

NMR Protein Structure Determination in a Structural Genomics Context – Developments, Methods and Applications

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

eingereicht im Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin

vorgelegt von

Diplom Biochemiker Christoph Brockmann

aus Herdecke / Ruhr

Dezember 2005

1. Gutachter: Prof. Dr. H. Oschkinat
NMR-unterstützte Strukturforschung
Forschungsinstitut für Molekulare Pharmakologie
Campus Berlin-Buch
Robert-Rössle-Str. 10
13125 Berlin
Deutschland

2. Gutachter: Prof. Dr. U. Heinemann
FG Kristallographie
Max-Delbrück-Center für Molekulare Medizin
Robert-Rössle-Str. 10
13122 Berlin
Deutschland

Tag der Disputation am 15. 3. 2006

Acknowledgements

The work presented in this thesis was carried out in the department of NMR-supported structural biology at the Forschungsinstitut für Molekulare Pharmakologie (FMP) in Berlin under the supervision of Hartmut Oschkinat with financial support from the DFG Graduiertenkolleg #788 "Hydrogen bonding and hydrogen transfer". It would not have been possible to complete this thesis without the advice and support of many people.

I would like to thank Hartmut Oschkinat for the possibility to work in his lab, the outstanding environment and his constant support.

Annette Diehl, Kristina Rehbein and Martina Leidert for the expression and purification of all the protein samples. I would like to thank Volker Sievert and Konrad Büssow for the cloning of the BAG-domain construct and Bernhard Korn for the cloning of the B8 subunit.

Dietmar Leitner, Peter Schmieder and Dirk Labudde for introducing me to protein NMR spectroscopy, the assignment of protein resonances and the calculation of structures. I am grateful to Ronald Kühne for his continuous support and his help with all aspects of structure calculation and computer modelling of proteins.

Andrea Steuer for help with the bureaucracy and the administrative aspects of the work.

Thomas Sandmann, Ronald Kühne and Holger Strauss for reading the manuscript and helpful discussions on the structure of this thesis. I also wish to thank all other members of the Oschkinat group for the nice atmosphere, creative discussions and support in all the big and small questions of everyday work.

My wife Mirjam and my Mother for their motivation and patience.

Table of Content

1	Introduction	7
1.1	Structural Genomics	7
1.1.1	Aims of Structural Genomics	8
1.1.2	NMR in Structural Genomics	10
1.2	Protein NMR.....	13
1.2.1	Introduction to Protein Solution NMR Methods.....	13
1.2.2	Structure Calculation from NMR Data	24
1.3	Improvement of NMR Structure Determination by Automation and Integration	26
1.3.1	Sample Preparation.....	26
1.3.2	Data Acquisition.....	27
1.3.3	Assignment / Structure Calculation	29
1.4	Quality Assessment of NMR-Structures	32
1.4.1	Accuracy	33
1.4.2	Precision	35
1.4.3	Protein Normality	36
1.4.4	Discussion	38
1.5	Aims of this Thesis	38
1.6	References for Chapter 1.....	39
2	A New Method for the Assignment of Amino Acid Types	44
2.1	Difficulties with the assignment of large proteins	44
2.2	Assignment Based on Labelling Pattern.....	46
2.2.1	Principle of the Method.....	46
2.2.2	Results on α -Spectrin SH3	50
2.2.3	Discussion and Outlook.....	59
2.3	References for Chapter 2.....	60
3	Structures of Small Domains and Proteins.....	62
3.1	SODD – BAG-Domain.....	62
3.1.1	SODD biology.....	62
3.1.2	Structure of the SODD-BAG Domain	64
3.1.3	Computer Model of the SODD-BAG/HSP70 Complex	68
3.1.4	Discussion	72
3.2	Complex I – Subunit B8	73
3.2.1	Introduction to Complex I	73
3.2.2	Structure of CI-B8.....	77

3.2.3	Comparison to other Structures.....	83
3.2.4	Biophysical Experiments	85
3.2.5	Discussion	87
3.3	Zinc Finger Domain from the Human Hypothetical Protein BC018415.....	89
3.3.1	Available Biological Information.....	89
3.3.2	Strategy for Structure Determination	90
3.3.3	Structure of the An1-like Zinc Finger	92
3.3.4	Quality of the Obtained Ensemble	93
3.4	Towards an Efficient NMR Structure Calculation Strategy	98
3.4.1	Comparison of the Structure Determination Processes	98
3.4.2	A Practical Strategy for Structure Calculation.....	101
3.4.3	Considerations on the Additional Experiments.....	105
3.4.4	Conclusion	106
3.5	References for Chapter 3	106
4	Experimental Procedures.....	111
4.1	NMR Experiments.....	111
4.1.1	NMR Spectra Used in Chapter 2	112
4.1.2	Sample and NMR Experiments SODD.....	113
4.1.3	Sample and NMR Experiments: CI-B8	113
4.1.4	Sample and NMR Experiments: An1 like Zinc Finger	114
4.1.5	Relaxation rate measurements.....	114
4.2	Resonance Assignments	114
4.2.1	Assignment of the SODD-BAG domain.....	115
4.2.2	Assignment of the Zinc finger domain.....	115
4.3	Secondary structure Prediction	115
4.4	Structure Calculation	116
4.4.1	Structure Calculation: SODD BAG-Domain.....	116
4.4.2	NOE-Assignment and Structure Calculation: CI-B8	116
4.4.3	NOE-Assignment and Structure Calculation: An1-like Zinc Finger Domain.....	117
4.5	Modelling of the SODD-BAG / HSP70 complex	118
4.6	Confirmation of the Disulfide Bond in CI-B8.....	118
4.7	Surface Accessability of methyl-groups in CI-B8.....	118
4.8	Determination of the Redox Potential of CI-B8.....	119
4.9	References for the Experimental Section.....	119
5	Summary	122
6	Appendix	125
6.1	Protein Sequences.....	125

6.1.1	SODD-BAG Domain.....	125
6.1.2	CI-B8	125
6.1.3	An1-like Zinc Finger.....	125
6.2	Chemical Shift Assignments	126
6.2.1	Assignment of the SODD-BAG Domain.....	126
6.2.2	Assignment of the CI-B8	129
6.2.3	Assignment of the An1-like Zinc Finger.....	132
6.3	Publications	134
6.4	Structures	134
6.5	CV	135

5 Summary

The main topic of this thesis is the application of NMR for protein structure determination in the context of a structural genomics project. The structures of three small protein domains are presented within their biological context together with additional experiments that relate the structures to the possible functions of the proteins.

After an introduction to the field in chapter 1 covering recent developments in structural genomics and protein structure determination by NMR, a method that allows the assignment of amino acid types in larger proteins is presented in chapter 2. This method relies on different ^{13}C -labelling patterns obtained from biosynthetic labelling with either 1,3- ^{13}C -glycerol or 2- ^{13}C -glycerol. First results obtained on a small test protein, the SH3-domain from α -spectrin, are reported. It is possible to assign all N-H crosspeaks from this protein to the correct pair of amino acid-types. Furthermore, unambiguous sequence specific assignments for 19 of the 62 amino acids of the α -spectrin SH3 domain could be obtained using this method.

In chapter 3, the structures of three small protein domains are presented. The BAG-domain from "Silencer of Death Domains" (SODD) shows a three helix bundle and is a representative of a new shorter subclass of BAG-domains. In a homology model of the complex with its interaction partner HSP70, new subclass-specific interactions were identified. The structure of the oxidized Complex I subunit B8 (CI-B8), presented in the following section, is the first structure at atomic resolution obtained from a Complex I subunit. It is a single domain protein showing a mixed parallel anti-parallel β -sheet accompanied by three α -helices, that is identified as a thioredoxin-fold. The homology to thioredoxins includes the overall position of the disulfide bond. The redox-potential that was measured to be $-251,6$ mV is comparable to that of thioredoxins. This suggests a thioredoxin-like function of subunit B8 within Complex I. The third protein described in this thesis is an An1-like Zinc finger domain from the hypothetical protein

BC018415. Apart from two very short stretches of β -sheet, it does not display any regular secondary structure, while the core of the protein is dominated by its two zinc centers.

In the last section of chapter 3, the structure determination process of the three proteins is compared and an optimized strategy of structure calculation and refinement is presented. This strategy allows taking advantage of the increased speed of automated NOE-assignment algorithms without sacrificing the quality of the obtained structures.

5. Zusammenfassung

Im Mittelpunkt dieser Arbeit steht die Anwendung der Proteinstrukturbestimmung mittels NMR im Rahmen eines Structural Genomics Projektes. Die Strukturen von drei kleinen Proteindomänen und Proteinen werden zusammen mit ihrem jeweiligen biologischen Hintergrund präsentiert. Darüberhinaus wurden weitere Experimente zur Untersuchung der Struktur-Funktions-Beziehungen durchgeführt und vervollständigen die jeweiligen Abschnitte.

Nach einer kurzen Einführung in das Gebiet der Structural Genomics und der neueren Entwicklungen im Bereich der NMR-Proteinstrukturbestimmung in Kapitel 1 wird in Kapitel 2 eine neue Methode zur Zuordnung von Aminosäure-Typen präsentiert. Diese Methode beruht auf unterschiedlichen ^{13}C -Markierungsmustern, die biosynthetisch entweder durch den Einsatz von 1,3- ^{13}C -Glycerin oder 2- ^{13}C -Glycerin erzeugt werden können. Die Ergebnisse, die mit dieser Methode an einem kleinen Testprotein, der SH3-Domäne aus α -Spektrin, erzielt werden können, sind vielversprechend. So ist es möglich, alle N-H Kreuzsignale des Proteins dem korrekten Paar von Aminosäuretypen zuzuordnen. Darüberhinaus können im Fall von α -Spektrin SH3 mit dieser Methode 19 der 62 im Protein enthaltenen Aminosäuren eindeutig zugeordnet werden.

In Kapitel 3 werden die NMR-Strukturen dreier kleiner Domänen und Proteine beschrieben, welche im Rahmen dieser Arbeit bestimmt wurden. Die BAG-Domäne aus dem Protein ‚Silencer of Death Domains‘ (SODD) besteht aus einem Drei-Helix-Bündel und repräsentiert eine kürzere Unterfamilie der BAG-Domänen. Im Modell des Komplexes der SODD BAG-Domäne mit HSP70, welches mittels Homologiemodelling erstellt wurde, konnten neue Interaktionen identifiziert werden, die für die kurze Unterfamilie der BAG-Domänen spezifisch sind. Die Struktur der oxidierten B8-Untereinheit aus Komplex I (CI-B8), welche im darauffolgenden Abschnitt beschrieben wird, zeigt ein gemischtes β -Faltblatt, dessen eine Seite mit drei α -Helices wechselwirkt. Die Struktur der einzigen Proteindomäne zeigt strukturelle Homologie zu Thioredoxinen, und das Redox-Potential von $-251,6$ mV liegt ebenfalls im Bereich der Redox-Potentiale typischer Vertreter der Thioredoxin-Familie. Da außerdem die Disulfid-Brücke von CI-B8 in einem ähnlichen Bereich der Struktur lokalisiert ist, kann davon ausgegangen werden, dass CI-B8 innerhalb von Komplex I eine Thioredoxin-ähnliche Funktion haben könnte. Als drittes wird in diesem Kapitel die Struktur einer An1-ähnlichen Zink-Finger-Domäne aus dem hypothetischen Protein BC018415 beschrieben. Diese zeigt nur in zwei sehr kurzen β -Faltbättern eine reguläre Sekundärstruktur und wird von den beiden Zink-Komplexen im Kern dominiert.

Den Abschluss von Kapitel 3 bildet eine vergleichende Betrachtung des Strukturbestimmungsprozesses der drei Proteine. Daraus wird eine Strategie zur Berechnung von NMR Strukturen entwickelt, welche die Geschwindigkeitsvorteile neuer automatischer NOE-Zuordnungsroutinen nutzt, ohne jedoch von den Qualitätsstandards manueller NMR-Strukturbestimmung abzuweichen.

6.5 CV

Name	Christoph Brockmann
Date of Birth	22. 12. 1975 in Herdecke (Ruhr)
Nationality	German

Education:

March 2002 – Present	Doctoral thesis on „NMR Protein Structure Determination in a Structural Genomics Context - Developments Methods and Applications“ in the department of NMR-Supported Structural Biology at the Forschungsinstitut für Molekulare Pharmakologie supervised by Prof. Dr. H. Oschkinat.
Feb 2002	Degree: Diplom
June 2001 – Feb 2002	Diploma thesis on „Erprobung neuer Strategien zur Zuordnung von Protein NMR-Spektren an der BAG-Domäne von SODD (Silencer of Death Domains)“ in the NMR division of the german structural genomics project „Protein Structure Factory“, supervised by Prof. Dr. H. Oschkinat.
April 1999 – March 2002	Main studies in biochemistry at the „Freie Universität Berlin“.
March 1999	Degree: Vordiplom
Okt 1996 – March 1999	Undergraduate studies in biochemistry at the „Eberhard Karls Universität Tübingen“