

1. Einleitung

1.0. Die Bedeutung der Auflösung bei Kristallstrukturanalysen

Die überwiegende Mehrzahl kristallographischer Studien von Protein- bzw. Nukleinsäurestrukturen, basieren auf experimentellen Daten bis zu einer maximalen Auflösung von etwa 3.5 Å bis 2.0 Å. Die relativ schlechten Diffraktionseigenschaften makromolekularer Einkristalle resultieren aus der Größe und der hohen Flexibilität der den Kristall aufbauenden Moleküle. Zusätzlich besitzen diese Kristalle einen Wasseranteil von etwa 30% bis 80%, der von den Packungseigenschaften der Makromoleküle die das dreidimensionale Kristallgitter aufbauen abhängig ist. Bis vor wenigen Jahren waren nur wenige makromolekulare Kristallstrukturen von relativ kleinen, hochgeordneten Molekülen mit einem deutlich geringeren Solvensanteil im Kristall, die bis zu einer atomaren Auflösung von < 1.2 Å diffraktierten, bekannt. Innerhalb der vergangenen Jahre haben wichtige Entwicklungen kristallographischer Techniken sowie verbesserte Detektortechnologien und die wachsende Verfügbarkeit hochintensiver Synchrotronstrahlungsquellen zu einem starken Anwachsen atomar aufgelöster Protein- und Nukleinsäurestrukturen geführt (Dauter *et al.*, 1995). Bei atomar aufgelösten Strukturen übersteigt die Anzahl der gemessenen Reflexe die Zahl der Parameter, die für die Beschreibung des dreidimensionalen Strukturmodell benutzt werden um ein Vielfaches. Dabei bedingt der Grad der Überbestimmung die erhöhte Genauigkeit des Strukturmodells. Durch die bei der Verfeinerung atomar aufgelöster Strukturmodelle verwendeten speziellen Protokolle ist man in der Lage, ein detaillierteres Bild der geometrischen und konformationellen Eigenschaften der Makromoleküle zu erzeugen (Sheldrick & Schneider, 1997). Die speziellen Eigenschaften dieser Protokolle beruhen hauptsächlich auf der Reduzierung stereochemischer Einschränkungen "restraints" und der Beschreibung atomarer, anisotroper Temperaturfaktoren während des Verfeinerungsprozesses. Die auf dieser Grundlage gewonnenen Strukturmodelle erweitern wiederum die Datenbasis hochaufgelöster Strukturmodelle, die zu einer stetigen Verbesserung stereochemischer Zielwerte führen. Diese Zielwerte werden für die Verfeinerung und Validierung geringer aufgelöster Strukturmodelle benötigt. Zusätzlich ermöglichen atomar aufgelöste Strukturmodelle neue Einblicke in die Solvens-, Kationen- und Ligandenstruktur von Makromolekülen, wie auch der Beobachtung mikroskopischer Protonierungszustände von Protonendonoren bzw. -akzeptoren innerhalb von Proteinstrukturen die für daß Verständnis der biologischen Funktionen dieser Moleküle wichtig sein können.

1.1. Der Akzeptorstamm der tRNA^{Ala} aus *Escherichia coli*

1.1.1. Nukleinsäuren

Nukleinsäuren sind lineare Polymere, die aus monomeren Untereinheiten, den Nukleotiden, aufgebaut sind. Nukleotide bestehen aus drei Teilfragmenten, Zucker, aromatischer Heterozyklus und dem Phosphat. In der DNA ist der Zucker, die Desoxyribose, ein furanoides System und ist über eine β -glykosidische Bindung zwischen C1' und N9 mit einer der vier heterozyklischen Basen Adenin, Guanin, Thymin oder Cytosin verbunden. Bei der RNA ist das Thymin durch Uracil ersetzt, dessen N1 kovalent mit dem anomeren C1' des Zuckers Ribose verknüpft ist. Die aromatischen Heterozyklen von Adenin und Guanin werden durch 2,6-substituierte Purine (Imidazo[4,5-d]pyrimidin), die von Thymin, Uracil und Cytosin durch 2,4,5-substituierte Pyrimidine (1,3-Diazin) gebildet (siehe **Abschnitt 7.3**). Durch die Phosphorylierung der 3'-OH und 5'-OH Gruppe des Zuckers wird die Polymerenkette aufgebaut.

Der bahnbrechende Durchbruch, der mit der Beschreibung der basengepaarten doppelhelikalen Struktur der DNA durch Watson und Crick im Jahre 1953 gelang, eröffnete den Weg für die Entwicklung der modernen Molekularbiologie (Watson & Crick, 1953). Die Basis für diese Entwicklung ist die Beschreibung der Struktur des genetischen Materials, der DNA (Saenger, 1984). Die hauptsächlichen Konformationstypen der DNA sind die A-, B-, und Z-Form. Die wichtigsten Merkmale zur Unterscheidung dieser Typen sind in **Tab. 1.1** dargestellt. Native DNA kommt in der B-Form vor. Die Ausbildung der anderen DNA-Konformationen ist hauptsächlich von den Umgebungsbedingungen abhängig, wie etwa der Salz- und Alkoholkonzentration, welche die Wasseraktivität beeinflussen. Hinzu kommen noch Sequenzeigenschaften, die auch die Konformation speziell bei der Ausbildung von Z-DNA beeinflussen.

Ribonukleinsäuren (RNA), die sich durch die zusätzliche 2'-OH Gruppe des Zuckers und der Base Uracil als Analogon zur DNA-Base Thymin von der DNA unterscheiden, kommen je nach ihrer biologischen Funktion in unterschiedlichen Strukturen vor. Dabei können sie doppelhelikale wie auch globuläre Formen annehmen. Beispiele für doppelhelikale Bereiche findet man in tRNA (Kim *et al.*, 1974), 5S-rRNA (Correll *et al.*, 1997) und auch in mRNA (Proudfoot & Brownlee, 1974). Dabei wird fast immer die A-Form ausgebildet, deren Konformation in doppelhelikalen, basengepaarten RNA-Strukturen energetisch eindeutig bevorzugt ist. Hierbei ist die Zuckerwellung, die Konformation des furanoiden Ringes, im C3'-endo Bereich des Konformationsraumes der Ribose lokalisiert. Interessant ist, daß die Fähigkeit zur Ausbildung globulärer Strukturen von der

einen zusätzlichen C2'-OH Gruppe bestimmt wird. Der Grund dafür ist die einzigartige Fähigkeit der C2'-OH Gruppe, als Donor von Wasserstoffbrückenbindungen im Rückgrat komplexe Tertiärstrukturen zu stabilisieren. Die Ausbildung von Nicht-Watson-Crick Basenpaarwechselwirkungen stellt ein weiteres stabilisierendes strukturelles Merkmal dar (Cate *et al.*, 1996). Die strukturanalytische Aufklärung komplexer RNA-Strukturen bei oft begrenzter Auflösung, wie auch die komplementäre Untersuchung kleinerer, doppelhelikaler Fragmente bei hoher Auflösung bildet somit eine Grundlage für das Verständnis von biologisch relevanten Prozessen der RNA-Proteinwechselwirkung.

Tab. 1.1 Strukturelle Eigenschaften kanonischer DNA-Formen (Saenger, 1984)

	A	B	Z
Gangrichtung	rechtshändig	rechtshändig	linkshändig
Helikale Symmetrie	11 ₁	10 ₁	6 ₅ ¹
Helikaler Twist	32.7°	36.0°	-30.0°
Steighöhe pro Windung	28.2 Å	33.8 Å	45.0 Å
Steighöhe pro Bp²-Schritt	2.56 Å	3.38 Å	3.7 Å
Inklination³	20°	6°	7°
Große Furche	schmal, tief	weit, tief	flach
Kleine Furche	weit, seicht	schmal, tief	schmal, tief
Zuckerwellung	C3'-endo	C2'-endo	C2'-endo (Pyr. ⁴) C3'-endo (Pur. ⁵)
Glykosidische Bindung	Anti	Anti	Anti (Pyr. ⁴) Syn (Pur. ⁵)

- 1: Symmetrie bezogen auf Dinukleotidschritt
- 2: Basenpaar
- 3: Basenpaar-Verkipfung zur Helixachse
- 4: Pyrimidine
- 5: Purine

1.1.2. Die tRNA-Erkennung durch die Aminoacyl-tRNA Synthetasen

Der erste Schritt bei der Proteinbiosynthese ist die Aktivierung der Aminosäuren unter Verbrauch von ATP zum Aminoacyladenylat. Danach wird die so aktivierte Aminosäure spezifisch an die unbeladene tRNA gebunden. Beide Reaktionen werden von Aminoacyl-tRNA-Synthetasen (aaRS) katalysiert (Schimmel, 1990). Die Übertragung der aktivierten Aminosäure auf ihre spezifische tRNA verlangt die Unterscheidung der in ihrer Tertiärstruktur sehr ähnlichen tRNA-Moleküle. Im

Normalfall gehört zu jeder Aminosäure eine entsprechende aaRS. Ausnahmen bilden zum Beispiel tRNA^{Lys} aus *E. coli*, die in ungeklärter Weise zwei aaRS für diesen Schritt benutzt. *B. subtilis* benutzt die GluRS zur Fehlbeladung von tRNA^{Gln} mit Glu zu Glu-tRNA^{Gln}. Danach modifiziert eine Amidotransferase das Molekül zu Gln-tRNA^{Gln}. Ähnliche Ausnahmen wurden in Chloroplasten, Mitochondrien und Archaeobakterien gefunden (Rogers & Söll, 1995).

Die Aminoacylierung der tRNA mit der Aminosäure führt zu einer Bindung mit einer großen negativen freien Standardenthalpie der Hydrolyse. Das zweite Ergebnis ist die Beladung der tRNA mit der zugehörigen Aminosäure. Ein Korrekturlesen auf Richtigkeit der Aminosäure erfolgt durch die aaRS aber nicht mehr im Ribosom. Der genetische Code, der bei der Proteinbiosynthese zuverlässig in die einzig richtige Polypeptidsequenz übersetzt wird, verlangt also, daß die aaRS hoch spezifisch die Aminosäure, wie auch die dazugehörige tRNA erkennt. Diese Funktion der aaRS wird deshalb als der "zweite genetische Code" bezeichnet (Schimmel *et al.*, 1993).

Ein besonderes Interesse gilt den Determinanten der Erkennung der tRNA-Moleküle. Diese konnten durch gezielte Einzelmutationen der Nukleotidsequenz teilweise ermittelt werden. Obwohl tRNAs im Komplex mit den entsprechenden aaRS eine große Wechselwirkungsfläche von etwa 2500 Å² besitzen (Arnez & Moras, 1997), sind die Haupterkennungssignale in der Antikodonschleife wie auch im Akzeptorstamm lokalisiert. Einige aaRS benötigen die Informationen beider Bereiche, andere hingegen nutzen nur die Nukleotidsequenz im Akzeptorstamm oder nur die Erkennung innerhalb der Antikodonschleife (Arnez & Moras, 1997). Diese Eigenschaften variieren stark unter den verschiedenen aaRS sowie je nach Organismus. Ein einfaches universelles Schema dieser Reaktion ist nicht vorhanden, und die Mechanismen sind aktiver Forschung unterworfen.

1.1.3. Die tRNA^{Ala} aus *Escherichia coli*

Das am Besten im Hinblick auf die spezifische Erkennung durch die aaRS untersuchte System stellt die tRNA^{Ala} aus *Escherichia coli* dar (**Abb. 1.1-A**). Die Alanyl-tRNA-Synthetase (AlaRS) ist eine Klasse-II Synthetase, die lediglich den Akzeptorstamm der tRNA zur Identifizierung benutzt. Das beweisen *in vitro* Beladungsexperimente von Minimalstrukturen der tRNA^{Ala}, den sogenannten Mini- und Mikrohelices (**Abb. 1.1-C**), die trotz der extrem starken Vereinfachung mit einer nur 50-fach reduzierten Effizienz verglichen mit der nativen tRNA^{Ala} aminoacyliert werden können (Francklyn & Schimmel, 1989). Der Grund hierfür ist trotz des Erhalts der Struktur des Akzeptorstamms der Verlust von stabilisierenden Kontakten, die für die Bindung und Aminoacylierung von Bedeutung sind. Protektionsexperimente am tRNA^{Ala}-AlaRS-Komplex

zeigen, daß nur Phosphate des Akzeptorstamms von Position 64-70 geschützt werden (Park & Schimmel, 1988). Mutationsstudien am Akzeptorstamm dieser Moleküle, die im Zusammenhang mit *in vitro* Beladungsexperimenten ausgeführt wurden, führten zu der Hypothese, daß das Vorhandensein des G3•U70 "wobble"-Basenpaars das Hauptidentitätssignal darstellt. Zusätzlich ist dieses Basenpaar in allen untersuchten Organismen konserviert. Verändert man die Sequenz nach G3-C70, A3-U70 oder U3-G70 verliert das Enzym nahezu jegliche Aktivität (Musier Forsyth *et al.*, 1991). Die Analyse der G3•U70- abhängigen Beladung durch die AlaRS führte zur Aufstellung der Hypothese, daß an der freien exozyklischen Aminogruppe von G3, die in die kleine Furche der Akzeptorstammhelix ragt, das Erkennungssignal lokalisiert ist. Dieser Bereich ist weit und seicht und kann leicht von den Seitengruppen der Aminosäurekette der AlaRS erreicht werden.

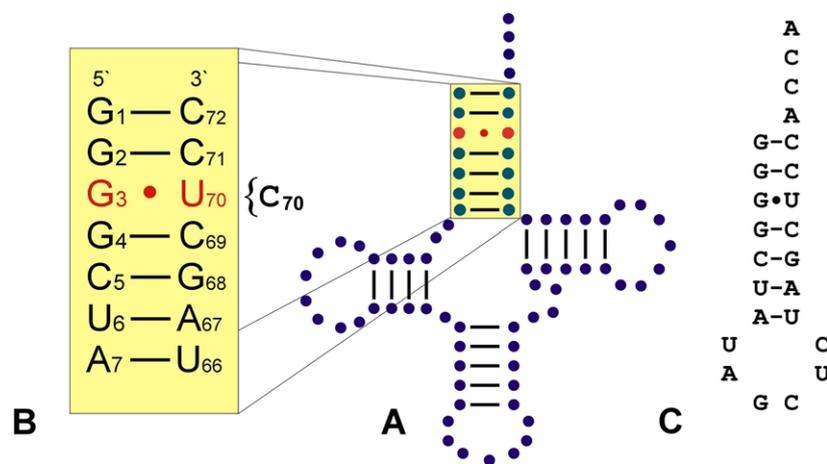


Abb. 1.1 Die tRNA^{Ala} aus *Escherichia coli*. A) Kleeblattstruktur der tRNA^{Ala} mit hervorgehobenem Akzeptorstamm. B) Für die Strukturlösung verwendeter Abschnitt des Akzeptorstamms. C) Ala-Mikrohelix wird von AlaRS aminoacyliert.

Zur Unterstützung dieser Hypothese untersuchte man G3•U70-Analoga um die spezifische sequenzabhängige Wechselwirkung der AlaRS mit dem Akzeptorstamm zu prüfen (Musier Forsyth *et al.*, 1991). Verändert man zum Beispiel G3•U70 in I3•U70, welches isosterisch zum G•U-"wobble"-Basenpaar ist, jedoch nicht über die freie Aminogruppe verfügt, wird dieses Substrat nicht beladen. Der gleiche Effekt wurde mit einem 2-Aminopurin an der Position 3 erzielt, welches die exozyklische Aminogruppe für die Basenpaarung benutzt. Unterstützend dazu wurde die Insertion von G3•U70 in den Akzeptorstamm anderer tRNAs untersucht. Dadurch wurde die ursprüngliche tRNA-Identität erfolgreich maskiert und die mutierte tRNA konnte somit effizient von der AlaRS beladen werden (Hou *et al.*, 1989).

Dem Konzept der "direkten Erkennung" des G3•U70-Basenpaar durch die AlaRS, welches auf den beschriebenen *in vitro* Experimenten basiert, steht das der "indirekten Erkennung" gegenüber (McClain *et al.*, 1988). Im Rahmen dieser Hypothese wird neben der Bedeutung des G3•U70-Basenpaars als Erkennungssignal eine mögliche Modulation und Destabilisierung der Duplexstruktur des Akzeptorstamms durch dieses Nicht-Watson-Crick Basenpaar beschrieben. Diese Annahme stützt sich vor allem auf Daten von *in vivo* Experimenten, bei denen trotz einer Mutation der 3-70 Position zum Beispiel nach A3•C70 und G3•A70 an der vollständigen "amber suppressor" tRNA^{Ala}_{CUA} eine meßbare Aminoacylierungsreaktion nachweisbar ist, obwohl kein strukturelles Merkmal des G3•U70-Basenpaar an dieser Position verbleibt (Gabriel *et al.*, 1996). Die Wiederholung der Experimente unter Verwendung derselben Mutationen, die *in vivo* benutzt wurden, zeigte keine Beladungsaktivität *in vitro* (Beuning *et al.*, 1997), was die großen Unterschiede der verschiedenen experimentellen Methoden unterstreicht. Zwei NMR-Strukturstudien des tRNA^{Ala}-Akzeptorstamms aus *E. coli* unterstützen die Hypothese der "indirekten Erkennung", da sie eine veränderte Zuckerwellung bzw. ein diskretes Verhalten der Torsionswinkel des Phosphatrückgrats in der Nähe des G3•U70-Basenpaars beschreiben (Limmer *et al.*, 1993; Ramos & Varani, 1997).

1.1.4. Aufgabenstellung

Im Rahmen dieser Arbeit sollen die Heptamere des tRNA^{Ala}-Akzeptorstamms (ALA^{wt}) und der C70-Mutante (ALA^{C70}) kristallisiert und die Kristallstrukturen bei hoher Auflösung ermittelt werden (**Abb. 1.1-B**). Die Interpretation der Kristallstrukturen dient der Charakterisierung der Unterschiede und Gemeinsamkeiten zwischen Wildtyp und inaktiver Ala^{C70}-Mutante im Hinblick auf die Erkennung durch die AlaRS.

1.2. DNA-RNA Hybridstrukturen im Replikationszyklus des HIV-Virus

Intensive Anstrengungen zur umfassenden Charakterisierung des HIV-Virus werden auf vielen Teilgebieten der biologischen Forschung durchgeführt. Das Hauptziel dabei ist die Entwicklung selektiv wirksamer Pharmaka, die diese Krankheit der Moderne einzudämmen vermögen. Ein Hauptangriffspunkt dieser Bestrebungen sind enzymatisch aktive pol-Proteine des Partikels. Das PR-Protein ist eine Protease, die während des Reifeprozesses von HIV für das Schneiden des Gag-Pol Polyproteins zur Erlangung der vollen Pathogenität des Virus benutzt wird. Es stellt ein primäres Zielmolekül der Wirkstoffforschung dar (Condra *et al.*, 1995; Ridky & Leis, 1995; Schock *et al.*, 1996). Die Reverse Transkriptase (RT) stellt das zweite Haupt-Zielmolekül dar. Die RT überträgt die genomische Information des Virus in DNA, die im weiteren Verlauf der Infektion in das Wirts-Chromosom eingeführt wird. Weiterhin beherbergt die RT die RNase-H Funktionalität, die für die Abtrennung der RNA vom neu synthetisierten DNA-Strang verantwortlich ist. Das Enzym ist strukturell charakterisiert worden. Es existieren Kristallstrukturen von RT-DNA wie auch RT-Inhibitorkomplexen (Jacob-Molina *et al.*, 1993; Esnouf *et al.*, 1995; Rogers & Söll, 1995; Hsiou *et al.*, 1996). Als Inhibitoren werden Nukleosid-Analoga wie 3'-Azido-3'-Deoxythymidin (AZT) und Di-deoxyinosin (ddI) verwendet, die irreversibel an das enzymatische Zentrum der Polymerase binden und das Enzym dadurch blockieren. Die Untersuchung der Mechanismen der RNase-H Aktivität bilden die Voraussetzung zur Aufklärung weiterer Inhibierungsmöglichkeiten. Im Rahmen dieser Anstrengungen stellt die strukturelle Charakterisierung des 3'-DNA-RNA-5' Transkriptionsprodukts der RT, welches spezifisch von der RNase-H Funktion geschnitten wird, einen Teilbetrag dar.

1.2.1. Die 3'-DNA-RNA-5' Hybridstruktur bei der Replikation des HIV-I Virus

Die Replikation des HIV-I Virus wird von der Reversen-Transkriptase (RT) in zwei Stufen katalysiert. Zuerst wird ein DNA-Strang, der komplementär zum viralen RNA-Genom ist, synthetisiert (Minus-Strang Synthese) (**Abb. 1.2-A**). Dieser Strang dient als Matrize für die Synthese des zweiten DNA-Strangs (Plus-Strang Synthese). Beide Prozesse werden von RNA-Primern initiiert (Skalka & Goff, 1993). Die Primerverlängerung durch die RT führt zur Ausbildung einer 3'-DNA-RNA-5' Hybridbindung, im Falle der Minus-Strang-Synthese mit basengepaartem RNA-Gegenstrang.

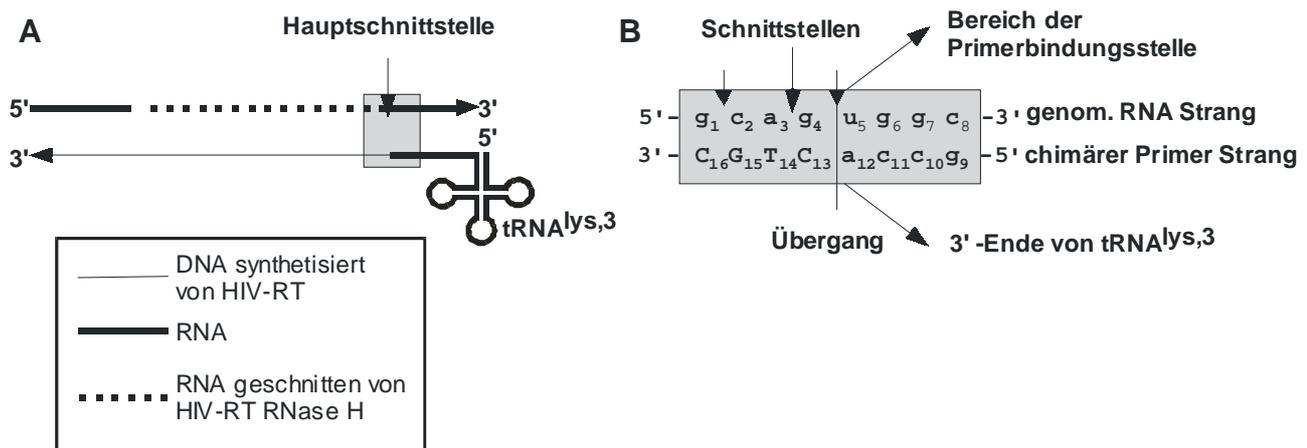


Abb. 1.2 Hybridstruktur bei der Replikation des HIV-I Virus A) Schematische Darstellung der Minus-Strang Synthese des HIV-I Virus, B) chimäres Oktamer welches zur Strukturlösung benutzt wurde.

1.2.2. Aufgabenstellung

Das Ziel des Projekts ist die Kristallisation und Strukturlösung eines chimären Oktameren (HIV^{Chi}) als Modell für die Beschreibung der 3'-DNA-RNA-5' Verknüpfung. Die durch die Verknüpfung möglicherweise induzierten Abweichungen von der kanonischen A-Form sollen als mögliche strukturelle Signale für die RNase-H Aktivität identifiziert werden. Anhaltspunkte für eine modifizierte Konformation ist die durchgeführte NMR-Strukturbestimmung des Plus-Strang Produkts aus dem *Moloney Murine* Leukämie-Virus. Hier wird eine veränderte Geometrie der kleinen Furche als strukturelle Determinante diskutiert (Salazar *et al.*, 1996).

1.3. Das Kälteschockprotein *Bc-Csp* aus *Bacillus caldolyticus*

1.3.1. Proteine

Die Aufklärung der dreidimensionalen Strukturen von Proteinmolekülen, stellt eine grundlegende Voraussetzung für das Verständnis der Funktion dieser universellen Moleküle dar. Obwohl alle Proteine aus nur 20 α -L-Aminosäuren aufgebaut sind (siehe **Abschnitt 7.4**), zeichnen sie sich durch große strukturelle und funktionelle Mannigfaltigkeit aus. Die Polypeptidkette ordnet sich unter Ausbildung von lokalen Sekundärstrukturen, α -Helices, β -Faltblättern und Schleifenbereichen. Die dreidimensionale Organisation dieser lokalen Strukturelemente zueinander bildet die Tertiärstruktur

des Proteins. Ist ein Protein aus mehr als einer Polypeptidkette aufgebaut, bilden die Anordnungen der multimeren Tertiärstrukturen die vierte Ordnungsebene, die Quartärstruktur.

1.3.2. Proteinstabilität

Vergleicht man die Stabilität von Proteinen mit anderen Makromolekülen stellt man fest, daß sie unter physiologischen Bedingungen nur marginal stabil sind. Die freie Enthalpie der Entfaltung beträgt bei einem Protein, bestehend aus einer 100 Aminosäuren langen Polypeptidkette, nur etwa $40 \text{ kJ}\cdot\text{mol}^{-1}$. Verglichen dazu beträgt der Energieaufwand zur Spaltung einer Wasserstoffbrückenbindung nur etwa $20 \text{ kJ}\cdot\text{mol}^{-1}$. Proteine existieren deshalb unter physiologischen Bedingungen, weil sie durch kooperative Wechselwirkungen, die einzeln betrachtet sehr schwach sind, in ihrer Summe den Erhalt der nativen Struktur gewährleisten. Diese Wechselwirkungen kann man nach fallender Stärke in folgende Gruppen einteilen.

Elektrostatische Kräfte wirken zwischen elektrischen Ladungen, die zum Beispiel an Karboxyl- und Aminogruppen auftreten können, und können je nach Vorzeichen attraktiv ($+ \rightarrow -$) oder repulsiv ($+ \leftarrow +$; $- \leftarrow -$) ausfallen. Ihr Betrag ist indirekt proportional zum Quadrat des Abstandes und kann nach dem Coulomb'schen Gesetz beschrieben werden [**Glg. 1.1**]. Diese elektrostatischen Kräfte wirken weitreichend. Die elektrostatischen Wechselwirkungen werden aber durch die Umgebung beeinflusst. So beträgt die Dielektrizitätskonstante ϵ_0 im Vakuum 1, im Wasser 80 und im hydrophoben Kern des Proteins 3-5.

$$F = \frac{1}{4\pi\epsilon_0} \cdot \frac{q_1 q_2}{r^2} \quad [\text{Glg. 1.1}]$$

q_1, q_2 : Ladungen der Wechselwirkungspartner

ϵ_0 : Dielektrizitätskonstante

r : Ladungsabstand

Mit anderen Worten werden die elektrostatischen Kräfte besonders im Inneren des Proteins wirksam, wogegen sie an der Oberfläche durch die Solvation abgeschwächt werden. Tatsächlich befinden sich aber die meisten geladenen Seitenketten an der Oberfläche von Proteinen. Ihre elektrostatische Anziehung besitzt deshalb nur einen kleinen stabilisierenden Beitrag in Hinblick auf die Ausbildung des nativen Zustands. Zusätzlich besitzen bestimmte Gruppen permanente oder

induzierte Dipole, die durch die Polarisierung entlang von Bindungen aber auch durch räumliche Nähe zu anderen geladenen Bereichen entstehen. Zum Beispiel bilden Peptidbindungen permanente Dipole, die bei der Anordnung in α -Helices zu einem Makrodipol führen (Hol, 1978).

Wasserstoffbrückenbindungen (D-H \cdots A) sind elektrostatische Wechselwirkungen von Protonendonoren mit -akzeptoren. Donoren wie Akzeptoren können Sauerstoff und Stickstoff sein. Akzeptoren müssen über mindestens ein freies Elektronenpaar verfügen. Diese kurzreichweitigen Wechselwirkungen sind stark gerichtet und für die Ausbildung der Sekundärstruktur verantwortlich.

Van-der-Waals Wechselwirkungen treten durch wechselseitige Polarisierungen sich sehr dicht annähernder Moleküle bzw. Atomgruppen auf und sind stark abstandsabhängig. Die Energie der van-der-Waals Anziehung zwischen Atomen in einem Abstand r ist proportional zu r^{-6} . Aufgrund ihrer Vielzahl tragen van-der-Waals Wechselwirkungen stark zur Proteinstabilisierung bei.

Der hydrophobe Effekt trägt sehr stark zur weiteren Stabilisierung des gefalteten Zustandes der Polypeptidkette bei. Er wird in der Regel über ein entropisches Modell erklärt. In nativen, gefalteten Proteinen liegen polare und geladene Seitenketten meistens an der solvenszugängigen Oberfläche, hydrophobe Seitenketten bilden den Kern des Proteins. Der bei der Faltung der ungeordneten Polypeptidkette stattfindende Ausschluß dieser hydrophoben Seitenketten aus der wässrigen Lösung ist eine Konsequenz aus deren Unfähigkeit, Wasserstoffbrückenbindungen auszubilden. Sie bilden den hydrophoben Kern und werden durch van-der-Waals- und Dispersionswechselwirkungen stabilisiert. Die bestimmende Triebkraft der Faltung ist die große positive Entropieänderung ($\Delta S = S_N - S_U \gg 0$) des Solvens (Kauzmann, 1959). Im ungefalteten Zustand U werden unpolare Gruppen durch hochgeordnete Wasserstrukturen stabilisiert (Clathrate). Diese Strukturen gehen bei der Ausbildung der gefalteten Polypeptidkette verloren, die Unordnung des Solvens nimmt stark zu. Die isotherme freie Enthalpie (ΔG) der Faltung ist nach der Gibbs-Gleichung [**Glg. 1.2**] also negativ, wenn $|\Delta H| < |-T\Delta S|$ ist. Der hydrophobe Kollaps verläuft freiwillig (Schulz & Schirmer, 1978).

$$\Delta G = \Delta H - T\Delta S \quad (T=\text{const}) \quad [\text{Glg. 1.2}]$$

1.3.3. Das Hauptkälteschockprotein Csp

Durch eine Verringerung der Temperatur von 37 °C auf 10 °C wird in *E. coli* eine Kälteschockantwort ausgelöst (Jones *et al.*, 1987; Wolffe, 1995). Diese führt zu einer starken Expression des für das Hauptkälteschockprotein, CspA kodierenden Gens. CspA ist ein kleines, 69 Aminosäuren umfassendes β -Faltblattprotein. Es besitzt eine hohe Sequenz- und Strukturübereinstimmung zum Hauptkälteschockprotein CspB aus *Bacillus subtilis*, so daß eine ähnliche Funktion der beiden Proteine naheliegt (Schindelin *et al.*, 1993; Schindelin *et al.*, 1994). Deren detaillierte Beschreibung ist gegenwärtig noch Gegenstand aktiver Forschungen (Bae *et al.*, 1997; Graumann & Marahiel, 1997; Graumann *et al.*, 1997). Prinzipiell kann bakteriellen Kälteschockproteinen eine Schalterfunktion bei der Kälteschockantwort dieser Organismen zugeordnet werden. Beide Proteine binden DNA und RNA und haben eine hohe Sequenzübereinstimmung mit der Kälteschockdomäne CSD, die in Proteinen aus höheren Eukaryoten für die Bindung an die Y-Box verantwortlich ist (Schnuchel *et al.*, 1993). Y-Box bindende Proteine haben die Funktion als Transkriptions- und Translationsregulatoren. Es werden Modelle diskutiert, die eine Bindung von CspB als Dimer oder auch als Monomer an die DNA oder mRNA realisieren und somit die Transkription bzw. Translation bei tiefen Temperaturen fördert (Graumann & Marahiel, 1994). Die Nukleinsäurebindungsfähigkeit von CspB ist essentiell für seine *in vivo* Funktion (Schröder *et al.*, 1995).

Eine weitere hypothetische Funktion des Protein CspA ist die eines "RNA-Chaperon". Dabei soll eine Sekundärstrukturbildung der mRNA verhindert werden, um bei tiefen Temperaturen eine effiziente Translation zu ermöglichen (Jiang *et al.*, 1997).

Die dreidimensionalen Strukturen von CspB und CspA wurden durch Kristallstrukturanalyse und NMR-Experimente aufgeklärt und haben eine große Ähnlichkeit (Schindelin *et al.*, 1993; Schnuchel *et al.*, 1993; Newkirk *et al.*, 1994; Schindelin *et al.*, 1994). Beide Proteine organisieren ihre Tertiärstruktur als 5-strängiges β -Faß. Die β -Stränge 1-3 und 4-5 bilden zwei Subdomänen, deren antiparallele Faltblätter ca. 90° zueinander stehen (**Abb. 1.3**). Es handelt sich hierbei um eine neue Faltungsklasse, den sogenannten OB-fold (oligosaccharide/oligonucleotide binding-fold) (Chothia & Murzin, 1993; Murzin, 1993). Tatsächlich stellt der OB-fold ein ubiquitäres Faltungsmotiv dar, welches bei einer Vielzahl strukturell untersuchter, nukleinsäurebindender Proteine gefunden wird. Beispiele sind dafür die B-Untereinheit des Verotoxin-1 aus *E. coli*, die Antikodonbindungsdomäne der aspartyl-tRNA Synthetase im Komplex mit tRNA (Ruff *et al.*, 1991), die Staphylokokken

Nuklease (Hynes & Fox, 1991) und die Struktur des Telomeren-Enden bindenden Proteins aus *Oxytricha nova* (OnTEBP) (Horvath *et al.*, 1998).

Innerhalb des Kristalls bilden zwei CspB-Monomere über ein intermolekulares, antiparalleles Faltblatt ein Dimer. Interessant ist, daß das Protein nur aus β -Strängen, Haarnadelschleifen und Schleifenbereichen aber nicht aus α -helikalen Sekundärstrukturelementen zusammengesetzt ist. Es besitzt keine cis-Proline und keine Disulfidbrücken. Außerdem besitzt es an der Position 8 ein schon im gefalteten Zustand relativ solvensexponiertes Tryptophan, welches als molekulare Sonde bei der Änderung der Tertiärstruktur des Protein eingesetzt werden kann, um Faltungs- und Entfaltungsreaktionen spektroskopisch zu beobachten.

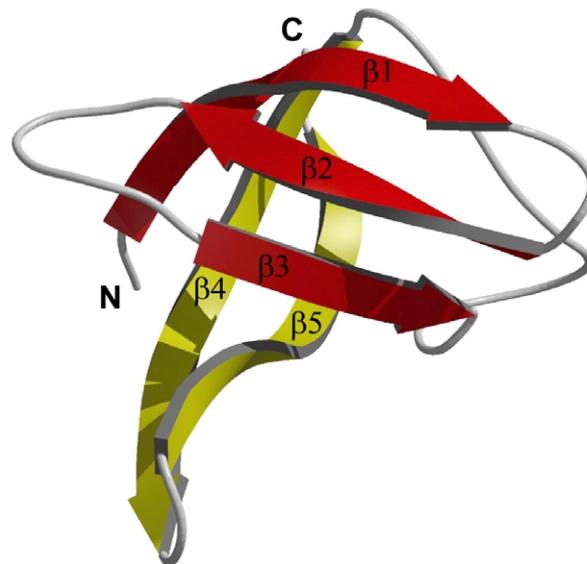


Abb. 1.3 Tertiärstruktur von CspB

CspB-Homologe wurden unter anderen auch in thermophilen Bakterien wie *Bacillus caldolyticus* (*Bc-Csp*) gefunden. Dieser Organismus hat seine optimalen Wachstumstemperatur bei 72 °C. Die Polypeptidkette von *Bc-Csp* umfaßt 66 Aminosäuren, daraus resultiert ein Molekulargewicht von 7333 Da. Die Sequenz ist zu 83% identisch mit der von CspB (**Tab. 1.2**). 11 Aminosäureaustausche und die Deletion einer Aminosäure am C-Terminus überführen die *Bc-Csp*-Sequenz von in die von CspB. *Bc-Csp* ist mit einem berechneten pI von 5.25 weniger sauer als CspB das einen pI von 4.54 besitzt. Aus der großen Sequenzübereinstimmung mit CspB und CD-Untersuchungen von *Bc-Csp* kann auf eine ebenfalls sehr ähnliche β -Faltblatt-Struktur geschlossen werden.

Tab. 1.2 Primärstrukturvergleich von *Bc-Csp* und *CspB*

<i>B. subtilis</i>	1-M	L	E	G	K	V	K	W	F	N	S	E	K	G	F	G	F
<i>B. caldolyticus</i>	1-M	Q	R	G	K	V	K	W	F	N	N	E	K	G	Y	G	F
<i>B. subtilis</i>	18-I	E	V	E	G	Q	D	D	V	F	V	H	F	S	A	I	Q
<i>B. caldolyticus</i>	18-I	E	V	E	G	G	S	D	V	F	V	H	F	T	A	I	Q
<i>B. subtilis</i>	35-G	E	G	F	K	T	L	E	E	G	Q	A	V	S	F	E	I
<i>B. caldolyticus</i>	35-G	E	G	F	K	T	L	E	E	G	Q	E	V	S	F	E	I
<i>B. subtilis</i>	52-V	E	G	N	R	G	P	Q	A	A	N	V	T	K	E	A	
<i>B. caldolyticus</i>	52-V	Q	G	N	R	G	P	Q	A	A	N	V	V	K	L	-	

1.3.4. Faltungskinetiken und der Konformationsstabilität der Kälteschockproteine

Den Ausgangspunkt für die in dieser Arbeit durchgeführten Strukturuntersuchungen bilden die thermodynamischen und kinetischen Daten, die für die Wildtyp-Csp Proteine beider Organismen und für die R3E-Mutante (*Bc-Csp*^{R3E}) aus *B. caldolyticus* vorliegen (Tab. 1.3) (Perl, 1997; Perl *et al.*, 1998; Schindler *et al.*, 1998). Diese Arbeiten stellen den Anfang einer systematischen Studie über die molekulare und strukturelle Basis der Thermostabilität dieser kleinen, "einfach" gefalteten Proteine dar.

Die thermischen und Guanidiniumhydrochlorid (GdmHCl) -induzierten Entfaltungen der Kälteschockproteine beider Organismen sind reversible $N \leftrightarrow U$ Zweizustandsreaktionen, die sehr schnell verlaufen. Bei der Untersuchung der Faltungskinetik wurden keine Faltungsintermediate gefunden. Die konformationelle Stabilität $\Delta G_{\text{Ef}}^{\text{H}_2\text{O}}$ von *Bc-Csp* ist mit 20.7 kJ/mol (T=25 °C, pH 7.0) fast doppelt so hoch wie die von *CspB*. Die Rückfaltungsgeschwindigkeit ist gegenüber *CspB* unverändert, die Entfaltung hingegen stark verlangsamt. Darin liegt die erhöhte Stabilität von *Bc-Csp* begründet. Dieses Ergebnis bedeutet auch, daß durch die unterschiedliche Sequenz nur der gefaltete Zustand stabilisiert wird, die denaturierte Form und der Übergangszustand werden davon nicht beeinflusst. Der geschwindigkeitsbestimmende Schritt der Faltung wird also durch die Sequenzunterschiede nicht verändert. Die Einführung der Punktmutation R3E destabilisiert *Bc-Csp* stark. Die konformationelle Stabilität $\Delta G_{\text{Ef}}^{\text{H}_2\text{O}}$ fällt auf einen Wert von 14.8 kJ/mol. Interessanterweise wirkt sich diese einzige Punktmutation drastisch auf die Rückfaltungsgeschwindigkeit aus, die im Vergleich zum Wildtyp *Bc-Csp* verlangsamt ist. Die Entfaltungsgeschwindigkeit bleibt davon unbeeinflusst. Dieses Ergebnis wird dahingehend interpretiert, daß die Mutation gleichmäßig den Übergangszustand wie auch das gefaltete Protein destabilisiert (Perl, 1997). Aus diesen Ergebnissen und durchgeführten Strukturvorhersagen wird geschlossen, daß R3 eine Stabilisierung der Wildtypstruktur von *Bc-Csp* durch die Ausbildung einer

intramolekularen Salzbrücke zu E46 induziert, die schon im Übergangszustand vorhanden sein muß.

Tab. 1.3 Thermodynamische und kinetische Daten der Csp-Proteine (Perl, 1997)

	CspB	Bc-Csp	Bc-Csp^{R3E}
$\Delta G_{\text{Ef}}^{\text{H}_2\text{O}}$ (kJ/mol) ¹	11.8	20.7	14.8
$k_{\text{Rf},0}$ (s ⁻¹) ²	600	1370	297
$k_{\text{Ef},0}$ (s ⁻¹) ³	5.4	0.64	0.6
T_m (°C) ⁴	49.9	71.6	56.1

- 1: Freie Enthalpie der Entfaltung in Abwesenheit von GdmCl (T=298K)
- 2: Geschwindigkeitskonstante der Rückfaltung in Abwesenheit von GdmCl (T=298K)
- 3: Geschwindigkeitskonstante der Entfaltung in Abwesenheit von GdmCl (T=298K)
- 4: Mittelpunkt der thermischen Entfaltung, $\Delta G_{\text{Ef}}^{\text{H}_2\text{O}}(T) = 0$ (Schindler, 1994)

1.3.5. Aufgabenstellung

Ausgehend von den durchgeführten thermodynamischen und kinetischen Untersuchungen sollen die bereitgestellten, gereinigten Proteine, *Bc-Csp* und die Einzelmutante *Bc-Csp^{R3E}*, kristallisiert und ihre Kristallstrukturen bei hoher Auflösung ermittelt werden. Die verfeinerten Strukturmodelle sind entsprechend der biophysikalischen Fragestellungen, in Hinblick auf die strukturellen Veränderungen welche die Thermostabilität beeinflussen können, zu interpretieren. Dabei ist auch die Kristallstruktur von *CspB* mit einzubeziehen.