}</i>	-0,76/-0,52	-0,60(0,09)	-0,62	-0,57	-0,54	0,09	0,06

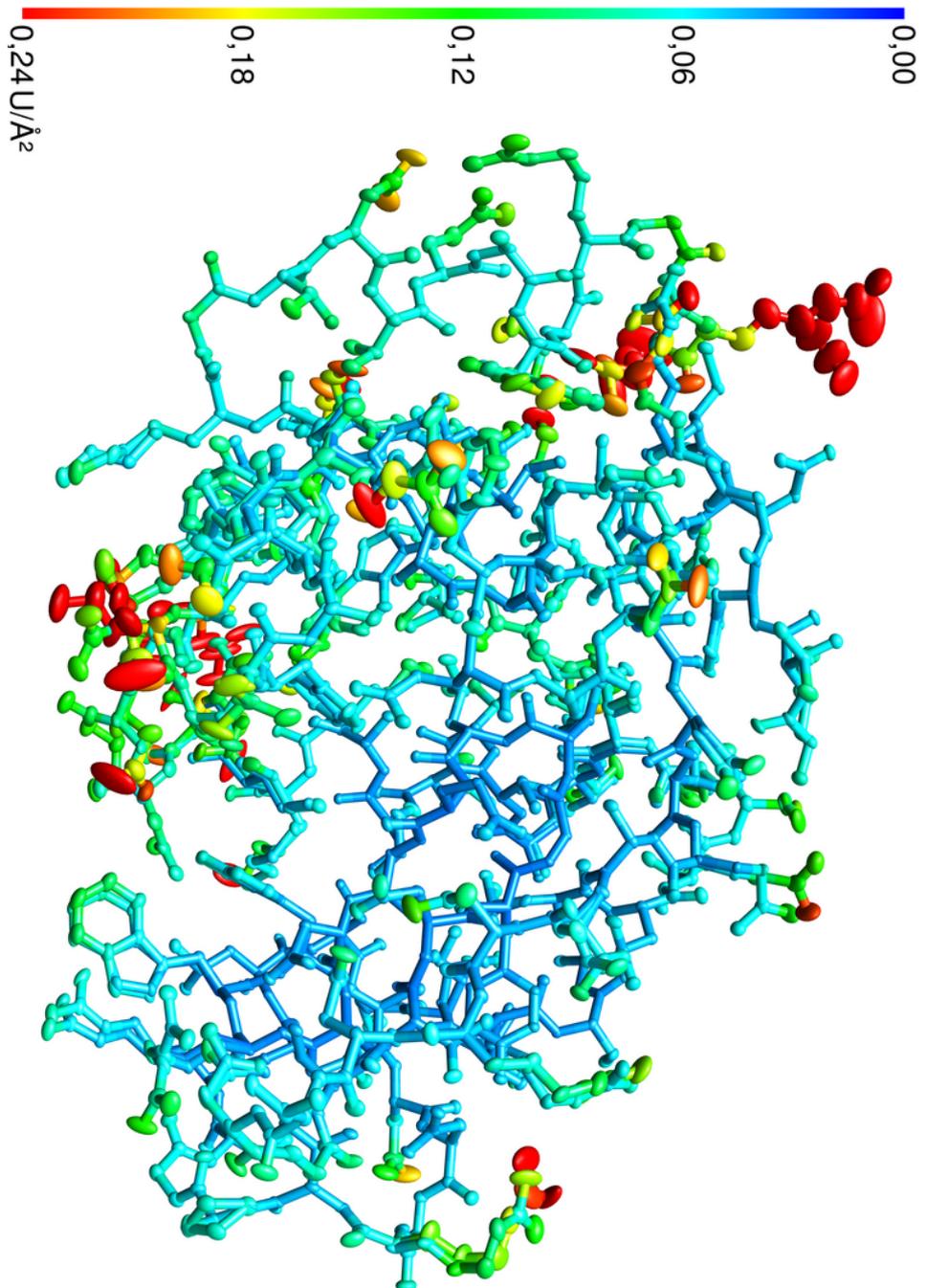


Abbildung B.1: Anisotrope Verschiebungsparameter P1LYSO nach asphärischer Verfeinerung – Farbcode: U_{eq} .

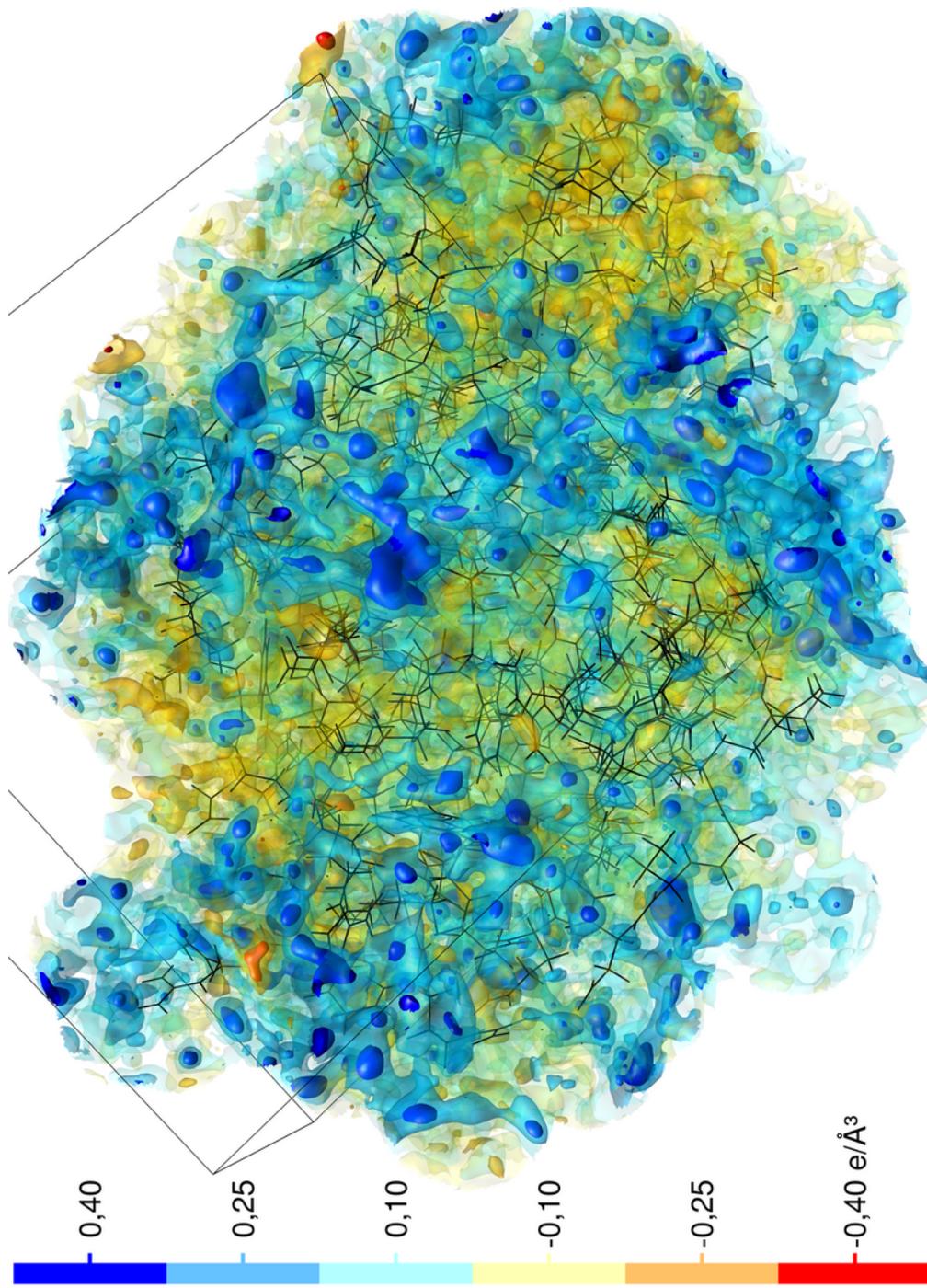


Abbildung B.2: Restdichten von PLYSO im Modell iam_{rB} , Darstellung durch sechs Isooberflächenwerten bis zu 4 \AA um die Struktur.

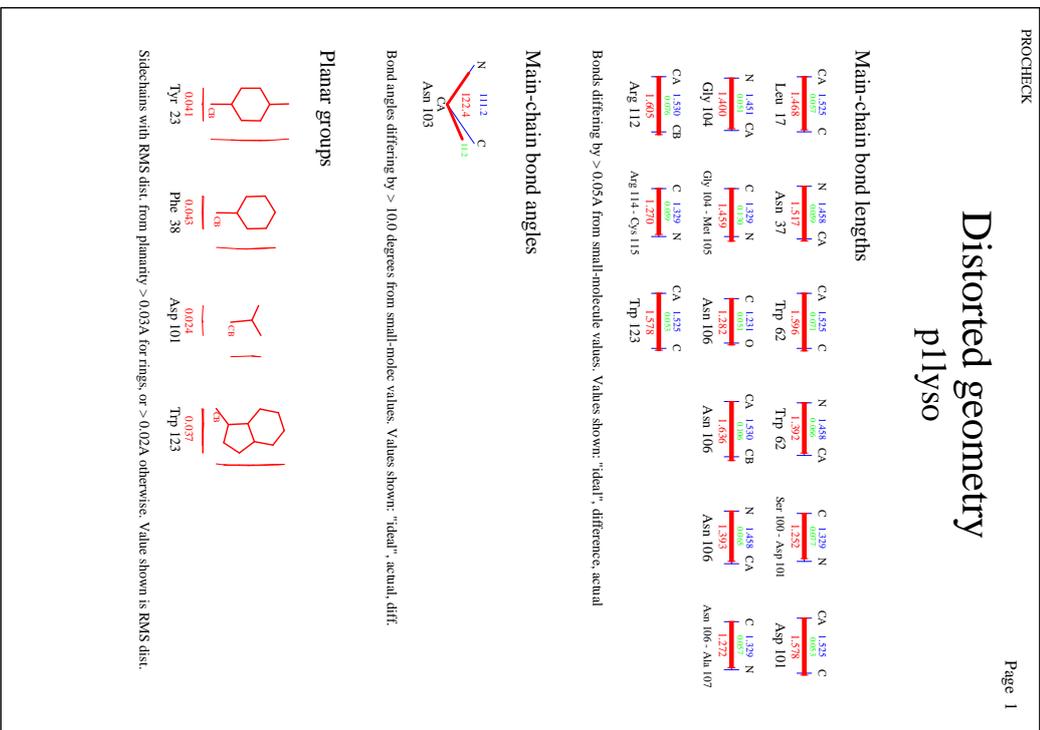
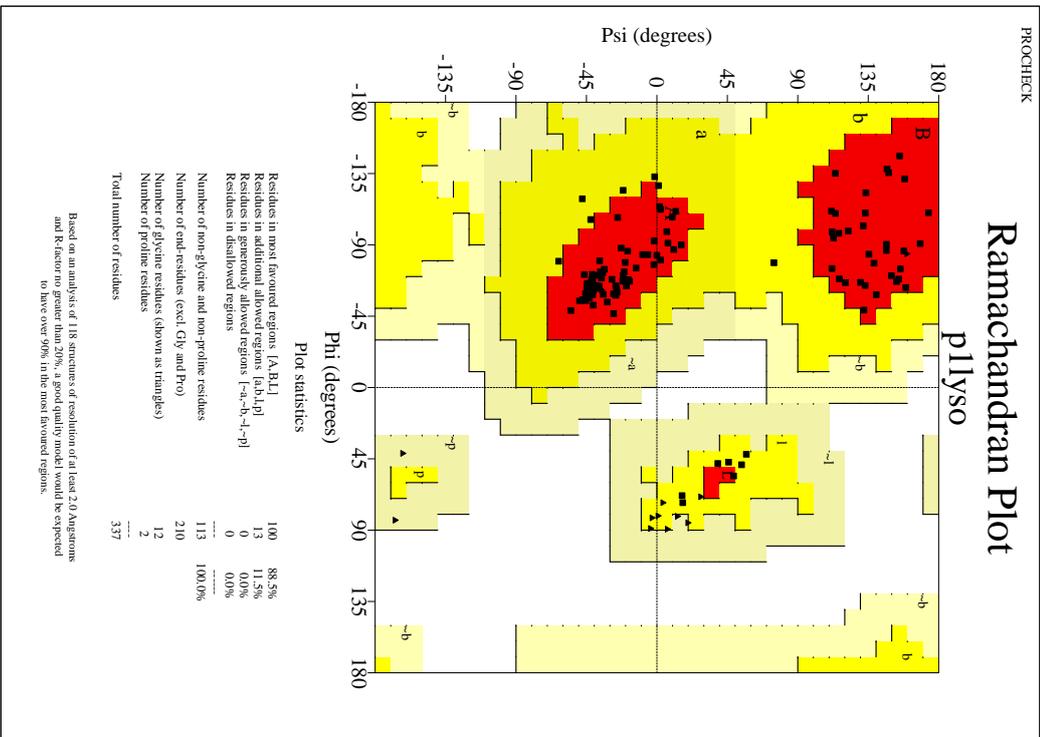


Abbildung B.3: PROCHECK-Ausgabe pIlyso.

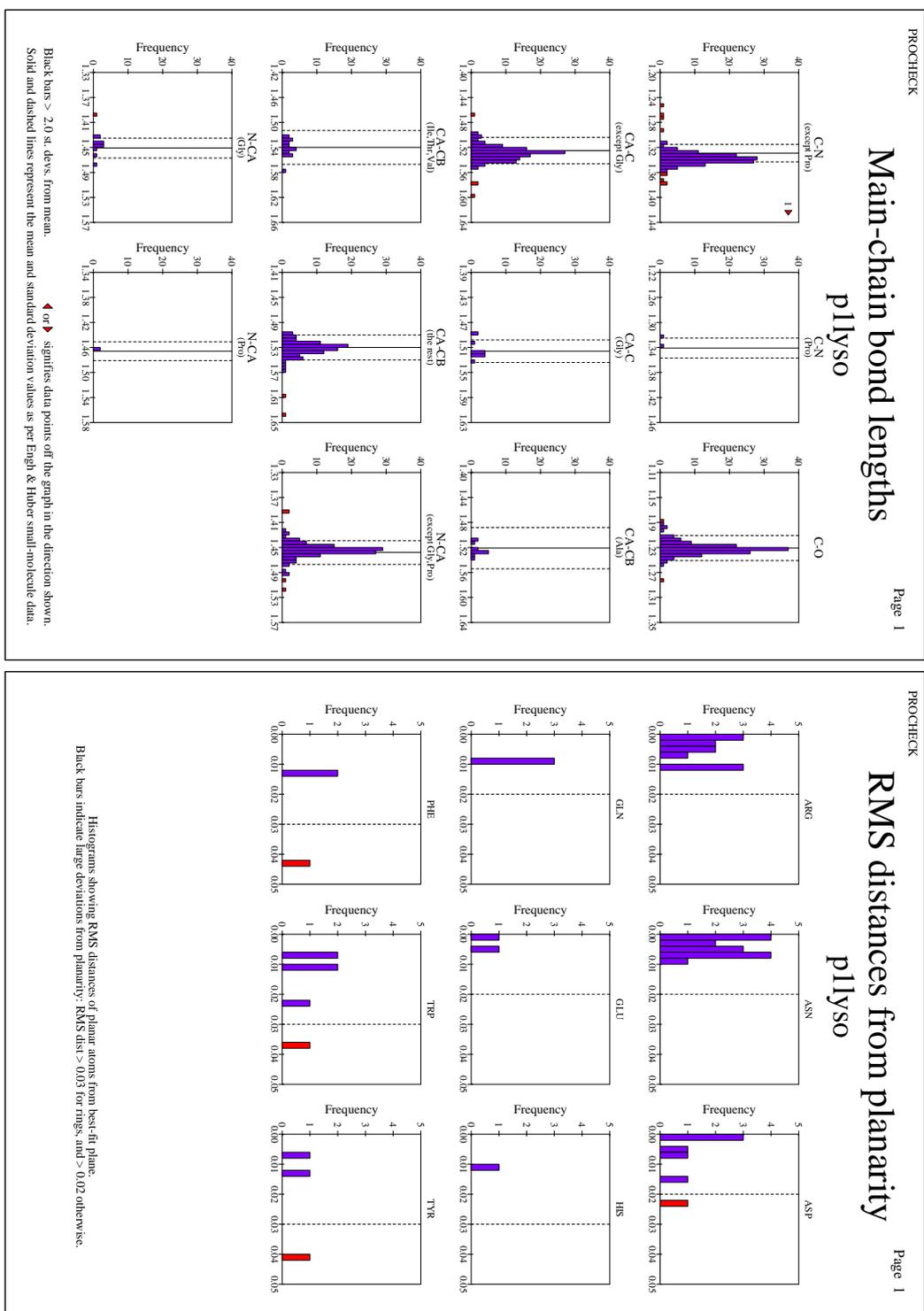


Abbildung B.3: PROCHECK-Ausgabe PLYSO (Forts.).

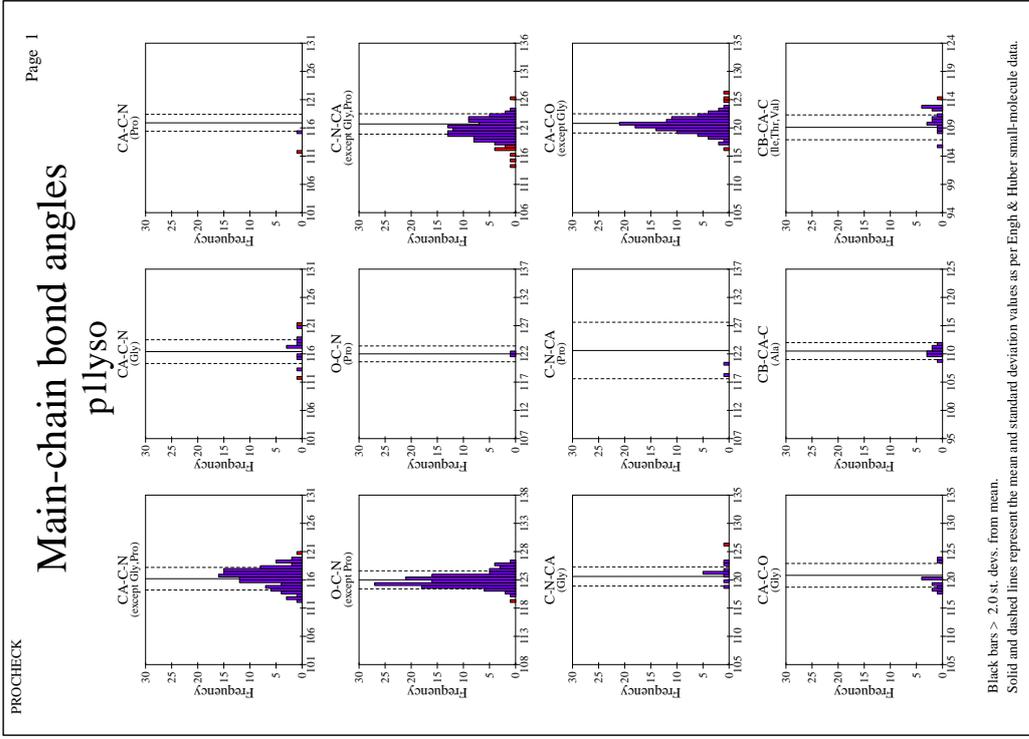
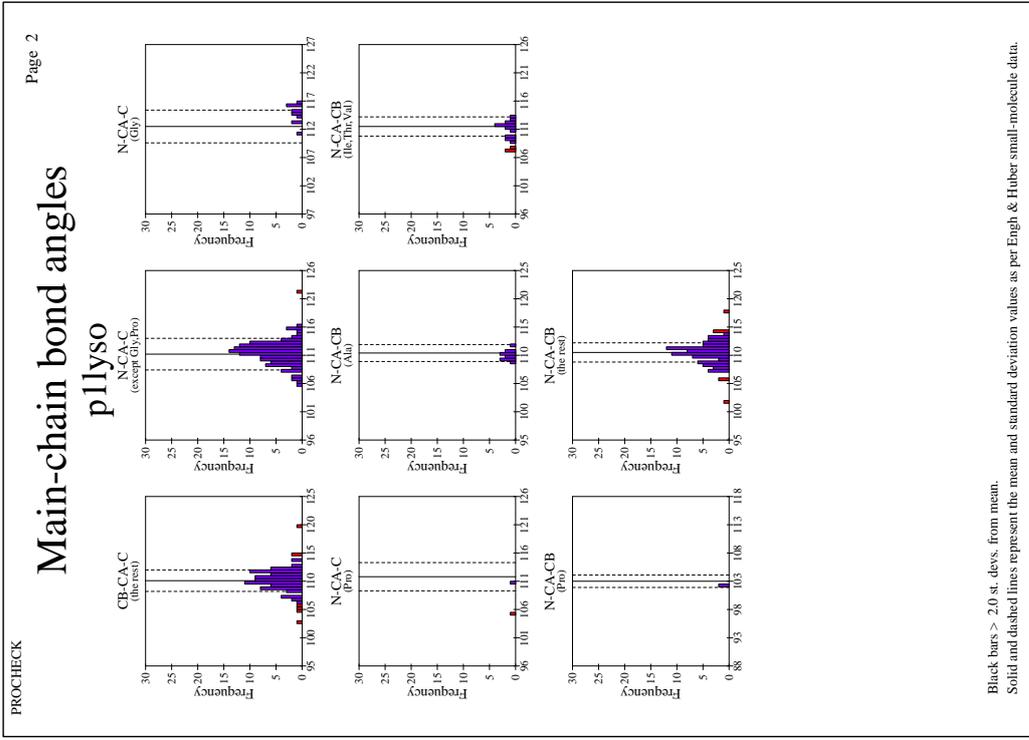


Abbildung B.3: PROCHECK-Ausgabe P1LYSO (Forts.).

1 p1lyso.pdb

1.1 Sect 0) Symmetry related problems

1.1.1 Warning: Triclinic cell with mixed acute and obtuse angles

The CRYST1 cell dimensions
 $A = 27.070$ $B = 31.250$ $C = 33.760$
 $\alpha = 87.980$ $\beta = 108.000$ $\gamma = 112.110$

1.1.2 Warning: Conventional cell

The CRYST1 cell dimensions
 $A = 27.070$ $B = 31.250$ $C = 33.760$
 $\alpha = 87.980$ $\beta = 108.000$ $\gamma = 112.110$

Dimensions of a reduced cell
 $A = 27.070$ $B = 31.250$ $C = 33.760$
 $\alpha = 87.980$ $\beta = 72.000$ $\gamma = 67.890$

Dimensions of the conventional cell
 $A = 27.070$ $B = 31.250$ $C = 33.760$
 $\alpha = 87.980$ $\beta = 72.000$ $\gamma = 67.890$

Transformation to conventional cell
 $\begin{bmatrix} 1.000000 & 0.000000 & 0.000000 \\ 0.000000 & -1.000000 & 0.000000 \\ 0.000000 & 0.000000 & -1.000000 \end{bmatrix}$

1.1.3 Note: Matthews coefficient OK

Molecular weight of all polymer chains: 14311.160
 Volume of the Unit Cell $V = 29057.057$
 Cell multiplicity: 1
 Matthews coefficient for observed atoms $V_{m} = 1.751$

1.2 Sect 3) Coordinate problems, unexpected atoms, B-factor and occupancy checks

1.2.1 Warning: Strange distribution of occupancy values

1.2.2 Warning: Occupancies atoms do not add up to 1.0.

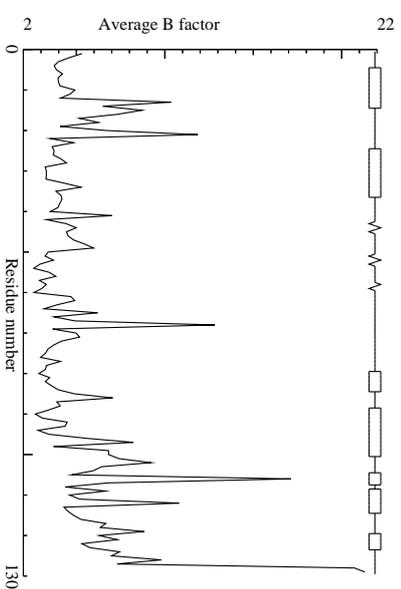
1.2.3 Warning: Average B-factor out of normal range

Average B-factor for buried atoms : 4.137

1.2.4 Warning: More than 5 percent of buried atoms has low B-factor

Percentage of buried atoms with B less than 5 : 81.25

1.2.5 Note: B-factor plot



1.3 Sect 5) Geometric checks

1.3.1 Warning: Unusual bond lengths

Residue	Atom pair	Distance	Z-value
15	HIS (15-) - CEE1	1.26	-4.4
65	ASN (65-) - CG	1.11	-6.1
68	ARG (68-) - CZ	1.43	5.9
105	MEP (105-) - N	1.46	6.5
106	ASN (106-) - CA	1.64	5.3
113	ASN (113-) - CG	1.32	4.3

1.3.2 Note: Normal bond length variability

RMS Z-score for bond lengths: 0.943

RMS-deviation in bond distances: 0.019

1.3.3 Note: No bond length directionality

1.3.4 Warning: Unusual bond angles

1

Abbildung B.4: Gekürzte WHATCHECK-Ausgabe P1LYSO.

2

Residue	Atom_Triplet	Bond Angle	Z-value
15 HIS (15-)	- CB CG ND1	139.17	11.7
15 HIS (15-)	- CB CG CD2	114.24	-11.4
21 ARG (21-)	- CD NE CZ	128.96	4.0
34 PHE (34-)	- CA CB CG	109.74	-4.1
37 ASN (37-)	- N CA CB	117.64	4.2
65 ASN (65-)	- CB CG ND2	109.74	-4.4
66 ASP (66-)	- CA CB CG	116.74	4.1
68 ARG (68-)	- CD NE CZ	148.45	14.9
68 ARG (68-)	- NE CZ NH1	127.60	4.0
68 ARG (68-)	- NE CZ NH2	109.89	-5.3
103 ASN (103-)	- CA CB CG	107.02	-5.6
106 ASN (106-)	- CA CB CG	106.81	-5.8
111 TRP (111-)	- C N CA	114.29	-4.1
112 ARG (112-)	- N CA CB	101.87	-5.1
112 ARG (112-)	- C CA CB	119.90	5.2
112 ARG (112-)	- CD NE CZ	134.60	7.2
112 ARG (112-)	- NE CZ NH1	109.85	-5.3
112 ARG (112-)	- NE CZ NH2	129.01	4.7
125 ARG (125-)	- CD NE CZ	129.76	4.5
128 ARG (128-)	- CG CD NE	118.09	4.5
128 ARG (128-)	- CD NE CZ	134.06	6.9

1.3.5 Note: Normal bond angle variability

RMS Z-score for bond angles: 1.172

RMS-deviation in bond angles: 2.150

1.3.6 Note: Chirality OK

All protein atoms have proper chirality.

1.3.7 Note: Improper dihedral angle distribution OK

Improper dihedral RMS Z-score : 0.964

1.3.8 Note: Side chain planarity OK

1.3.9 Note: Atoms connected to aromatic rings OK

1.4 Sect 6) Torsion-related checks

1.4.1 Note: PRO puckering amplitude OK

1.4.2 Note: PRO puckering phases OK

1.4.3 Note: Torsion angles OK

1.4.4 Warning: Backbone torsion angle evaluation shows unusual conformations

Residue	Description
3 PHE (3-)	- ω poor
19 ASN (19-)	- Poor ϕ/ψ
38 PHE (38-)	- Poor ϕ/ψ
53 TYR (53-)	- ω poor
62 TRP (62-)	- ω poor
77 ASN (77-)	- Poor ϕ/ψ
104 GLY (104-)	- Poor ϕ/ψ
105 MET (105-)	- ω poor

1.4.5 Note: Ramachandran Z-score OK

Ramachandran Z-score : -0.114

1.4.6 Note: Omega angle restraint OK

Standard deviation of ω values : 1.246

1.4.7 Note: chi-1/chi-2 angle correlation Z-score OK

χ -1/ χ -2 correlation Z-score : 0.884

1.4.8 Note: Backbone oxygen evaluation OK

1.4.9 Note: Rotamers checked OK

1.4.10 Warning: Unusual backbone conformations

Residue	# hits
38 PHE (38-)	1
103 ASN (103-)	1
105 MET (105-)	1
21 ARG (21-)	2
55 ILE (55-)	2
63 TRP (63-)	2

1.4.11 Note: Backbone conformation Z-score OK

Backbone conformation Z-score : 0.490

1.5 Sect 7) Bump checks

1.5.1 Error: Abnormally short interatomic distances

	Atom 1			Atom 2		Bump Dist	Status		
145	HOH (1176)	-	W	<-> 145	HOH (1179)	-	W	0.68 2.32 INTRA	BF
68	ARG (68)	-	A NH2	<-> 145	HOH (1054)	-	O	0.33 2.37 INTRA	
145	HOH (1153)	-	O	<-> 145	HOH (1176)	-	W	0.28 2.32 INTRA	BF
129	LEU (129)	-	N	<-> 145	HOH (1191)	-	W	0.25 2.85 INTRA	
145	HOH (1087)	-	O	<-> 145	HOH (1183)	-	W	0.24 2.36 INTRA	
145	HOH (1152)	-	O	<-> 145	HOH (1195)	-	W	0.23 2.57 INTRA	
121	GLN (121)	-	NE2	<-> 145	HOH (1172)	-	W	0.22 2.88 INTRA	
145	HOH (1174)	-	W	<-> 145	HOH (1183)	-	W	0.21 2.79 INTRA	BF
145	HOH (1087)	-	O	<-> 145	HOH (1174)	-	W	0.20 2.40 INTRA	
145	HOH (1065)	-	A NE	<-> 145	HOH (1174)	-	W	0.19 2.41 INTRA	
5	ARG (5)	-	A NE	<-> 145	HOH (1171)	-	W	0.18 2.92 INTRA	
15	HIS (15)	-	A CE1	<-> 138	NO3 (403)	-	O2	0.16 2.64 INTRA	
101	ASP (101)	-	OD1	<-> 145	HOH (1177)	-	W	0.16 2.64 INTRA	BF
145	HOH (1170)	-	O	<-> 145	HOH (1192)	-	W	0.16 2.44 INTRA	
14	ARG (14)	-	A NH1	<-> 145	HOH (1138)	-	O	0.16 2.54 INTRA	
64	CYS (64)	-	C	<-> 80	CYS (80)	-	SG	0.15 3.25 INTRA	
43	THR (43)	-	A O	<-> 145	HOH (1165)	-	O	0.14 2.26 INTRA	BF
145	HOH (1189)	-	W	<-> 145	HOH (1187)	-	W	0.14 2.86 INTRA	BF
145	HOH (1089)	-	O	<-> 145	HOH (1185)	-	W	0.12 2.48 INTRA	
145	HOH (1166)	-	O	<-> 145	HOH (1181)	-	W	0.10 2.50 INTRA	BF
15	HIS (15)	-	A CE1	<-> 145	HOH (1089)	-	O	0.09 2.71 INTRA	BF
129	LEU (129)	-	CD1	<-> 145	HOH (1190)	-	W	0.07 3.13 INTRA	BF
138	NO3 (403)	-	O3	<-> 145	HOH (1137)	-	O	0.07 2.33 INTRA	
30	CYS (30)	-	C	<-> 145	CYS (115)	-	SG	0.06 3.54 INTRA	
113	ASN (113)	-	A ND2	<-> 145	HOH (1160)	-	O	0.05 2.65 INTRA	
103	ASN (103)	-	A O	<-> 145	HOH (1157)	-	O	0.03 2.37 INTRA	
73	ARG (73)	-	A NE	<-> 145	HOH (1041)	-	O	0.02 2.28 INTRA	
114	ARG (114)	-	A C	<-> 115	CYS (115)	-	CA	0.01 2.69 INTRA	B2
33	LYS (33)	-	NZ	<-> 37	ASN (37)	-	A OD1	0.01 2.69 INTRA	
5	ARG (5)	-	A NH2	<-> 145	HOH (1183)	-	W	0.00 3.10 INTRA	

1.6 Sect 8) Packing, accessibility and threading

1.6.1 Note: Inside/Outside residue distribution normal

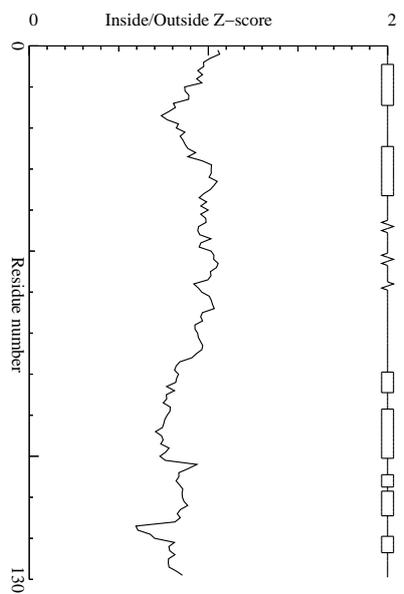
inside/outside RMS Z-score : 0.894

1.6.2 Note: Inside/Outside RMS Z-score plot

The Inside/Outside distribution normality RMS Z-score over a 15 residue window is plotted as function of the residue number. High areas in the plot (above 1.5) indicate unusual inside/outside patterns.

5

Abbildung B.4: Gekürzte WHATCHECK-Ausgabe PLYSO (Forts.).



1.6.3 Note: Second generation packing environment OK

1.6.4 Warning: Abnormal packing Z-score for sequential residues

Start residue 82 ALA (82) - - - 85 SER (85) - - -1.97
Av. Z-score

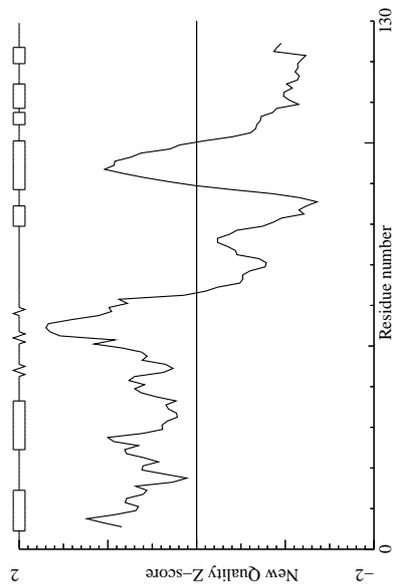
1.6.5 Note: Structural average packing Z-score OK

All contacts : Average = 0.027 Z-score = 0.08
BB-BB contacts : Average = -0.079 Z-score = -0.54
BB-SC contacts : Average = 0.088 Z-score = 0.59
SC-BB contacts : Average = 0.217 Z-score = 1.35
SC-SC contacts : Average = -0.009 Z-score = 0.07

1.6.6 Note: Second generation quality Z-score plot

The second generation quality Z-score smoothed over a 10 residue window is plotted as function of the residue number. Low areas in the plot (below -1.3) indicate "unusual" packing.

6



Residue	Atom
101 ASP (101-) -	OD2
106 ASN (106-) -	A ODI
113 ASN (113-) -	A ODI

1.8 Final summary

1.8.1 Note: Summary report for users of a structure

This is an overall summary of the quality of the structure as compared with current reliable structures. This summary is most useful for biologists seeking a good structure to use for modelling calculations.

The second part of the table mostly gives an impression of how well the model conforms to common refinement constraint values. The first part of the table shows a number of constraint-independent quality indicators.

Structure Z-scores, positive is better than average:

1st generation packing quality : -1.478

2nd generation packing quality : 0.083

Ramachandran plot appearance : -0.114

χ^1/χ^2 rotamer normality : 0.884

Backbone conformation : 0.400

RMS Z-scores, should be close to 1.0:

Bond lengths : 0.943

Bond angles : 1.172

Omega angle restraints : 0.227 (tight)

Side chain planarity : 1.014

Improper dihedral distribution : 0.964

Inside/Outside distribution : 0.894

1.7 Sect 9) Water and hydrogenbond related checks

1.7.1 Note: Water contacts OK

1.7.2 Note: Water hydrogen bonds OK

1.7.3 Error: HIS, ASN, GLN side chain flips

Residue
41 GLN (41-) -

1.7.4 Note: Histidine type assignments

Residue	Geometry	Z-score	H-bond	Z-score
15 HIS (15-) -	HIS-E	2.38	HIS-D	2.81

1.7.5 Warning: Buried unsatisfied hydrogen bond donors

Residue	Atom
45 ARG (45-) -	A NH2
104 GLY (104-) -	N

1.7.6 Warning: Buried unsatisfied hydrogen bond acceptors

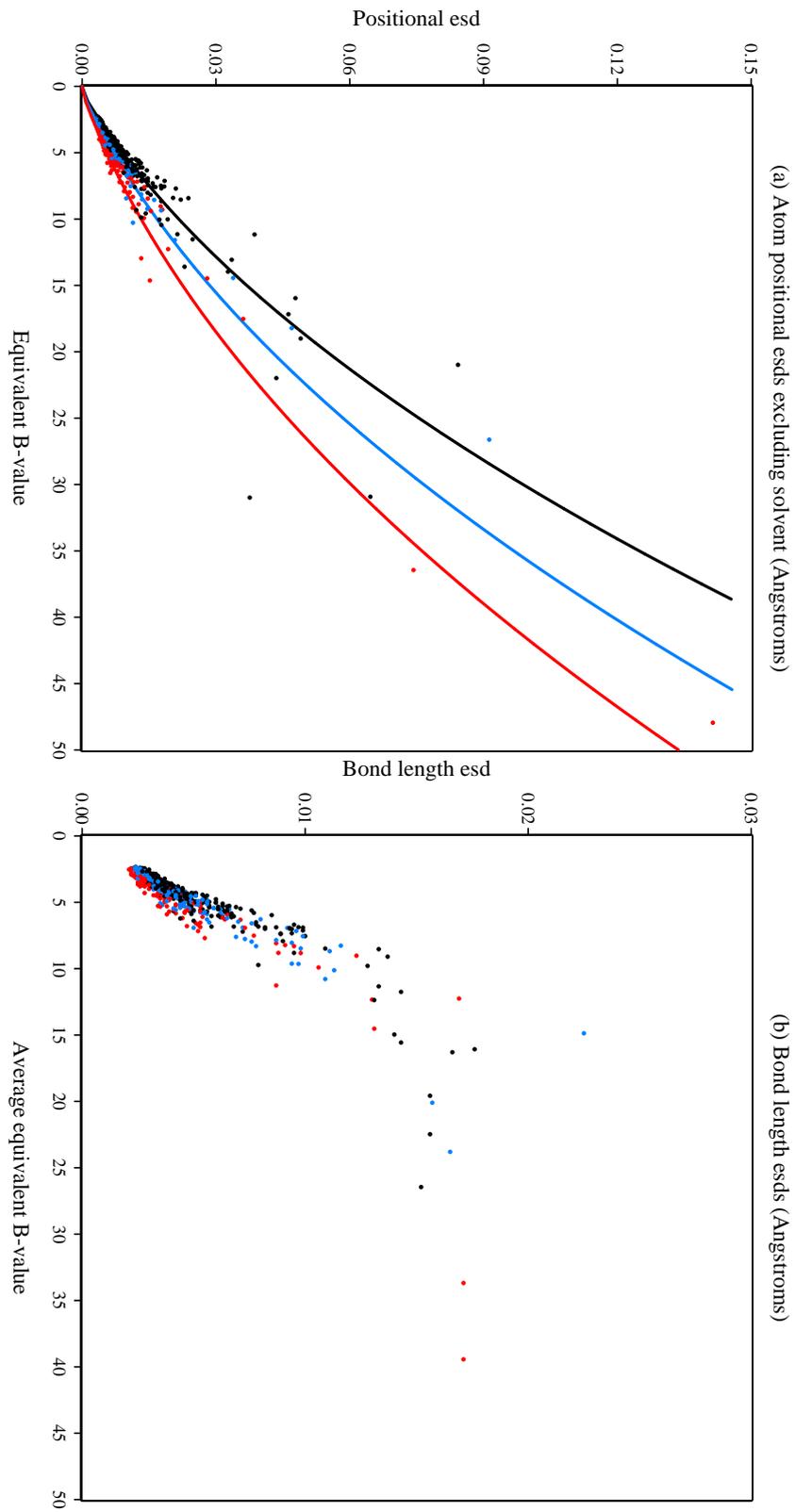


Abbildung B.5: SHELXPRO-Ausgabe PLIYSO.

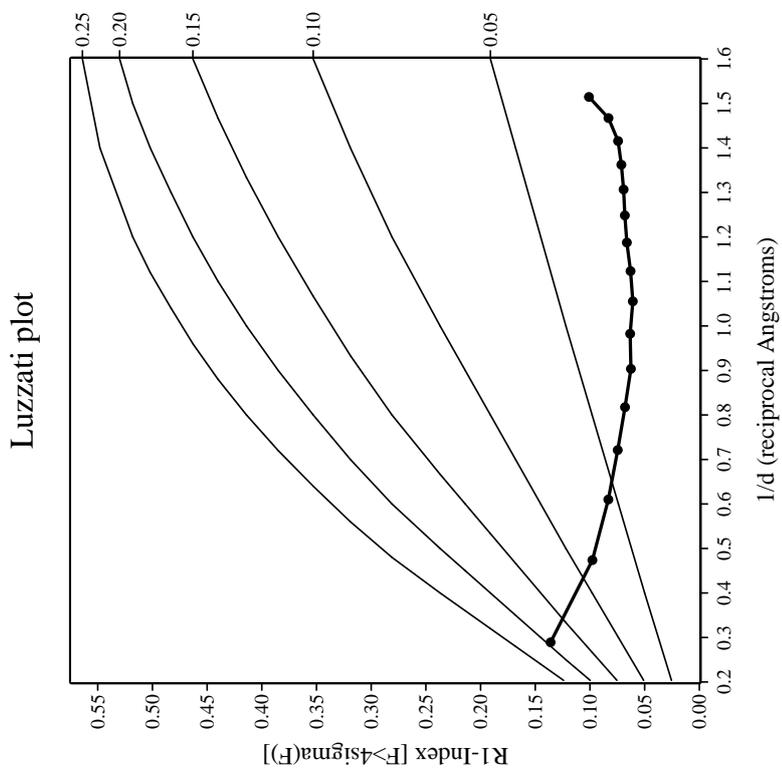
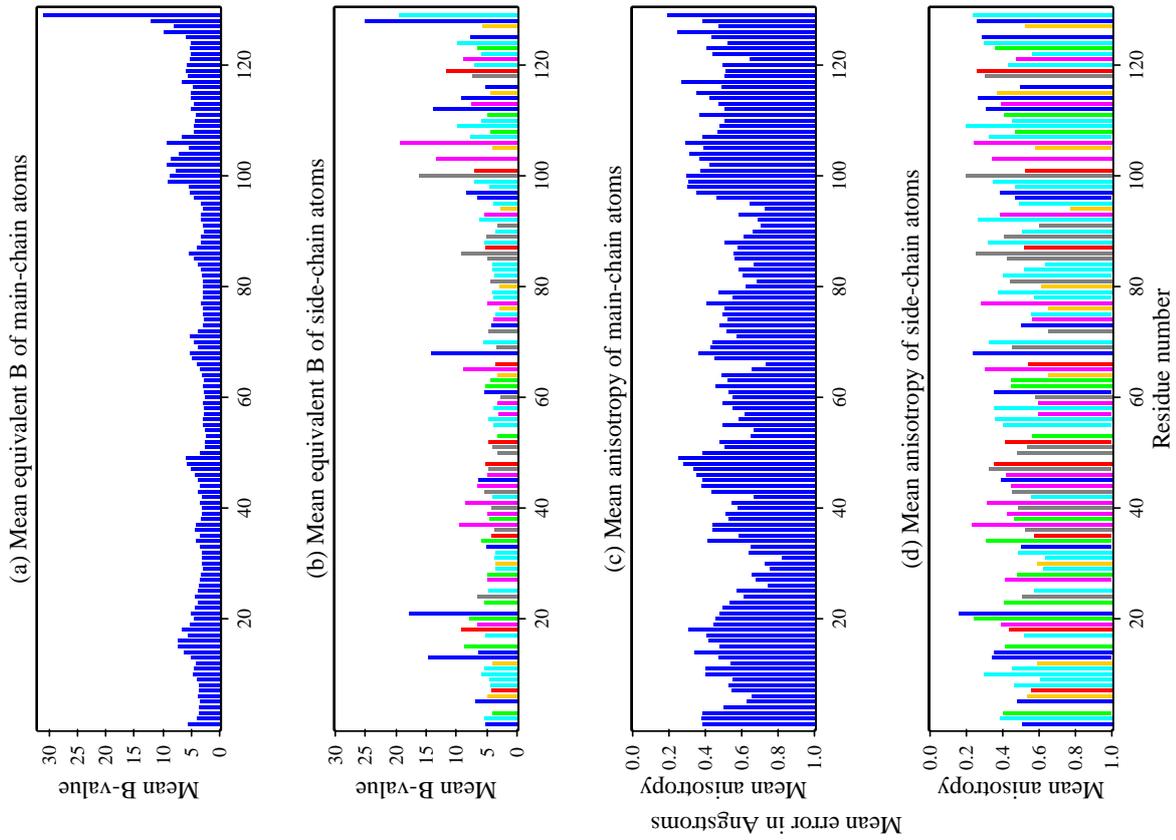


Abbildung B.5: SHELXPRO-Ausgabe P1LYSO (Forts.).

C Tetragonales Lysozym

Tabelle C.1: Statistik der U_{eq} verschiedener Teilmengen von TETRALYSO. ([†]Minimum, Maximum und Mittelwert mit Standardabweichung der U_{eq} und ΔU_{eq} ; [‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-3}Å^2)

	$x_{\text{min}/\text{max}}^{\dagger}$	$\bar{x}(\sigma_x)^{\dagger}$	$Q1^{\ddagger}$	$Q2^{\ddagger}$	$Q3^{\ddagger}$	IQR^{\ddagger}	MAD^{\ddagger}
alle Atome ($n = 2393$)							
<i>mul</i>	57,73/2000,0	163,10(106,28)	96,24	133,40	194,20	97,95	64,23
<i>iam</i>	58,95/2000,0	163,40(106,07)	96,68	133,50	194,20	97,56	63,81
<i>iam_{rb}</i>	59,02/2000,0	163,50(106,04)	96,65	133,80	194,20	97,54	64,06
ohne Solvens ($n = 2104$)							
<i>mul</i>	57,73/544,20	147,00(74,97)	92,86	126,00	177,80	84,97	56,13
<i>iam</i>	58,95/544,20	147,40(74,71)	93,42	126,10	177,90	84,44	55,81
<i>iam_{rb}</i>	59,02/544,20	147,40(74,68)	93,69	126,10	177,90	84,16	55,67
ohne Solvens, ohne H ($n = 1070$)							
<i>mul</i>	57,73/449,20	125,00(58,24)	82,37	106,80	149,10	66,68	43,10
<i>iam</i>	58,95/449,20	125,70(57,84)	83,61	107,80	149,50	65,88	42,82
<i>iam_{rb}</i>	59,02/449,20	125,80(57,80)	83,85	108,00	149,50	65,66	42,88
ohne Solvens, nur Hauptpopulation ($n = 1961$)							
<i>mul</i>	57,73/544,20	144,60(73,84)	91,99	123,50	173,80	81,81	54,09
<i>iam</i>	58,95/544,20	144,90(73,57)	92,53	123,90	173,80	81,24	54,03
<i>iam_{rb}</i>	59,22/544,20	145,00(73,53)	92,62	123,90	173,80	81,18	53,98
ohne Solvens, Hauptpopulation, exkl. H ($n = 1001$)							
<i>mul</i>	57,73/449,20	123,00(57,23)	81,75	105,80	145,40	63,68	41,81
<i>iam</i>	58,95/449,20	123,80(56,83)	82,76	106,70	145,80	63,07	41,19
<i>iam_{rb}</i>	59,22/449,20	123,90(56,78)	83,13	106,60	145,80	62,65	40,97
nur verfeinerte Atome ($n = 751$)							
<i>mul</i>	57,73/370,00	103,90(35,44)	77,02	96,18	122,20	45,13	31,11
<i>iam</i>	58,95/379,40	104,90(35,12)	78,40	97,23	122,30	43,91	30,89
<i>iam_{rb}</i>	59,22/385,70	105,00(35,10)	78,72	97,16	122,20	43,52	30,47
Differenzen der verfeinerten Atome							
<i>mul-iam</i>	-9,46/10,49	-1,02(1,22)	-1,63	-1,18	-0,64	0,99	0,73
<i>mul-iam_{rb}</i>	-15,74/19,12	-1,17(1,61)	-1,89	-1,33	-0,72	1,16	0,86

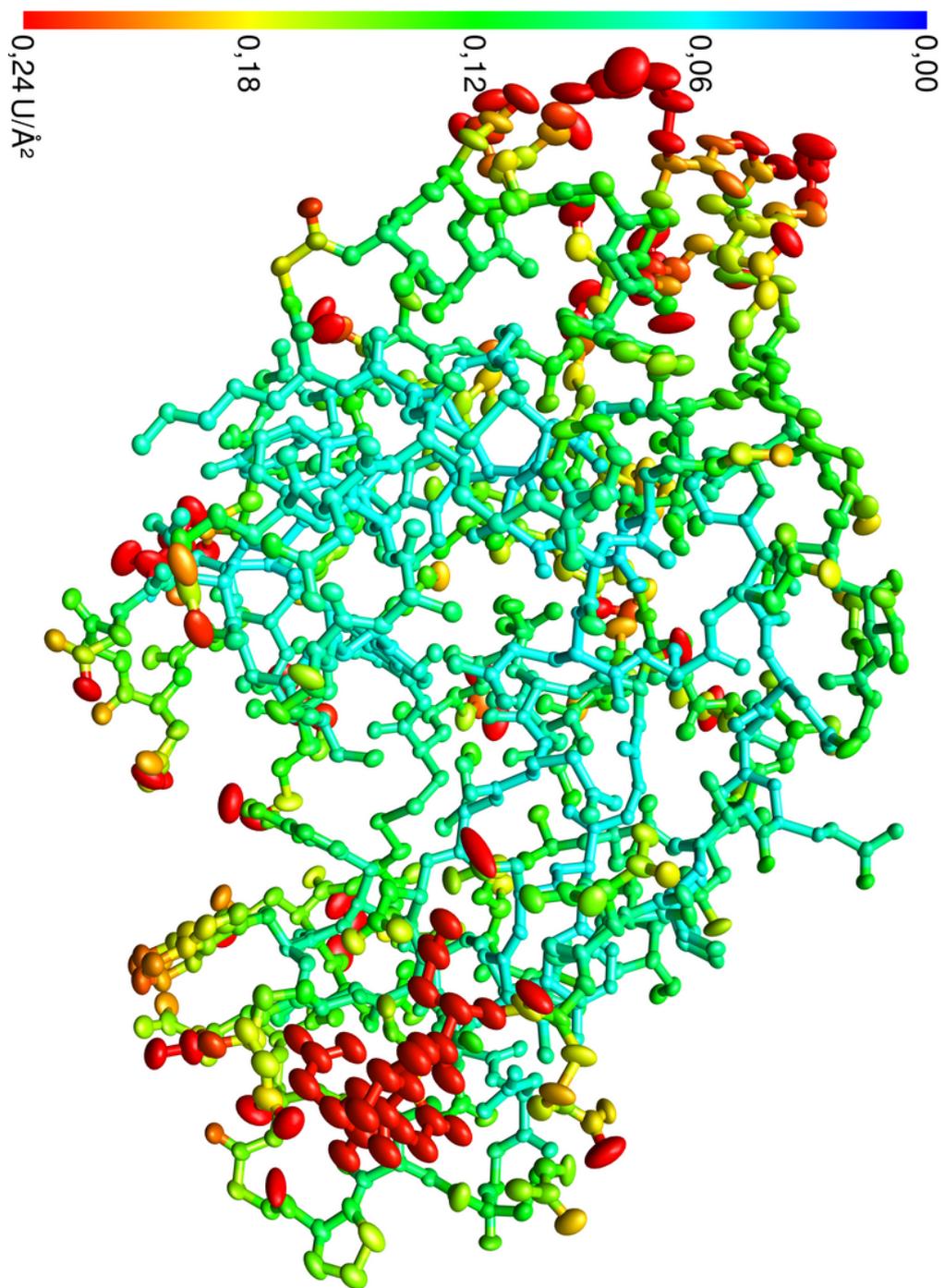


Abbildung C.1: Anisotrope Verschiebungsparameter der Nichtwasserstoffatome von TETRALYSSO nach asphärischer Verfeinerung – Farbcode: U_{eq}.

Tabelle C.2: Statistik verfeinerter U_{eq} von TETRALYSO– nach Elementen. ([†]Minimum, Maximum und Mittelwert mit Standardabweichung der U_{eq} und ΔU_{eq} ; [‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-3}\AA^2 .)

	$x_{\min/\max}^{\dagger}$	$\bar{x}(\sigma_x)^{\dagger}$	$Q1^{\ddagger}$	$Q2^{\ddagger}$	$Q3^{\ddagger}$	IQR^{\ddagger}	MAD^{\ddagger}
verfeinerte C-Atome ($n = 461$)							
<i>mul</i>	57,73/328,50	100,90(31,18)	75,17	94,75	119,50	44,33	31,32
<i>iam</i>	58,95/320,80	102,10(30,73)	76,84	95,72	119,90	43,09	30,96
<i>iam_{rb}</i>	59,22/322,00	102,20(30,58)	76,84	95,98	120,60	43,77	30,69
<i>mul-iam</i>	-5,61/7,71	-1,21(1,11)	-1,77	-1,41	-0,92	0,85	0,62
<i>mul-iam_{rb}</i>	-4,47/6,49	-1,34(1,25)	-2,04	-1,60	-0,96	1,08	0,73
verfeinerte N-Atome ($n = 132$)							
<i>mul</i>	59,40/268,50	101,50(34,17)	76,61	94,29	115,10	38,49	27,24
<i>iam</i>	60,67/258,00	101,90(33,48)	77,26	95,21	115,00	37,74	27,48
<i>iam_{rb}</i>	60,92/249,30	101,80(32,88)	77,40	95,24	115,30	37,95	27,15
<i>mul-iam</i>	-3,09/10,49	-0,34(1,40)	-0,92	-0,64	-0,13	0,79	0,61
<i>mul-iam_{rb}</i>	-2,58/19,12	-0,24(2,05)	-1,10	-0,70	-0,01	1,09	0,72
verfeinerte O-Atome ($n = 148$)							
<i>mul</i>	64,32/370,00	115,60(45,74)	83,39	102,70	135,10	51,70	35,48
<i>iam</i>	65,54/379,40	116,60(45,91)	84,49	103,20	136,40	51,89	34,84
<i>iam_{rb}</i>	65,69/385,70	117,00(46,43)	84,61	103,40	137,40	52,75	34,57
<i>mul-iam</i>	-9,46/4,48	-0,98(1,22)	-1,35	-1,03	-0,63	0,72	0,52
<i>mul-iam_{rb}</i>	-15,74/8,37	-1,36(1,91)	-1,54	-1,27	-0,94	0,61	0,47
verfeinerte S-Atome ($n = 10$)							
<i>mul</i>	69,40/159,70	99,31(30,92)	79,52	86,74	105,90	26,35	14,84
<i>iam</i>	71,17/162,10	101,40(31,07)	81,52	88,66	108,10	26,61	15,18
<i>iam_{rb}</i>	71,38/162,60	101,60(31,17)	81,64	88,92	108,50	26,89	15,22
<i>mul-iam</i>	-2,40/-1,77	-2,06(0,21)	-2,21	-1,98	-1,92	0,28	0,17
<i>mul-iam_{rb}</i>	-2,89/-1,98	-2,33(0,31)	-2,56	-2,19	-2,10	0,46	0,24

Tabelle C.3: Statistik der DMSDA-Beträge der verfeinerten Bindungen der Hauptpopulation von TETRALYSO und deren Differenzen – nach Elementen. (†Minimum, Maximum und Mittelwert mit Standardabweichung der DMSDA-Beträge; ‡0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-4}\AA^2 .)

	$x_{min/max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
C–C ($n = 373$)							
<i>mul</i>	0/941	70,33(94,52)	19,00	40,00	83,00	64,00	41,51
<i>iam</i>	0/973	70,29(97,66)	16,00	39,00	84,00	68,00	41,51
<i>iam_{rb}</i>	0/947	69,91(98,21)	15,00	42,00	82,00	67,00	45,96
<i>mul-iam</i>	–136/115	0,04(29,36)	–13,00	0,00	15,00	28,00	22,24
<i>mul-iam_{rb}</i>	–197/135	0,43(33,51)	–13,00	1,00	17,00	30,00	22,24
C–N ($n = 258$)							
<i>mul</i>	1/623	69,55(93,27)	19,00	37,00	75,00	56,00	35,58
<i>iam</i>	0/589	69,22(88,10)	20,25	40,00	83,75	63,50	40,03
<i>iam_{rb}</i>	0/629	70,57(90,29)	19,00	41,50	80,00	61,00	40,77
<i>mul-iam</i>	–146/146	0,33(30,32)	–16,00	–1,00	15,00	31,00	23,72
<i>mul-iam_{rb}</i>	–167/159	–1,01(33,40)	–17,00	–3,00	13,75	30,75	23,72
C–O ($n = 148$)							
<i>mul</i>	0/437	86,92(95,57)	19,00	47,00	113,00	94,00	49,67
<i>iam</i>	0/381	91,05(92,07)	27,75	57,00	125,20	97,50	54,11
<i>iam_{rb}</i>	0/396	90,98(91,13)	25,00	60,50	127,00	102,00	61,53
<i>mul-iam</i>	–118/114	–4,13(33,41)	–18,00	–3,00	9,00	27,00	20,02
<i>mul-iam_{rb}</i>	–136/166	–4,06(37,77)	–17,25	–3,00	10,00	27,25	20,02
C–S ($n = 12$)							
<i>mul</i>	0/148	43,83(40,91)	16,00	36,00	62,00	46,00	40,03
<i>iam</i>	3/107	36,17(29,48)	15,00	31,00	50,00	35,00	28,17
<i>iam_{rb}</i>	5/104	36,92(29,43)	12,50	33,50	52,75	40,25	31,88
<i>mul-iam</i>	–23/41	7,67(18,51)	–5,00	6,50	18,00	23,00	17,05
<i>mul-iam_{rb}</i>	–31/44	6,92(23,77)	–12,25	6,00	25,25	37,50	28,17
S–S ($n = 4$)							
<i>mul</i>	8/57	37,25(21,75)	27,50	42,00	51,75	24,25	17,05
<i>iam</i>	3/65	39,25(27,96)	24,75	44,50	59,00	34,25	24,46
<i>iam_{rb}</i>	3/64	38,25(27,94)	22,50	43,00	58,75	36,25	25,95
<i>mul-iam</i>	–15/5	–2,00(8,91)	–3,75	1,00	2,75	6,50	3,71
<i>mul-iam_{rb}</i>	–14/5	–1,00(8,98)	–3,50	2,50	5,00	8,50	3,71
alle ($n = 795$)							
<i>mul</i>	0/941	72,60(93,70)	19,00	40,00	84,00	65,00	40,03
<i>iam</i>	0/973	73,14(93,06)	18,50	42,00	91,00	72,50	43,00
<i>iam_{rb}</i>	0/947	73,39(93,82)	17,00	44,00	87,50	70,50	45,96
<i>mul-iam</i>	–146/146	–0,54(30,28)	–15,00	–1,00	15,00	30,00	22,24
<i>mul-iam_{rb}</i>	–197/166	–0,78(34,10)	–15,00	–1,00	15,00	30,00	22,24

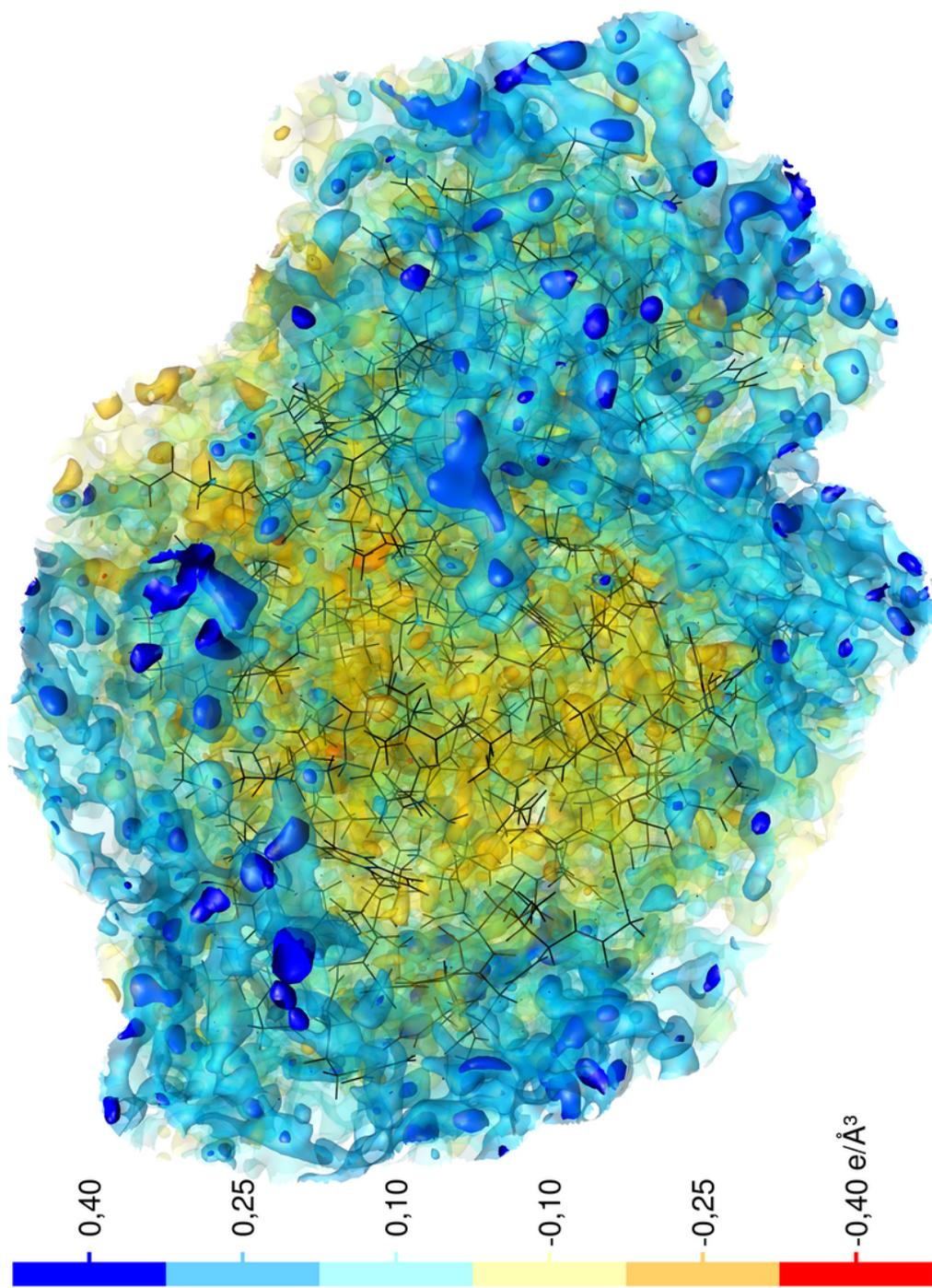


Abbildung C.2: Restdichten im Modell *iam_{7b}*, 4 Å um TETRALYSO.

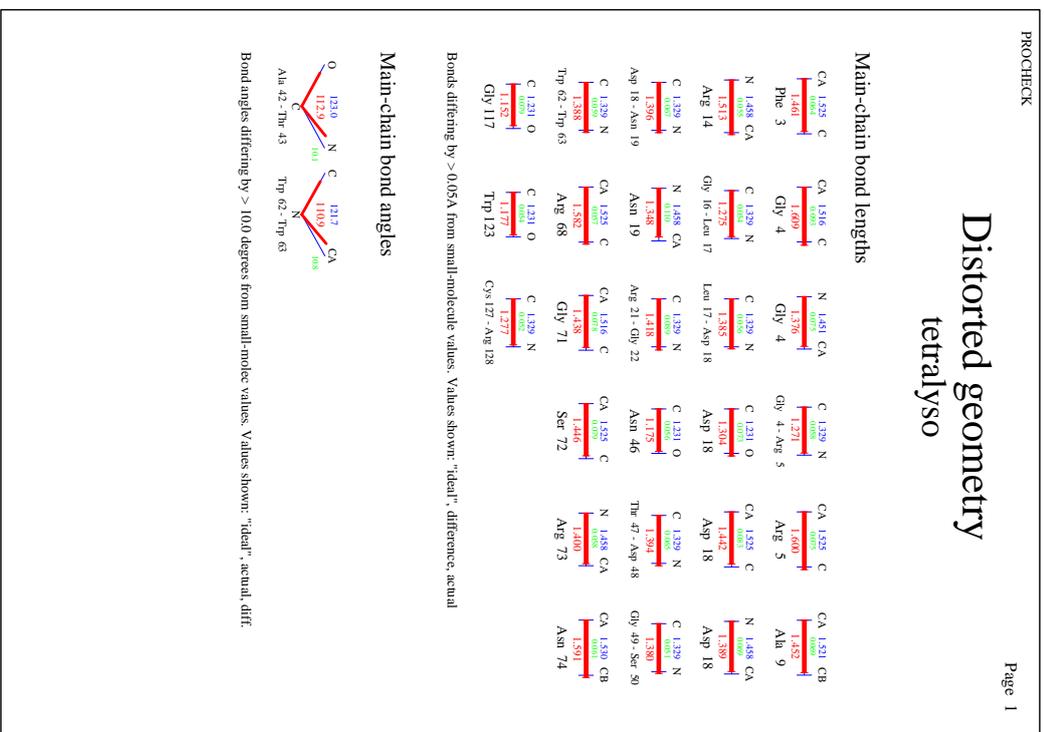
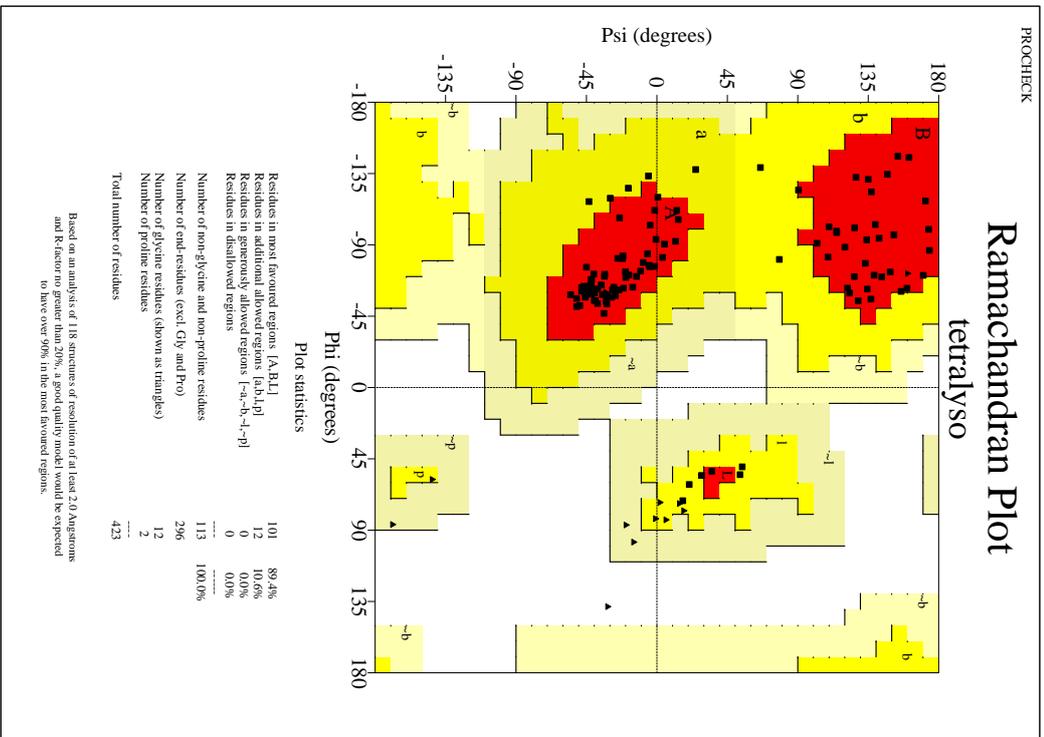


Abbildung C.3: PROCHECK-Ausgabe TETRALYSO.

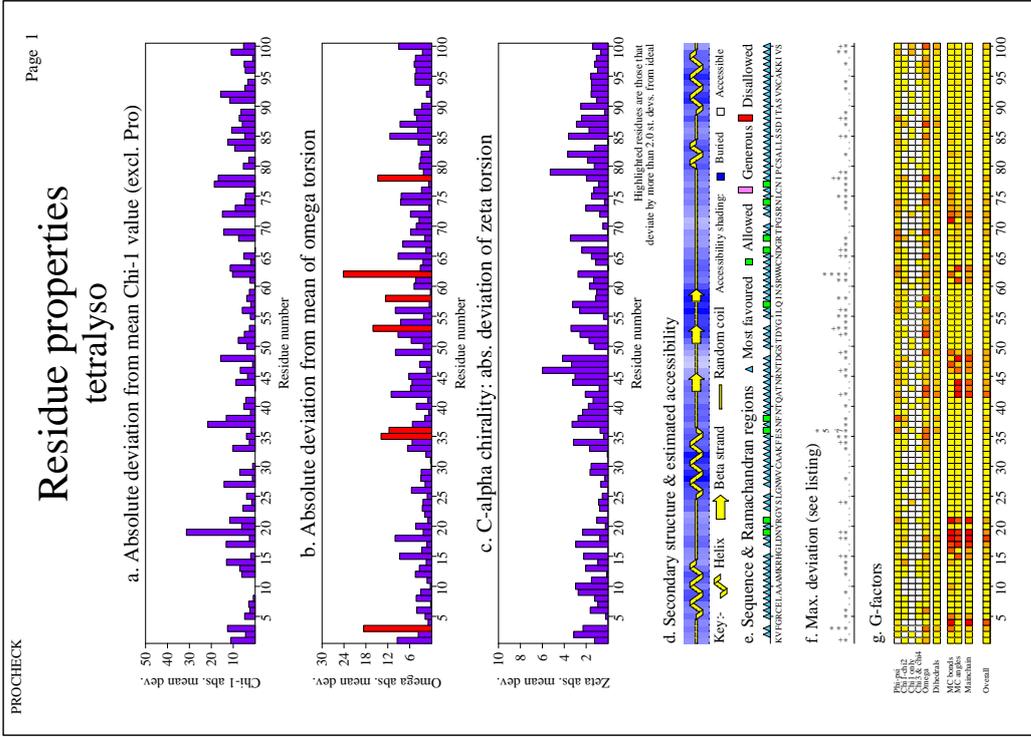
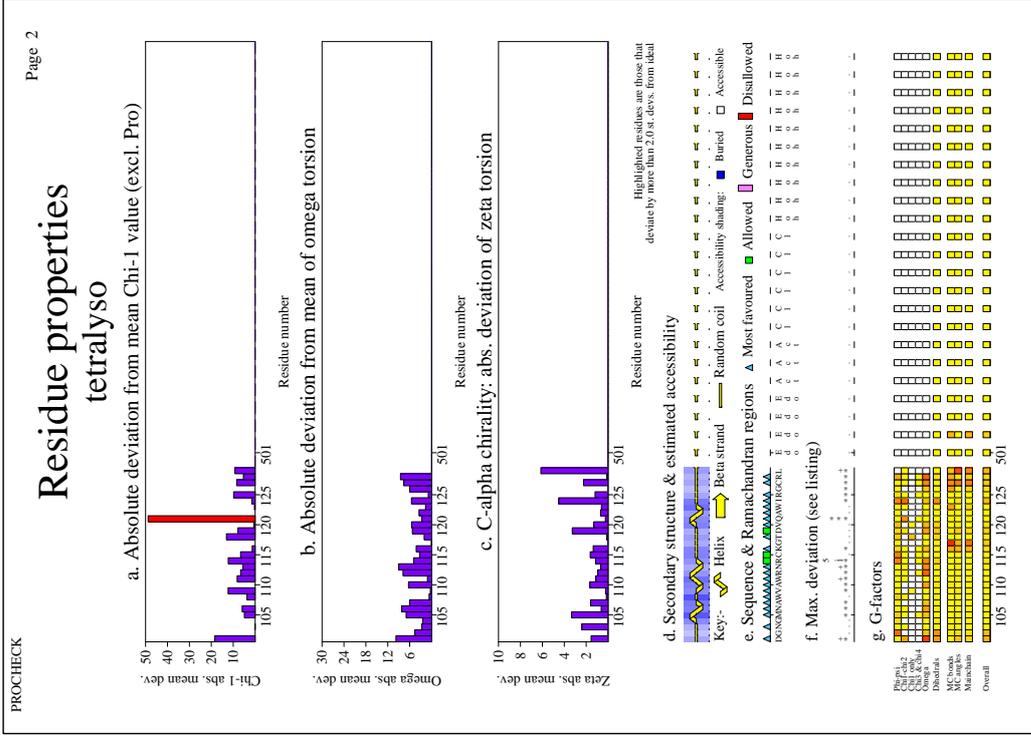


Abbildung C.3: PROCHECK-Ausgabe TETRALYSO (Forts.).

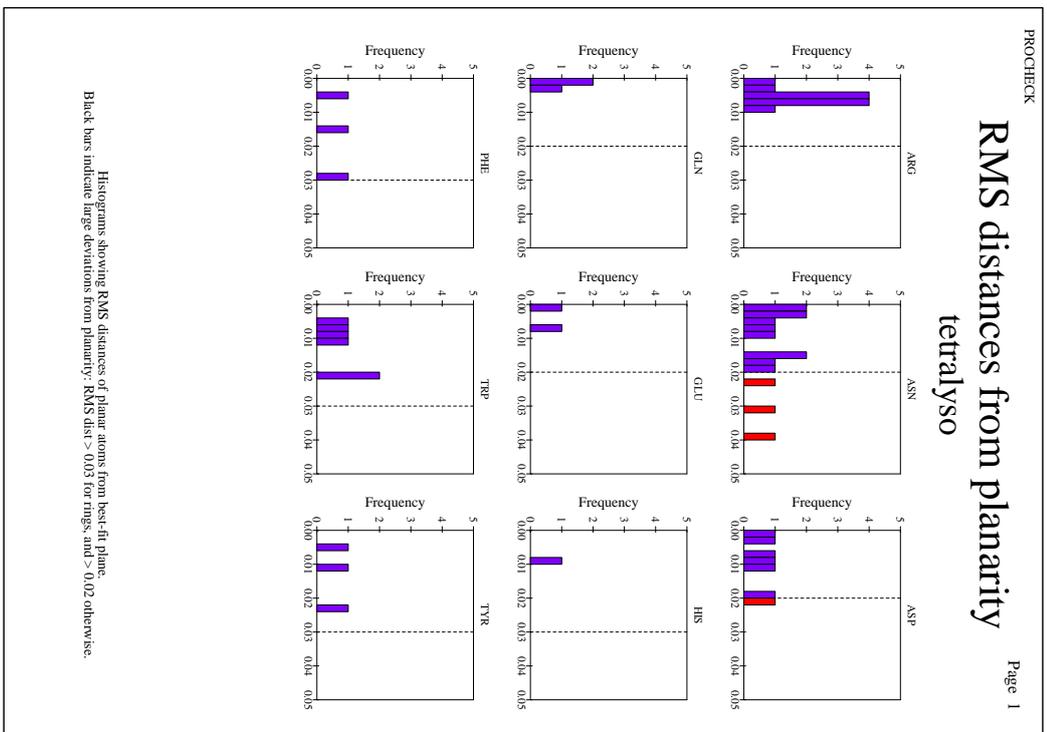
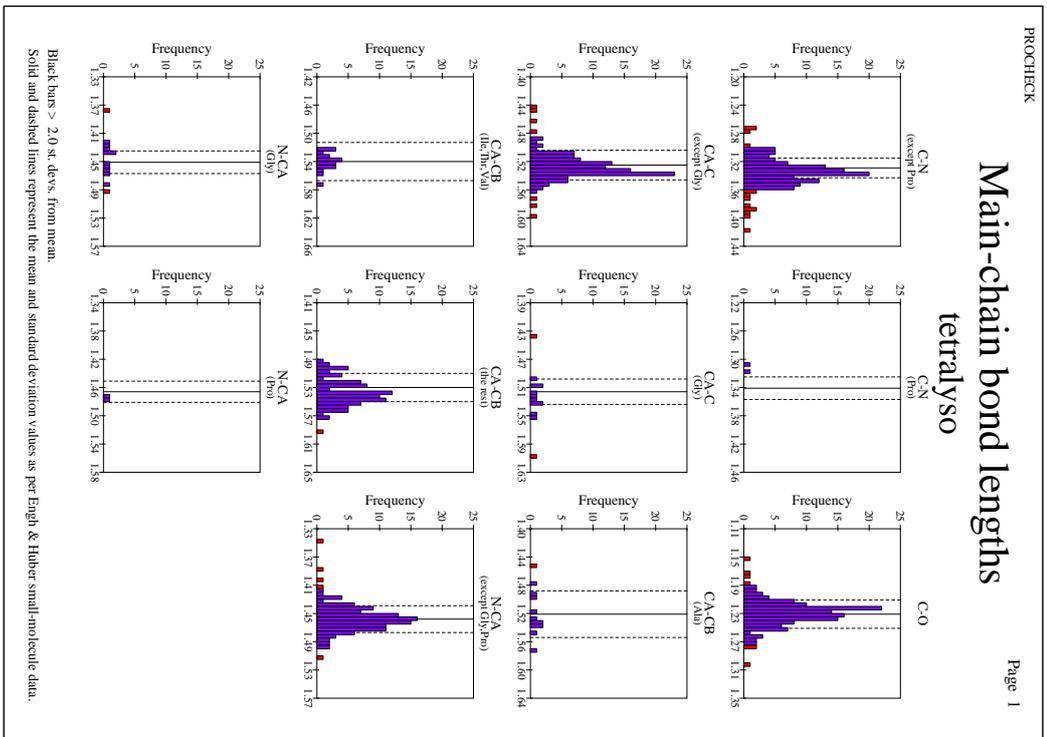


Abbildung C.3: PROCHECK-Ausgabe TETRALYSO (Forts.).

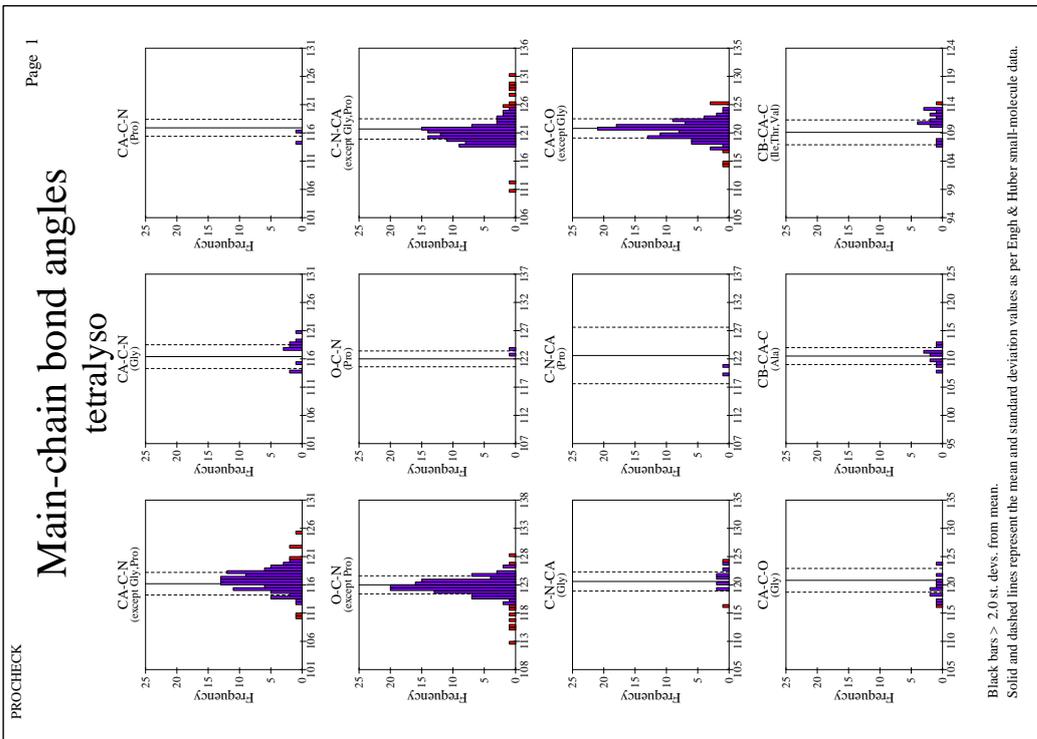
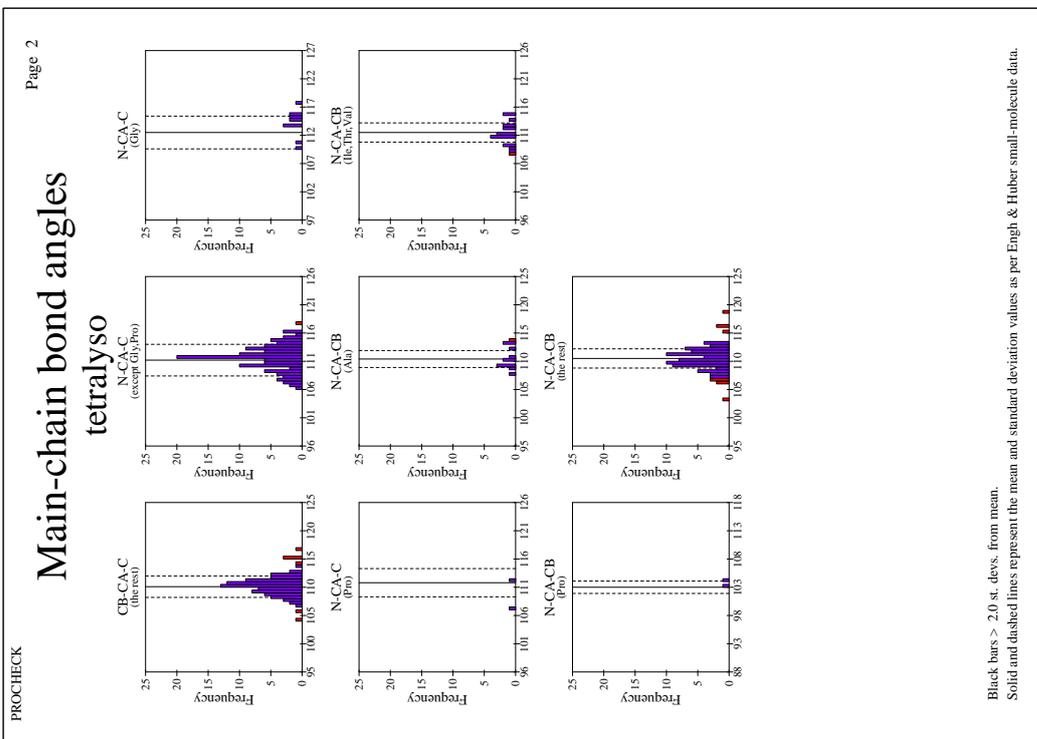


Abbildung C.3: PROCHECK-Ausgabe TETRALYSO (Forts.).

1 tetralyso.pdb

1.1 Sect 0) Symmetry related problems

1.1.1 Note: Matthews coefficient OK

Molecular weight of all polymer chains: 14507.880
 Volume of the Unit Cell V = 229846.641
 Cell multiplicity: 8
 Matthews coefficient for observed atoms V_m = 1.980

1.1.2 Error: Atoms too close to symmetry axes

Residue	Atom
134	EDO(504) - - - Cl

1.2 Sect 3) Coordinate problems, unexpected atoms, B-factor and occupancy checks

1.2.1 Note: Normal distribution of occupancy values

1.2.2 Warning: Occupancies atoms do not add up to 1.0.

1.2.3 Warning: Average B-factor out of normal range

Average B-factor for buried atoms : 7.759

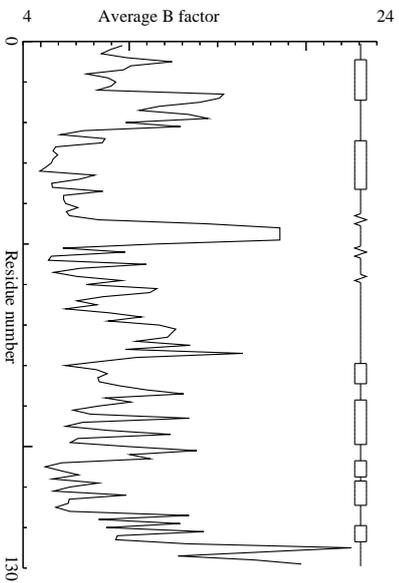
1.2.4 Warning: More than 2 percent of buried atoms has low B-factor

Percentage of buried atoms with B less than 5 : 2.03

1.2.5 Note: B-factor distribution normal

RMS Z-score : 0.769 over 892 bonds
 Average difference in B over a bond : 1.17
 RMS difference in B over a bond : 2.00

1.2.6 Note: B-factor plot



1.3 Sect 5) Geometric checks

1.3.1 Warning: Unusual bond lengths

Residue	Atom pair	Distance	Z-value
4	GLY (4) - N CA	1.38	-4.7
4	GLY (4) - CA C	1.61	6.2
14	ARG (14) - NE CZ	1.27	-4.0
15	HIS (15) - CE1 NE2	1.25	-4.8
19	ASN (19) - N CA	1.35	-5.8
22	GLY (22) - N -C	1.42	4.4
35	GLU (35) - CD OE2	1.33	4.5
45	ARG (45) - CZ NH2	1.44	6.5
71	GLY (71) - CA C	1.44	4.4
89	THR (89) - CB OG1	1.36	-4.6

1.3.2 Note: Normal bond length variability

RMS Z-score for bond lengths: 1.162

RMS-deviation in bond distances: 0.024

1.3.3 Note: No bond length directionality

1.3.4 Warning: Unusual bond angles

1

Abbildung C.4: Gekürzte WHATCHECK-Ausgabe TETRALYSSO.

2

Residue	Atom Triplet			Bond Angle	Z-value
5 ARG (5-) - - CG CD NE	117.78			4.3	
5 ARG (5-) - - CD NE CZ	129.30			4.2	
15 HIS (15-) - - CG ND1 CE1	111.95			6.3	
15 HIS (15-) - - ND1 CE1 NE2	106.37			-4.1	
18 ASP (18-) - - O -C N	115.77			-4.5	
19 ASN (19-) - - O -C N	115.09			-4.9	
19 ASN (19-) - - C N CA	129.00			4.1	
19 ASN (19-) - - CA CB CG	105.56			-7.0	
19 ASN (19-) - - ND2 CG OD1	116.76			-5.8	
21 ARG (21-) - - NH1 CZ NH2	126.91			4.0	
43 THR (43-) - - O -C N	112.95			-6.3	
43 THR (43-) - - CA -C N	125.30			4.5	
43 THR (43-) - - C N CA	131.41			5.4	
44 ASN (44-) - - C N CA	129.62			4.4	
44 ASN (44-) - - N CA CB	118.66			4.8	
45 ARG (45-) - - CD NE CZ	134.89			7.3	
46 ASN (46-) - - N CA CB	108.11			-4.3	
46 ASN (46-) - - CB CG ND2	122.54			4.1	
46 ASN (46-) - - ND2 CG OD1	112.71			-9.9	
48 ASP (48-) - - CA CB CG	121.41			8.8	
48 ASP (48-) - - CB CG OD2	135.80			7.6	
48 ASP (48-) - - OD2 CG OD1	97.76			-10.5	
49 GLY (49-) - - O -C N	116.58			-4.0	
55 ILE (55-) - - CG2 CB CG1	128.45			5.9	
59 ASN (59-) - - CA CB CG	119.02			6.4	
61 ARG (61-) - - C N CA	112.38			-5.2	
61 ARG (61-) - - CD NE CZ	130.49			4.9	
61 ARG (61-) - - NE CZ NH1	130.01			5.3	
63 TRP (63-) - - C N CA	110.92			-6.0	
66 ASP (66-) - - CA CB CG	116.62			4.0	
87 ASP (87-) - - CA CB CG	117.55			5.0	
124 ILE (124-) - - CB CG1 CD1	126.04			5.8	
128 ARG (128-) - - CB CG CD	105.36			-4.4	

1.3.5 Note: Normal bond angle variability

RMS Z-score for bond angles: 1.451

RMS-deviation in bond angles: 2.573

1.3.6 Warning: Chirality deviations detected

Residue	Atom	Z-value	Value	Expected	Notes
55 ILE (55-) - - CB	-18.6	7.40	32.30		
117 GLY (117-) - - C	6.9	9.08	0.09		

1.3.7 Note: Improper dihedral angle distribution OK

Improper dihedral RMS Z-score : 1.662

1.3.8 Error: Side chain planarity problems

Residue	Z-score
59 ASN (59-) - -	6.69
44 ASN (44-) - -	5.22

1.3.9 Note: Atoms connected to aromatic rings OK

1.4 Sect 6) Torsion-related checks

1.4.1 Note: PRO puckering amplitude OK

1.4.2 Note: PRO puckering phases OK

1.4.3 Warning: Torsion angle evaluation shows unusual residues

Residue	Z-score
38 PHE (38-) - -	-2.5
68 ARG (68-) - -	-2.2
124 ILE (124-) - -	-2.1

1.4.4 Warning: Backbone torsion angle evaluation shows unusual conformations

Residue	Description
3 PHE (3-) - - ω poor	
19 ASN (19-) - - Poor ϕ/ψ	
21 ARG (21-) - - Poor ϕ/ψ	
38 PHE (38-) - - Poor ϕ/ψ	
53 TYR (53-) - - ω poor	
62 TRP (62-) - - ω poor	
114 ARG (114-) - - Poor ϕ/ψ	

1.4.5 Note: Ramachandran Z-score OK

Ramachandran Z-score : -0.322

1.4.6 Note: Omega angle restraint OK

Standard deviation of ω values : 1.124

1.4.7 Note: chi-1/chi-2 angle correlation Z-score OK

χ^{-1}/χ^{-2} correlation Z-score : -0.370

1.4.8 Note: Backbone oxygen evaluation OK

1.4.9 Note: Rotamers checked OK

1.4.10 Warning: Unusual backbone conformations

Residue	# H-b
101 ASP (101) -	0
103 ASN (103) -	0
105 MET (105) -	0
108 TRP (108) -	0
38 PHE (38) -	1
68 ARG (68) -	1
21 ARG (21) -	2
55 ILE (55) -	2
63 TRP (63) -	2

1.4.1.1 Note: Backbone conformation Z-score OK

Backbone conformation Z-score : 0.208

1.5 Sect 7) Bump checks

1.5.1 Error: Abnormally short interatomic distances

Atom 1	Atom 2	Bump Dist	Status
61 ARG (61) -	A NH2 <-> 144 HOH (1207) -	W 0.71 2.39	INTRA
144 HOH (1206) -	W <-> 144 HOH (1277) -	W 0.69 2.31	INTRA
144 HOH (1226) -	W <-> 144 HOH (1276) -	W 0.69 2.31	INTRA
144 HOH (1265) -	W <-> 144 HOH (1275) -	W 0.65 2.35	INTRA
144 HOH (1221) -	W <-> 144 HOH (1267) -	W 0.62 2.38	INTRA
46 ASN (46) -	ND2 <-> 144 HOH (1268) -	W 0.61 2.49	INTRA
144 HOH (1027) -	O <-> 144 HOH (1193) -	O 0.59 1.61	INTRA
144 HOH (1230) -	W <-> 144 HOH (1235) -	W 0.59 2.41	INTRA
144 HOH (1088) -	O <-> 144 HOH (1198) -	O 0.58 1.62	INTRA
144 HOH (1234) -	W <-> 144 HOH (1258) -	W 0.57 2.43	INTRA
135 ACT (505) -	A CH3 <-> 144 HOH (1181) -	O 0.57 2.23	INTRA
144 HOH (1253) -	W <-> 144 HOH (1269) -	W 0.57 2.43	INTRA
144 HOH (1151) -	O <-> 144 HOH (1156) -	O 0.49 1.71	INTRA
121 GLN (121) -	A CD <-> 144 HOH (1170) -	O 0.48 2.32	INTRA
133 EDO (503) -	C2 <-> 144 HOH (1119) -	O 0.44 2.36	INTRA
21 ARG (21) -	A NH2 <-> 144 HOH (1240) -	W 0.42 2.68	INTRA
144 HOH (1100) -	O <-> 144 HOH (1174) -	O 0.41 1.79	INTRA
121 GLN (121) -	A NE2 <-> 144 HOH (1225) -	W 0.40 2.70	INTRA
121 GLN (121) -	A OE1 <-> 144 HOH (1170) -	O 0.38 2.02	INTRA
144 HOH (1210) -	W <-> 144 HOH (1276) -	W 0.36 2.64	INTRA
144 HOH (1210) -	W <-> 144 HOH (1226) -	W 0.36 2.64	INTRA
144 HOH (1250) -	W <-> 144 HOH (1252) -	W 0.35 2.65	INTRA
47 THR (47) -	N <-> 144 HOH (1174) -	O 0.34 2.36	INTRA
144 HOH (1185) -	O <-> 144 HOH (1251) -	W 0.28 2.32	INTRA
144 HOH (1110) -	O <-> 144 HOH (1266) -	W 0.26 2.34	INTRA

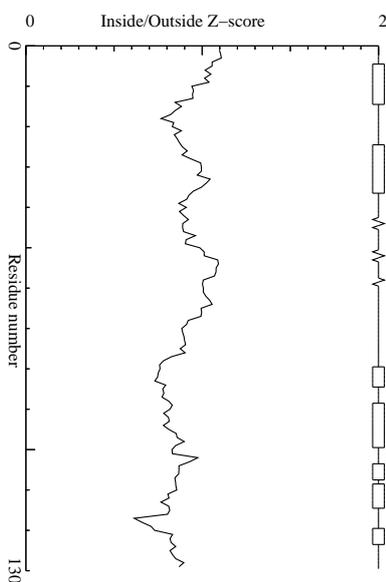
And so on for a total of 64 lines.

1.6 Sect 8) Packing, accessibility and threading

1.6.1 Note: Inside/Outside residue distribution normal
inside/outside RMS Z-score : 0.904

1.6.2 Note: Inside/Outside RMS Z-score plot

The Inside/Outside distribution normality RMS Z-score over a 15 residue window is plotted as function of the residue number. High areas in the plot (above 1.5) indicate unusual inside/outside patterns.



1.6.3 Warning: Low packing Z-score for some residues

Residue	Z-score
68 ARG (68) -	-2.78
73 ARG (73) -	-2.59

1.6.4 Warning: Abnormal packing Z-score for sequential residues

Start residue	End residue	Av. Z-score
123 TRP (123) -	126 GLY (126) -	-1.85

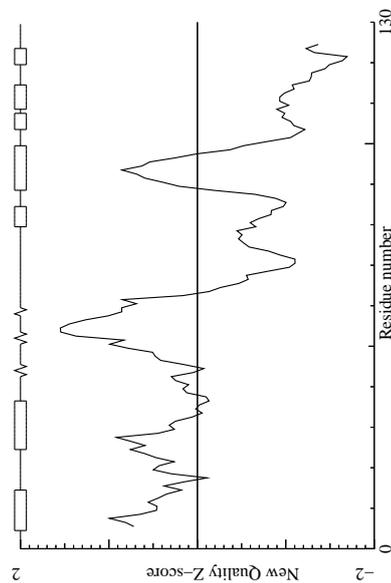
1.6.5 Note: Structural average packing Z-score OK

All contacts : Average = -0.158 Z-score = -0.93
 BB-BB contacts : Average = -0.205 Z-score = -1.30
 BB-SC contacts : Average = -0.044 Z-score = -0.34
 SC-BB contacts : Average = 0.035 Z-score = 0.26
 SC-SC contacts : Average = -0.089 Z-score = -0.46

Abbildung C.4: Gekürzte WHATCHECK-Ausgabe TETRALYSO (Forts.).

1.6.6 Note: Second generation quality Z-score plot

The second generation quality Z-score smoothed over a 10 residue window is plotted as function of the residue number. Low areas in the plot (below -1.3) indicate "unusual" packing.



Residue	Atom
57	GLN (57) - - NE2
64	CYS (64) - - N
68	ARG (68) - - NHI

1.7.6 Note: Buried hydrogen bond acceptors OK

1.8 Final summary

1.8.1 Note: Summary report for users of a structure

This is an overall summary of the quality of the structure as compared with current reliable structures. This summary is most useful for biologists seeking a good structure to use for modelling calculations.

The second part of the table mostly gives an impression of how well the model conforms to common refinement constraint values. The first part of the table shows a number of constraint-independent quality indicators.

Structure Z-scores, positive is better than average:

1st generation packing quality : -1.335
 2nd generation packing quality : -0.933
 Ramachandran plot appearance : -0.322
 χ -1/ χ -2 rotamer normality : -0.370
 Backbone conformation : 0.208

RMS Z-scores, should be close to 1.0:

Bond lengths : 1.162
 Bond angles : 1.451
 Omega angle restraints : 0.204 (tight)
 Side chain planarity : 1.839
 Improper dihedral distribution : 1.662 (loose)
 B-factor distribution : 0.769
 Inside/Outside distribution : 0.904

1.7 Sect 9) Water and hydrogenbond related checks

1.7.1 Note: Water contacts OK

1.7.2 Note: Water hydrogen bonds OK

1.7.3 Error: HIS, ASN, GLN side chain flips

Residue	ASN (44)	-
44	ASN (44) - -	
93	ASN (93) - -	

1.7.4 Note: Histidine type assignments

Residue	Geometry	Z-score	H-bond	Z-score
15	HIS (15) - - HIS-E	2.19	HIS-D	2.49

1.7.5 Warning: Buried unsatisfied hydrogen bond donors

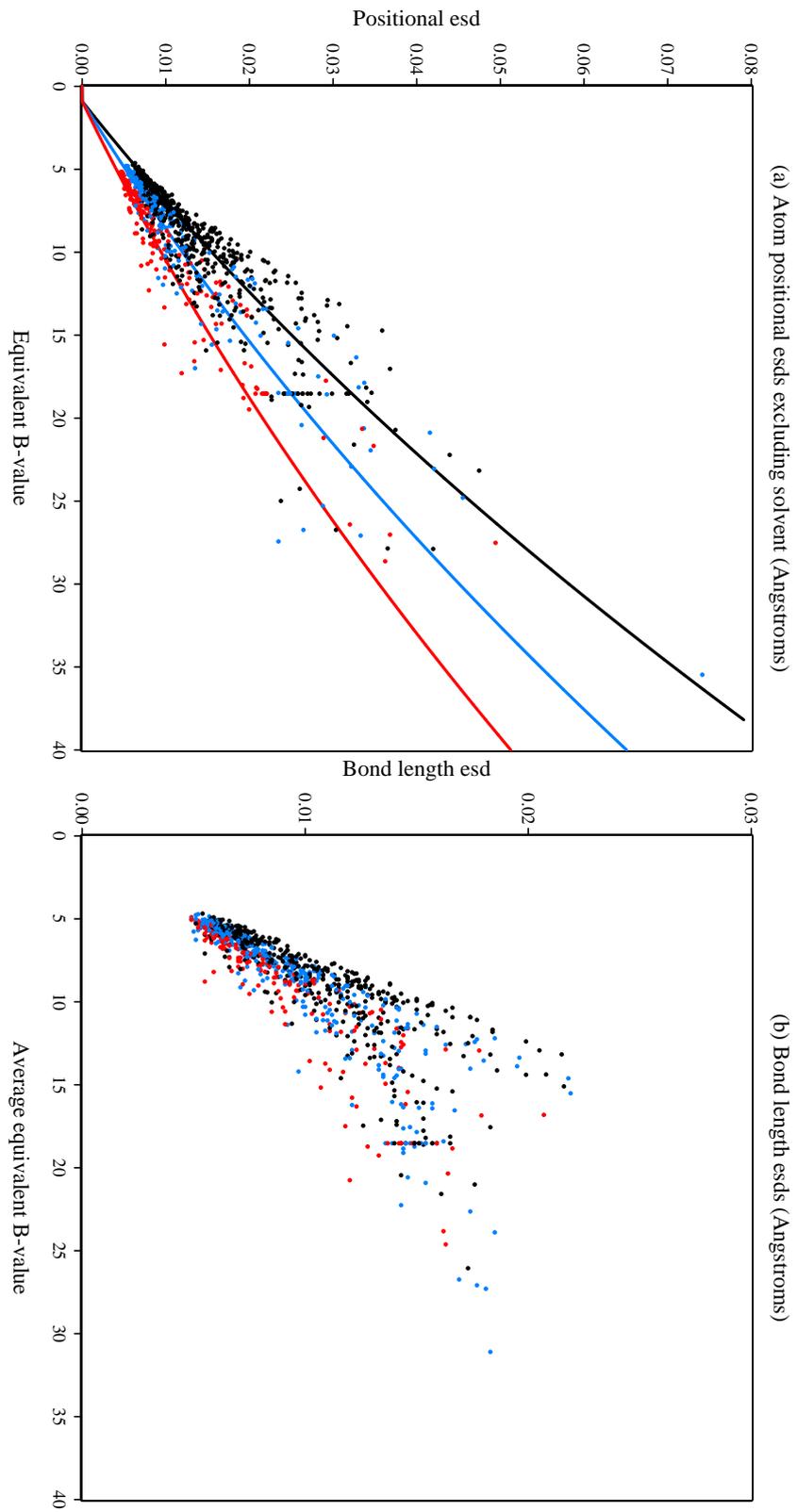


Abbildung C.5: SHELXPRO-Ausgabe TETRAIYSO.

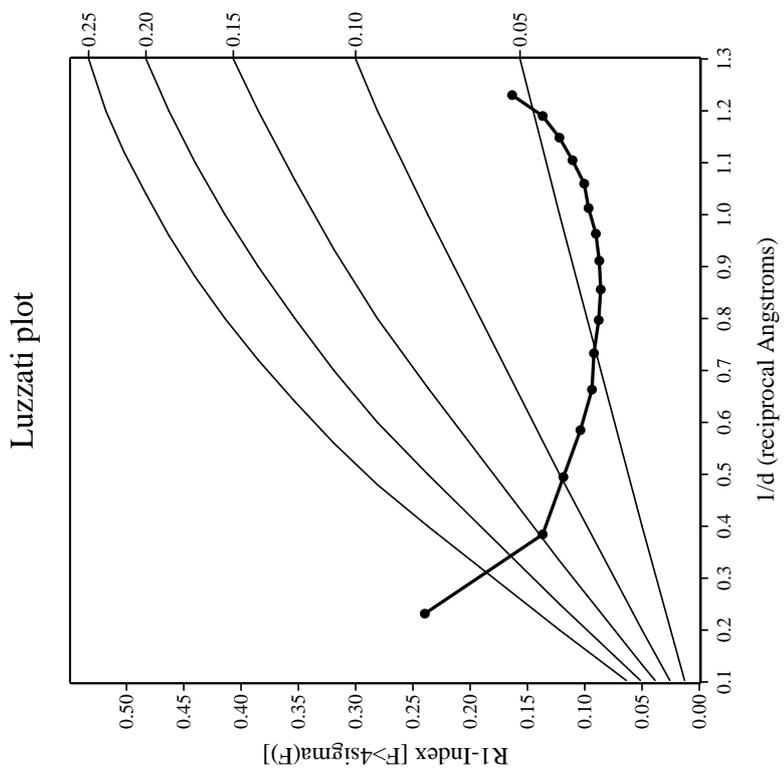
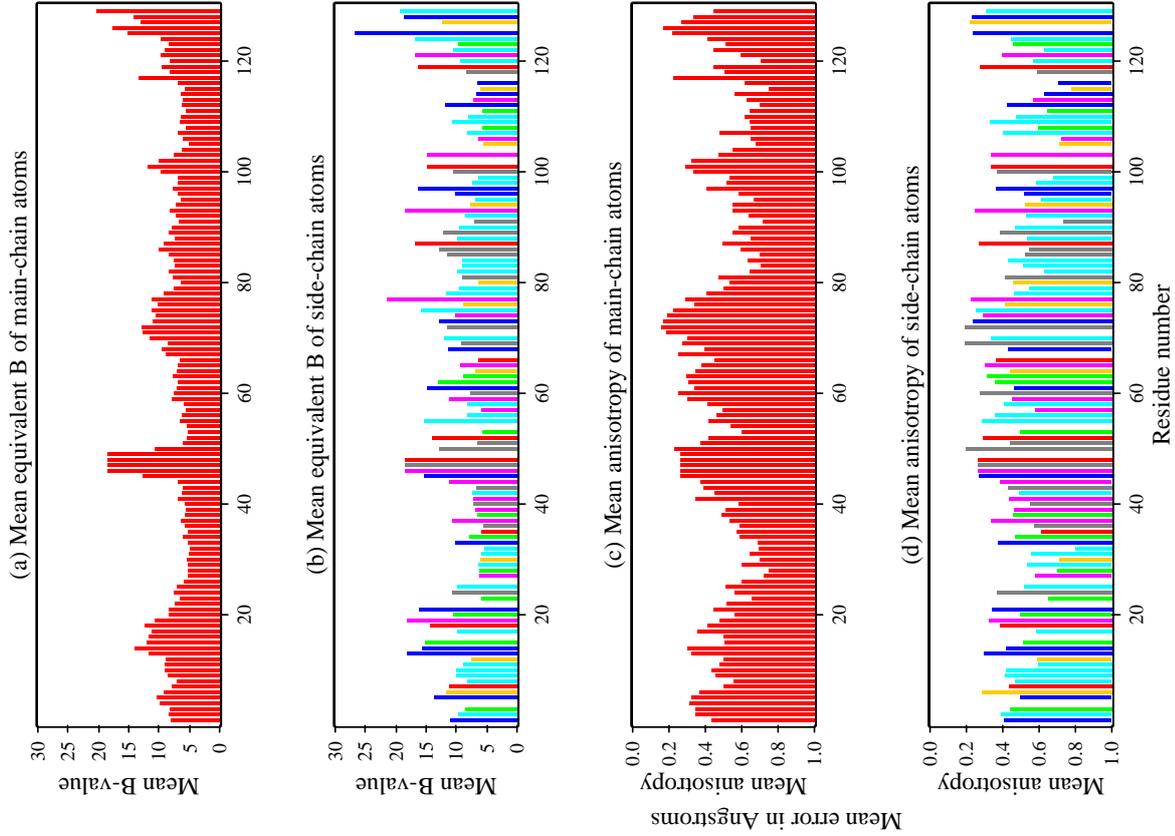


Abbildung C.5: SHELXPRO-Ausgabe TETRALYSO (Forts.).

D Insulin

Tabelle D.1: Statistik der U_{eq} verschiedener Teilmengen von INSULIN. († Minimum, Maximum und Mittelwert mit Standardabweichung der U_{eq} und ΔU_{eq} ; ‡ 0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-3}\AA^2)

	$x_{\min/\max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
alle Atome ($n = 2238$)							
<i>mul</i>	80,97/1252,0	216,10(128,49)	138,90	180,70	248,50	109,59	74,16
<i>iam</i>	80,97/1252,0	216,30(128,38)	138,90	180,70	248,50	109,57	73,58
<i>iam_{rb}</i>	80,97/1252,0	216,30(128,34)	139,10	180,70	248,00	108,88	73,39
ohne Solvens ($n = 2055$)							
<i>mul</i>	80,97/1252,0	205,80(116,63)	136,70	175,90	235,60	98,96	69,20
<i>iam</i>	80,97/1252,0	206,00(116,51)	136,90	175,90	236,40	99,52	68,85
<i>iam_{rb}</i>	80,97/1252,0	206,00(116,47)	136,90	175,70	236,00	99,11	68,39
ohne Solvens, ohne H ($n = 1039$)							
<i>mul</i>	80,97/834,40	176,00(88,62)	120,90	153,40	198,10	77,21	52,23
<i>iam</i>	80,97/834,40	176,20(88,40)	121,30	153,60	198,50	77,20	52,15
<i>iam_{rb}</i>	80,97/834,40	176,30(88,30)	121,40	153,60	198,40	76,99	51,76
ohne Solvens, nur Hauptpopulation ($n = 1572$)							
<i>mul</i>	81,19/1252,0	207,30(126,95)	133,20	172,80	233,60	100,48	68,56
<i>iam</i>	84,44/1252,0	207,50(126,81)	133,80	172,70	233,60	99,88	68,43
<i>iam_{rb}</i>	84,81/1252,0	207,50(126,75)	133,80	173,00	233,60	99,88	68,61
ohne Solvens, Hauptpopulation, exkl. H ($n = 810$)							
<i>mul</i>	81,19/834,40	175,60(94,97)	119,30	150,40	196,30	76,96	52,23
<i>iam</i>	84,44/834,40	176,00(94,71)	119,60	150,60	196,30	76,66	51,65
<i>iam_{rb}</i>	84,81/834,40	176,10(94,60)	119,80	150,50	195,90	76,12	51,65
nur verfeinerte Atome ($n = 403$)							
<i>mul</i>	81,19/328,10	135,10(38,16)	105,80	125,20	156,70	50,89	34,12
<i>iam</i>	84,44/315,80	135,80(37,64)	107,20	125,70	157,30	50,08	34,14
<i>iam_{rb}</i>	84,81/309,70	136,00(37,27)	107,40	125,90	157,60	50,21	33,94
Differenzen der verfeinerten Atome							
<i>mul-iam</i>	-9,99/12,35	-0,73(1,50)	-1,42	-0,88	-0,22	1,20	0,92
<i>mul-iam_{rb}</i>	-9,70/18,43	-0,93(1,91)	-1,80	-1,25	-0,48	1,32	1,00

Tabelle D.2: Statistik verfeinerter U_{eq} von INSULIN- nach Elementen. (\dagger Minimum, Maximum und Mittelwert mit Standardabweichung der U_{eq} und ΔU_{eq} ; \ddagger 0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-3}\AA^2)

	$x_{\min/\max}^\dagger$	$\bar{x}(\sigma_x)^\dagger$	$Q1^\ddagger$	$Q2^\ddagger$	$Q3^\ddagger$	IQR^\ddagger	MAD^\ddagger
verfeinerte C-Atome ($n = 253$)							
<i>mul</i>	81,19/328,10	135,80(39,21)	105,90	126,30	159,40	53,58	35,79
<i>iam</i>	84,44/315,80	136,60(38,43)	107,50	127,00	159,70	52,17	35,14
<i>iam_{rb}</i>	84,81/309,70	136,70(37,82)	107,90	127,10	159,80	51,92	34,56
<i>mul-iam</i>	-9,99/12,35	-0,77(1,73)	-1,61	-1,05	-0,21	1,40	1,00
<i>mul-iam_{rb}</i>	-9,70/18,43	-0,88(2,27)	-2,02	-1,42	-0,39	1,63	1,04
verfeinerte N-Atome ($n = 66$)							
<i>mul</i>	88,02/203,80	123,30(29,26)	97,80	115,60	140,40	42,56	28,57
<i>iam</i>	88,85/204,40	123,50(28,81)	98,56	116,20	140,10	41,54	28,46
<i>iam_{rb}</i>	89,14/204,60	123,70(28,73)	98,63	116,20	139,90	41,30	28,25
<i>mul-iam</i>	-1,61/2,00	-0,22(0,73)	-0,73	-0,39	0,20	0,93	0,64
<i>mul-iam_{rb}</i>	-2,22/2,13	-0,37(0,87)	-1,01	-0,61	0,02	1,03	0,69
verfeinerte O-Atome ($n = 76$)							
<i>mul</i>	90,17/252,30	142,70(40,35)	112,00	133,80	164,70	52,79	35,54
<i>iam</i>	91,44/253,60	143,50(40,33)	113,30	133,80	164,90	51,57	35,48
<i>iam_{rb}</i>	91,81/253,90	144,00(40,40)	113,60	134,60	166,30	52,71	35,34
<i>mul-iam</i>	-3,76/1,73	-0,84(0,82)	-1,23	-0,89	-0,43	0,80	0,60
<i>mul-iam_{rb}</i>	-3,24/0,63	-1,31(0,73)	-1,70	-1,36	-0,96	0,75	0,56
verfeinerte S-Atome ($n = 8$)							
<i>mul</i>	106,10/189,80	135,50(30,76)	112,00	122,50	157,90	45,82	20,94
<i>iam</i>	108,50/194,30	138,30(31,32)	114,50	125,00	160,70	46,17	20,94
<i>iam_{rb}</i>	108,90/194,90	138,70(31,41)	114,80	125,40	160,90	46,15	21,00
<i>mul-iam</i>	-4,51/-2,43	-2,81(0,70)	-2,67	-2,55	-2,48	0,18	0,10
<i>mul-iam_{rb}</i>	-5,16/-2,72	-3,21(0,80)	-3,08	-2,93	-2,86	0,22	0,14

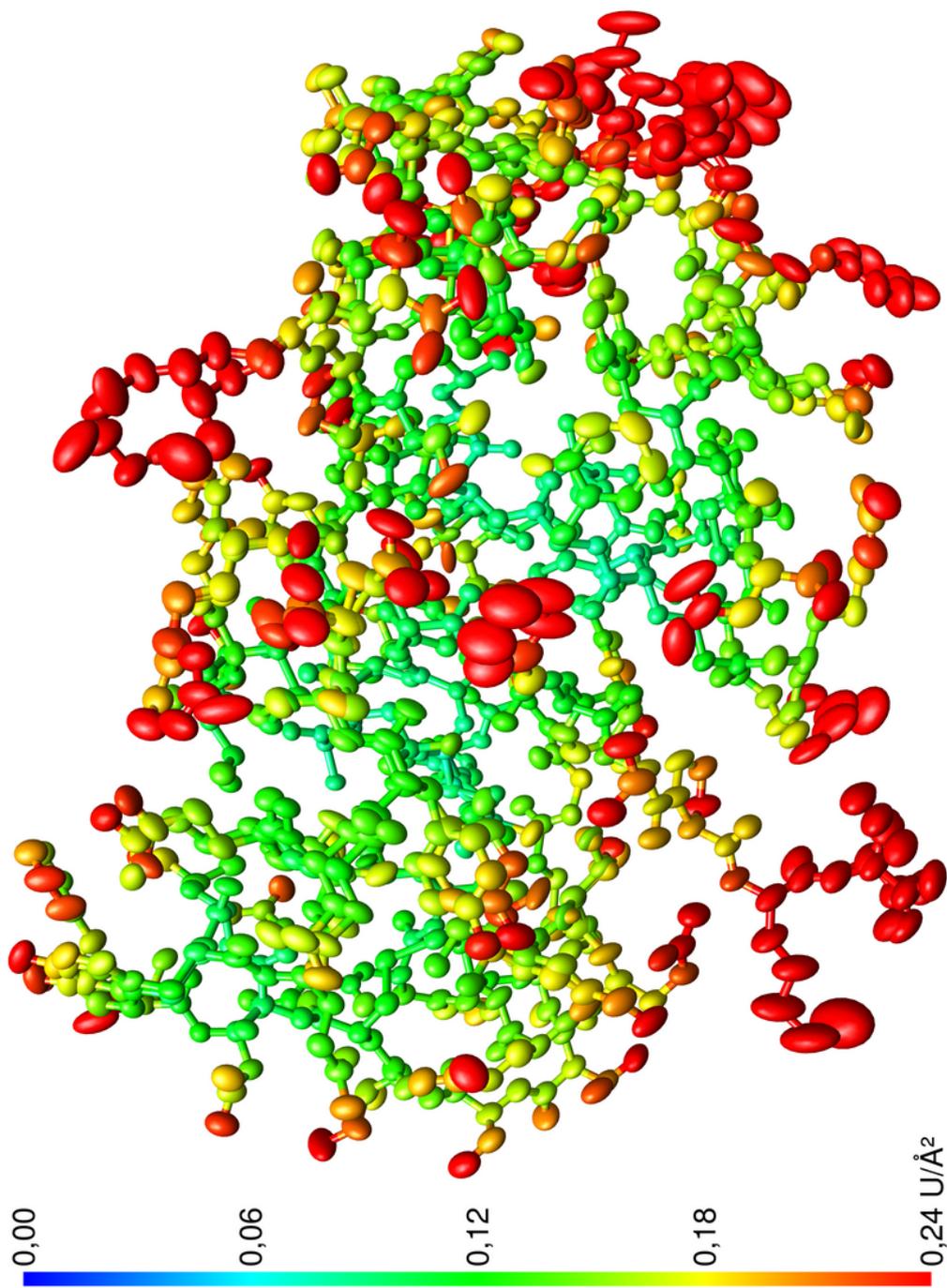


Abbildung D.1: Anisotrope Verschiebungsparameter INSULIN der Nichtwasserstoffatome nach asphärischer Verfeinerung – Farbcode: U_{eq} .

Tabelle D.3: Statistik der DMSDA-Beträge der verfeinerten Bindungen der Hauptpopulation von INSULIN und deren Differenzen – nach Elementen. ([†]Minimum, Maximum und Mittelwert mit Standardabweichung der DMSDA-Beträge; [‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-4}Å^2 .)

	$x_{min/max}^{\dagger}$	$\bar{x}(\sigma_x)^{\dagger}$	$Q1^{\ddagger}$	$Q2^{\ddagger}$	$Q3^{\ddagger}$	IQR^{\ddagger}	MAD^{\ddagger}
C–C ($n = 217$)							
<i>mul</i>	1/712	117,90(129,75)	32,00	80,00	140,00	108,00	78,58
<i>iam</i>	2/833	123,90(133,53)	40,00	82,00	152,00	112,00	75,61
<i>iam_{rb}</i>	0/823	122,60(136,97)	37,00	77,00	150,00	113,00	75,61
<i>mul-iam</i>	-143/204	-5,95(47,92)	-32,00	-5,00	18,00	50,00	37,06
<i>mul-iam_{rb}</i>	-181/163	-4,64(48,23)	-31,00	-5,00	28,00	59,00	44,48
C–N ($n = 144$)							
<i>mul</i>	0/583	80,76(80,82)	26,50	57,50	107,20	80,75	56,34
<i>iam</i>	1/638	85,28(89,03)	27,00	58,00	105,80	78,75	54,86
<i>iam_{rb}</i>	1/566	84,68(83,96)	25,75	62,50	107,20	81,50	57,82
<i>mul-iam</i>	-124/72	-4,51(33,63)	-25,00	-4,00	15,25	40,25	29,65
<i>mul-iam_{rb}</i>	-154/81	-3,92(36,32)	-29,00	-5,00	22,00	51,00	37,81
C–O ($n = 76$)							
<i>mul</i>	2/575	120,60(117,39)	38,50	83,50	158,50	120,00	88,21
<i>iam</i>	0/642	116,40(117,18)	35,00	73,00	167,20	132,25	74,13
<i>iam_{rb}</i>	0/698	117,30(115,98)	38,75	80,00	169,50	130,75	77,10
<i>mul-iam</i>	-100/108	4,17(36,44)	-20,00	3,00	22,00	42,00	28,91
<i>mul-iam_{rb}</i>	-123/126	3,30(41,01)	-17,00	-0,50	23,25	40,25	28,17
C–S ($n = 8$)							
<i>mul</i>	7/213	78,25(76,23)	33,50	46,00	109,00	75,50	54,86
<i>iam</i>	5/219	73,25(78,80)	21,50	43,00	93,00	71,50	39,29
<i>iam_{rb}</i>	6/223	63,25(78,35)	11,25	33,00	71,50	60,25	33,36
<i>mul-iam</i>	-42/38	5,00(28,01)	-11,00	12,00	25,50	36,50	24,46
<i>mul-iam_{rb}</i>	-37/43	15,00(30,34)	-10,75	31,50	35,00	45,75	15,57
S–S ($n = 5$)							
<i>mul</i>	10/122	54,60(41,16)	41,00	47,00	53,00	12,00	8,90
<i>iam</i>	5/186	64,00(71,12)	24,00	51,00	54,00	30,00	40,03
<i>iam_{rb}</i>	12/201	67,20(76,93)	20,00	50,00	53,00	33,00	44,48
<i>mul-iam</i>	-64/17	-9,40(31,71)	-7,00	2,00	5,00	12,00	13,34
<i>mul-iam_{rb}</i>	-79/21	-12,60(38,41)	-3,00	-2,00	0,00	3,00	2,97
alle ($n = 450$)							
<i>mul</i>	0/712	105,10(113,75)	28,00	71,50	137,50	109,50	73,39
<i>iam</i>	0/833	108,70(117,81)	33,00	71,50	143,80	110,75	65,98
<i>iam_{rb}</i>	0/823	107,90(118,37)	32,00	72,00	142,80	110,75	69,68
<i>mul-iam</i>	-143/204	-3,62(41,51)	-25,00	-4,00	18,00	43,00	32,62
<i>mul-iam_{rb}</i>	-181/163	-2,81(43,16)	-28,00	-3,00	24,00	52,00	38,55

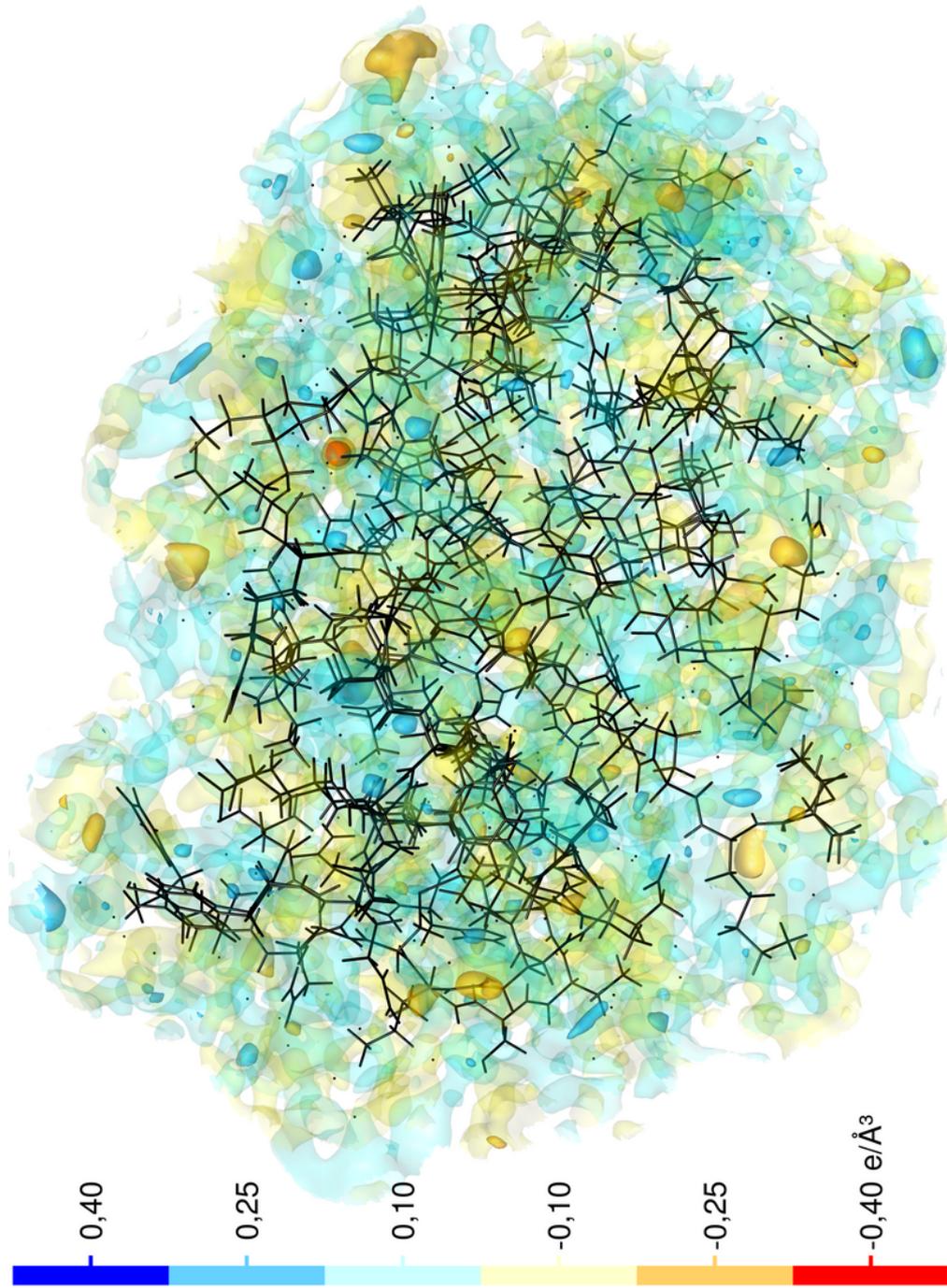


Abbildung D.2: Restdichten im Modell *iam_r,b* 4 Å um INSULIN.

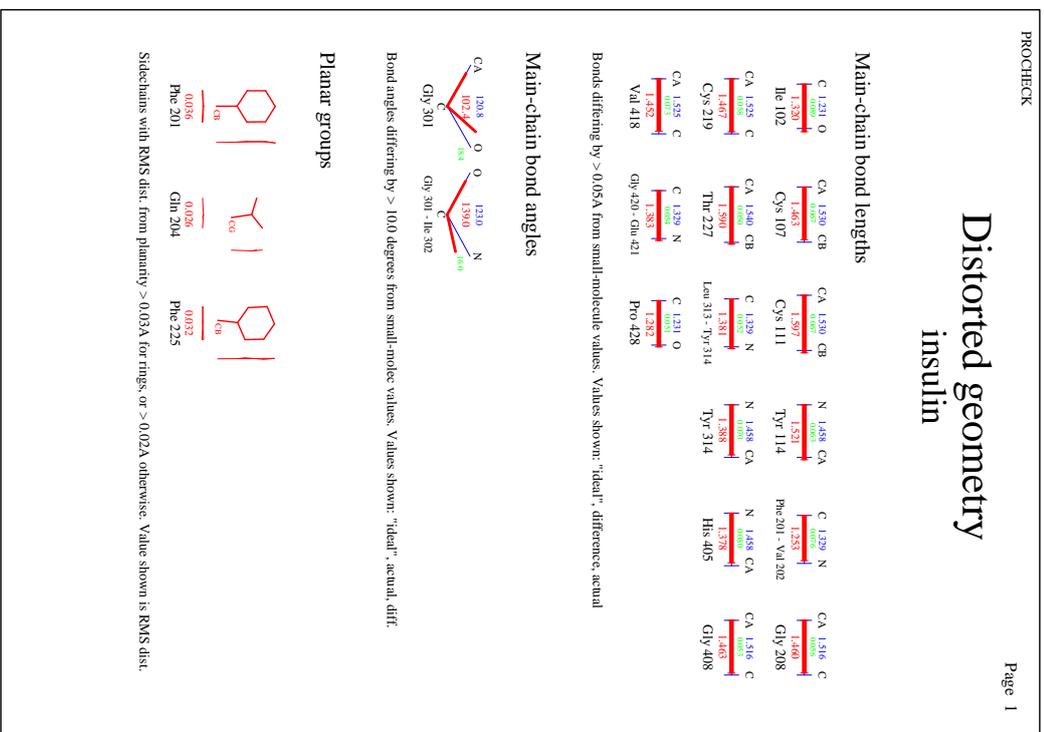
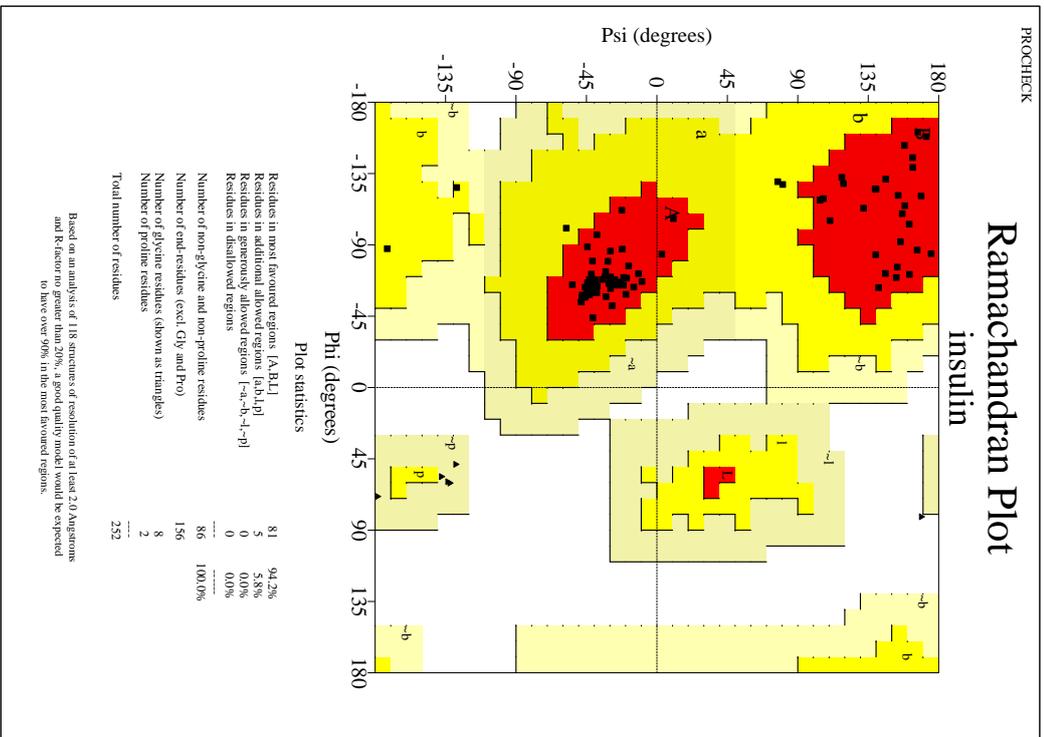
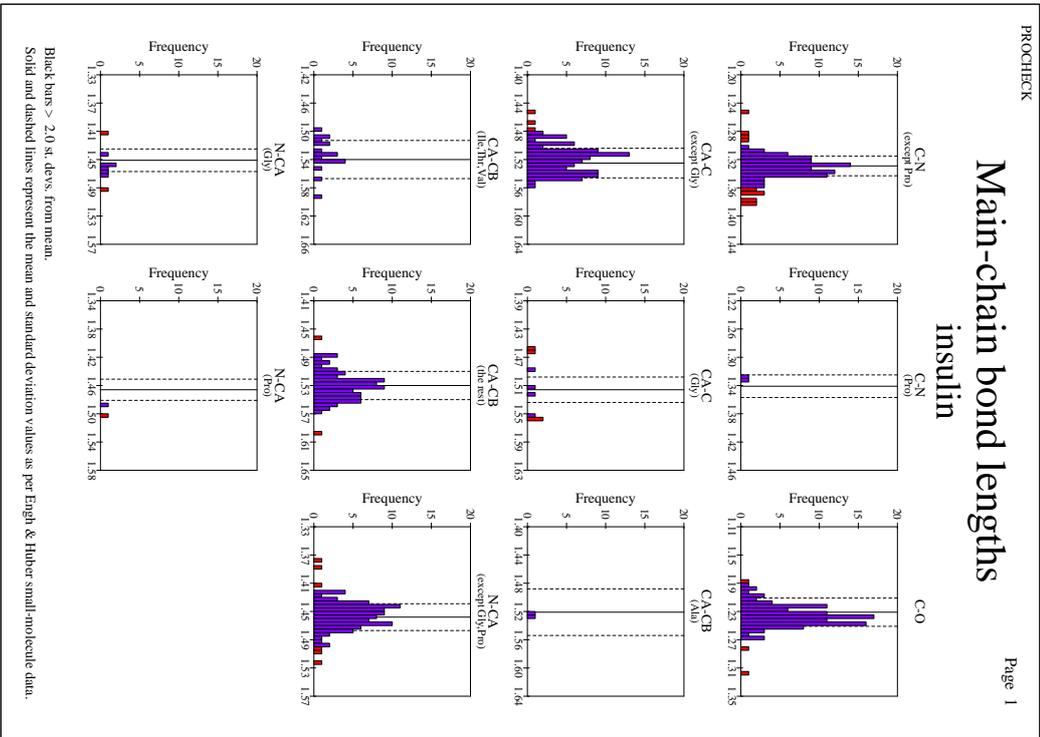


Abbildung D.3: PROCHECK-Ausgabe INSULIN.

Main-chain bond lengths insulin



RMS distances from planarity insulin

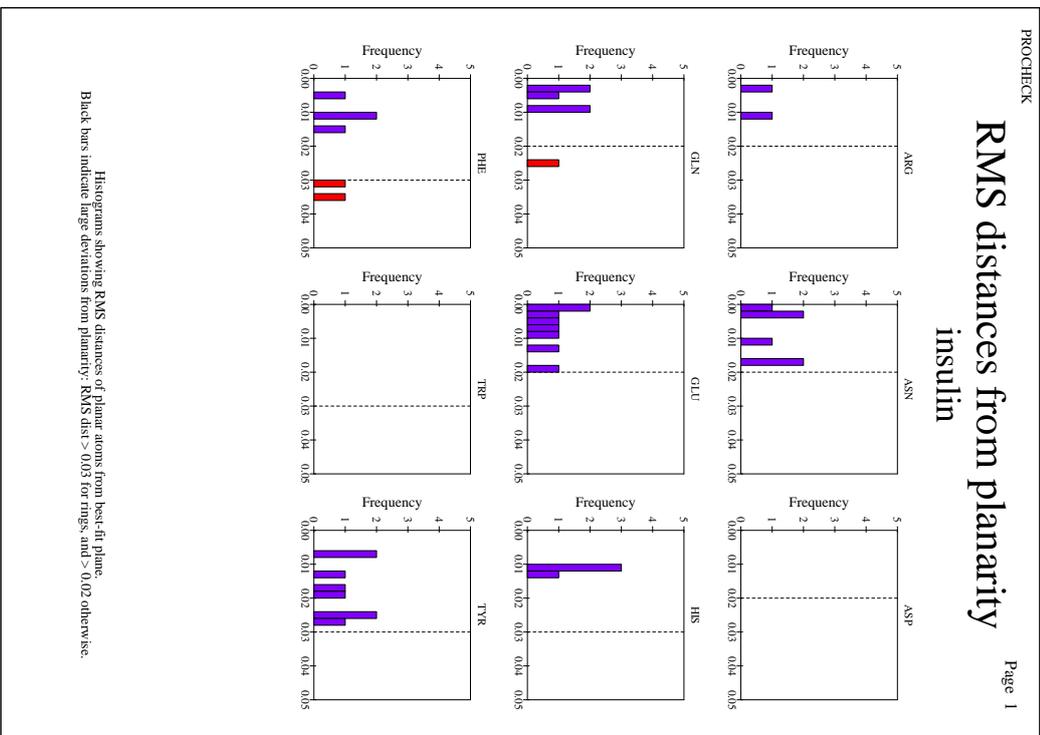
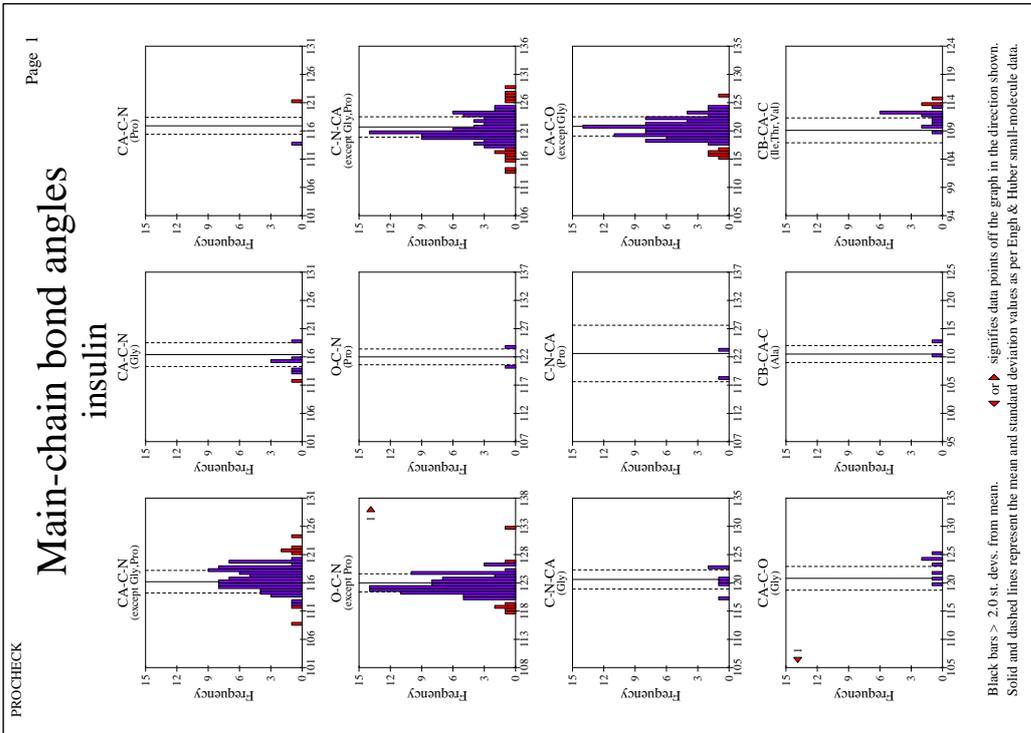
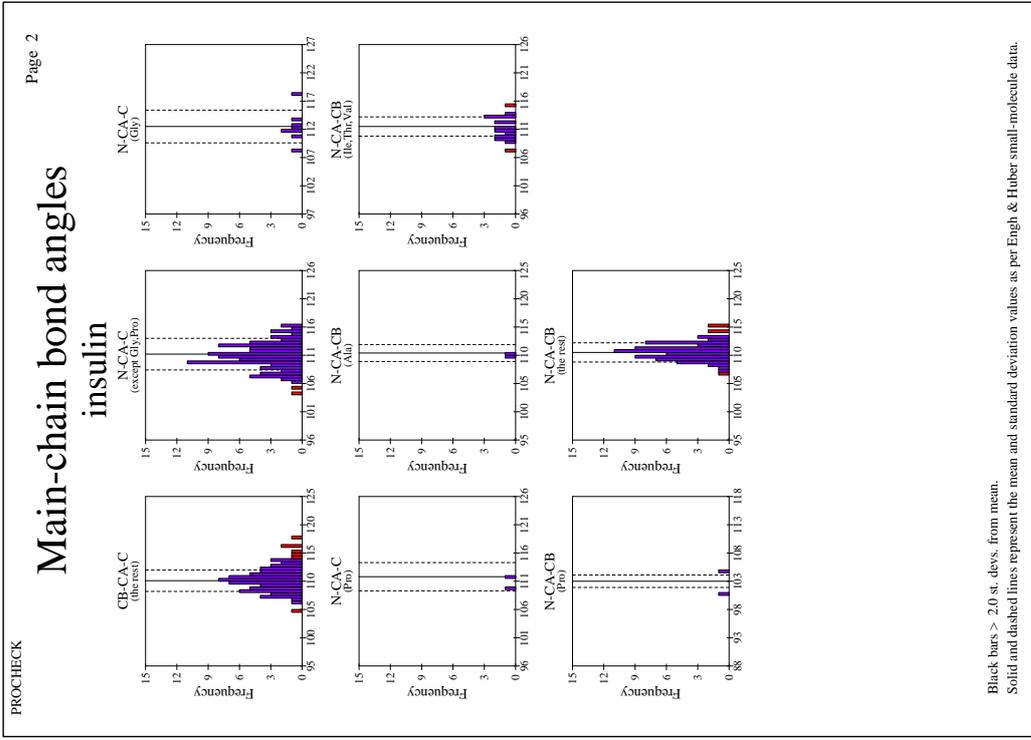


Abbildung D.3: PROCHECK-Ausgabe INSULIN (Forts.).



Black bars > 2.0 st. devs. from mean.
Solid and dashed lines represent the mean and standard deviation values as per Engh & Huber small-molecule data.

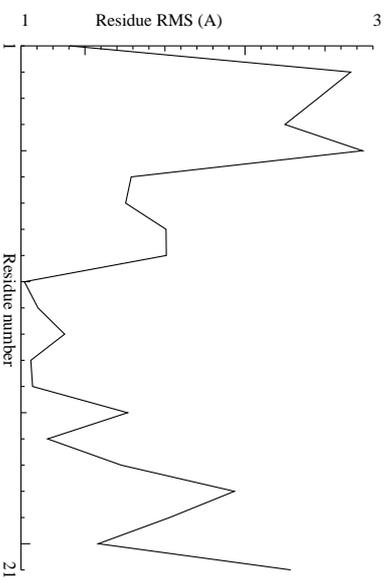
Black bars > 2.0 st. devs. from mean.
◄ or ► signifies data points off the graph in the direction shown.
Solid and dashed lines represent the mean and standard deviation values as per Engh & Huber small-molecule data.

Abbildung D.3: PROCHECK-Ausgabe INSULIN (Forts.).

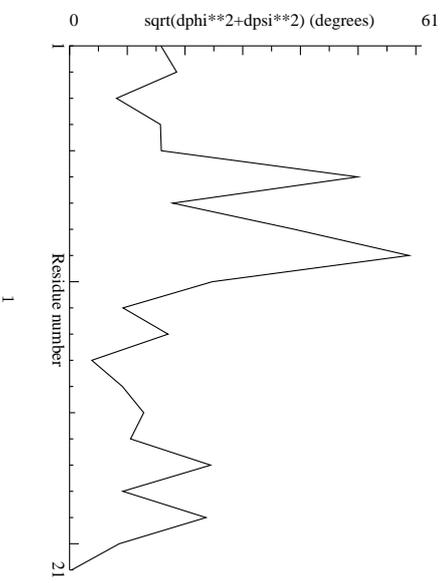
1 insulin.pdb

1.1 Sect 0) Symmetry related problems

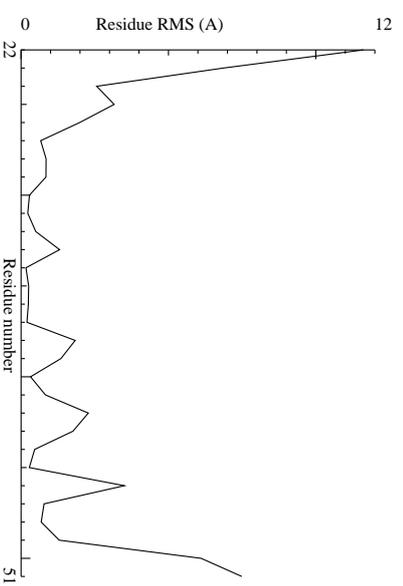
1.1.1 Note: Non crystallographic symmetry RMS plot



1.1.2 Note: Non crystallographic symmetry backbone difference plot



1.1.3 Note: Non crystallographic symmetry RMS plot



1.1.4 Note: Non crystallographic symmetry backbone difference plot

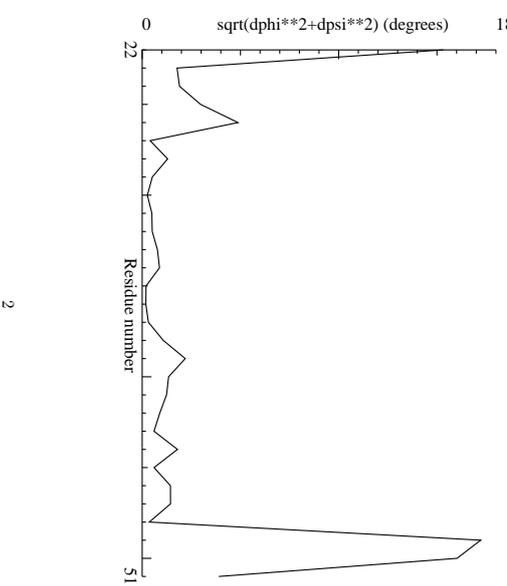


Abbildung D.4: Gekürzte WHATCHECK-Ausgabe INSULIN.

1.1.5 Note: No crystallographic symmetry between molecules

1.1.6 Note: Matthews coefficient OK

Molecular weight of all polymer chains: 11748.933

Volume of the Unit Cell V = 194924.438

Cell multiplicity: 9

Matthews coefficient for observed atoms $V_m = 1.843$

1.1.7 Note: All atoms are sufficiently far away from symmetry axes

1.2 Sect 3) Coordinate problems, unexpected atoms, B-factor and occupancy checks

1.2.1 Warning: Strange distribution of occupancy values

1.2.2 Warning: Occupancies atoms do not add up to 1.0.

1.2.3 Note: Average B-factor OK

Average B-factor for buried atoms : 10.518

1.2.4 Note: Number of buried atoms with low B-factor is OK

Percentage of buried atoms with B less than 5 : 0.00

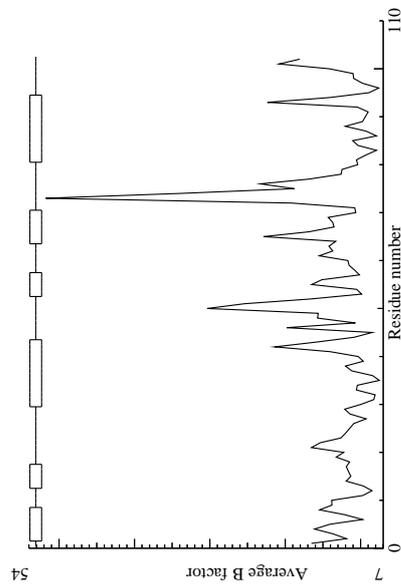
1.2.5 Note: B-factor distribution normal

RMS Z-score : 0.904 over 724 bonds

Average difference in B over a bond : 1.77

RMS difference in B over a bond : 2.73

1.2.6 Note: B-factor plot



3

1.3 Sect 5) Geometric checks

1.3.1 Warning: Unusual bond lengths

Residue	Atom pair	Distance	Z-value
2	ILE (102) - C	1.32	4.4
77	HIS (405) - N	1.38	-4.2
77	HIS (405) - CG	1.30	-4.9

1.3.2 Note: Normal bond length variability

RMS Z-score for bond lengths: 1.106

RMS-deviation in bond distances: 0.025

1.3.3 Note: No bond length directionality

1.3.4 Warning: Unusual bond angles

Residue	Atom Triplet	Bond Angle	Z-value
3	VAL (103) - CA -C	124.25	4.0
18	ASN (118) - CA CB	117.18	4.6
19	TYR (119) - CA CB	104.17	-5.0
19	TYR (119) - CB CG	126.97	4.1
31	HIS (210) - NE2 CD2	111.75	5.2
33	VAL (212) - CG1 CB	120.67	4.5
46	PHE (225) - CG CD1	128.05	4.3
46	PHE (225) - CG CD2	111.58	-5.4
46	PHE (225) - CD1 CE1	109.86	-5.6
52	GLY (301) - CA C	102.42	-8.8
53	ILE (302) - O -C	139.03	10.0
53	ILE (302) - C N	113.76	-4.4
55	GLU (304) - CB CG	122.35	5.7
62	CYS (311) - C N	114.20	-4.2
72	ASN (321) - CA CB	117.95	5.4
73	PHE (401) - C CA	117.99	4.2
77	HIS (405) - O -C	132.90	6.2
77	HIS (405) - CB CG	130.83	6.2
77	HIS (405) - NE2 CD2	114.22	7.7
82	HIS (410) - CA CB	109.62	-4.2
82	HIS (410) - NE2 CD2	111.77	5.3
93	GLU (421) - CA CB	104.82	-4.6
93	GLU (421) - CB CG	133.87	12.5

1.3.5 Note: Normal bond angle variability

RMS Z-score for bond angles: 1.449

RMS-deviation in bond angles: 2.614

1.3.6 Warning: Chirality deviations detected

4

Abbildung D.4: Gekürzte WHATCHECK-Ausgabe INSULIN (Forts.).

Residue	Atom	Z-value	Value	Expected	Notes	
33	VAL (212)-	-	CB	6.6	-24.52	-32.97
41	GLY (220)-	-	C	7.1	9.39	0.09

1.3.7 Note: Improper dihedral angle distribution OK
Improper dihedral RMS Z-score : 1.352

1.3.8 Error: Side chain planarity problems

Residue	Z-score
25	GLN (204)-

1.3.9 Note: Atoms connected to aromatic rings OK

1.4 Sect 6) Torsion-related checks

1.4.1 Note: PRO puckering amplitude OK

1.4.2 Note: PRO puckering phases OK

1.4.3 Warning: Torsion angle evaluation shows unusual residues

Residue	Z-score
76	GLN (404)-

1.4.4 Warning: Backbone torsion angle evaluation shows unusual conformations

Residue	Description
9	SER (109)- - - Poor ϕ/ψ
24	ASN (203)- - - ω poor

1.4.5 Note: Ramachandran Z-score OK

Ramachandran Z-score : -2.710

1.4.6 Note: Omega angle restraint OK

Standard deviation of ω values : 1.093

1.4.7 Note: chi-1/chi-2 angle correlation Z-score OK

χ -1/ χ -2 correlation Z-score : 0.238

1.4.8 Note: Backbone oxygen evaluation OK

position.

1.4.9 Note: Rotamers checked OK

1.4.10 Warning: Unusual backbone conformations

Residue	# hits
20	CYS (120)-
21	ASN (121)-
22	PHE (201)-
23	VAL (202)-
43	ARG (222)-
50	LYS (229)-
51	THR (230)-
52	GLY (301)-
53	ILE (302)-
58	CYS (307)-
71	CYS (320)-
72	ASN (321)-
73	PHE (401)-
74	VAL (402)-
91	CYS (419)-
94	ARG (422)-
9	SER (109)-
28	CYS (207)-
29	CYS (407)-
40	CYS (219)-

1.4.11 Warning: Backbone conformation Z-score low

Backbone conformation Z-score : -2.012

1.5 Sect 7) Bump checks

1.5.1 Error: Abnormally short interatomic distances

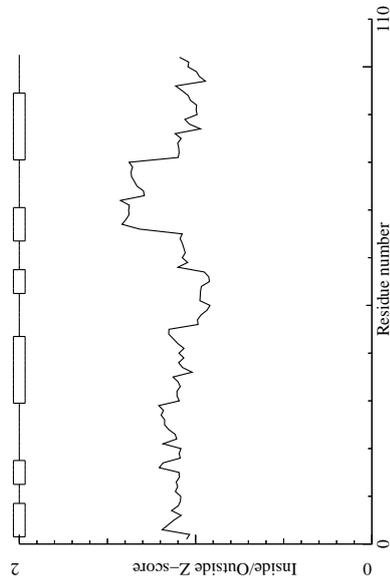
Atom 1	Atom 2	Bump Dist	Status
53	ILE (302)-	-	-
25	GLN (204)-	-	-
56	GLN (305)-	-	-
51	THR (230)-	-	-
52	GLY (301)-	-	-
19	TYR (119)-	-	-
76	GLN (404)-	-	-
4	GLU (104)-	-	-
76	GLN (404)-	-	-
76	GLN (301)-	-	-
76	GLN (404)-	-	-
1	GLY (101)-	-	-
76	GLN (404)-	-	-
101	LYS (429)-	-	-
29	GLY (208)-	-	-
20	CYS (120)-	-	-
21	ASN (121)-	-	-
22	PHE (201)-	-	-
23	VAL (202)-	-	-
43	ARG (222)-	-	-
50	LYS (229)-	-	-
51	THR (230)-	-	-
52	GLY (301)-	-	-
53	ILE (302)-	-	-
58	CYS (307)-	-	-
71	CYS (320)-	-	-
72	ASN (321)-	-	-
73	PHE (401)-	-	-
74	VAL (402)-	-	-
91	CYS (419)-	-	-
94	ARG (422)-	-	-
9	SER (109)-	-	-
28	CYS (207)-	-	-
29	CYS (407)-	-	-
40	CYS (219)-	-	-
53	ILE (302)-	-	-
25	GLN (204)-	-	-
56	GLN (305)-	-	-
51	THR (230)-	-	-
52	GLY (301)-	-	-
19	TYR (119)-	-	-
76	GLN (404)-	-	-
4	GLU (104)-	-	-
76	GLN (404)-	-	-
76	GLN (301)-	-	-
76	GLN (404)-	-	-
1	GLY (101)-	-	-
76	GLN (404)-	-	-
101	LYS (429)-	-	-
29	GLY (208)-	-	-
20	CYS (120)-	-	-
21	ASN (121)-	-	-
22	PHE (201)-	-	-
23	VAL (202)-	-	-
43	ARG (222)-	-	-
50	LYS (229)-	-	-
51	THR (230)-	-	-
52	GLY (301)-	-	-
53	ILE (302)-	-	-
58	CYS (307)-	-	-
71	CYS (320)-	-	-
72	ASN (321)-	-	-
73	PHE (401)-	-	-
74	VAL (402)-	-	-
91	CYS (419)-	-	-
94	ARG (422)-	-	-
9	SER (109)-	-	-
28	CYS (207)-	-	-
29	CYS (407)-	-	-
40	CYS (219)-	-	-
53	ILE (302)-	-	-
25	GLN (204)-	-	-
56	GLN (305)-	-	-
51	THR (230)-	-	-
52	GLY (301)-	-	-
19	TYR (119)-	-	-
76	GLN (404)-	-	-
4	GLU (104)-	-	-
76	GLN (404)-	-	-
76	GLN (301)-	-	-
76	GLN (404)-	-	-
1	GLY (101)-	-	-
76	GLN (404)-	-	-
101	LYS (429)-	-	-
29	GLY (208)-	-	-
20	CYS (120)-	-	-
21	ASN (121)-	-	-
22	PHE (201)-	-	-
23	VAL (202)-	-	-
43	ARG (222)-	-	-
50	LYS (229)-	-	-
51	THR (230)-	-	-
52	GLY (301)-	-	-
53	ILE (302)-	-	-
58	CYS (307)-	-	-
71	CYS (320)-	-	-
72	ASN (321)-	-	-
73	PHE (401)-	-	-
74	VAL (402)-	-	-
91	CYS (419)-	-	-
94	ARG (422)-	-	-
9	SER (109)-	-	-
28	CYS (207)-	-	-
29	CYS (407)-	-	-
40	CYS (219)-	-	-
53	ILE (302)-	-	-
25	GLN (204)-	-	-
56	GLN (305)-	-	-
51	THR (230)-	-	-
52	GLY (301)-	-	-
19	TYR (119)-	-	-
76	GLN (404)-	-	-
4	GLU (104)-	-	-
76	GLN (404)-	-	-
76	GLN (301)-	-	-
76	GLN (404)-	-	-
1	GLY (101)-	-	-
76	GLN (404)-	-	-
101	LYS (429)-	-	-
29	GLY (208)-	-	-
20	CYS (120)-	-	-
21	ASN (121)-	-	-
22	PHE (201)-	-	-
23	VAL (202)-	-	-
43	ARG (222)-	-	-
50	LYS (229)-	-	-
51	THR (230)-	-	-
52	GLY (301)-	-	-
53	ILE (302)-	-	-
58	CYS (307)-	-	-
71	CYS (320)-	-	-
72	ASN (321)-	-	-
73	PHE (401)-	-	-
74	VAL (402)-	-	-
91	CYS (419)-	-	-
94	ARG (422)-	-	-
9	SER (109)-	-	-
28	CYS (207)-	-	-
29	CYS (407)-	-	-
40	CYS (219)-	-	-
53	ILE (302)-	-	-
25	GLN (204)-	-	-
56	GLN (305)-	-	-
51	THR (230)-	-	-
52	GLY (301)-	-	-
19	TYR (119)-	-	-
76	GLN (404)-	-	-
4	GLU (104)-	-	-
76	GLN (404)-	-	-
76	GLN (301)-	-	-
76	GLN (404)-	-	-
1	GLY (101)-	-	-
76	GLN (404)-	-	-
101	LYS (429)-	-	-
29	GLY (208)-	-	-
20	CYS (120)-	-	-
21	ASN (121)-	-	-
22	PHE (201)-	-	-
23	VAL (202)-	-	-
43	ARG (222)-	-	-
50	LYS (229)-	-	-
51	THR (230)-	-	-
52	GLY (301)-	-	-
53	ILE (302)-	-	-
58	CYS (307)-	-	-
71	CYS (320)-	-	-
72	ASN (321)-	-	-
73	PHE (401)-	-	-
74	VAL (402)-	-	-
91	CYS (419)-	-	-
94	ARG (422)-	-	-
9	SER (109)-	-	-
28	CYS (207)-	-	-
29	CYS (407)-	-	-
40	CYS (219)-	-	-

Abbildung D.4: Gekürzte WHATCHECK-Ausgabe INSULIN (Forts.).

1.6 Sect 8) Packing, accessibility and threading

1.6.1 Note: Inside/Outside residue distribution normal inside/outside RMS Z-score : 1.119

1.6.2 Note: Inside/Outside RMS Z-score plot



1.6.3 Warning: Low packing Z-score for some residues

Residue	Z-score
76 GLN (404)	-3.11

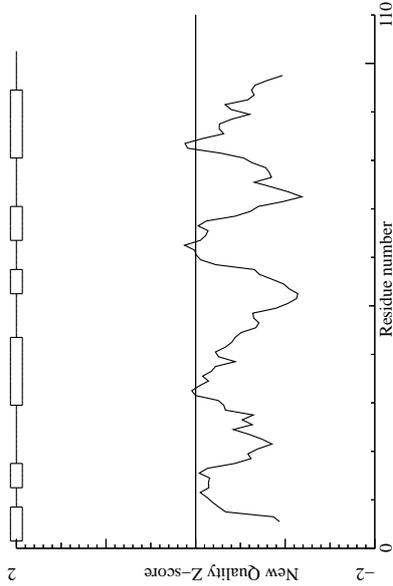
1.6.4 Note: No series of residues with abnormal new packing environment

1.6.5 Warning: Structural average packing Z-score a bit worrystone

All contacts : Average = -0.594 Z-score = -3.34
 BB-BB contacts : Average = -0.500 Z-score = -3.08
 BB-SC contacts : Average = -0.187 Z-score = -1.34
 SC-BB contacts : Average = -0.579 Z-score = -3.39
 SC-SC contacts : Average = -0.406 Z-score = -2.53

1.6.6 Note: Second generation quality Z-score plot

The second generation quality Z-score smoothed over a 10 residue window is plotted as function of the residue number. Low areas in the plot (below -1.3) indicate "unusual" packing.



1.7 Sect 9) Water and hydrogenbond related checks

1.7.1 Error: Water clusters without contacts with non-water atoms

Residue	Atom
111 HOH (7052)	O

1.7.2 Error: Water molecules without hydrogen bonds

Residue	Atom
111 HOH (7052)	O

1.7.3 Error: HIS, ASN, GLN side chain flips

Residue
25 GLN (204)
76 GLN (404)
77 HIS (405)

1.7.4 Note: Histidine type assignments

Residue	Geometry	Z-score	H-bond	Z-score
26 HIS (205)	- HIS-D	1.56	HIS-E	1.63
31 HIS (210)	- HIS-D	1.22		
77 HIS (405)	- HIS-D	2.38	HIS-H	2.61
82 HIS (410)	- HIS-E	0.92	HIS-D	1.07

1.7.5 Note: Checked buried hydrogen bond donors OK

1.7.6 Note: Buried hydrogen bond acceptors OK

1.8 Final summary

1.8.1 Note: Summary report for users of a structure

This is an overall summary of the quality of the structure as compared with current reliable structures. This summary is most useful for biologists seeking a good structure to use for modelling calculations.

The second part of the table mostly gives an impression of how well the model conforms to common refinement constraint values. The first part of the table shows a number of constraint-independent quality indicators.

Structure Z-scores, positive is better than average:

1st generation packing quality : -1.485

2nd generation packing quality : -3.335 (poor)

Ramachandran plot appearance : -2.710

χ^1/χ^2 rotamer normality : 0.238

Backbone conformation : -2.012

RMS Z-scores, should be close to 1.0:

Bond lengths : 1.106

Bond angles : 1.449 (tight)

Omega angle restraints : 0.199

Side chain planarity : 1.823

Improper dihedral distribution : 1.352

B-factor distribution : 0.904

Inside/Outside distribution : 1.119

9

Abbildung D.4: Gekürzte WHATCHECK-Ausgabe INSULIN (Forts.).

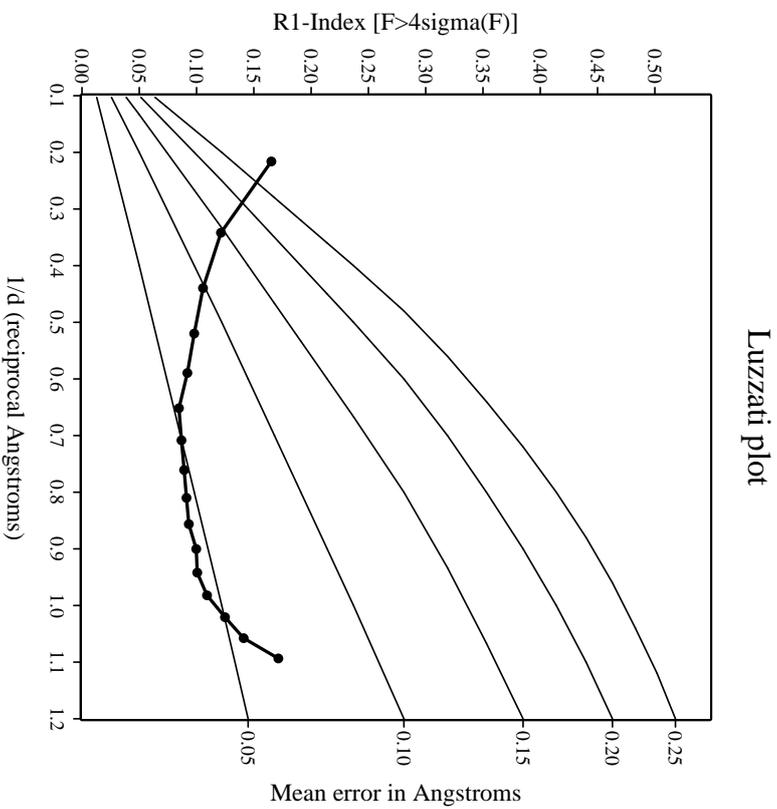


Abbildung D.5: SHELXPRO-Ausgabe INSULIN.

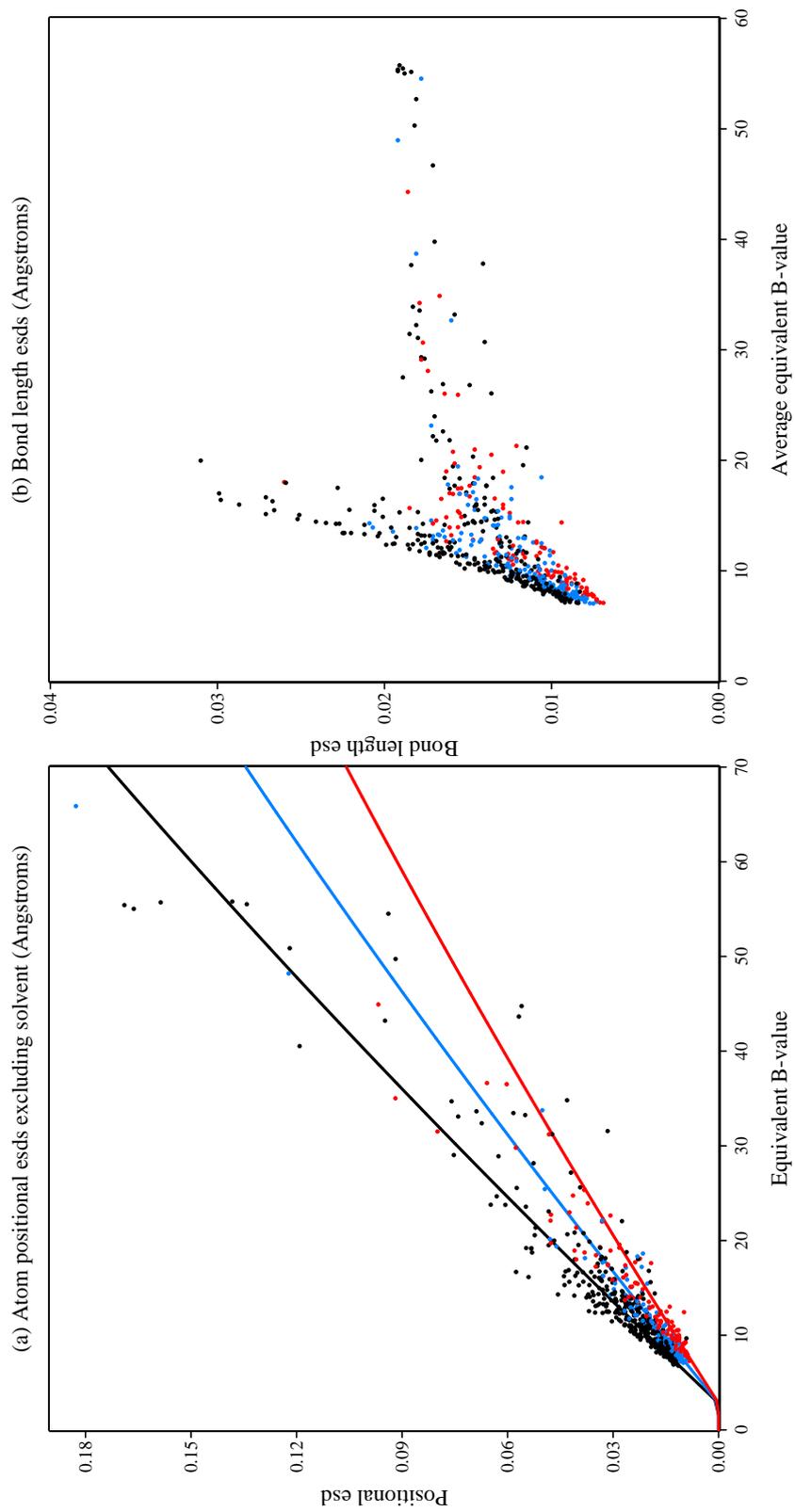


Abbildung D.5: SHELXPRO-Ausgabe INSULIN (Forts.).

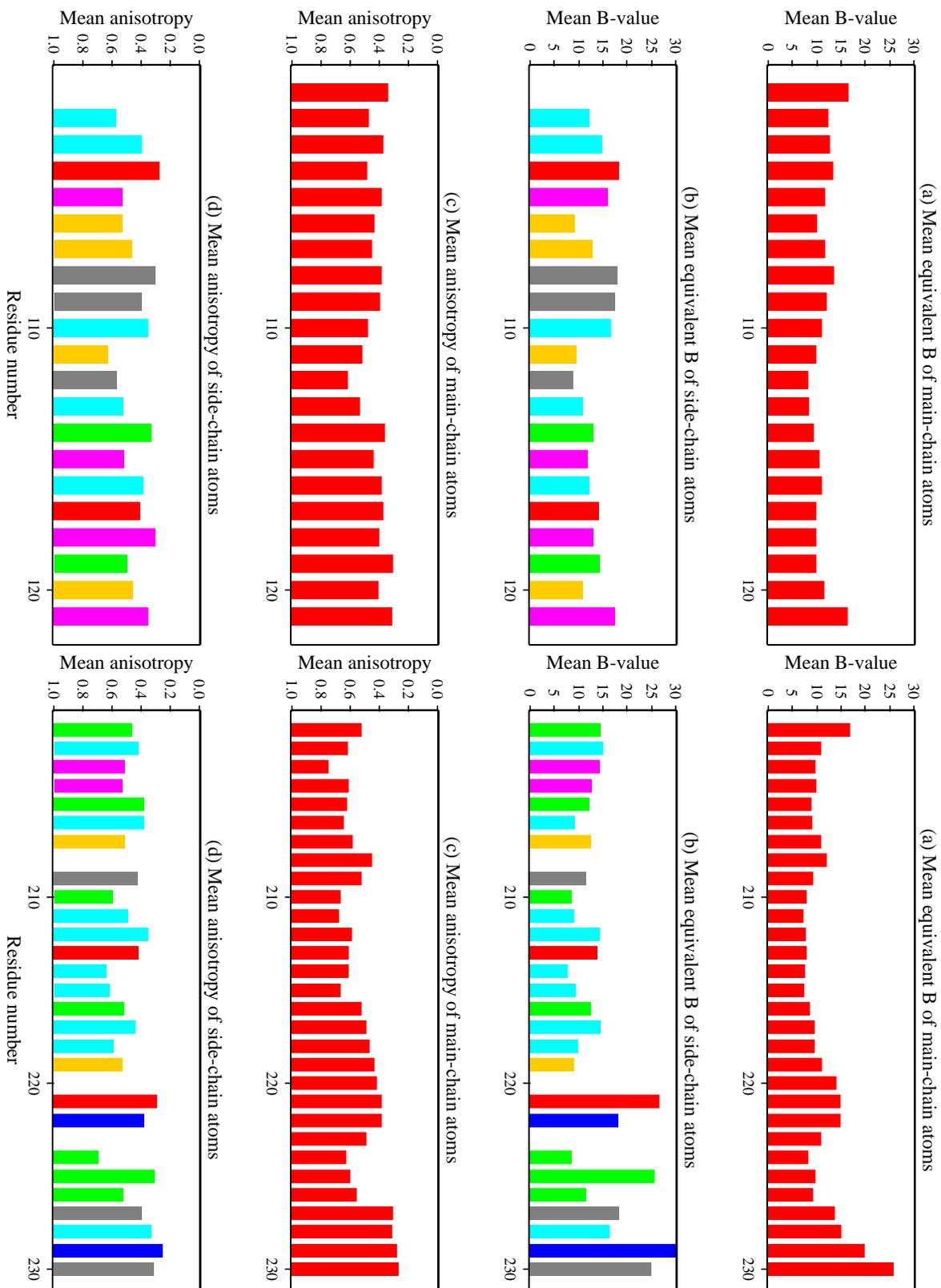


Abbildung D.5: SHELXPRO-Ausgabe INSULIN (Forts.).

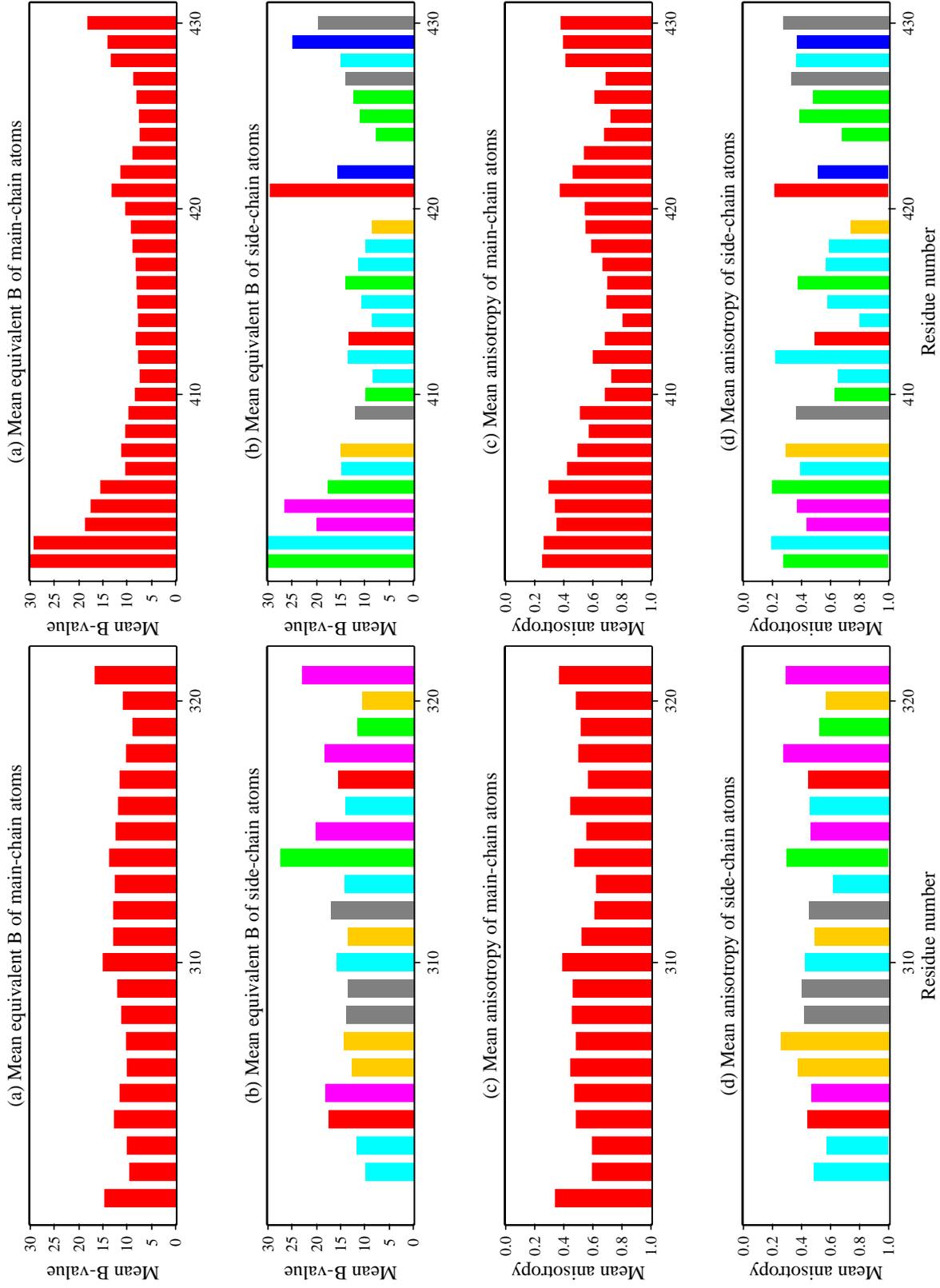


Abbildung D.5: SHELXPRO-Ausgabe INSULIN (Forts.).

E Trichotoxin

Tabelle E.1: Statistik der U_{eq} von TRICHOTOXIN. ([†]Minimum, Maximum und Mittelwert mit Standardabweichung der U_{eq} und ΔU_{eq} ; [‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-3}\AA^2 .)

	$x_{\min/\max}^{\dagger}$	$\bar{x}(\sigma_x)^{\dagger}$	$Q1^{\ddagger}$	$Q2^{\ddagger}$	$Q3^{\ddagger}$	IQR^{\ddagger}	MAD^{\ddagger}
alle Atome ($n = 529$)							
<i>mul</i>	21,15/398,50	73,40(56,56)	39,89	56,29	82,08	42,19	29,95
<i>iam</i>	23,90/398,50	74,50(56,06)	41,52	57,70	82,62	41,10	28,87
<i>iam_{rb}</i>	24,78/398,50	74,95(55,78)	42,24	58,35	82,62	40,38	28,06
ohne Solvens ($n = 510$)							
<i>mul</i>	21,15/323,40	69,45(48,92)	39,36	54,38	80,83	41,47	27,99
<i>iam</i>	23,90/323,40	70,60(48,41)	41,06	56,42	80,94	39,88	28,20
<i>iam_{rb}</i>	24,78/323,40	71,06(48,12)	42,01	57,05	80,97	38,97	27,35
ohne Solvens, ohne H ($n = 238$)							
<i>mul</i>	21,15/252,40	49,37(30,63)	32,48	41,31	53,81	21,33	14,72
<i>iam</i>	23,90/256,10	51,83(30,45)	34,76	43,82	56,36	21,60	14,93
<i>iam_{rb}</i>	24,78/254,70	52,82(30,03)	36,16	44,74	57,51	21,34	14,55
nur verfeinerte Atome ($n = 233$)							
<i>mul</i>	21,15/252,40	46,86(25,22)	31,83	41,11	52,31	20,48	14,31
<i>iam</i>	23,90/256,10	49,37(25,26)	34,50	43,26	55,42	20,92	14,57
<i>iam_{rb}</i>	24,78/254,70	50,38(24,84)	36,05	44,56	55,86	19,81	14,25
Differenzen der verfeinerten Atome							
<i>mul-iam</i>	-5,11/0,42	-2,51(0,69)	-2,96	-2,50	-2,10	0,86	0,66
<i>mul-iam_{rb}</i>	-5,59/3,68	-3,52(1,02)	-4,16	-3,68	-3,05	1,11	0,82

Tabelle E.2: Statistik verfeinerter U_{eq} von TRICHOTOXIN– nach Elementen. ([†]Minimum, Maximum und Mittelwert mit Standardabweichung der U_{eq} und ΔU_{eq} ; [‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-3}\AA^2)

	$x_{\min/\max}^{\dagger}$	$\bar{x}(\sigma_x)^{\dagger}$	$Q1^{\ddagger}$	$Q2^{\ddagger}$	$Q3^{\ddagger}$	IQR^{\ddagger}	MAD^{\ddagger}
verfeinerte C-Atome ($n = 152$)							
<i>mul</i>	21,15/170,00	46,95(22,02)	32,72	41,72	53,75	21,03	14,97
<i>iam</i>	23,90/175,10	49,60(21,94)	35,33	44,39	56,40	21,07	15,56
<i>iam_{rb}</i>	24,78/171,70	50,54(21,29)	36,78	45,68	57,79	21,01	14,70
<i>mul-iam</i>	-5,11/0,42	-2,65(0,71)	-3,11	-2,67	-2,24	0,88	0,65
<i>mul-iam_{rb}</i>	-5,40/3,68	-3,59(1,10)	-4,22	-3,77	-3,30	0,92	0,69
verfeinerte N-Atome ($n = 40$)							
<i>mul</i>	22,94/252,40	42,90(36,06)	28,56	35,24	45,25	16,69	12,98
<i>iam</i>	24,51/256,10	44,87(36,31)	30,49	36,99	47,31	16,81	13,32
<i>iam_{rb}</i>	25,91/254,70	45,59(35,92)	31,12	37,77	47,45	16,33	12,07
<i>mul-iam</i>	-3,72/-0,44	-1,98(0,58)	-2,25	-1,93	-1,69	0,56	0,46
<i>mul-iam_{rb}</i>	-4,24/-0,82	-2,69(0,54)	-3,00	-2,62	-2,47	0,53	0,48
verfeinerte O-Atome ($n = 41$)							
<i>mul</i>	29,52/141,70	50,38(23,75)	35,15	42,30	54,10	18,95	13,75
<i>iam</i>	32,18/143,50	52,90(23,76)	37,89	44,49	56,87	18,98	13,32
<i>iam_{rb}</i>	33,64/145,60	54,43(23,85)	39,33	46,11	57,71	18,38	13,62
<i>mul-iam</i>	-4,11/-1,76	-2,52(0,42)	-2,73	-2,49	-2,37	0,36	0,36
<i>mul-iam_{rb}</i>	-5,59/-3,08	-4,05(0,48)	-4,32	-4,01	-3,70	0,62	0,46

Tabelle E.3: Statistik der DMSDA-Beträge der verfeinerten Bindungen von TRICHOTOXIN und deren Differenzen – nach Elementen. ([†]Minimum, Maximum und Mittelwert mit Standardabweichung der DMSDA-Beträge; [‡]0,25-Quartil, Median, 0,75-Quartil, Interquartilabstand und Median der absoluten Abweichungen; alle Einheiten in 10^{-4}\AA^2)

	$x_{\min/\max}^{\dagger}$	$\bar{x}(\sigma_x)^{\dagger}$	$Q1^{\ddagger}$	$Q2^{\ddagger}$	$Q3^{\ddagger}$	IQR^{\ddagger}	MAD^{\ddagger}
C–C ($n = 116$)							
<i>mul</i>	1/899	60,27(102,78)	16,00	36,00	68,25	52,25	35,58
<i>iam</i>	1/878	69,81(102,10)	23,75	49,50	82,50	58,75	40,03
<i>iam_{rb}</i>	0/867	65,75(99,59)	22,75	43,00	74,75	52,00	34,10
<i>mul-iam</i>	-90/62	-9,54(35,61)	-35,00	-9,00	12,25	47,25	37,06
<i>mul-iam_{rb}</i>	-105/120	-5,48(38,42)	-25,25	-9,00	15,00	40,25	31,13

C–N ($n = 78$)							
<i>mul</i>	2/343	54,64(51,00)	19,50	44,00	75,50	56,00	38,55
<i>iam</i>	1/303	66,18(54,90)	22,25	55,00	92,00	69,75	49,67
<i>iam_{rb}</i>	0/258	74,14(59,94)	29,25	67,00	103,00	73,75	55,60
<i>mul-iam</i>	-84/76	-11,54(31,55)	-30,75	-10,50	4,75	35,50	25,95
<i>mul-iam_{rb}</i>	-139/89	-19,50(44,61)	-52,75	-13,00	7,50	60,25	45,22
C–O ($n = 41$)							
<i>mul</i>	3/224	79,88(53,92)	45,00	74,00	119,00	74,00	65,23
<i>iam</i>	6/214	72,05(50,46)	28,00	62,00	100,00	72,00	50,41
<i>iam_{rb}</i>	6/299	73,71(59,24)	31,00	63,00	96,00	65,00	48,93
<i>mul-iam</i>	-55/66	7,83(24,56)	-8,00	10,00	23,00	31,00	23,72
<i>mul-iam_{rb}</i>	-93/88	6,17(32,51)	-14,00	9,00	29,00	43,00	34,10
alle ($n = 235$)							
<i>mul</i>	1/899	61,82(81,36)	19,50	45,00	77,00	57,50	41,51
<i>iam</i>	1/878	69,00(80,96)	24,00	54,00	88,50	64,50	45,96
<i>iam_{rb}</i>	0/867	69,92(81,69)	24,00	50,00	87,50	63,50	44,48
<i>mul-iam</i>	-90/76	-7,17(33,20)	-28,00	-6,00	13,00	41,00	29,65
<i>mul-iam_{rb}</i>	-139/120	-8,10(40,54)	-33,00	-6,00	16,00	49,00	35,58

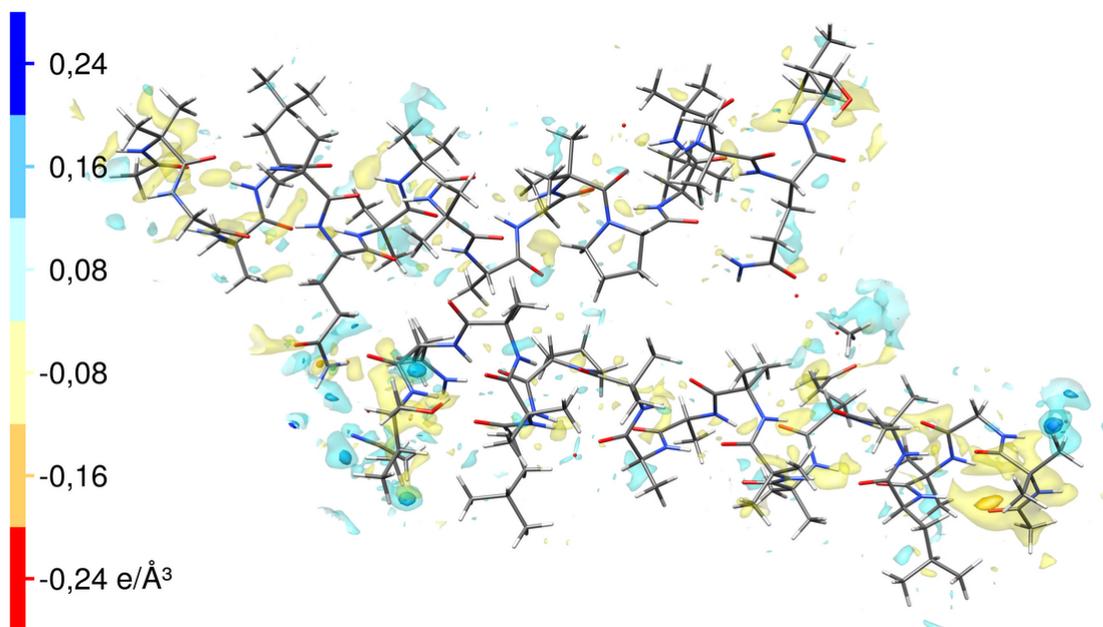
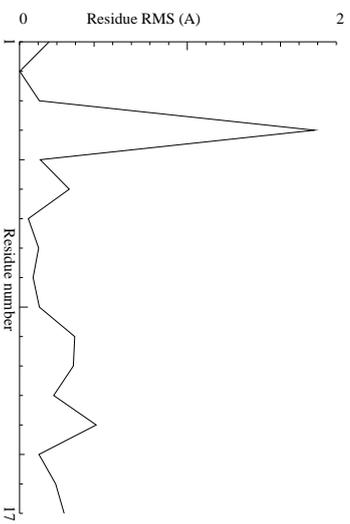


Abbildung E.1: Restdichte des Modells *iam_{rb}* von TRICHOTOXIN. Darstellung der Isooberflächen im Umkreis von 2 Å.

1 1m24.pdb

1.1 Sect 0) Symmetry related problems

1.1.1 Note: Non-crystallographic symmetry RMS plot



All-atom RMS fit for the two chains : 0.494
CA-only RMS fit for the two chains : 0.233

1.1.2 Note: Low non-crystallographic symmetry phi and psi differences

All comparable residues in the two chains mentioned below have comparable backbone torsion angles

1.1.3 Note: No crystallographic symmetry between molecules

1.1.4 Error: Matthews Coefficient (Vm) too low

Molecular weight of all polymer chains: 3462.229

Volume of the Unit Cell V = 4907.767

Cell multiplicity: 1

Matthews coefficient for observed atoms Vm=1.418

1.2 Sect 3) Coordinate problems, unexpected atoms, B-factor and occupancy checks

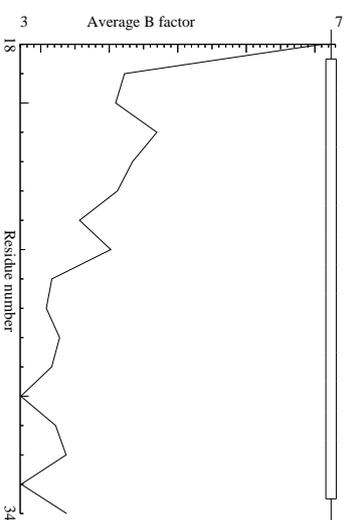
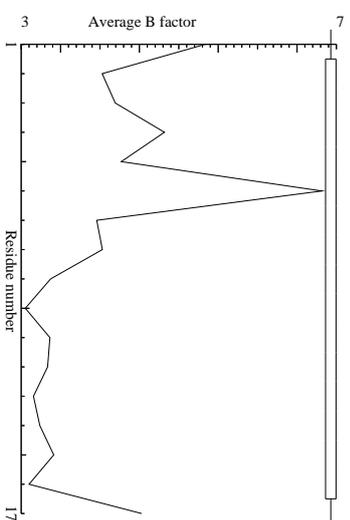
1.2.1 Note: Normal distribution of occupancy values

1.2.2 Note: All occupancies seem to add up to 1.0.

1.2.3 Warning: Average B-factor problem

Average B-factor for buried atoms : 2.712

1.2.4 Note: B-factor plot



1.3 Sect 5) Geometric checks

1.3.1 Note: All bond lengths OK

1.3.2 Warning: Low bond length variability

RMS Z-score for bond lengths: 0.604

RMS-deviation in bond distances: 0.014

1

2

Abbildung E.2: Gekürzte WHATCHECK-Ausgabe TRICHOTOXIN.

1.3.3 Note: No bond length directionality

1.3.4 Note: All bond angles OK

1.3.5 Note: Normal bond angle variability

RMS Z-score for bond angles: 0.711

RMS-deviation in bond angles: 1.762

1.3.6 Note: Chirality OK

1.3.7 Note: Improper dihedral angle distribution OK

Improper dihedral RMS Z-score : 0.961

1.3.8 Note: Side chain planarity OK

1.3.9 Note: Atoms connected to aromatic rings OK

1.4 Sect 6) Torsion-related checks

1.4.1 Note: PRO puckering amplitude OK

1.4.2 Note: PRO puckering phases OK

1.4.3 Note: Torsion angles OK

1.4.4 Note: Backbone torsion angles OK

1.4.5 Error: Ramachandran Z-score very low

Ramachandran Z-score : -4.461

1.4.6 Note: Omega angle restraint OK

Standard deviation of ω values : 0.896

1.4.7 Note: chi-1/chi-2 angle correlation Z-score OK

χ -1/ χ -2 correlation Z-score : 0.954

1.4.8 Note: Backbone oxygen evaluation OK

1.4.9 Note: Rotamers checked OK

1.4.10 Warning: Unusual backbone conformations

Residue	# hits
3 AIB (103-) A	0
5 AIB (105-) A	0
7 AIB (107-) A	0
8 AIB (108-) A	0
9 AIB (109-) A	0
12 AIB (112-) A	0

15	AIB (115-) A	-	0
16	AIB (116-) A	-	0
17	GLN (117-) A	-	0
18	AIB (201-) B	-	0
19	GLY (202-) B	-	0
20	AIB (203-) B	-	0
22	AIB (205-) B	-	0
24	AIB (207-) B	-	0
25	AIB (208-) B	-	0
26	AIB (209-) B	-	0
29	AIB (212-) B	-	0
32	AIB (215-) B	-	0

1.4.11 Note: Backbone conformation Z-score OK

Backbone conformation Z-score : -0.001

1.5 Sect 7) Bump checks

1.5.1 Error: Abnormally short interatomic distances

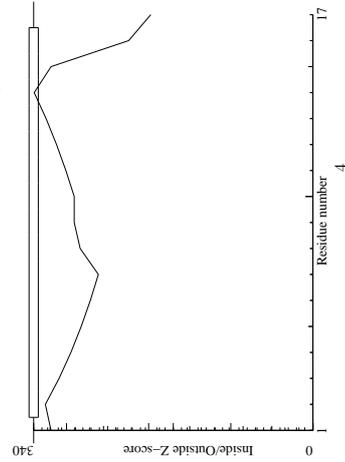
Atom 1	Atom 2	Bump Dist	Status
39 CCN (339-) -	N <-> 41 HOH (306) -	O	1.32 1.38 INTRA
39 CCN (339-) -	C1 <-> 41 HOH (306) -	O	0.50 2.30 INTRA
29 AIB (212-) B	N <-> 30 PRO (213-) B	CD	0.12 2.88 INTRA
12 AIB (112-) A	N <-> 13 PRO (113-) A	CD	0.12 2.88 INTRA
18 AIB (201-) B	C <-> 37 ACE (200-) B	O	0.06 2.74 INTRA
11 ALA (111-) A	C <-> 13 PRO (113-) A	CD	0.05 3.15 INTRA

1.6 Sect 8) Packing, accessibility and threading

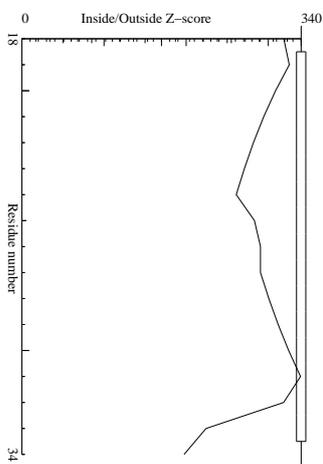
1.6.1 Warning: Inside/Outside residue distribution unusual

inside/outside RMS Z-score : 1.292

1.6.2 Note: Inside/Outside RMS Z-score plot



3



1.6.3 Note: Too many strange amino acid residues

Since more than 5 percent of the amino acids is non-natural, the Quality Control check has been skipped.

1.6.4 Warning: Low packing Z-score for some residues

Residue	Z-score
4 LEU (104-) A	-3.08
21 LEU (204-) B	-3.07
27 ALA (210-) B	-2.60
10 ALA (110-) A	-2.59
28 ALA (211-) B	-2.55

1.6.5 Note: No series of residues with abnormal new packing environment

1.6.6 Error: Abnormal structural average packing Z-score

All contacts : Average = -1.989 Z-score = -11.02
 BB-BB contacts : Average = -1.730 Z-score = -10.49
 BB-SC contacts : Average = -1.331 Z-score = -9.35
 SC-BB contacts : Average = -1.216 Z-score = -7.18
 SC-SC contacts : Average = -1.679 Z-score = -10.83

1.7 Sect 9) Water and hydrogenbond related checks

1.7.1 Note: Water contacts OK

1.7.2 Note: Water hydrogen bonds OK

1.7.3 Note: HIS, ASN, GLN side chains OK

1.7.4 Warning: Buried unsatisfied hydrogen bond donors

Residue	Atom
4 LEU (104-) A	N
21 LEU (204-) B	N
23 GLN (206-) B	NE2

1.7.5 Note: Buried hydrogen bond acceptors OK

1.8 Final summary

1.8.1 Note: Summary report for users of a structure

This is an overall summary of the quality of the structure as compared with current reliable structures. This summary is most useful for biologists seeking a good structure to use for modelling calculations.

The second part of the table mostly gives an impression of how well the model conforms to common refinement constraint values. The first part of the table shows a number of constraint-independent quality indicators.

Structure Z-scores, positive is better than average:

2nd generation packing quality : -11.019 (bad)

Ramachandran plot appearance : -4.461 (bad)

χ^1/χ^2 rotamer normality : 0.954

Backbone conformation : -0.001

RMS Z-scores, should be close to 1.0:

Bond lengths : 0.604 (tight)

Bond angles : 0.711 (tight)

Omega angle restraints : 0.163 (tight)

Side chain planarity : 2.183 (loose)

Improper dihedral distribution : 0.961

Inside/Outside distribution : 1.292 (unusual)

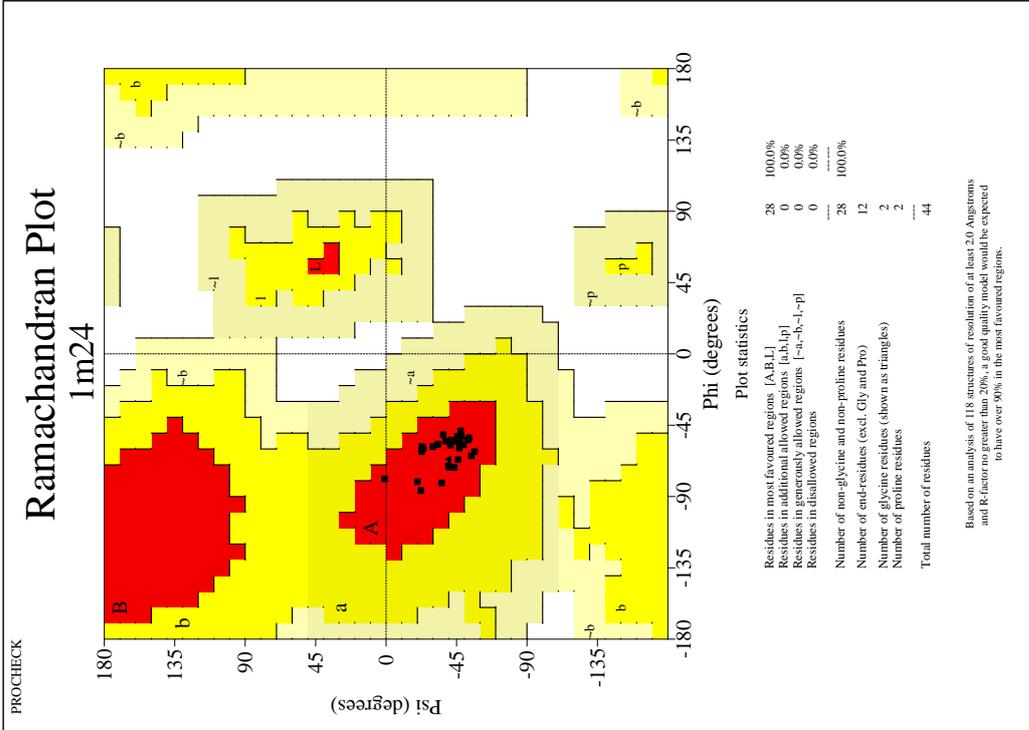
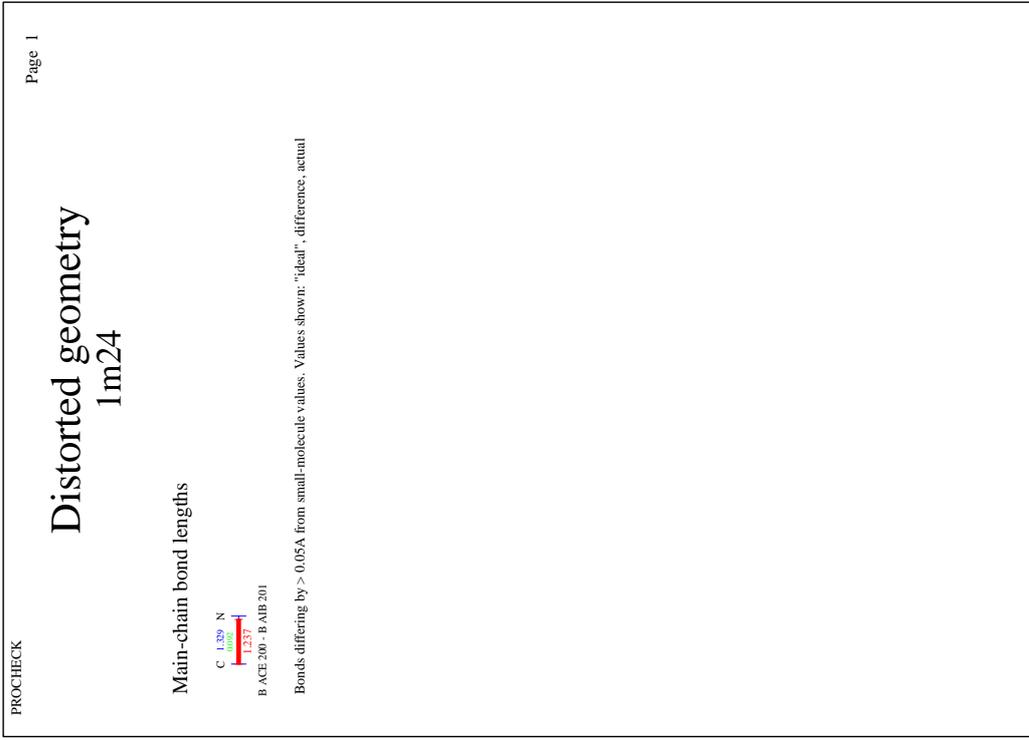


Abbildung E.3: PROCHECK-Ausgabe TRICHOTOXIN.

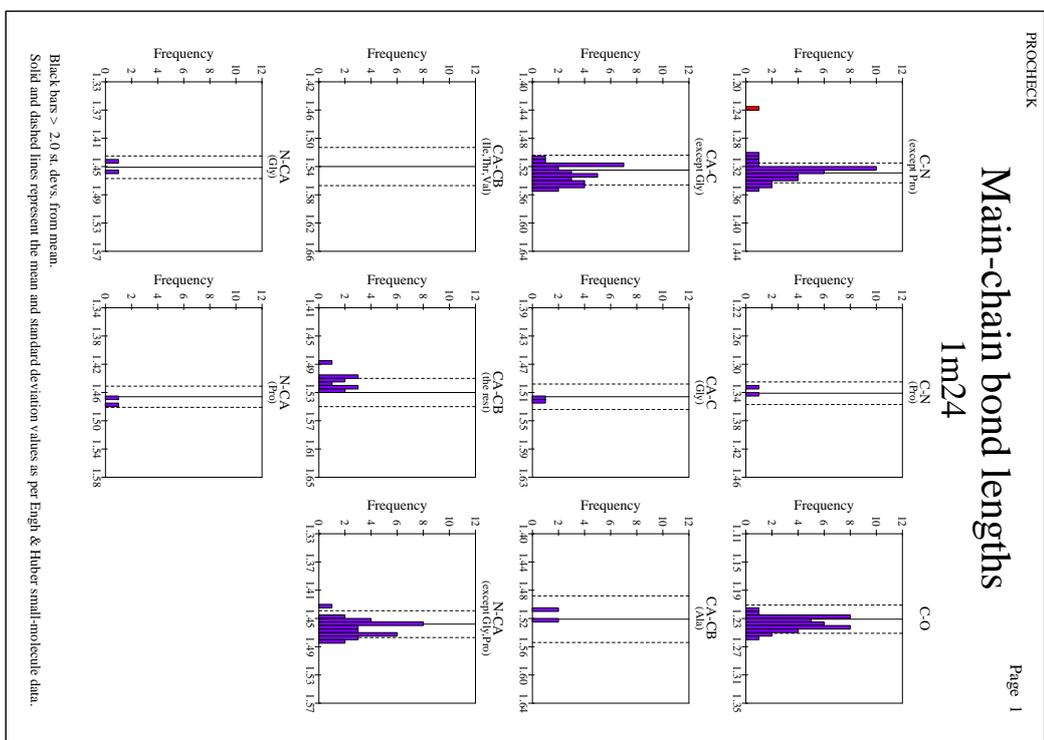


Abbildung E.3: PROCHECK-Ausgabe TRICHOTOXIN (Forts.).

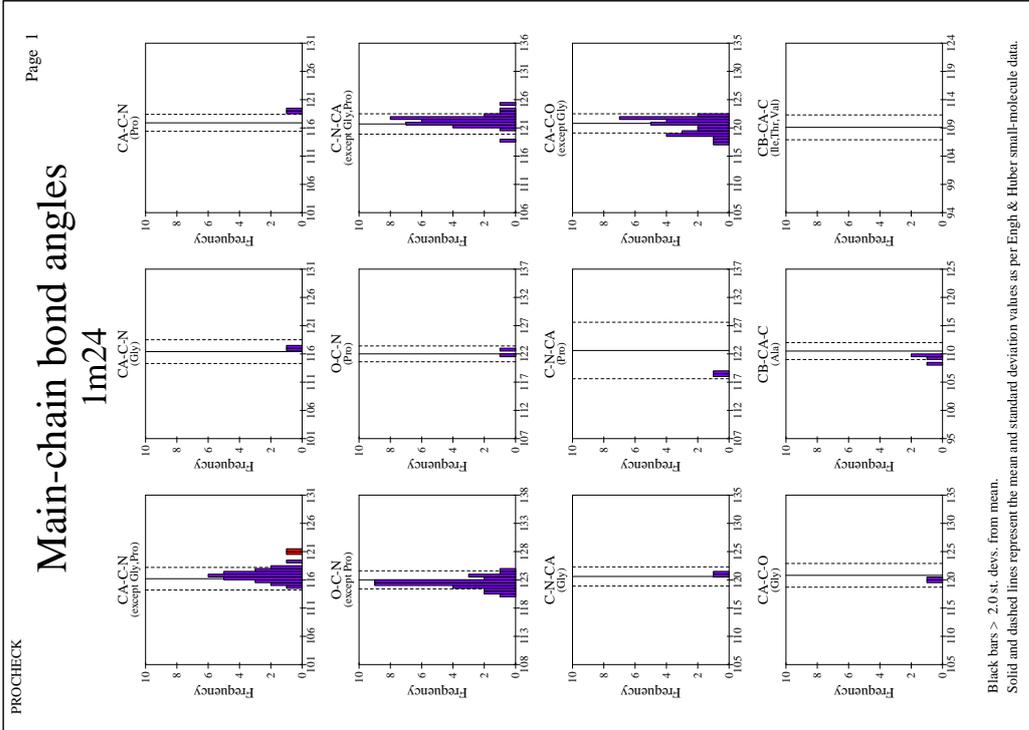
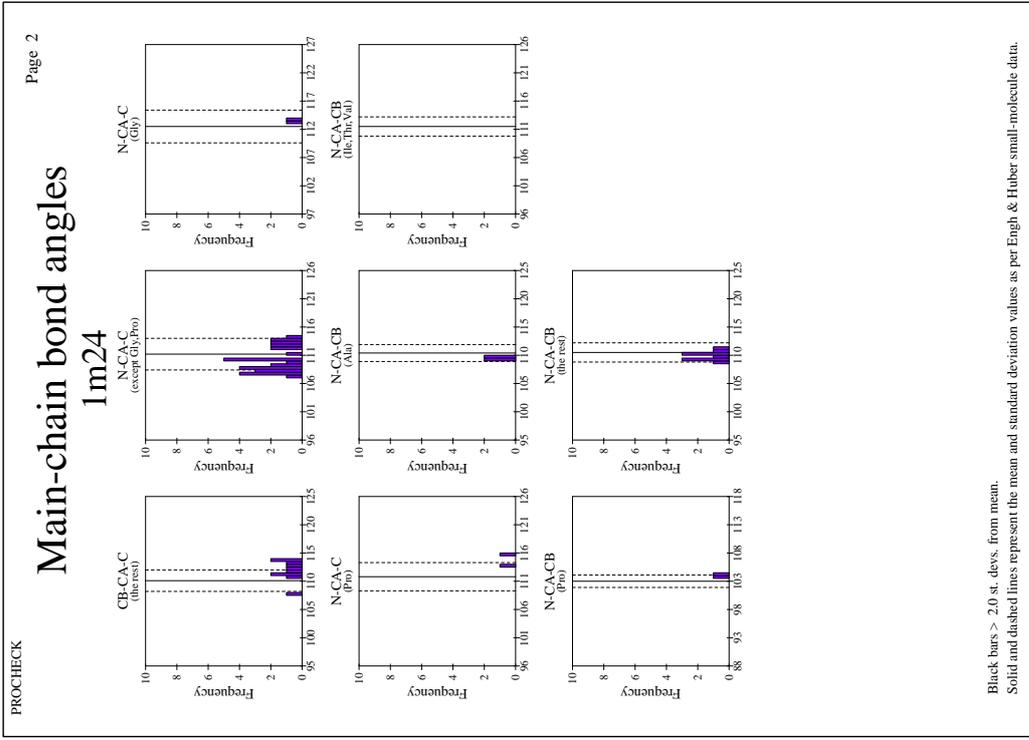


Abbildung E.3: PROCHECK-Ausgabe TRICHOTOXIN (Forts.).

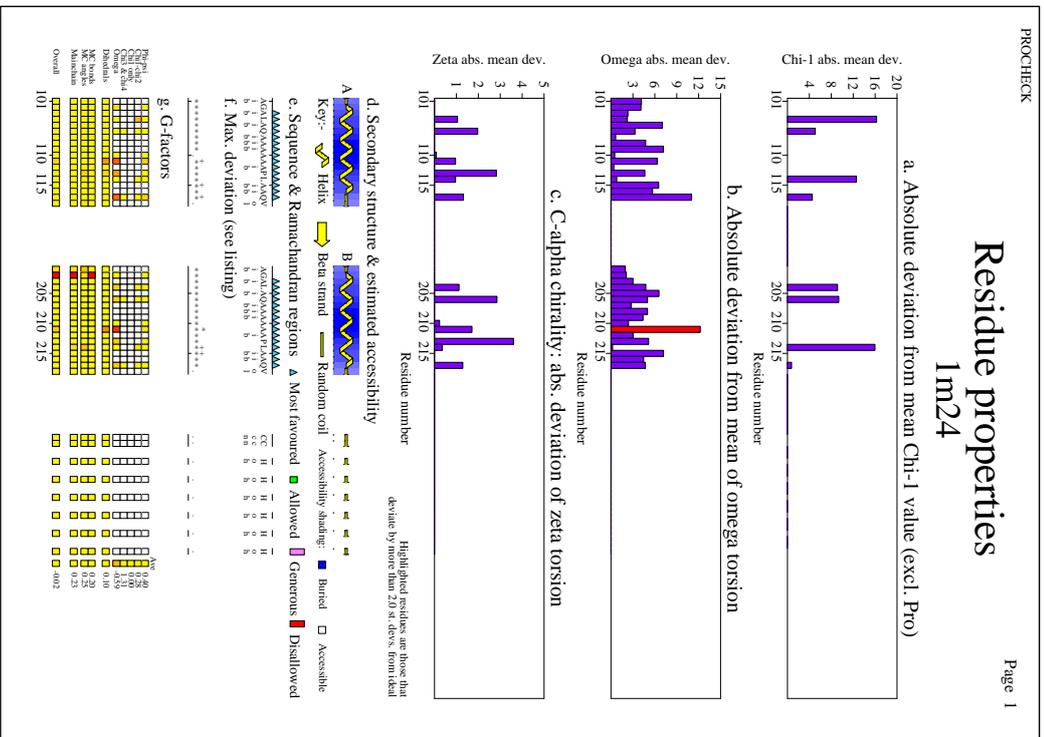


Abbildung E.3: PROCHECK-Ausgabe TRICHOTOXIN (Forts.).

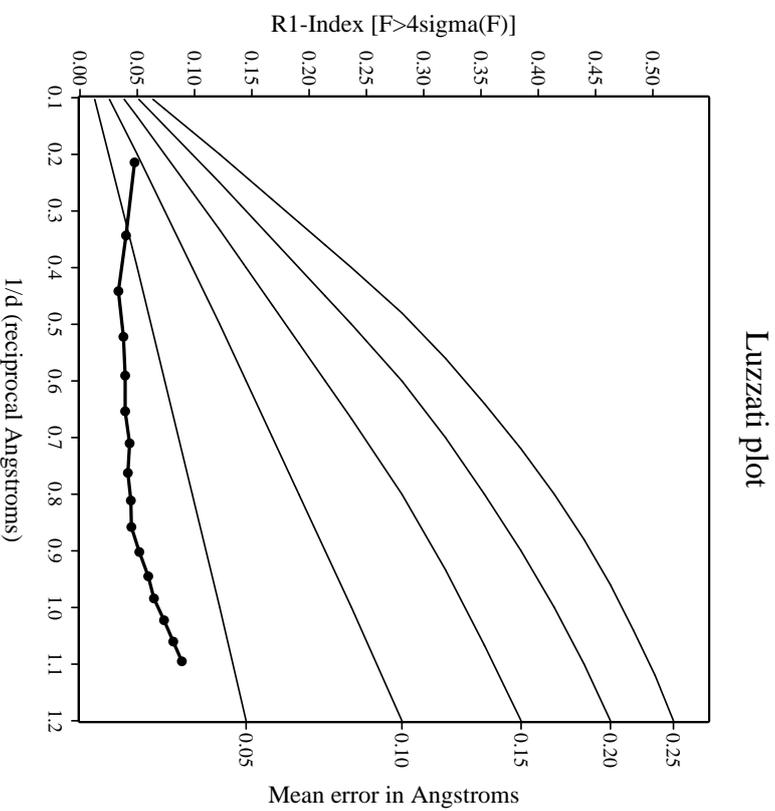


Abbildung E.4: SHELLXPRO-Ausgabe TRICHOTOXIN.

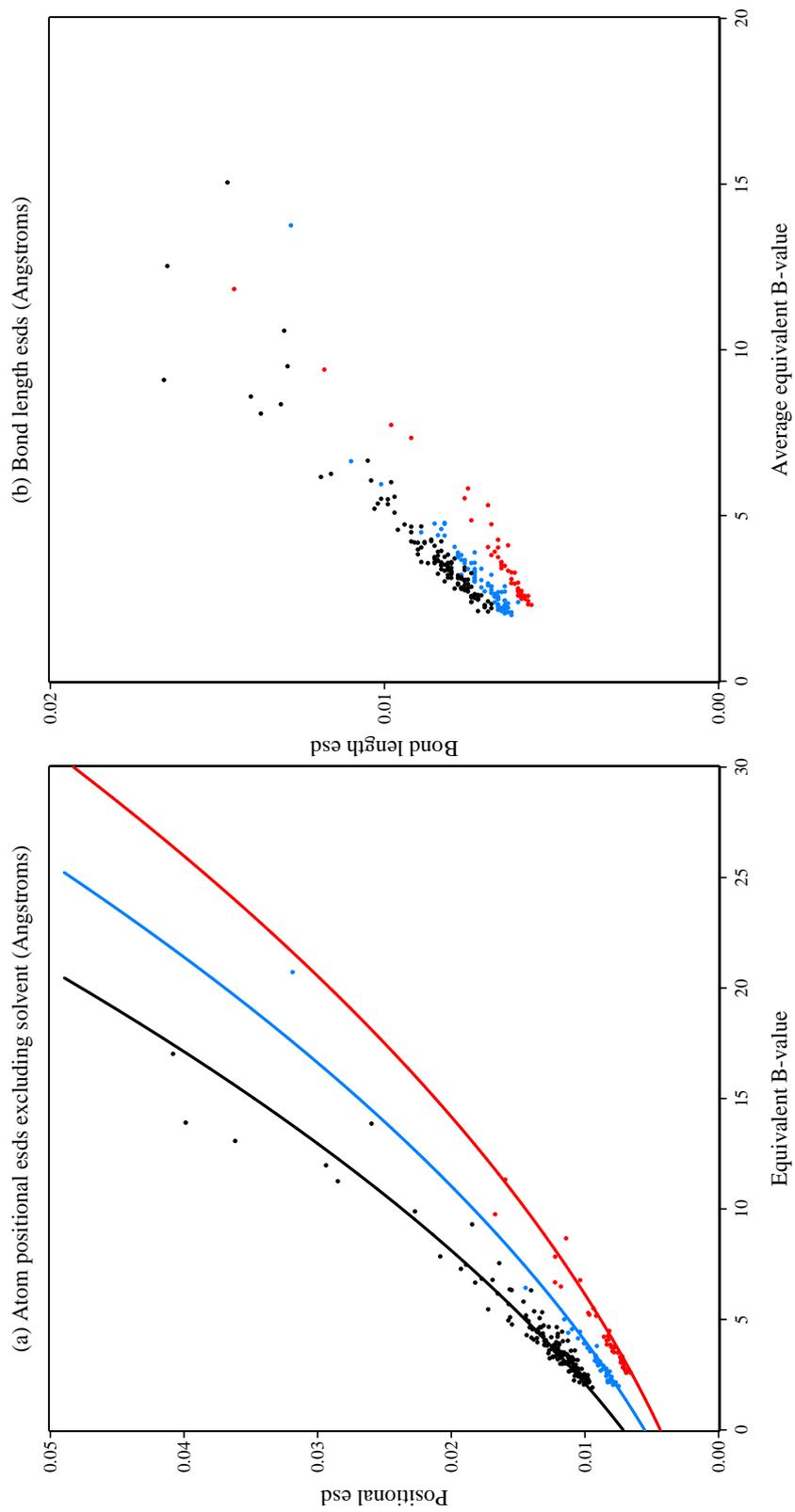


Abbildung E.4: SHELXPRO-Ausgabe TRICHOTOXIN (Forts.).

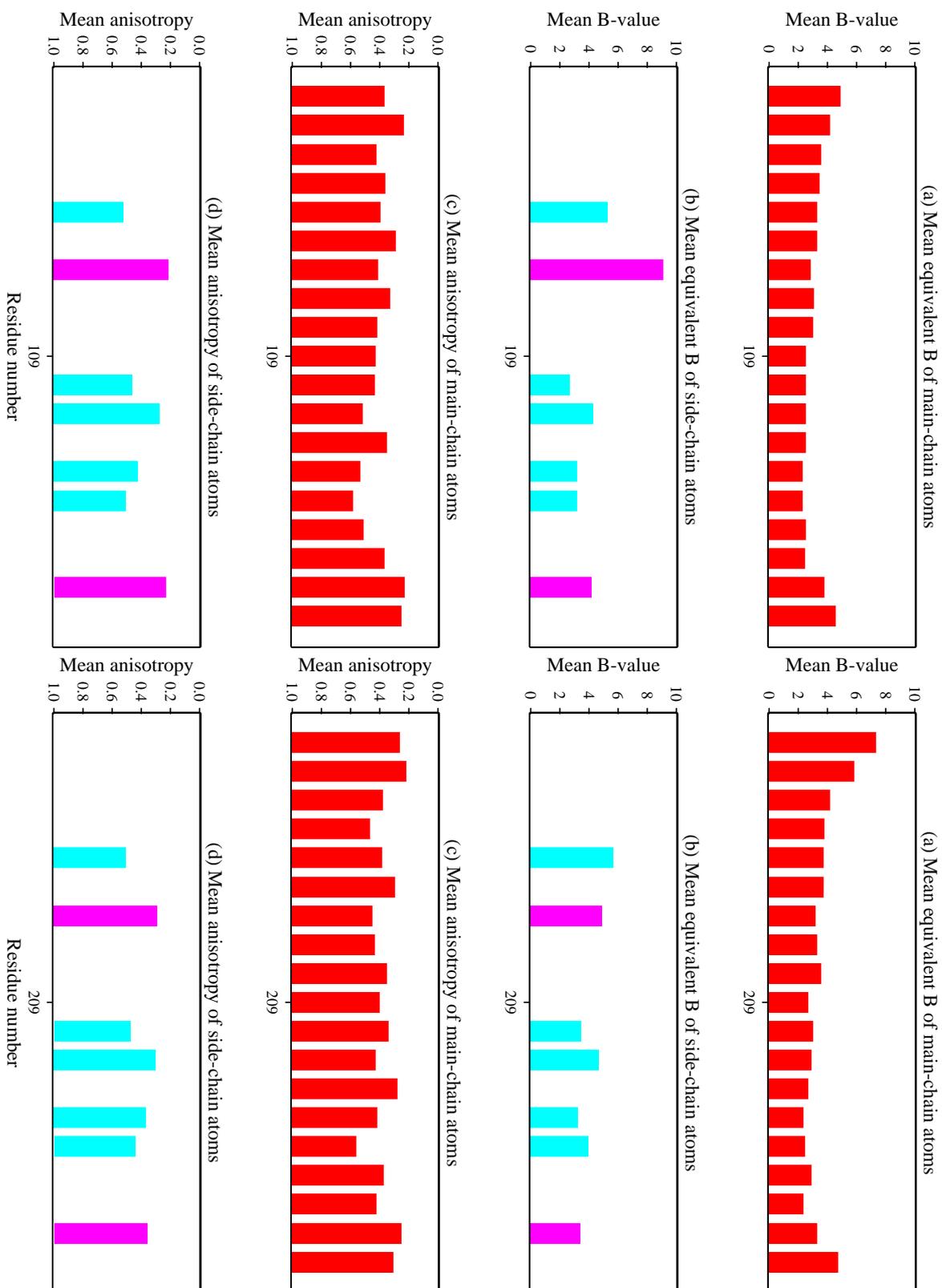


Abbildung E.4: SHELLXPRO-Ausgabe TRICHOTOXIN (Forts.).

Weitere Arbeiten

Experimentelle Elektronendichtebestimmung von Ala-Pro-Ala – topologische Analyse intra- und intermolekularer Eigenschaften

Anhand des Vergleichs bindungstopologischer und atomarer Eigenschaften von L-Ala-L-Pro-L-Ala mit vier weiteren Tripeptidmolekülen, wurde die durchschnittliche Streuung bestimmt, die für ähnliche Moleküle aus Multipolverfeinerungen verschiedener Experimente resultiert: $0,06 \text{ e}\text{\AA}^{-3}/2,4\%$ für die Elektronendichte ρ am bindungskritischen Punkt, $2,94 \text{ e}\text{\AA}^{-5}/14,2\%$ für den Laplacian $\nabla^2\rho$, $0,56 \text{ \AA}^3/5,3\%$ für atomares Volumen V und $0,08 \text{ e}/16,1\%$ für atomare Ladung Q . Hierzu wurden Durchschnittswerte der Standardabweichungen aus Mittelung von in Nächster Nachbar Näherung gleichen Bindungen und Atomen als Streuungsmaß herangezogen. Nach analogen topologischen Analysen der Invariomdichte und *ab initio*-Dichte von L-Ala-L-Pro-L-Ala, wurden für die Schwankungen relativ zum multipolverfeinerten Modell ähnliche Werte gefunden. Zusätzliche Vergleiche mit literaturbekannten Studien weiterer Peptide bestätigten die Ergebnisse.

Die Wasserstoffbrücken von L-Ala-L-Pro-L-Ala wurden anhand der acht Koch&Popelier-Parameter (J. Phys. Chem. 1999, 99, 9747-9754) und den Espinosa-Relationen (Chem. Phys. Lett. 1998, 285, 170-173; Acta Cryst. 1999, B55, 563-572; J. Chem. Phys. 2002, 117, 5529-5542) detailliert untersucht. Die verfeinerte Multipoldichte erfüllte fast alle der acht theoretisch abgeleiteten Parameter für das Vorhandensein einer Wasserstoffbrücke, ebenso wurden die exponentiellen Relationen bindungstopologischer Parameter an diesem Molekül bestätigt. Es wurde aber auch gefunden, daß die wechselwirkungsfreien Dichten des sphärischen und Invariom-Modells etliche der Parameter für Wasserstoffbrückenbindung erfüllten und den Espinosa-Relationen folgten.

(Veröffentlicht in: Kalinowski u. a., *Acta Cryst.* B63 (2007), 753–767.)

Untersuchungen zur Transferierbarkeit

Später wurden die Studien zu Transferierbarkeit und Reproduzierbarkeit topologischer Parameter aus experimentell bestimmten Elektronendichten auf eine breitere Basis gestellt. Nach äquivalentem Schema wurden acht Tripeptidmoleküle

untersucht und die Analyse auf theoretische Daten aller 20 Tripeptide Ala-Xxx-Ala ausgedehnt. Als Grenzen der Reproduzierbarkeit experimentell bestimmter QTAM-Größen, wurden Werte von $0,07 \text{ e}\text{\AA}^{-3}$ für ρ , $3,3 \text{ e}\text{\AA}^{-5}$ für $\nabla^2\rho$, $0,4 \text{ \AA}^3$ für V und $0,12 \text{ e}$ für Q gefunden. Die Transferierbarkeitslimits ergaben sich zu $0,09 \text{ e}\text{\AA}^{-3}$ für ρ , $2,8 \text{ e}\text{\AA}^{-5}$ für $\nabla^2\rho$, $0,7 \text{ \AA}^3$ für V und $0,11 \text{ e}$ für Q . Analoge Betrachtungen für theoretische Dichten lieferten $0,005 \text{ e}\text{\AA}^{-3}$ für ρ , $0,14 \text{ e}\text{\AA}^{-5}$ für $\nabla^2\rho$, $0,3 \text{ \AA}^3$ für V und $0,02 \text{ e}$ für Q .

(Veröffentlicht in: Grabowsky u. a., *Acta Cryst.* B65 (2009), 749–756.)

Erstmalige Elektronendichteuntersuchung an einem C_{70} -Fulleren: $\text{C}_{70}(\text{C}_2\text{F}_5)_{10}$

Der Versuch, hochaufgelöste Datensätze von C_{70} -Fullerenen zu gewinnen, erwies sich als außerordentlich schwierig: Sieben verschiedene Derivate, zur Verfügung gestellt von Prof. Dr. Sergej Troyanov (Lomonosov Universität Moskau, Humboldt Universität Berlin), wurden in mehreren Meßreihen an Synchrotronen untersucht. Fast alle wiesen zu geringe Streukraft auf oder zeigten einen hohen Anteil fehlgeordneter Atome, was eine experimentelle Ladungsdichteanalyse unmöglich machte.

Eine Struktur ($\text{C}_{70}(\text{C}_2\text{F}_5)_{10}$) konnte am Meßplatz F1 des HASYLAB/DESY bei 100 K bis zu einer Auflösung von $d = 0,5 \text{ \AA}$ gemessen werden, die zudem nur eine fehlgeordnete Perfluorethylengruppe aufwies. Durch Multipolverfeinerung ($R(F) = 4,03 \%$) wurde die erste experimentell bestimmte Elektronendichte eines C_{70} -Derivats gewonnen. Untersuchungen der bindungstopologischen Eigenschaften und Bestimmung atomarer Ladungen und Volumina wurden durchgeführt und mit Ergebnissen theoretischer Rechnungen an diesem Molekül und freiem C_{70} verglichen.

Statt der acht unterscheidbaren Bindungen des freien C_{70} wurde eine kontinuierliche Verteilung von Atomabständen und Werten der Elektronendichte an den bindungskritischen Punkten gefunden, was gut mit theoretischen Resultaten übereinstimmt. Bindungen vom Käfig zu den Perfluorethylengruppen und innerhalb dieser sind signifikant höher mit Elektronendichte populiert. An Integrationen atomarer Volumina und Ladungen wird der elektronenziehende Effekt der Fluoratome deutlich, ein Befund ist beispielsweise der Transfer von $0,62 \text{ e}$ aus dem Käfig auf jede C_2F_5 -Gruppe.

(Veröffentlicht in: Kalinowski u. a., *Z. Naturforsch.* 65b (2010), 1–7.)

Veröffentlichungen

Artikel

- [1] DITTRICH, B. ; HÜBSCHLE, C.B. ; MESSERSCHMIDT, M. ; KALINOWSKI, R. ; GIRNT, D. ; LUGER, P.: The invariom model and its application: refinement of D,L-serine at different temperatures and resolution. In: *Acta Cryst.* A61 (2005), Nr. 3, S. 314–320
- [2] HOFMANN, A. ; KALINOWSKI, R. ; LUGER, P. ; VAN SMAALEN, S.: Accurate charge density of the tripeptide Ala-Pro-Ala with the maximum entropy method (MEM): influence of data resolution. In: *Acta Cryst.* B63 (2007), Nr. 4, S. 633–643
- [3] KALINOWSKI, R. ; DITTRICH, B. ; HÜBSCHLE, C.B. ; PAULMANN, C. ; LUGER, P.: Experimental charge density of L-alanyl-L-prolyl-L-alanine hydrate: classical multipole and invariom approach, analysis of intra- and intermolecular topological properties. In: *Acta Cryst.* B63 (2007), Nr. 5, S. 753–767
- [4] GRABOWSKY, S. ; KALINOWSKI, R. ; WEBER, M. ; FOESTER, D. ; PAULMANN, C. ; LUGER, P.: Transferability and reproducibility in electron-density studies - bond-topological and atomic properties of tripeptides of the type L-alanyl-X-L-alanine. In: *Acta Cryst.* B65 (2009), Nr. 4, S. 488–501
- [5] DITTRICH, B. ; WEBER, M. ; KALINOWSKI, R. ; GRABOWSKY, S. ; HÜBSCHLE, C.B. ; LUGER, P.: How to easily replace the independent atom model - the example of bergenin, a potential anti-HIV agent of traditional Asian medicine. In: *Acta Cryst.* B65 (2009), Nr. 6, S. 749–756
- [6] KALINOWSKI, R. ; WEBER, M. ; TROYANOV, S.I. ; PAULMANN, C. ; LUGER, P.: A First Experimental Electron Density Study on a C70 Fullerene: C₇₀(C₂F₅)₁₀. In: *Z. Naturforsch.* 65b (2010), Nr. 1, S. 1–7
- [7] DITTRICH, B. ; BOND, C.S. ; KALINOWSKI, R. ; SPACKMAN, M.A. ; JAYATILAKA, D. : Revised Electrostatics from Invariom Refinement of the 18-Residue Peptaibol Antibiotic Trichotoxin A50E. In: *Cryst. Eng. Comm.* eingereicht

Vorträge

- [1] KALINOWSKI, R. ; GRABOWSKY, S. ; FÖRSTER, D. ; PAULMANN, C. ; MORGENROTH, W. ; WEBER, M. ; LUGER, P. : Comparative Charge Density Study

on Tripeptides of the Type ALA-XXX-ALA. Auf: *5th European Charge Density Meeting in Conjunction with DFG 1178 Annual Meeting*, Gravedona/Italien, Juni 2008

- [2] KALINOWSKI, R. ; LUGER, P. ; MEBS, S. ; WEBER, M. ; TROYANOV, S.I. ; CHECINSKA, L. ; MORGENROTH, W. ; PAULMANN, C. : Atomic and bond topological properties of some C₆₀ and C₇₀ fullerenes. Auf: *16. Jahrestagung der Deutschen Gesellschaft für Kristallographie*, Erlangen, März 2008
- [3] KALINOWSKI, R. ; MEENTS, A. ; LUGER, P. : Asphärische Verfeinerung von HEW-Lysozym: Vorläufige Ergebnisse. Auf: *16. Jahrestagung der Deutschen Gesellschaft für Kristallographie*, Erlangen, März 2008

Poster

- [1] CHECINSKA, L.; FÖRSTER, D. ; RÖDEL, E. ; SCHEINS, S. ; KALINOWSKI, R. ; MEBS, S. ; HÜBSCHLE, C.B. ; DITTRICH, B. ; MORGENROTH, W. ; PAULMANN, C. ; LUGER, P.: Recent experimental and methodical advances in the charge density studies on tripeptides of the type Ala-Xxx-Ala. Auf: *DGK-Workshop AK14 Computational Crystallography*, Aachen, Februar 2006
- [2] KALINOWSKI, R. ; DITTRICH, B. ; HÜBSCHLE, C.B. ; PAULMANN, C. : LUGER, P. : Topologische Analyse von Wasserstoffbrückenbindungen im Multipol-, Invariom- und IAM-Modell. Auf: *15. Jahrestagung der Deutsche Gesellschaft für Kristallographie*, Bremen, März 2007
- [3] KALINOWSKI, R. ; MEENTS, A. ; WAGNER, A. ; LUGER, P. : Application of the Aspherical Scattering Formalism on the Refinement of Macromolecules. Auf: *21st Congress and General Assembly of the International Union of Crystallography*, Osaka/Japan, August 2008 (*Acta Cryst.* (2008). A64, C570–C571)
- [4] LUGER, P. ; WEBER, M. ; KALINOWSKI, R. ; GRABOWSKY, S. ; DITTRICH, B. ; HÜBSCHLE, C.B. : Teaching how to simply replace the independent atom model - the example of Bergenin. Auf: *21st Congress and General Assembly of the International Union of Crystallography*, Osaka/Japan, August 2008 (*Acta Cryst.* A64 (2008), C641-C642)
- [5] LUGER, P. ; KALINOWSKI, R. ; MEBS, S. ; WEBER, M. ; TROYANOV, S.I. ; CHECINSKA, L. ; MORGENROTH, W. ; PAULMANN, C. : Intra- and intermolecular electron-density properties of fullerene derivatives: first C₇₀ examples. Auf: *21st Congress and General Assembly of the International Union of Crystallography*, Osaka/Japan, August 2008 (*Acta Cryst.* A64 (2008), C117)