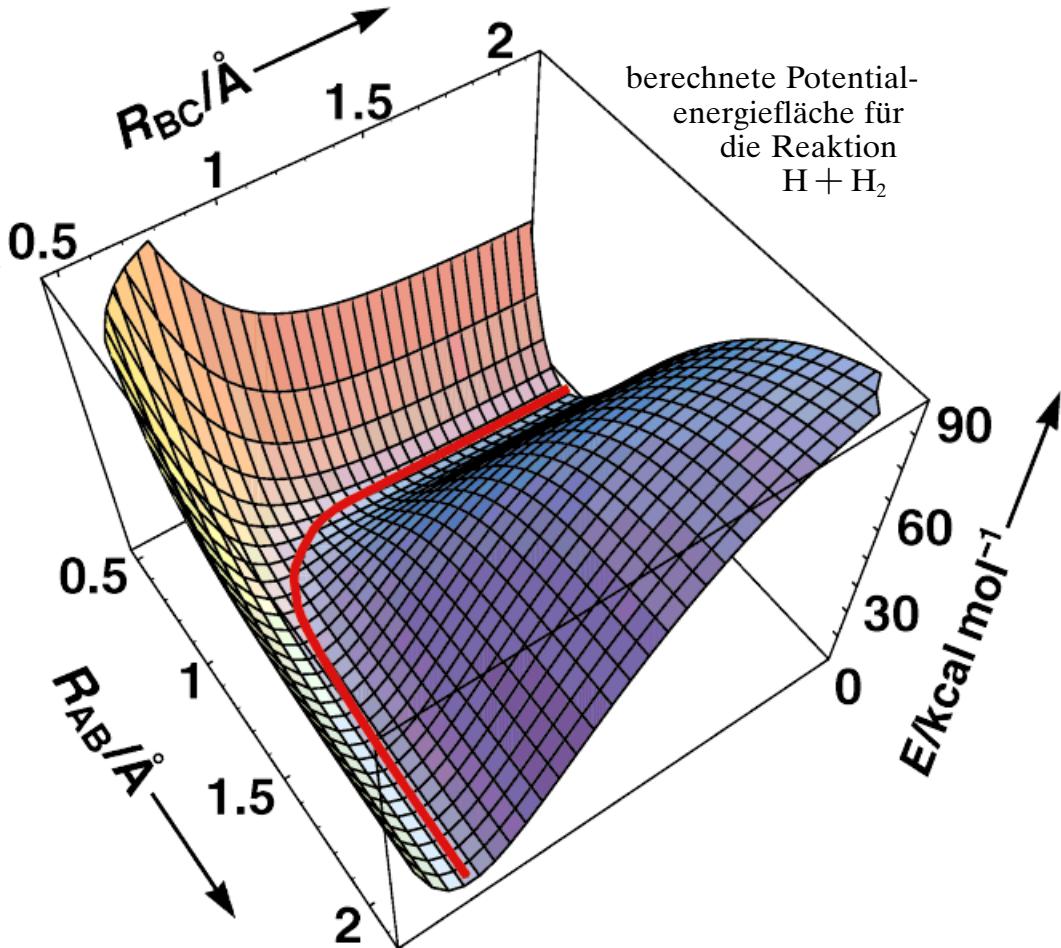


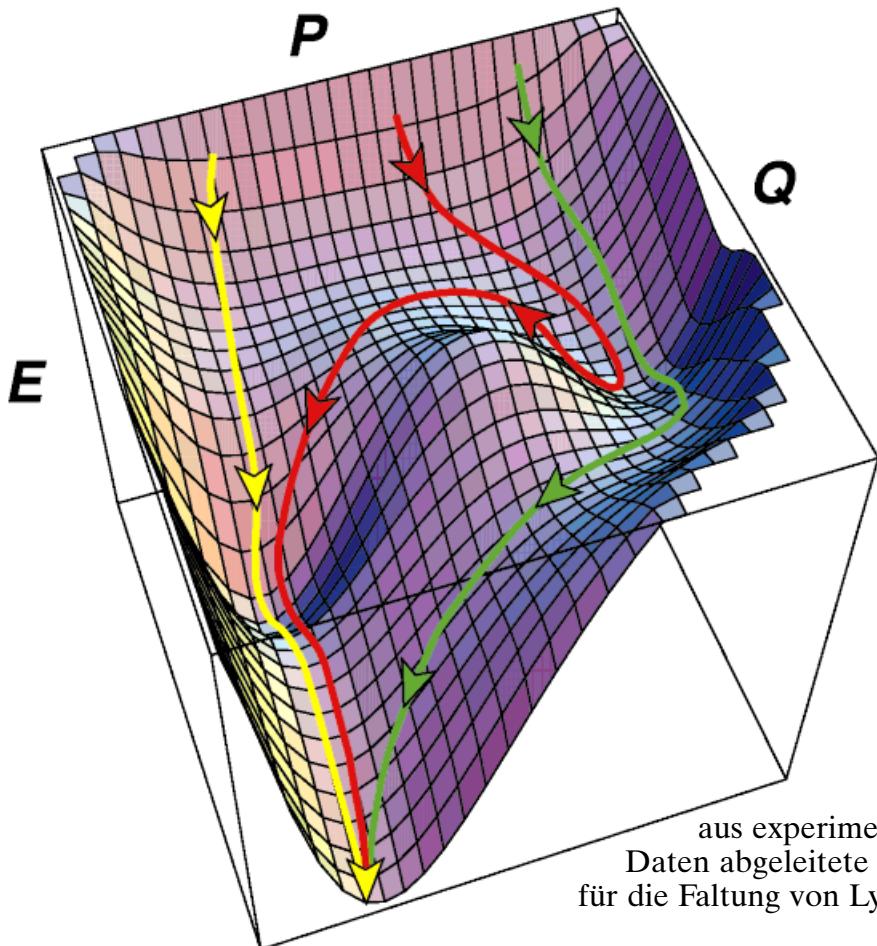
Ganz analog zur bekannten Beschreibung der $\text{H} + \text{H}_2$ -Austauschreaktion mit einer Potentialenergiefläche, die durch zwei AbstandsvARIABLEN definiert wird, lässt sich für die ungleich komplexere Faltung einer Polypeptidkette eine Potentialener-

giefläche in Abhängigkeit von geeigneten Fortschrittsvariablen errechnen. Diese energetische Beschreibung der Proteinfaltung ermöglicht es, immer

berechnete Potentialenergiefläche für die Reaktion $\text{H} + \text{H}_2$



tiefergehende Einblicke in diesen lange Zeit unverstandenen Prozeß zu erhalten.



aus experimentellen Daten abgeleitete Energiefläche für die Faltung von Lysozym

Proteinfaltung aus theoretischer und experimenteller Sicht

Christopher M. Dobson,* Andrej Šali und Martin Karplus*

Proteine sind an nahezu jedem biologischen Prozeß beteiligt, wobei ihre Funktionen von der Katalyse chemischer Reaktionen bis zur Aufrechterhaltung elektrochemischer Potentiale an Zellmembranen reichen. Sie werden an den Ribosomen anhand der Information, die in der zellulären DNA codiert ist, als lineare Ketten von Aminosäuren in einer definierten Reihenfolge hergestellt und müssen sich, um ihre Funktion auszuüben, zu einer einzigartigen, dreidimensionalen nativen Struktur falten, die für jedes Protein charakteristisch ist. Dieses komplexe molekulare Erkennungsphänomen beruht auf dem Zusammenwirken vieler relativ schwacher nichtbindender Wechselwirkungen. Da die Zahl der möglichen Konformationen einer Polypeptidkette astronomisch groß ist, würde eine systematische Suche nach der nativen Struktur (mit der niedrigsten Energie) eine nahezu unendliche Zeitspanne erfordern. Kürzlich wurden auf dem Weg zur Lösung die-

ses Paradoxons und dem Verstehen des Faltungsmechanismus wesentliche Fortschritte erzielt. Dies kam durch Verbesserungen der experimentellen Strategien zur Untersuchung von Faltungsreaktionen im Labor durch bio-physikalische Techniken zustande sowie durch Fortschritte bei den theoretischen Ansätzen, welche den Faltungsvorgang mit vereinfachten Modellen simulieren. Die jüngsten Fortschritte in diesem Bereich ähneln in ihrer Bedeutung denen, die vor 30 Jahren bezüglich des Verständnisses der Reaktionen kleiner Moleküle gemacht wurden. In diesem Beitrag fassen wir den gegenwärtigen Stand des Wissens über das Faltungsproblem der Proteine zusammen und vergleichen die neuen Konzepte mit denen, die für einfachere Reaktionen inzwischen etabliert sind. Ein wichtiger Unterschied zwischen der Proteinfaltung und den Reaktionen kleiner Moleküle liegt in der Heterogenität des Faltungsvorgangs und dem daraus resultieren-

den komplexen Wechselspiel zwischen entropischen und enthalpischen Beiträgen zur Freien Energie des Systems während der Reaktion. Ein vereinheitlichtes Modell der Proteinfaltung wird skizziert, das auf der effektiven Energiefläche einer Polypeptidkette beruht, auf der der native Zustand eine begünstigte Lage einnimmt. Das Verständnis der Faltung ist wichtig für die Analyse der vielen Ereignisse, die an der zellulären Regulation beteiligt sind, für das Design von Proteinen mit neuartigen Funktionen, für die Nutzung der Sequenzinformation aus unterschiedlichen Genomprojekten und für die Entwicklung neuer Strategien für die Therapie oder die Prävention schwerer menschlicher Erkrankungen, die mit Fehlern bei der Proteinfaltung zusammenhängen.

Stichwörter: Moleküldynamik • Proteine • Proteinfaltung • Reaktionskinetik • Strukturaufklärung

[*] Prof. C. M. Dobson
Oxford Centre for Molecular Sciences, New Chemistry Laboratory
University of Oxford
South Parks Road, Oxford OX13QT (Großbritannien)
Fax: (+44) 1-865-275-921
E-mail: Chris.Dobson@cl.ox.ac.uk

Prof. M. Karplus
Department of Chemistry and Chemical Biology, Harvard University
12 Oxford Street, Cambridge, MA 02138 (USA)
Fax: (+1) 617-496-3204
E-mail: marci@tammy.harvard.edu

und
Laboratoire de Chimie Biophysique
Institut Le Bel, Université Louis Pasteur
F-67000 Strasbourg (Frankreich)
Fax: (+33) 388-60-83-63
E-mail: marci@brel.u-strasbg.fr

Prof. A. Šali
The Rockefeller University, New York, NY (USA)

1. Einleitung

Reaktionen sind in der Chemie von zentraler Bedeutung. Sie reichen von einfachen Austauschprozessen weniger Atome bis zur Proteinfaltung, an der Tausende von Atomen beteiligt sind. Das Verständnis einer Reaktion basiert auf der Kenntnis der dominierenden Wechselwirkungen, welche die Potentialenergiefläche bestimmen, und einer Beschreibung der dynamischen Prozesse, die von den Reaktanten zum Produkt führen. Wesentliche Fortschritte beim Verständnis einfacher Reaktionen konnten in den letzten 30 Jahren durch eine Kombination von experimentellen Neuerungen und theoretischen Entwicklungen erzielt werden.^[1] Im Unterschied dazu ist der Vorgang der Proteinfaltung so komplex, daß selbst die Methoden, die zu ihrer Charakterisierung verwendet wer-

den, nicht vollständig definiert sind.^[2, 3] Zu Beginn dieses Aufsatzes werden wir einige Ähnlichkeiten und Unterschiede zwischen einfachen chemischen Reaktionen und der Proteinfaltung diskutieren. Abgesehen von den zahlreichen komplexen Eigenschaften der Faltungsreaktion, welche die Einführung vereinfachter theoretischer Modelle und neuer experimenteller Vorgehensweisen erfordert, wird dabei auch eine Reihe von allgemeineren Themen behandelt. Ein wichtiges davon ist die Anwesenheit eines Ensembles von Molekülen, die sich in ihrer Struktur und Dynamik erheblich unterscheiden, in verschiedenen Stadien der Faltungsreaktion.

Globuläre Proteine werden als lineare Ketten von Aminosäuren synthetisiert. Um ihre Funktion auszuüben, müssen sie sich schnell und zuverlässig zu einer spezifischen Struktur fal-

ten, die durch die Evolution für die Ausführung einer bestimmten Aufgabe maßgeschneidert ist (Abb. 1). Obwohl der Faltungsvorgang innerhalb der Zelle eine Reihe von katalytischen und steuernden Systemen erfordert,^[4] ist die für die Faltung notwendige Information bei den meisten, wenn nicht sogar bei allen Proteinen in der Sequenz enthalten.^[5] Eine Proteinsequenz muß hierfür zwei Kriterien erfüllen: ein thermodynamisches und ein kinetisches.^[2, 6] Das thermodynamische Kriterium besteht darin, daß die Moleküle eine einzellige gefaltete Konformation einnehmen (den nativen Zustand), die unter physiologischen Bedingungen stabil ist. Das kinetische Kriterium besteht darin, daß sich die denaturierte Polypeptidkette in einer annehmbaren Zeit zum nativen Zustand falten kann. Eine Polypeptidkette kann so viele

Christopher Martin Dobson, 1949 in Rinteln geboren, studierte in Oxford Chemie und promovierte dort 1975 bei R. J. P. Williams zum D.Phil. mit einer Arbeit zur NMR-Spektroskopie von Proteinen. Nach Assistententätigkeiten in Oxford ging er 1977 als Assistant Professor für Chemie an die Harvard University, Boston. 1980 kehrte er an die University of Oxford zurück, wo er zunächst Lecturer, dann Reader war. Gegenwärtig ist er Professor für Chemie und designierter Direktor des Oxford Centre for Molecular Sciences. 1992 wurde er zum Howard Hughes International Research Scholar ernannt und 1996 zum Fellow of the Royal Society gewählt. Seine Forschungsinteressen konzentrieren sich auf die Entwicklung und Anwendung von NMR-spektroskopischen und anderen biophysikalischen Methoden im Hinblick auf biologische und chemische Fragestellungen. Insbesondere richten sich seine Forschungsaktivitäten auf die Aufklärung des Mechanismus der Proteinfaltung und deren Bedeutung für biologische Abläufe und damit verbundene Fehlfunktionen.



C. M. Dobson



A. Šali



M. Karplus

Andrej Šali, 1963 in Kranj (Slowenien) geboren, studierte an der Universität von Ljubljana Chemie. 1991 promovierte er zum Ph.D. am Birkbeck College der University of London bei T. L. Blundell auf dem Gebiet der Modellierung von Proteinstrukturen. Danach arbeitete er als Postdoktorand an der Harvard University bei M. Karplus an Proteinfaltungs-Simulationen. Seit 1995 ist er Assistant Professor an der Rockefeller University, New York, wo er an der Modellierung von Proteinstrukturen und ihrer Anwendung in der Genomforschung arbeitet.

Martin Karplus, 1930 in Wien geboren, erhielt 1950 seinen B.A. am Harvard College und promovierte 1953 am Caltech, Pasadena, zum Ph.D. Danach arbeitete er an der University of Oxford als Postdoktoranden-Stipendiat der National Science Foundation und ging 1955 als Assistant Professor an die University of Illinois. 1960 bekam er eine Professur an der Columbia University, New York, 1966 an der Harvard University. Dort wurde er 1979 zum Theodore William Richards Professor of Chemistry ernannt. Außerdem ist er Professeur Conventionné an der Universität Straßburg, Mitglied der National Academy of Sciences (USA), der American Academy of Arts and Sciences sowie der Niederländischen Akademie der Künste und Wissenschaften. Zu Beginn seiner Forscherlaufbahn beschäftigte er sich mit der Kernresonanzspektroskopie, wobei sein besonderes Interesse der theoretischen Analyse von Kernspin-Kopplungskonstanten galt. Er leistete grundlegende Beiträge zur Theorie der reaktiven Zusammenstöße zwischen kleinen Molekülen auf der Basis von Trajektorien-Berechnungen. Außerdem wendete er als einer der ersten die Vielkörper-Störungstheorie auf atomare und molekulare Systeme an. Er untersuchte Vorgänge bei biologischen Molekülen, beispielsweise die Bindung von Sauerstoff an Hämoglobin sowie die Photochemie der Sehpigmente. In neueren Forschungsarbeiten beschäftigte er sich mit der Entwicklung von Moleküldynamik- und anderen Simulationsmethoden zur Aufklärung der Eigenschaften von Proteinen und Nucleinsäuren. Gegenwärtig untersucht er enzymatische Reaktionen, die Proteinfaltung und die Anwendung von Rechenverfahren zum kombinatorischen Ligandendesign.

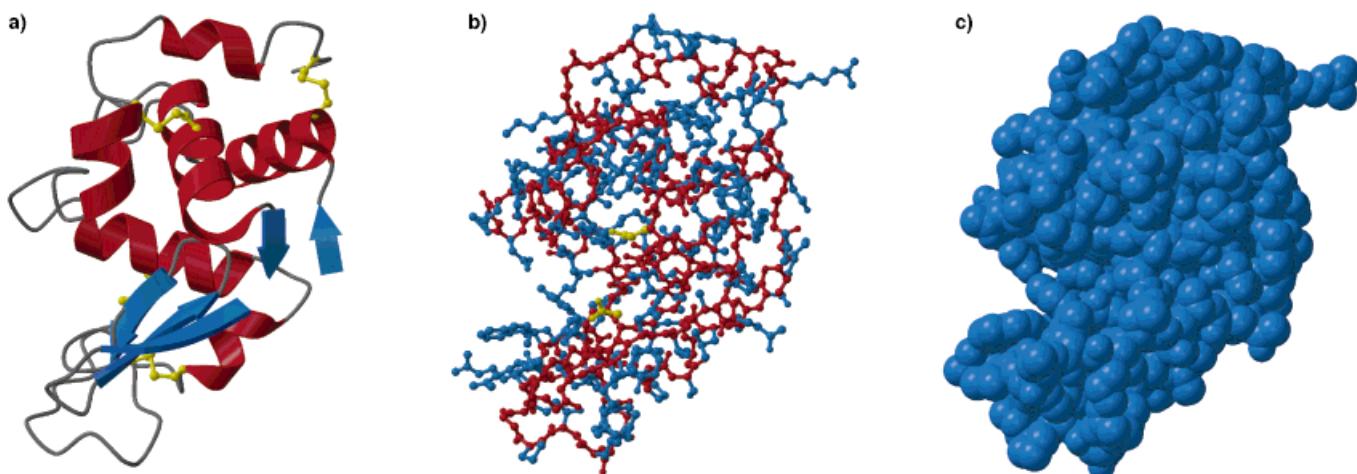


Abb. 1. Unterschiedliche Darstellungen des Proteins Lysozym (PDB-Code: 8LZT), angefertigt mit den Programmen Molscript^[163] und Render.^[164] a) Bänderdarstellung des Faltungsusters der aus 129 Aminosäuren gebildeten Hauptkette. α -Helices sind als rote Zylinder und β -Faltblätter als blaue Pfeile dargestellt. Quervernetzungen durch Disulfidbrücken sind gelb gezeichnet. b) Kugel-Stab-Darstellung, bei der die Hauptkette rot, die Seitenketten blau und die am katalytischen Mechanismus beteiligten Seitenketten gelb gezeichnet sind. c) Raumfüllendes Modell, das die dichtgepackte Struktur des Proteins zeigt. Laut Rechnungen auf der Grundlage der üblichen van-der-Waals-Radien sind mehr als 75 % des Raums ausgefüllt.

Konformationen einnehmen, daß es einen Weg geben muß, den nativen Zustand in einer Zeit zu erreichen, die um mehrere Größenordnungen kürzer ist als die für eine statistische Suche erforderliche.^[7] Um einen Schätzwert für diese Größenordnung zu erhalten, geht man davon aus, daß es für jeden Aminosäurerest fünf Konformationen gibt. Für ein Protein, das aus einer Polypeptidkette von 100 Aminosäuren besteht, gibt es somit insgesamt 10^{70} Konformationen. Wenn nur 10^{-11} s benötigt würden, um eine Konformation in eine andere umzuwandeln, so würde eine statistische Suche aller Konformationen immer noch 10^{59} s oder 10^{52} Jahre erfordern. Dies veranlaßte Levinthal, „Wege“ der Proteinfaltung vorzuschlagen.^[8] Viele Versuche sind unternommen worden, Proteine auf der Grundlage des Konzepts der Faltungswege theoretisch zu „falten“, doch trotz einiger gegensätzlicher Behauptungen ist keiner davon gelungen^[9]. Der Unterschied zwischen den Grenzen dessen, was man mit einem Computer in Tausenden von Stunden machen kann, und dem, was die Natur in weniger als einer Sekunde erreicht, wirft deutlich die Frage nach dem Mechanismus der Proteinfaltung auf, eine Frage, welche die Forscher über Jahrzehnte hinweg beschäftigt hat.^[10, 11]

Trotz des langanhaltenden Interesses an der Proteinfaltung haben die Experimente zum detaillierten Studium der Vorgänge, die während des Faltungsprozesses ablaufen, erst vor relativ kurzer Zeit begonnen. Ähnlich wie es vor 30 Jahren für kleine Moleküle der Fall war, ist die Untersuchung der Proteinfaltung nun in einem Stadium, in dem Theorie und Experiment gemeinsam schnelle Fortschritte in Richtung auf das vollständige Verständnis dieses komplexen Prozesses machen können. Zu diesem Zweck muß ein Rahmen für die Interpretation der Ergebnisse neuer experimenteller Techniken und für die Untersuchung spezifischer Eigenschaften der Faltungsreaktion entwickelt werden. In diesem Aufsatz skizzieren wir kurz einige Aspekte der Reaktionen kleiner Moleküle, um sie dann der Reaktion der Proteinfaltung gegenüberstellen zu können. Danach zeigen wir, wie die Proteinfaltung durch die gegenwärtigen theoretischen und experi-

mentellen Ergebnisse beschrieben werden kann. Obwohl die Verschiedenheit von Proteinen oft betont wird, weisen die sich abzeichnenden Eigenschaften des Faltungsmechanismus auf dessen grundlegende Universalität hin, wenn dieser durch eine durchschnittliche effektive Energie und die damit verbundene Konformationsentropie beschrieben wird.

Aufgrund der jüngsten Fortschritte bei den theoretischen und experimentellen Methoden für die Untersuchung des Mechanismus der Proteinfaltung ist ein für Chemiker geschriebener Aufsatz zu diesem sich schnell weiterentwickelnden Forschungsgebiet besonders zeitgemäß. Einige andere Übersichtsartikel, die sich eher an Biologen oder Biophysiker richten,^[6, 14–17] sind kürzlich erschienen und mögen für den Leser ebenfalls von Interesse sein, da sie den vorliegenden Aufsatz ergänzen. Die hier präsentierte Sichtweise betont unsere eigene Arbeit und Erfahrungen, doch haben wir in diesem aufregenden und schnell fortschreitenden Forschungsgebiet auch von den Beiträgen vieler anderer, die in den Literaturverweisen genannt werden, stark profitiert.

2. Die Energiefläche für Reaktionen kleiner Moleküle

Die Natur jeder Reaktion, sei es die eines kleinen Moleküls oder die eines Proteins, wird von der Potentialenergiefläche und den Gesetzen der Dynamik bestimmt.^[1, 12, 13] Die detaillierte experimentelle und theoretische Analyse einfacher chemischer Reaktionen beruht auf wohldefinierten Vorgehensweisen. Diese wollen wir hier darstellen, um die Situation der für Proteine gültigen gegenüberzustellen. Bei experimentellen Studien bestimmt man die Potentialenergiefläche gewöhnlich zuerst in einer recht groben Weise, indem man die Reaktionsgeschwindigkeit in Abhängigkeit von der Temperatur mißt, um dann die Aktivierungsenergie E_A und den Faktor $A(T)$ aus der Arrhenius-Gleichung zu ermitteln. Dazu

schreibt man die temperaturabhängige Geschwindigkeitskonstante in Form der Gleichung (1).

$$k(T) = A(T) \exp(-E_A/RT) \quad (1)$$

Der größte Teil der Temperaturabhängigkeit röhrt vom exponentiellen Faktor E_A/RT (Arrhenius-Faktor) her, während der präexponentielle Faktor $A(T)$ im allgemeinen nur eine schwache Temperaturabhängigkeit aufweist. Messungen von $A(T)$ und E_A können durch Daten über die Streuung nach reaktiven Stößen aus Experimenten mit gekreuzten Molekülstrahlen^[12] und durch Messungen, welche die direkte Charakterisierung des Übergangszustands ermöglichen,^[18] ergänzt werden. Solche Untersuchungen liefern die Details, die für die Charakterisierung der Potentialenergiefläche notwendig sind. Theoretische Berechnungen können ebenfalls dazu verwendet werden, die Potentialenergiefläche, die Reaktionskoordinate und den Übergangszustand zu bestimmen.^[19] Schließlich kann man Trajektorienberechnungen oder ihr quantendynamisches Analogon dazu verwenden, die Gesamtreaktionsgeschwindigkeit zu ermitteln und die Streuung nach reaktiven Stößen im Detail zu simulieren.^[13, 20] Für eine kleine Zahl einfacher Reaktionen sind die verfügbaren theoretischen Beschreibungen ebenso genau wie die experimentellen Ergebnisse.^[1, 20]

Wir wollen kurz die einfachste und am besten verstandene chemische Reaktion betrachten, den bimolekularen Austausch zwischen einem Wasserstoffatom und einem Wasserstoffmolekül. Nur drei innere Koordinaten sind erforderlich, um die Reaktion zu beschreiben (Abb. 2a). Die Atome bewegen sich auf einer Potentialenergiefläche, welche die Energie des Systems als eine Funktion der Koordinaten angibt. Diese Fläche entspricht der „Energielandschaft“ für die Reaktion $H + H_2$. Abbildung 2b und 2c stellen die Fläche für eine geradlinige Kollision und die Reaktionskoordinate dar. Diese ist eine Kombination der Freiheitsgrade, die für die Beschreibung des Weges mit der geringsten Energie von den Reaktanten zu den Produkten essentiell sind. Die Energie entlang der Reaktionskoordinate ist in Abbildung 2d dar-

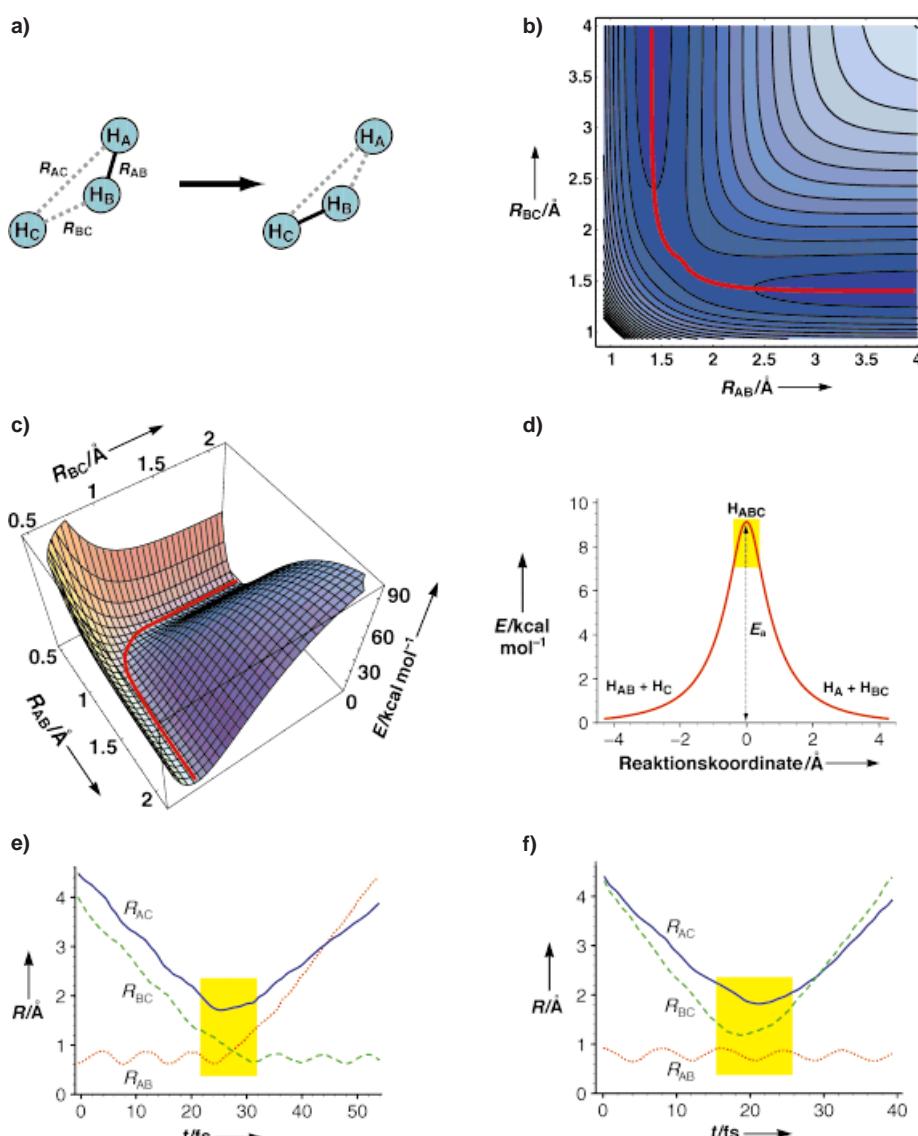


Abb. 2. Die Austauschreaktion zwischen einem Wasserstoffatom und einem Wasserstoffmolekül. a) Schematische Darstellung der Reaktion mit Definition der Abstände R_{AB} , R_{AC} und R_{BC} . b) Konturdiagramm der Potentialenergiefläche für einen geradlinigen Zusammenstoß als Funktion der Abstände R_{AB} und R_{BC} mit $R_{AC} = R_{AB} + R_{BC}$; der Weg mit der niedrigsten Energie ist rot eingezeichnet. c) Wie b), aber in dreidimensionaler Darstellung. d) Energie entlang der Reaktionskoordinate, die dem Weg der geringsten Energie in a) und b) entspricht (der Bereich für den Übergangszustand ist gelb unterlegt). e) Eine typische Trajektorie für den reaktiven dreidimensionalen Zusammenstoß; dargestellt sind die drei Abstände R_{AB} , R_{AC} und R_{BC} als Funktion der Zeit. f) Eine typische Trajektorie für den nichtreaktiven Zusammenstoß, dargestellt wie in e.). Sowohl in e) als auch in f) sind die Wechselwirkungen zwischen den drei Atomen auf einen sehr kurzen Zeitraum (gelb unterlegt) beschränkt. Dies spiegelt die schmale Barriere der potentiellen Energie in d) wider. (Die Ergebnisse wurden aus Lit.^[13, 19] mit Veränderungen übernommen.)

gestellt. Bei einer Temperatur T ist die kinetische Energie des Wasserstoffatoms und des Wasserstoffmoleküls so groß, daß die Reaktion nicht dem Weg der minimalen Energie folgt. Doch da das Tal zwischen Reaktanten und Produkten eng und wohldefiniert ist und es einen genau lokalisierten Übergangszustand am höchsten Punkt der Barriere gibt, bleiben die Trajektorien bei Raumtemperatur nahe am Minimalenergieweg. Wie wir später sehen werden, ist der Reaktionsweg für Proteine sehr viel schlechter definiert, und die sich faltenende Polypeptidkette bewegt sich während der Faltungsreaktion über einen weiten Bereich der Energiefläche. Dies bedeutet,

daß bei der Proteinfaltung – im Unterschied zur $H + H_2$ -Reaktion – die Entropie eine ebenso wichtige Rolle spielt wie die Energie. Der wohldefinierte Minimalenergieweg für $H + H_2$ führt zu den einfachen Trajektorien, die in Abbildung 2e und 2f für den reaktiven bzw. den nichtreaktiven Zusammenstoß dargestellt sind. Bei der reaktiven Kollision nähert sich ein Atom, kollidiert mit dem Molekül und bildet eine neue Bindung mit einem der Wasserstoffatome, während sich das andere schnell entfernt. Der nichtreaktive Zusammenstoß sieht sehr ähnlich aus, abgesehen davon, daß es keinen Austausch von Atomen gibt. Der gesamte Prozeß findet in nur 10 fs über einen sehr begrenzten Bereich von Abständen statt, der in den Abbildungen 2d–f gelb schattiert dargestellt ist. Die Anfangsbedingungen (Positionen und Geschwindigkeiten) und die Wechselwirkungen bei diesem ultrakurzen Ereignis bestimmen, ob der Zusammenstoß zur Reaktion führt oder nicht.

3. Die Proteinfaltungsreaktion

Proteinfaltungsreaktionen haben Eigenschaften, die von denen der Reaktionen kleiner Moleküle sehr verschieden sind. So weist der präexponentielle Faktor beispielsweise eine starke Temperaturabhängigkeit auf,^[21] obwohl die Arrhenius-Gleichung auch für die Proteinfaltung gilt und hierfür oft verwendet wird. Dies ist ein Hinweis darauf, daß die Proteinfaltung komplexe Eigenschaften aufweist, die den meisten einfachen Reaktionen fehlen. Des weiteren sind die Wechselwirkungen zwischen den verschiedenen Teilen der Polypeptidkette – im Unterschied zum Femtosekunden-Bereich der reaktiven Ereignisse im Fall $H + H_2$ (Abb. 2) – während der gesamten Faltungsdauer von Bedeutung, also für Millisekunden oder noch länger.

Ein besseres Analogon für die Proteinfaltung ist vielleicht die Kondensation von Atomclustern (Abb. 3), deren Thermodynamik und Kinetik sehr detailliert beschrieben wurden.^[22–24] Bei diesen Clustern ähnelt die Kondensation der Gas → Flüssigkeits-Umwandlung, die bei Proteinen dem Übergang vom denaturierten Random-coil- (statistisches Knäuel) zum Random-globule-Zustand (kompaktes, aber ungeordnetes Knäuel) entspricht. Wenn die Temperatur niedrig genug ist, ordnen sich die flüssigkeitsähnlichen Cluster zu einem organisierten Feststoff – ein Suchprozeß, der jenem vom kollabierten Knäuel zum nativen Zustand bei Proteinen analog ist. Natürlich hat ein Protein einen einzigen nativen Zustand, der von seiner Aminosäuresequenz bestimmt wird, während Cluster aus identischen Atomen viele symmetrieverwandte Festkörperstrukturen annehmen können. Bei den Phasenübergängen von Atomclustern (und Proteinen) sind

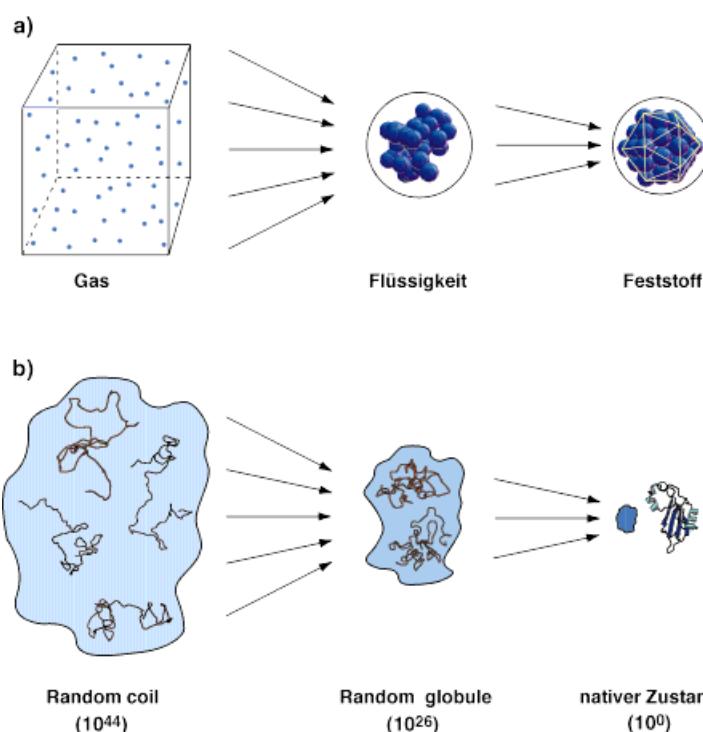


Abb. 3. a) Schematische Darstellung der Phasenübergänge in einem stabilen Cluster, der aus 55 Lennard-Jones-Teilchen besteht. Der Feststoff hat die Struktur eines Mackay-Ikosaeders.^[19] b) Schematische Darstellung der entsprechenden Übergänge bei der Proteinfaltung. Als Beispiel dient das Protein Barnase, das aus 110 Aminosäuren besteht. Der Verlauf der Kette aus den α -Kohlenstoffatomen ist wiedergegeben (erstellt mit dem Programm Molscript).^[163] Die Konformationen des Random-coil- und des Random-globule-Zustands wurden so konstruiert, daß ihre Gyrationsradien ca. 31 bzw. 15 Å betragen; für die native Struktur beträgt dieser 13.4 Å. Die Zahl der möglichen Konformationen der Hauptkette wurde aus Schätzungen der Hauptkettenentropie pro Aminosäure errechnet: Es gibt 2.57 und 1.7 Konformationen pro Aminosäure im Random-coil- bzw. im Random-globule-Zustand,^[55a] woraus sich $2.57^{109} = 10^{44}$ bzw. $1.7^{109} = 10^{26}$ mögliche Konformationen ergeben.

mehrere Reaktionswege die Regel, und der Übergang vom flüssigen (kompaktes Knäuel) zum festen (nativen) Grundzustand ist erster Ordnung, d. h., die beiden Zustände sind wohldefiniert und durch eine Barriere der Freien Energie getrennt. Da Proteine und Atomcluster eine mehrdimensionale Potentialenergiefläche mit einer großen Zahl von Freiheitsgraden aufweisen (anstatt nur drei wie im Fall $H + H_2$), kann es eine Vielzahl von lokalen Minima und Übergangsbereichen geben. Des weiteren betreffen die Strukturänderungen während der Faltung vom denaturierten Zustand (Reaktant) zum nativen Zustand (Produkt) das ganze Protein, ebenso wie die Umordnung zum Festkörper den ganzen Cluster betrifft.^[22, 23] Dies deutet darauf hin, daß die Ansätze, die auf der statistischen Mechanik und Dynamik großer Systeme beruhen,^[2, 25–28] besser Einsicht in bezug auf die Proteinfaltung vermitteln können als die Extrapolation des Verhaltens kleiner Moleküle.

Bei der Beschreibung der Reaktion der Proteinfaltung oder der Kondensation von Clustern ist es im Unterschied zur Beschreibung der Reaktionen kleiner Moleküle sinnvoll, die Potentialenergiefläche dieser kleinen Moleküle (Abb. 2) durch eine temperaturabhängige, gemittelte Fläche der effektiven Energie für die Proteine oder Cluster zu ersetzen. Diese Fläche entspricht einem nach Boltzmann gewichteten Mittelwert der zugänglichen Energien für jeden Wert einer geeigneten

Fortschrittsvariablen. Der Wert der Fortschrittsvariablen, welche die Annäherung an den nativen Zustand beschreibt, wird durch Mittelung über viele nichtessentielle Freiheitsgrade erhalten. Die resultierende Bewegung auf der effektiven Fläche sollte diffusiv sein. Deshalb kann sie durch ein Modell vom Kramers-Typ^[29] beschrieben werden, ähnlich jenem, das auf viele Reaktionen in Lösung anwendbar ist. Ein Beispiel ist der sehr einfache Konformationswechsel von *gauche* nach *trans* in flüssigem Butan.^[30] Natürlich ist die Proteinfaltungsreaktion sehr viel komplexer, da hierbei Konformationsübergänge in einer langen Polypeptidkette stattfinden. Insbesondere haben sowohl die Proteinfaltungsreaktion als auch die Bildung fester Cluster einen Vielteilchencharakter, der bei der nachfolgenden Analyse eine wichtige Rolle spielt.

Wie der Vergleich mit Atomclustern nahelegt, ist es nicht möglich, eine *einfache* Reaktionskoordinate auf der mehrdimensionalen Fläche der effektiven Energie zu definieren. Da es viele Wege gibt, den nativen Zustand zu erreichen, ist es unwahrscheinlich, daß ein einzelner Weg der minimalen Energie – der Faltungsweg in Levinthals Worten^[8] – die Faltungsreaktion beherrscht. Eine oder mehrere Fortschrittsvariablen müssen die Reaktionskoordinate in den Analysen der Proteinfaltung ersetzen. Diese Variablen beschreiben den Fortschritt der Reaktion vom Anfangs- zum Endzustand so wie eine Reaktionskoordinate, doch sie schließen die Möglichkeit vieler grundlegend verschiedener Wege ein. Das bedeutet, daß unser Wissen über die Faltung niemals so detailliert sein kann wie im Fall der kleinen Moleküle und daß wir vorsichtig sein müssen, wenn wir Konzepte der Reaktionen kleiner Moleküle oder auch derer, die von Enzymen katalysiert werden, auf die Proteinfaltung ausdehnen.

Wenn es in der Proteinfaltung keine einfache Reaktionskoordinate gibt, dann gibt es auch keinen einfachen Übergangszustand. Tatsächlich erfordert sogar die Beschreibung der Rotation einer einzelnen Seitenkette in einem Protein einen ausgedehnten Bereich der Potentialenergiefläche, einen Übergangs-, „Saum“.^[31] Wahrscheinlich werden sich die genauesten Aussagen, die für die Proteinfaltungsreaktion möglich sind, auf einen mehrdimensionalen Übergangsbereich der gemittelten Fläche der effektiven Energie beziehen, welche die Strukturen bestimmt, von denen aus die Polypeptidkette sich schnell und mit hoher Wahrscheinlichkeit zum nativen Zustand faltet. Vielleicht gibt es mehrere Übergangsbereiche mit ähnlichen Freien Energien aber deutlich verschiedenen Strukturen, die sich nicht schnell ineinander umwandeln, so daß parallele Reaktionswege zum Gesamtprozeß der Faltung beitragen. Auch muß die Frage gestellt werden, ob kinetische Standardansätze, welche die Konzentrationen weniger wohldefinierter Spezies verwenden, für die Proteinfaltungsreaktion angemessen sind.

Die Temperaturabhängigkeit der Stabilität der unterschiedlichen Zustände bei der Clusterbildung (oder bei der Proteinfaltung) wird durch das Gleichgewicht zwischen der effektiven Energie und der Konformationsentropie bestimmt. Die Zahl der verfügbaren Zustände nimmt auf dem Weg vom Gas (Random coil) über die Flüssigkeit (Random globule) zum festen (nativen) Zustand um viele Größenordnungen ab (Abb. 3b). Dies könnte nichts weiter als die notwendige Verringerung der Zahl der verfügbaren Zustände bedeuten,

wenn das System geordneter (dem nativen Zustand ähnlicher) wird. Eine alternative Erklärung wäre, daß die Funktion der effektiven Energie die Bewegung in Richtung auf den nativen Zustand beeinflußt und zu einer schnelleren Verringerung des zugänglichen Konformationsraums führt. Dies ist notwendig, um die schnelle Faltung von Proteinen oder die Verdichtung von Teilchen zu festen Clustern zu erreichen. Demnach ist das wichtigste zu lösende Problem die Frage, wie der Grundzustand „gefunden“ werden kann oder, in anderen Worten, welche Eigenschaft der effektiven Energiefläche und des aus ihr resultierenden Gleichgewichts zwischen Energie und Entropie bestimmt, ob das System den Grundzustand in einer annehmbaren Zeit erreichen kann oder nicht.

4. Die Thermodynamik der Proteinfaltung

Bevor wir das gegenwärtige Verständnis der Faltungsreaktion beschreiben, ist es nützlich, die Thermodynamik von Proteinen zu betrachten. Die Freie Energie des nativen Zustands eines Proteins ist unter physiologischen Bedingungen nur geringfügig niedriger als die des denaturierten Zustands, was auf eine annähernde wechselseitige Aufhebung großer enthalpischer und entropischer Beiträge zurückzuführen ist. Die wichtigsten Wechselwirkungen sind die nichtbindenden van-der-Waals- und elektrostatischen Terme der Potentialenergiefunktion.^[32] Sie wirken zwischen den Atomen des Proteins und zwischen den Proteinatomen und dem Lösungsmittel. Um die Dynamik und Thermodynamik der Faltung zu interpretieren, ist es nützlich, eine effektive Energiefunktion zu definieren, die einen Teil der Freien Energie einer Polypeptidkette in Lösung beinhaltet. Dies entspricht einem „Potential der mittleren Kraft“^[33] und kann aus einem Boltzmann-Mittelwert der Polypeptid-Lösungsmittel-Wechselwirkungen bei einer gegebenen Temperatur über alle Positionen der Lösungsmittelmoleküle für eine gegebene Polypeptidkonformation erhalten werden. Ein Potential der mittleren Kraft kann für das im Gleichgewicht befindliche System exakt definiert werden, so daß es dessen Thermodynamik beschreibt.^[2] Die Ausdehnung dieses Potentials der mittleren Kraft auf zeitabhängige Phänomene wie die Proteinfaltung setzt voraus, daß die Relaxation des Lösungsmittels im Vergleich zu den Bewegungen des beteiligten Proteins schnell ist. Dies scheint eine gute Näherung zu sein, wie die Simulationen der Entfaltung von Proteinen in Wasser verdeutlichen.^[34, 35] Eine solche effektive Energiefunktion, die den nativen Zustand stark begünstigt, enthält alle Teilbeträge der gesamten Freien Energie (z. B. den hydrophoben Effekt, die Solvatation polarer Gruppen sowie die Wasserstoffbrückenbindung, die sowohl van-der-Waals- als auch elektrostatische Wechselwirkungen enthält) ausgenommen die Konformationsentropie, die den denaturierten Zustand stark begünstigt. Dies ist auf die Tatsache zurückzuführen, daß der Konformationsraum, der der Polypeptidkette und ihren Seitenketten zugänglich ist, im nativen Zustand sehr viel stärker eingeschränkt ist als im entfalteten.

Die Analyse der Beiträge zur effektiven Energie (d. h. zum Potential der mittleren Kraft) ergibt, daß die unpolaren Gruppen den gefalteten Zustand stark begünstigen, was auf die

van-der-Waals-Anziehungskräfte in der dichtgepackten (feststoffartigen) nativen Struktur sowie auf den hydrophoben Effekt zurückzuführen ist, der bewirkt, daß die unpolaren Gruppen im Inneren des nativen Proteins verborgen sind.^[32, 36] Im Unterschied dazu leisten die polaren Gruppen (die Peptideneinheiten sowie die polaren und geladenen Seitenketten) einen viel kleineren Beitrag zur Stabilität des nativen Zustands, da ihre Wechselwirkungen im Inneren des Proteins gegen die mit dem Lösungsmittel aufzurechnen sind. Für Lysozym beispielsweise ergeben die Rechnungen, daß bei 25 °C die unpolaren Gruppen 450 kcal mol⁻¹ zur Freien Energie der Entfaltung beisteuern und die polaren Gruppen 87 kcal mol⁻¹.^[36] Die Stabilisierung des nativen Zustands durch den Term der effektiven Energie wird durch den Beitrag der Konformationsentropie von 523 kcal mol⁻¹ bei 25 °C fast aufgehoben. Dies ergibt für die Freie Entfaltungsenergie einen Nettobetrag von 14 kcal mol⁻¹, was mit dem daraus resultierenden Wert von 0.1 kcal mol⁻¹ pro Aminosäure typisch für globuläre Proteine ist.^[36] Im Unterschied dazu kann der Energie- oder Enthalpieunterschied zwischen nativem und entfaltetem Zustand wesentlich größer sein; für Lysozym beträgt die Entfaltungsenthalpie bei 25 °C 58 kcal mol⁻¹.^[36, 37]

Ein wichtiges Merkmal der Proteinfaltung ist, daß die Gesamtstruktur, z.B. die in Abbildung 1 dargestellte, von der Sequenz des Proteins bestimmt wird.^[5] Native Faltungsmuster von drei unterschiedlichen Proteinen sind in Abbildung 4 dargestellt. Der Unterschied in der Sequenz führt zu Unterschieden in der Sekundärstruktur (α -Helix- und β -Faltblattanteil) und in der Tertiärstruktur (Gesamtfaltungsmuster). Die dreidimensionalen Strukturen von ca. 6000 Proteinen sind durch Kristallstrukturanalyse und NMR-Spektroskopie bestimmt worden.^[38] Diese Strukturen können in ca. 350 Familien von Faltungsmustern eingeteilt werden, die aus Proteinen unterschiedlicher Sequenz, aber ähnlicher Struktur bestehen.^[39] Man schätzt, daß es insgesamt nur ca. 1000 unterschiedliche Faltungsmuster gibt.^[39–41] Dies ist zwar eine große Zahl, aber sie ist viel kleiner (ca. hundertfach) als die Gesamtzahl der verschiedenen Sequenzen im menschlichen Genom. Manche Faltungsmuster werden von einer großen Zahl von Sequenzen

ingenommen, was vielleicht an der Stabilität dieser Faltungsmuster^[42] oder der Robustheit und der Geschwindigkeit des Faltungsvorgangs liegt, doch für eine gegebene Sequenz gibt es fast immer nur ein Faltungsmuster. Die Einzigartigkeit des nativen Zustands ist auf den Umstand zurückzuführen, daß die Wechselwirkungen, welche die native Struktur stabilisieren, alternative Faltungsmuster derselben Sequenz erheblich destabilisieren. Das heißt, die Evolution hat eine Sequenz mit einem tiefen Energieminimum für den nativen Zustand hervorgebracht, um bei physiologischer Temperatur falsch oder partiell gefaltete Strukturen auszuschließen. Diese „Energilücke“ zwischen dem nativen Zustand und im wesentlichen falsch gefalteten Strukturen^[25, 55a] ist ein Schlüsselement in der Beziehung zwischen Sequenz und Struktur und spielt eine wichtige Rolle bei der Proteinfaltung, wie wir in Abschnitt 5 schildern werden.

Um die Freie Energie eines Proteins mit seinen Konformationen in Bezug zu setzen, muß man diese als Funktion eines Ordnungsparameters beschreiben. Qualitative Kurven der Freien Energie F einer Polypeptidkette im Gleichgewicht als Funktion des Gyurationsradius R_g , eines Ordnungsparameters für die Proteinfaltung, sind in Abbildung 5 für zwei Temperaturen gezeigt.^[2, 43] Für jeden Wert von R_g entspricht die Freie Energie dem Boltzmann-Mittelwert über alle Werte der übrigen inneren Koordinaten einschließlich derer des Lösungsmittels. Die Form der Kurve der Freien Energie hängt stark von den Umgebungsbedingungen ab, insbesondere von der Temperatur. Die Kurve in Abbildung 5a gilt für eine Temperatur, bei der der native Zustand stabil ist. Diese liegt weit unterhalb der Übergangstemperatur T_m für die Faltung, bei der der native und der denaturierte Zustand die gleiche Freie Energie aufweisen. Der native Zustand entspricht einem tiefen, engen Minimum und erfüllt somit das thermodynamische Kriterium für die Stabilität eines Proteins. Innerhalb dieses Minimums haben die Fluktuationen der Atome eine Amplitude von bis zu 2 Å, und diese Bewegungen finden auf einer Zeitskala von Pico- bis Nanosekunden statt. Alle Strukturen, die durch solche Fluktuationen eingenommen werden, sind sehr ähnlich und entsprechen ungefähr den Unterschieden

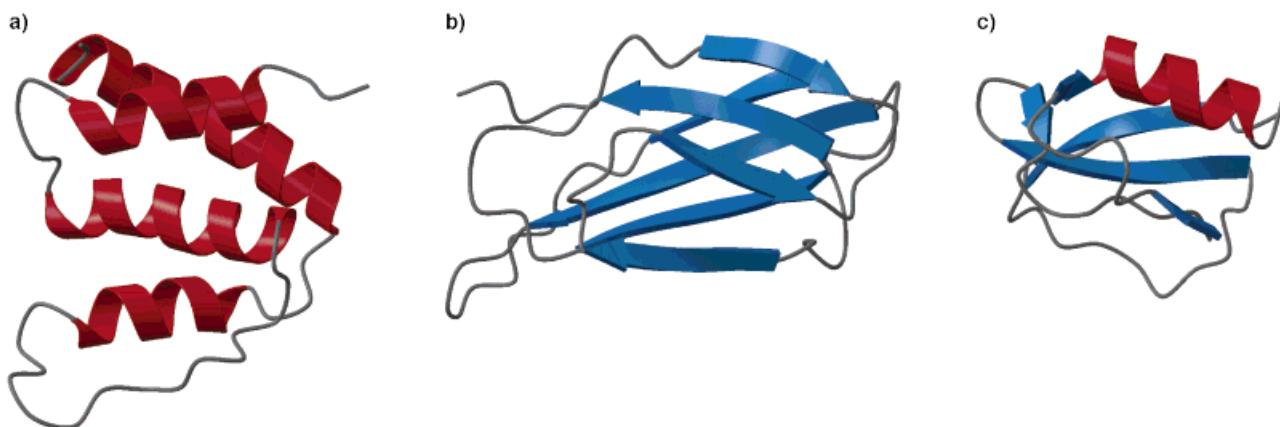


Abb. 4. Repräsentative Faltungsmuster (Tertiärstrukturen) kleiner Proteine. a) Acyl-Coenzym-A-bindendes Protein aus Rind (PDB-Code: 2ABD), ein rein α -helikales Protein. b) Typ-III-Modul der zehnten Domäne des menschlichen Fibronectins (10FNIII; PDB-Code: 1TTG), eines β -Faltblatt-Proteins. c) Chymotrypsininhibitor 2 aus Gerste (PDB-Code: 2CI2), ein α/β -Protein. Die Sekundärstrukturelemente sind rot (α -Helices) und blau (β -Faltblätter) gezeichnet. Diese Abbildungen wurden mit den Programmen Molscript^[163] und Render^[164] erstellt.

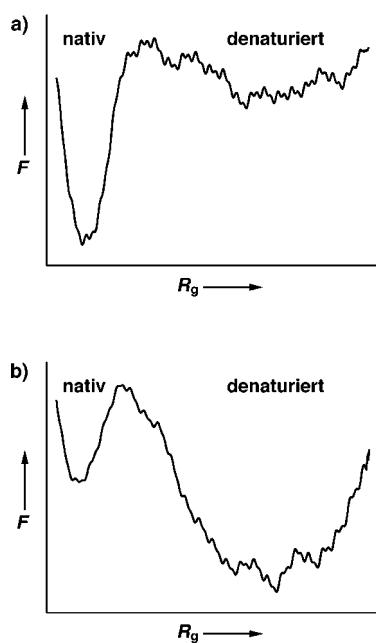


Abb. 5. Darstellung der Freien Energie einer Polypeptidkette im Gleichgewicht als Funktion des Gyurationsradius a) bei einer Temperatur unterhalb der Übergangstemperatur T_m , bei der der native Zustand stabil ist, und b) bei einer Temperatur oberhalb von T_m , bei der der denaturierte Zustand stabil ist. Die Bereiche, die dem nativen und dem denaturierten Zustand entsprechen, sind gekennzeichnet. Die Kurven sind schematisiert, sollten aber für reale Proteine repräsentativ sein.

zwischen den Kristallstrukturen von Proteinen, die Mitglieder derselben Familie sind (z.B. Myoglobin und Hämoglobin aus der Globin-Familie^[44]). Dieser Bereich des Konformationsraums, der für die Funktion der Proteine am relevantesten ist, wurde durch viele mit empirischen Potentialen durchgeführte Moleküldynamik-Simulationen untersucht.^[45] Der andere Bereich der Kurve der Freien Energie entspricht der großen Zahl von Konformationen des denaturierten Zustands. Im Laufe des Faltungsprozesses kann die Polypeptidkette von einem relativ langgestreckten Random-coil-Zustand mit großem R_g -Wert über einen Random-globule-Zustand mit flüssigkeitsähnlichen Eigenschaften in den nativen Zustand übergehen. Die Längenskala, auf der diese Strukturänderungen stattfinden, liegt in der Größenordnung von 10–20 Å, die Zeitskala liegt im Mikro- bis Millisekundenbereich. In anderen Fällen kann das Zufallsknäuel vielleicht zu einem Zwischenprodukt mit einer teilweisen Ordnung (z.B. mit einer dem nativen Zustand ähnlichen Sekundärstruktur) führen, das oft als „Molten globule“^[56] bezeichnet wird. Die Bereiche, die dem Random coil, dem Random globule und der nativen Struktur bei der Proteininfaltung entsprechen, haben Analoga im festen, flüssigen und gasförmigen Zustand bei den Atomclustern (siehe Abb. 3). Phasendiagramme für Proteine und einfache Proteinmodelle weisen auf die Existenz dieser Phasen hin.^[2, 47, 48] Abbildung 5b stellt die Freie Energie eines Proteins als Funktion von R_g bei einer Temperatur dar, die deutlich oberhalb von T_m liegt. Mit zunehmender Temperatur wird der denaturierte Zustand durch die Zunahme seiner Konformationsentropie gegenüber dem nativen Zustand stabilisiert. Die Auftragung in Abbildung 5b gilt auch für den

Fall, daß sich die Polypeptidkette in einem denaturierenden Lösungsmittel befindet. Wenn auch der Bereich des nativen Minimums fast unverändert sein mag, hat der Random-coil-Bereich eine niedrigere Freie Energie, vor allem wegen der energetischen Stabilisierung durch die bessere Wechselwirkung zwischen der entfalteten Polypeptidkette und dem Denaturierungsmittel.^[49] Daten für die Potentialenergiefläche und die für die verschiedenen Zustände charakteristischen Fluktuationen können experimentell erhalten werden, beispielsweise aus der Geschwindigkeit, mit der Amidgruppen des Proteins Wasserstoffatome mit dem Lösungsmittel austauschen.^[50]

5. Energieflächen für die Proteininfaltung: ein theoretischer Ansatz

Um die Besonderheiten der Proteininfaltungsreaktion und den Mechanismus zu bestimmen, mit dem das Levinthal-Paradoxon gelöst wird, muß man über die einfache eindimensionale Darstellung, die wir in Abbildung 5 verwendet haben, hinausgehen. Im Unterschied zur $(H + H_2)$ -Reaktion, bei der der Übergang von drei Variablen auf eine die Reaktionskoordinate liefert, erfordert eine Beschreibung, die für das Verständnis der Proteininfaltung nützlich sein soll, den Übergang von mehreren tausend Freiheitsgraden auf eine oder einige wenige Fortschrittsvariablen. Da die Zulässigkeit einer solchen Reduktion der Zahl der Variablen nicht a priori offensichtlich ist, muß man versuchen, sie a posteriori durch Vergleich der theoretischen Vorhersagen mit den experimentellen Messungen zu rechtfertigen.

Viele phänomenologische Modelle der Proteininfaltung sind im Laufe der Jahre vorgeschlagen worden, doch die meisten waren weder spezifisch genug noch ausreichend detailliert, um tatsächlich Rechnungen zu ermöglichen; eine neuere Übersicht über diese Modelle ist in Lit.^[51] zu finden. Die Einführung von Gittermodellen mit vereinfachten Potentialen (siehe Abschnitt 10.1) hat Faltungssimulationen für Systeme ermöglicht, die viele der wichtigsten komplexen Eigenschaften der Faltungsreaktion beinhalten (z.B. das Levinthal-Paradoxon) und doch einfach genug für eine vollständige Analyse sind (d.h., es ist möglich, beliebig genaue Informationen über die Faltung eines solchen Modellsystems zu erhalten). Aus den Gittersimulationen^[6, 14–16] und aus Modellen der statistischen Mechanik^[25] hat man Einsichten in mögliche Faltungszenarien erhalten. Sie liefern die Grundlage dafür, die allgemeinen Eigenschaften des Faltungsmechanismus bei realen Proteinen zu erkunden und Experiment und Theorie zueinander in Beziehung zu setzen. In den folgenden Abschnitten geben wir einen Überblick über die Ergebnisse der Studien mit Gittermodellen und zeigen, was man aus ihnen gelernt hat.

Obwohl Gittermodelle (siehe Abschnitt 10.1) eindeutig zu stark vereinfacht sind,^[52] spielen sie für die Untersuchung der Proteininfaltung eine Rolle, die derjenigen der $(H + H_2)$ -Reaktion für chemische Reaktionen analog ist. Die allgemeinen Konzepte, die aus der detaillierten Analyse der Wasserstoffaustauschreaktion entwickelt wurden, an der nur drei Atome beteiligt sind, dienen als Ausgangspunkt für die Be-

handlung von viel komplexeren Reaktionen mit Aktivierungsbarrieren. Außerdem wurden die Grundlagen der Theorie des Übergangszustands von Wigner^[53] für die ($H + H_2$)-Reaktion entwickelt. Einfache Gittermodelle behandeln eine Aminosäure als einzelnes Teilchen und berücksichtigen nicht die Seitenkette, verwenden Kontaktpotentiale anstelle der sehr viel detaillierteren Potentialfunktionen und beschränken die Bewegung auf ein Gitter anstelle des dreidimensionalen Raums. Trotz alledem können detaillierte Analysen von Gittermodellen nützliche Konzepte für die Faltungsreaktion realer Proteine liefern. Um über die thermodynamische Beschreibung (Abb. 5) hinauszugehen, muß man Fortschrittsvariablen finden, die für die Faltungsreaktion geeignet sind. Eine Beispiel hierfür ist der Gyurationsradius R_g (Abb. 5). Unter bestimmten Umständen ändert sich R_g jedoch im interessierenden Bereich sehr wenig – beispielsweise dann, wenn das System vom Random-globule- zum nativen Zustand fortschreitet und dabei kompakt bleibt. Eine nützlichere Fortschrittsvariable ist der Anteil Q der nativen Kontakte, der von einem Wert nahe Null für den hochgradig denaturierten Zustand (Random coil) bis hin zu einem Wert von Eins für den nativen Zustand variiert. So gibt es für das in Abschnitt 10.1 beschriebene 27mer innerhalb des $3 \times 3 \times 3$ -Würfels 156 mögliche Kontakte, von denen 28 native sind, während es für ein 125mer^[54a] 3782 mögliche und 176 native Kontakte innerhalb eines $5 \times 5 \times 5$ -Würfels gibt. In einem realen Protein kann man, obwohl es sehr viel mehr „native Kontakte“ gibt, eine Untergruppe von 50–100 Kontakten auswählen, die den nativen Zustand definieren sollen.^[54b] Dies spricht für die Verwendung der nativen Kontakte als Fortschrittsvariable für die Faltungsreaktion.

Um die Reaktion als Funktion von Q darzustellen, verwenden wir eine Fläche, welche die mittlere effektive Energie und Konformationsentropie der Polypeptidkette darstellt, die sich vom denaturierten zum nativen Zustand faltet (Abb. 6 und 7). Beide Größen werden durch Mittelwertbildung über eine Reihe von Simulationen für eine gegebene Sequenz ermittelt. Die Verwendung dieser Art von Mittelung empfiehlt sich aus der statistischen Vierteilchenmechanik und aus dem Vergleich zwischen der Proteinfaltung und der Bildung von Festkörpern aus Edelgasen (siehe Abb. 3). Die Ergebnisse hängen von der Temperatur des Systems ab, d.h., die Faltungstrajektorie berührt bei höheren Temperaturen mehr Konformationen mit

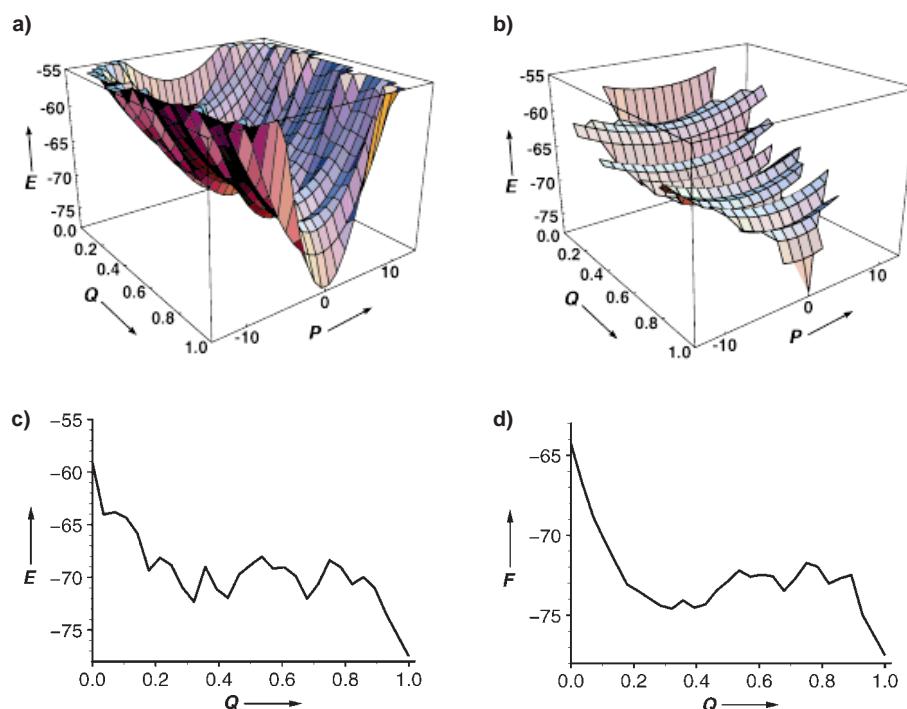


Abb. 6. Gittersimulation der Faltung eines 27mers bei niedriger Temperatur. a) Die mittlere effektive Energie und die Konformationsentropie sind als Funktion der Fortschrittsvariablen Q dargestellt (siehe Abschnitt 10.1). Bei jedem Wert von Q hat der Querschnitt der Fläche die Form einer Parabel, deren Scheitelpunkt der gemittelten effektiven Energie $E(Q)$ entspricht und deren Steilheit der Konformationsentropie $S(Q)$ proportional ist. Zur Veranschaulichung ist hier die Funktion $E(Q,P) = E(Q) + P^2[0.2(S(Q) + 1)]$ verwendet worden, in der P ein Maß für den verfügbaren Konformationsraum ist. b) Hier ist die zugängliche Energiefläche gezeigt, also der Teil der Fläche, der während der Reaktion in einem signifikantem Maß populiert ist. Man erhält diese Fläche, indem man die Ränder der Fläche in a) so abschneidet, daß die Länge der Parabel bei jedem Wert von Q den Betrag $4S(Q) + 0.2$ annimmt. c) Darstellung der gemittelten effektiven Energie E als Funktion von Q . d) Darstellung der effektiven Freien Energie F als Funktion von Q . Die Ergebnisse beruhen auf den Daten aus Lit.^[28] für $T = 0.7$ (willkürliche Einheiten).

höherer Energie. Daher ist es für die Erstellung und die Interpretation eines Energiediagramms wichtig, die Temperatur der Simulation zu berücksichtigen. Die Abbildungen 6 und 7 stellen Ergebnisse von Faltungssimulationen dar, die mit dem Gittermodell für ein 27mer durchgeführt wurden;^[50, 55a] die Ergebnisse dieses Modellsystem sollten auf kleine Proteine anwendbar sein. Abbildung 6 entspricht einer Simulation bei niedriger Temperatur, und Abbildung 7 einer bei hoher Temperatur. In beiden Fällen ist der native Zustand stabil (in Übereinstimmung mit dem Diagramm in Abbildung 5a), und die Faltung erfolgt in einer annehmbaren Zeit. Bei noch niedrigeren Temperaturen könnte die Faltung extrem langsam sein, weil die Kette nicht aus lokalen Minima entkommen kann; bei sehr hohen Temperaturen ist der native Zustand nicht stabil, und die Zahl der zugänglichen Konformationen ist so groß, daß das Levinthal-Paradoxon nicht gelöst werden kann. Die in den Abbildungen 6 und 7 dargestellten Szenarien könnten dem Verhalten unterschiedlicher Proteine unter gleichen Umgebungsbedingungen oder aber dem eines Proteins unter unterschiedlichen Bedingungen entsprechen (siehe unten). Da die resultierenden Flächen auf statistisch-mechanischen Mittelwerten über die von vielen Trajektorien erfaßten Energien beruhen, erscheinen sie glatter als die tatsächliche Schwankung der potentiellen Energie entlang ein-

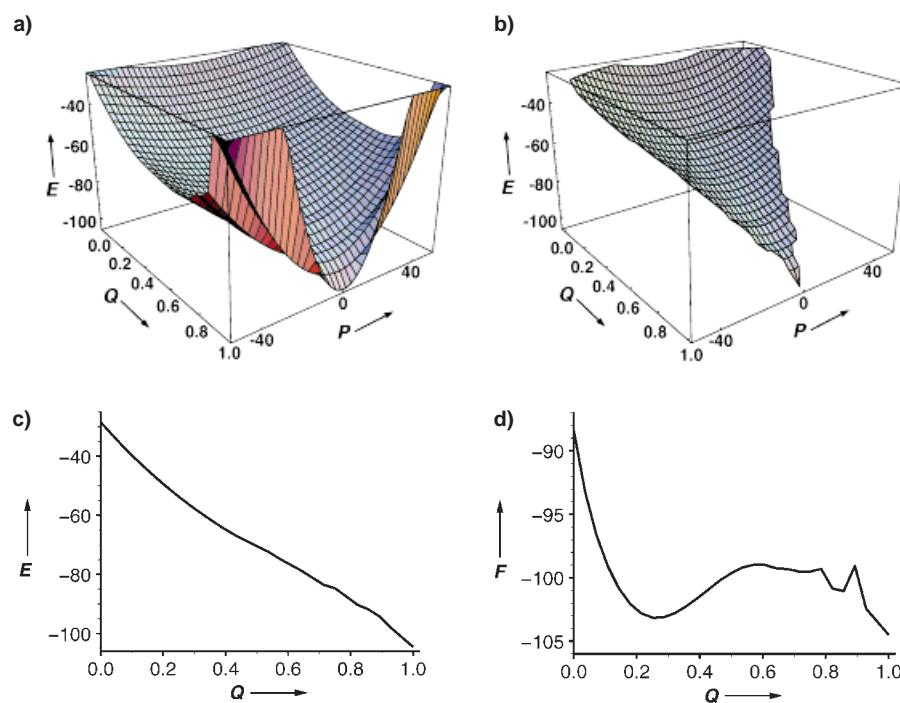


Abb. 7. Gittersimulation der Faltung desselben 27mers wie in Abbildung 6 bei hoher Temperatur. Die Abbildungen a)–d) entsprechen denen in Abbildung 6. Im Unterschied zu Abbildung 6 wurde der native Zustand durch eine Verschiebung der Energien aller 28 nativen Kontakte um einen Zufallswert aus einer Gauß-Verteilung (Mittelwert = $-0.8 T$, Standardabweichung = $0.1 T$) stabilisiert; $T = 2.0$. Der Peak bei $Q = 0.9$ ist möglicherweise ein durch die sehr geringe Zahl von Zuständen ausgelöster Gitterartefakt.

zelter Trajektorien wie der in Abschnitt 10.1 gezeigten. Für Schlußfolgerungen über das allgemeine Verhalten des Systems sind Ergebnisse der in den Abbildungen 6 und 7 gezeigten Art allerdings gut geeignet.

Aus der bei niedriger Temperatur ermittelten Fläche der mittleren Energie/Entropie (Abb. 6a, b) ist es offensichtlich, daß der dem Proteinmolekül zugängliche Konformationsraum selbst bei niedrigen Werten für Q begrenzt ist. Bei der Temperatur, die in Abbildung 6 zugrundeliegt, kollabiert die Proteinkette zu einem „mißgefalteten“ Random globule, dessen Q -Wert nahe dem eines Random coils liegt (siehe Abschnitt 10.1). Die Konformationsentropie bei dieser Temperatur ist so klein, daß ihr destabilisierender Beitrag zur Freien Energie beim Kollabieren auch ohne native Kontakte durch das Verbergen der hydrophoben Gruppen kompensiert wird. Wenn Q von 0.2 auf 0.8 ansteigt, nimmt die durchschnittliche Energie nicht wesentlich ab. In diesem Bereich führt das System eine langsame stochastische Suche innerhalb des kollabierten Zustands durch (siehe Abschnitt 10.1). Wie aus Abbildung 6c ersichtlich wird, ist die Fläche der mittleren effektiven Energie „rau“, d.h. es gibt Mulden und Grate entlang des Talbodens, die auf Energiebarrieren für die Reorganisation innerhalb des kollabierten Zustands zurückzuführen sind. Die gezackten Ränder der zugänglichen Energiefläche in Abbildung 6b spiegeln die „Rauheit“ der Entropie als Funktion von Q wider. Zwischenprodukte würden sich anreichern, wenn die Barrieren hoch genug oder die Temperatur hinreichend tief wären. Für die Sequenz des 27mers, für das die hier gezeigten Ergebnisse ermittelt wurden,^[30] gab es in den Gittersimulationen selbst bei den niedrigsten Temperaturen kei-

ne Abweichung vom einfachen exponentiellen kinetischen Verhalten.^[56] Dies deutet darauf hin, daß die Energiemulden keine wichtige Rolle spielen; ein komplexeres Verhalten zeigen andere Systeme (siehe unten). Der Übergangsbereich der Simulation liegt nahe dem nativen Zustand ($Q \approx 0.8\text{--}0.9$), wie das Profil der Freien Energie zeigt (Abb. 6d). Des Weiteren ergibt die Simulation, daß die Faltung hochkooperativ ist in dem Sinne, daß nur die Q -Werte nahe dem Zufallswert von 0.2 und die des nativen Zustands während der Reaktion nennenswert populiert sind.^[55b]

Abbildung 7a zeigt die Fläche der mittleren effektiven Energie/Entropie bei einer hohen Faltungstemperatur; die zugängliche Fläche ist in Abbildung 7b dargestellt. Zu Beginn der Faltung (z.B. für $Q = 0.2$) ist die Oberfläche sehr breit, was andeutet, daß sich das System über einen weiten Bereich von Konformationen hinwegbewegt; tatsächlich sind die

meisten der Random-coil-Konformationen, die das Levinthal-Paradoxon hervorrufen, zugänglich. Danach verengt sich die Oberfläche kontinuierlich, und die Entropie nimmt ab, während Q bis zum Wert Eins für die native Struktur zunimmt. Die Fläche der mittleren effektiven Energie wirkt demnach wie ein Trichter,^[57] in der Weise, daß die Moleküle ungeachtet ihrer Anfangsposition im Konformationsraum auf der Energiefläche abwärts, auf den nativen Zustand zu diffundieren, was eine Folge der zunehmenden Zahl der stabilisierenden Kontakte ist. Die Steilheit des Energiegradienten sollte mit dem Verhältnis der Energie des nativen Zustands zur durchschnittlichen Energie der Random-coil-Konformationen korreliert sein. Dies führt zum Energielücken-Kriterium (siehe Abschnitt 4) für die gleichzeitige Optimierung von thermodynamischer Stabilität und schneller Faltung.^[25, 55a] Es gilt für kleine Proteine, die durch das 27mer repräsentiert werden; für das 125mer findet man ein komplexeres Faltungsverhalten^[54a] (siehe Abschnitt 7).

Trotz der Glattheit der effektiven Energiefläche (Abb. 7c) kann es eine Übergangszustandsbarriere im Profil der Freien Energie geben. In den Simulationen für das 27mer entsteht diese Barriere (Abb. 7d) durch die Verringerung der Kettenentropie bei großen Q -Werten; d.h., die Zahl der zugänglichen Konformationen nimmt stark ab, wenn sich Q dem Wert für den nativen Zustand annähert (Abb. 7b). Dies führt zu einer „Verengung“ im Faltungstrichter und damit zu einem Entropie-„Engpaß“. Dies ähnelt den unimolekularen Dissoziationsreaktionen kleiner Moleküle, bei denen es aufgrund der Notwendigkeit, daß ein hinreichender Anteil der inneren Energie in dem mit der Dissoziation assoziierten Freiheits-

grad enthalten ist, ebenfalls einen Engpaß gibt.^[1, 58] Beim Hochtemperatur-Szenario aus Abbildung 7 nehmen sowohl die Energie als auch die Entropie im wesentlichen monoton ab, während das System sich dem nativen Zustand nähert. Demnach ist es das Verhältnis zwischen den Geschwindigkeiten der Abnahme der Energie und der Entropie, das bestimmt, ob es eine Barriere der Freien Energie gibt und wo diese auftritt. Eine anderes Verhältnis zwischen der gemittelten effektiven Energie und der Konformationsentropie könnte die Übergangsbarriere der Freien Energie zu kleineren Werten von Q als den in Abbildung 7d gezeigten bewegen. Die Kinetik für ein einfaches Modell, das auf einem Trichter mit einer „späten“ Barriere beruht, ist durch eine Mastergleichung beschrieben worden, die zusätzliche Einsicht verschafft.^[59]

Eine Fläche der in Abbildung 7 gezeigten allgemeinen Gestalt wird wahrscheinlich zu einer schnellen Faltung führen, und zwar viel schneller, als man es für die in Abbildung 6 gezeigte Fläche erwarten würde. Die Zahl der erforderlichen Monte-Carlo-Schritte liegt in der Größenordnung von 10^7 für Abbildung 7 (durch Simulationen erhalten) und von 10^{10} für Abbildung 6; diese Zahl wurde aus Abbildung 8 durch Extrapolation erhalten (siehe unten), da die für die Ermittlung eines genauen Werts erforderlichen Monte-Carlo-Simulationen zuviel Computerrechenzeit erfordert hätten. Nach den in Abschnitt 10.1 dargelegten Größenordnungsargumenten entspricht dies einem Wert von 10 ms für die schnelle Faltung (Abb. 7) und von 20 s für die langsamere Faltung (Abb. 6). Beide liegen im Bereich der tatsächlichen Faltungszeiten realer Proteine. Abbildung 7 stellt einen einfachen, aber nichttrivialen Trichter dar; er ist einfach, weil die mittlere effektive Energie im wesentlichen monoton in Richtung des nativen Zustands abnimmt, und nichttrivial, weil die Zahl der Zustände, die für die Faltungstrajektorien zugänglich sind, schneller abnimmt, als man aufgrund der Zunahme von Q erwarten würde. Faltungsszenarien, die auf Abbildung 7 beruhen, sind als Lösungen des Levinthal-Paradoxons einleuchtend und ansprechend. Sie entsprechen einer schnellen, vorbestimmten Diffusion der Fortschrittsvariablen zum nativen Zustand, wobei der geschwindigkeitsbestimmende Schritt die Überquerung einer Entropie-Barriere ist. Eine Beschreibung, die einem Modell vom Kramers-Typ^[31, 60] entspricht, wäre auf diese Reaktion anwendbar. Solch ein vereinfachter Trichter ist die idealisierte Darstellung einer möglichen Art von Faltungsszenarien. Gittersimulationen haben alternative Szenarien nahegelegt, z.B. das in Abbildung 6 gezeigte, und komplexere Szenarien ergeben sich, wenn man Experimente oder Simulationen längerer Gitterketten interpretiert (siehe Abschnitt 7). Ebenso wichtig ist, daß die verschiedenen Faltungsszenarien keine Auskunft darüber geben, wie die Fläche der mittleren effektiven Energie durch die Sequenz und die daraus resultierenden Wechselwirkungen bestimmt wird (siehe Abschnitt 8).

Bei den Szenarien der Abbildungen 6 und 7 ist die Wahrscheinlichkeit für das Erreichen des Übergangsbereichs gering, unabhängig davon, ob es eine entropische oder eine energetische Barriere gibt. Sie ist jedoch viel größer als bei der uneingeschränkten Zufallssuche, die Levinthal betrachtet hatte, und zwar in hinreichendem Ausmaß, so daß diese Modellsysteme sich in annehmbarer Zeit falten können. Wie wir

gesehen haben, ist die wichtigste Voraussetzung hierfür die Neigung der Energiefläche zum nativen Zustand hin, entweder durch einen Energiegradienten und/oder durch die Einschränkung des zugänglichen Konformationsraums.

Eine wichtige Schlüssefolgerung aus den Abbildungen 6 und 7 ist, daß die Flächen der mittleren effektiven Energie und der Freien Energie eines Proteins eine starke Temperaturabhängigkeit aufweisen, im Unterschied zur Situation bei einfachen chemischen Reaktionen wie $H + H_2$, bei denen diese Oberflächen nahezu unabhängig von der Temperatur sind. Dies ist auf die geringfügige Rolle der Entropie bei den meisten Reaktionen kleiner Moleküle und ihre essentielle Bedeutung für die Proteinfaltung zurückzuführen. Es gibt einen auffallenden Unterschied zwischen den Temperaturabhängigkeiten der Geschwindigkeitskonstanten beider Reaktionstypen (Abb. 8). Die Geschwindigkeitskonstante für die Reaktion $H + H_2$

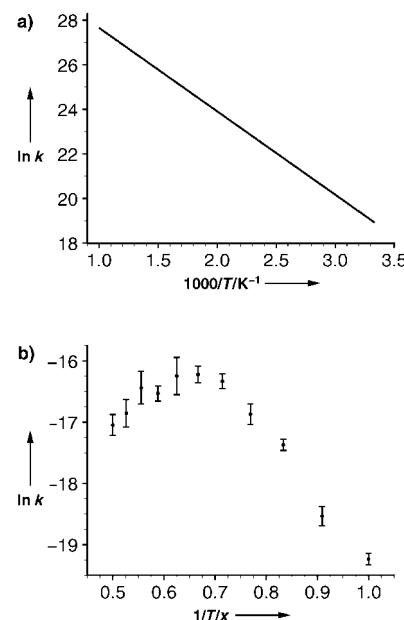


Abb. 8. Arrhenius-Auftragungen von $\ln k$ als Funktion von $1/T$: a) für die Reaktion $H + H_2$ (nach Lit.^[13]), hier wird für k die Einheit $\text{cm}^3 \text{mol}^{-1} \text{s}^{-1}$ verwendet; b) für die Proteinfaltungsreaktion am Beispiel der Sequenz aus Abbildung 6 (nach Lit.^[56]), hier ist k pro Monte-Carlo-Schritt angegeben. $x =$ willkürliche Einheit; für $x = 1.0$ ist $T \approx 300 \text{ K}$.

(Abb. 8a)^[13] zeigt ein Arrhenius-Verhalten, wie man es aufgrund des Reaktionsprofils in Abbildung 2d erwarten würde. Die Reaktion wird von der Aktivierungsenergie bestimmt, und die Aktivierungsentropie spielt nur eine kleine Rolle. Dies führt dazu, daß die Freie Aktivierungsenergie nahezu temperaturunabhängig ist. Für die Proteinfaltungsreaktion ist $\ln k$ als Funktion von $1/T$ in Abbildung 8b dargestellt.^[56] Bei niedrigen Temperaturen entspricht das Verhalten dem nach der Arrhenius-Gleichung erwarteten. Wenn jedoch die Temperatur zunimmt, weicht das Verhalten deutlich von dem nach der Arrhenius-Gleichung erwarteten ab: Die Reaktionsgeschwindigkeit erreicht ein Maximum und nimmt dann bei weiterer Temperaturerhöhung wieder ab. Dieses vorhergesagte Verhalten wurde bei Studien zur Temperaturabhängigkeit der Geschwindigkeit der Proteinfaltung experimentell bestätigt.^[23, 61] Eine solche Krümmung im Arrhenius-Plot

weist darauf hin, daß die Reaktion bei niedrigen Temperaturen von der Enthalpie und bei hohen von der Entropie beherrscht wird. Sowohl in der Simulation als auch im Experiment gibt es einen Beitrag der Aktivierungswärmekapazität, doch der Ursprung dieses Beitrags ist möglicherweise in beiden Fällen verschieden, da der Parameter für Anziehungskräfte im Gitter temperaturunabhängig ist, die hydrophobe Wechselwirkung hingegen nicht.^[23] Für die Entfaltung erwartet man wegen der großen Aktivierungsbarriere für das Entkommen aus dem nativen Zustand, die schematisch in Abbildung 5 a gezeigt ist, eine Temperaturabhängigkeit, die der Arrhenius-Beziehung entspricht. Entfaltungsexperimente mit Lysozym^[61] und Barnase^[62] haben ein solches einfaches Verhalten bestätigt; die Geschwindigkeitskonstante variiert mit der Temperatur wie in Abbildung 8 a dargestellt.

Die beiden in den Abbildungen 6 und 7 gezeigten Arten von Verhalten könnten für verschiedenartige Systeme bei unterschiedlichen Temperaturen eintreten, je nach dem Verhältnis zwischen energetischen und entropischen Faktoren und seiner Variation mit der Temperatur. Dies ist in Abbildung 8 b dargestellt. Für eine Aminosäuresequenz bei physiologischer Temperatur könnte das tatsächliche Verhalten zwischen den von den Abbildungen 6 und 7 beschriebenen Möglichkeiten liegen. Wahrscheinlich finden Faltungsszenarien statt, die Elemente von beiden aufweisen; d.h., für bestimmte Werte der Fortschrittsvariablen tritt ein trichterartiges Verhalten ein, für andere eine stochastische Suche. Es ist zu erwarten, daß die in Abbildung 8 b gezeigte Kurve für die Temperaturabhängigkeit der Geschwindigkeitskonstante der Faltungsreaktion für die meisten Proteine gilt, obwohl die optimale Faltungstemperatur und die maximale Faltungsgeschwindigkeit individuell verschieden sein werden.

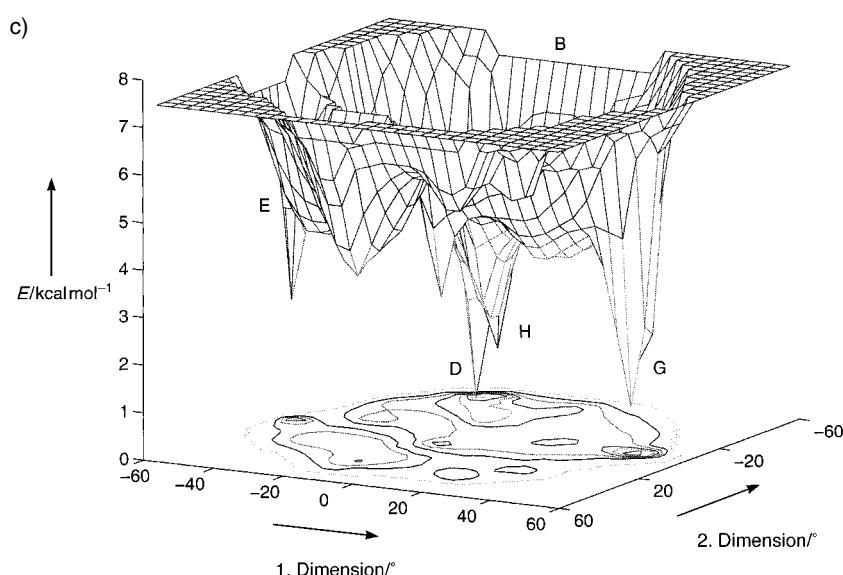
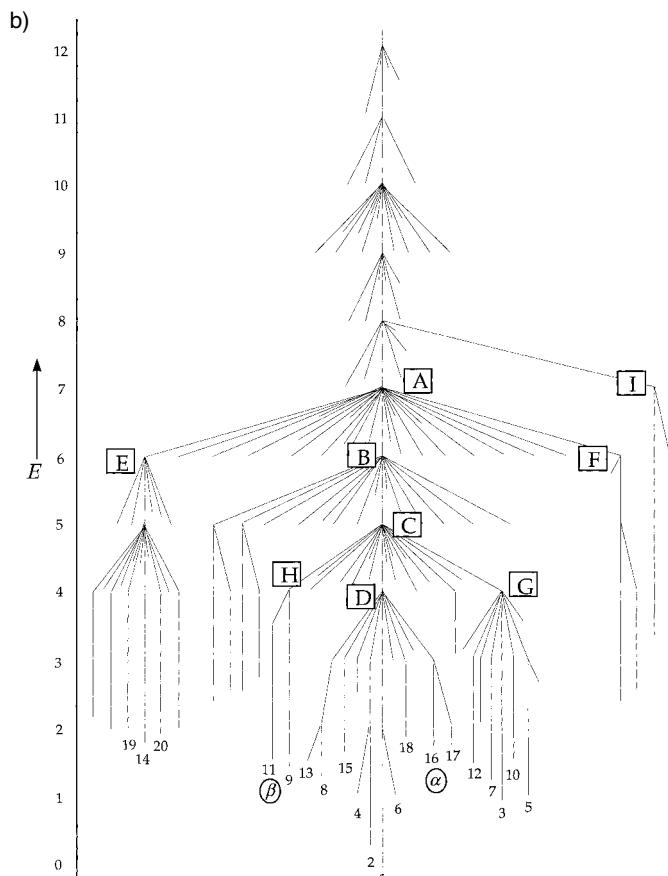
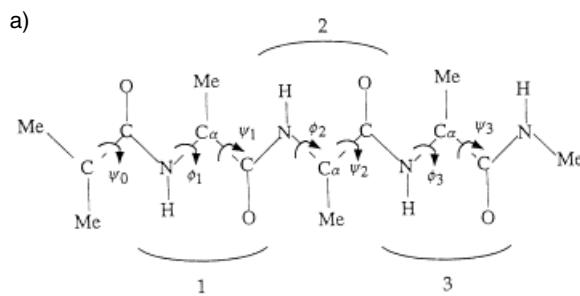
Die Ergebnisse der Gittersimulationen deuten auf die essentielle Bedeutung der Flächen der Enthalpie und der Freien Energie für reale Proteine hin. Obwohl die detaillierte Bestimmung der tatsächlichen Fläche für ein Protein jenseits gegenwärtiger Rechnerkapazitäten liegt, ist es interessant, hier die Ergebnisse zu betrachten, die für ein sehr kleines Modellsystem erhalten wurden: das durch ein detailliertes atomares Modell beschriebene Alanin-Tetrapeptid.^[63, 64] Abbildung 9 a zeigt die Struktur des Peptids und die wichtigen Diederwinkel seiner Konformation. In Abbildung 9 b ist die komplexe, baumartige Struktur der Potentialenergieläche mit mehreren Minima dargestellt. Eine Darstellung der Energie als Funktion der beiden orthogonalen Koordinaten (Hauptkomponenten), die die beste Beschreibung liefern, ist in Abbildung 9 c gezeigt. Die eher becken- als trichterförmige Gestalt der Gesamtfläche ist ebenso offensichtlich wie die zahlreichen tiefen lokalen Minima. Die Fläche beschreibt eine recht breite Grube bis hinab zu einer Energie (ca. 4 kcal mol⁻¹), bei der sich die Energieminima in einzelne Becken trennen. Da ein Protein hundertmal so viele veränderliche Diederwinkel hat wie dieses kleine Peptid und die Rechnungen ohne Berücksichtigung eines Lösungsmittels durchgeführt wurden (das möglicherweise einige der lokalen Minima „geglättet“ hätte), ist die Extrapolation auf das zu erwartende Verhalten schwierig. Trotzdem macht das Peptid deutlich, daß es viele Fallen geben kann, die vermieden werden müssen, wenn die Sequenz sich schnell zum Energieminimum des nativen Zustands fal-

ten soll. Wie dies geschieht, ist eine wichtige Frage, die noch beantwortet werden muß, obwohl es bei einer größeren Zahl von Freiheitsgraden eine erhöhte Wahrscheinlichkeit geben mag, diese Barrieren und Fallen zu vermeiden.

6. Experimentelle Befunde über Energiefächen

Ein wichtiges Ziel bei der experimentellen Untersuchung jeder Reaktion ist es, Informationen über die Strukturen der Reaktanten in Abhängigkeit von der Zeit zu erhalten und diese mit der Energiefäche in Beziehung zu setzen. Obwohl dies im Detail bisher nur für die Reaktion H + H₂ und die einiger anderer kleiner Moleküle (siehe Abschnitt 2) erreicht wurde, haben die aus diesen Untersuchungen gewonnenen Einsichten ein begriffliches Gerüst für einen weiten Bereich von Reaktionen bereitgestellt, einschließlich der meisten Reaktionen, die in der Organischen Chemie von Bedeutung sind.^[66] Bei diesen trifft es oft zu, daß nur ein kleiner Teil des Moleküls tatsächlich an der Reaktion teilnimmt, so daß die Anwendung der Theorie der Reaktionsgeschwindigkeit nicht viel komplexer ist als für sehr einfache Reaktionen.^[67] Das Problem ist bei Proteinen sehr viel anspruchsvoller, weil bis auf die letzten Stadien der Faltungsreaktion immer ein sehr großer Bereich von Struktureinheiten beteiligt ist, wobei jeweils die gesamte Polypeptidkette betroffen ist. Das bedeutet, daß man die Verteilungsfunktion, welche die beteiligten Strukturen beschreibt, als eine Funktion der Fortschrittsvariablen bestimmen muß. Nur in einigen einfachen Fällen, z.B. den in den Abbildungen 6 und 7 dargestellten, läßt sich die Zeitabhängigkeit der *mittleren* Struktur durch einfache kinetische Schemata interpretieren, so wie bei Reaktionen kleiner Moleküle. Wegen solcher Komplikationen war ein beträchtlicher Einfallreichum gefragt, um Methoden zu ersinnen, die für Strukturuntersuchungen bei der Proteininfaltung geeignet sind (siehe Abschnitt 10.2). Darüber hinaus finden viele Faltungsreaktionen im Bereich von Millisekunden bis Sekunden statt, was im Vergleich mit den Zeiten, die für die Bestimmung von Proteinstrukturen in Lösung mit konventionellen Techniken erforderlich sind, schnell ist. Bei den meisten Untersuchungen erfolgt eine schnelle Störung des Systems aus dem Gleichgewicht, so daß die Annäherung an den nativen Zustand verfolgt werden kann. Da man mit einer einzelnen experimentellen Methode nicht in der Lage ist, mehr als einen kleinen Teil der Einzelheiten beizusteuern, die für eine vollständige Charakterisierung der Faltungsreaktion erforderlich sind, hat man Kombinationen von Methoden verwendet, um die verschiedenen Aspekte der sich bildenden Struktur zu untersuchen (siehe Abschnitt 10.2).

Bei manchen Proteinen ergeben Untersuchungsmethoden für unterschiedliche Aspekte der Struktur sehr unterschiedliche Ergebnisse für das kinetische Verhalten während der Faltung. Einer der frühesten diesbezüglichen Befunde war, daß bei vielen Proteinen die extrinsische Fluoreszenz eines Farbstoffs (Ammoniumsalz der 8-Anilinonaphthalin-1-sulfonsäure, ANS), der zur Reaktionsmischung zugesetzt wurde, früher auftrat als die intrinsische Fluoreszenz der aromatischen Seitenketten.^[68] Da erstere die Bildung von Clustern aus hydrophoben Aminosäuren anzeigen und letztere die dem nativen



Zustand ähnliche Anordnung der Seitenketten, deutet dieses Ergebnis an, daß das Kollabieren zu einem recht ungeordneten Knäuel vor der Bildung einer nativ-ähnlichen Struktur stattfindet. Direkte Hinweise auf einen solchen Kollaps ergeben sich kürzlich in Untersuchungen des Myoglobins durch Röntgenkleinwinkelstreuung (small angle X-ray scattering, SAXS).^[69] Bei der Faltung unter sauren Bedingungen ist nach 6 ms der Gyrationradius nahezu ununterscheidbar von dem im nativen Zustand, der sich unter denselben Bedingungen nach ca. 1 s bildet.

Diese Ergebnisse sind weitgehend konsistent mit dem in Abbildung 6 gezeigten Faltungsszenario, bei dem der Random-globule-Zustand schnell gebildet wird. In der Simulation mit dem 27mer ist die Gesamtzahl der Kontakte groß (≥ 20 von 28), so wie man es für einen kollabierte Zustand erwarten würde, im Unterschied zum Random-coil-Zustand, der im Mittel sieben Kontakte aufweist, von denen man die meisten letztlich aber nicht in der nativen Struktur findet (der durchschnittliche Q -Wert beträgt ca. 0.25). Dieser ungeordnete kollabierte Zustand entspricht einer Situation, in der ein unspezifisches Verbergen der hydrophoben Seitenketten den Verlust an Konformationsentropie kompensiert. Dies erklärt den experimentellen Befund, daß die meisten tertiären Wechselwirkungen bei diesen Zuständen nicht beständig sind und daß Umordnungen der Struktur ein Teil des Faltungsprozesses sind.^[70, 71] Der Kollaps zu einem ungeordneten Knäuel tritt mit besonders hoher Wahrscheinlichkeit bei Proteinen auf, die einen nennenswerten hydrophoben Kern aufweisen, also bei größeren Proteinen, bei denen der Anteil der im Inneren verborgenen im Vergleich zu dem der oberflächenexponierten Aminosäuren groß ist. Andere Faktoren, wie etwa das Vorhandensein von Disulfidbrücken, mögen auch zu diesem Verhalten beitragen, da die Bildung von Quervernetzungen die Entropie-„Gebühr“ für den Kollaps dadurch verringert, daß sie die Unordnung der entfalteten Polypeptidkette begrenzt.^[72]

Die Eigenschaften der kollabierten, aber nichtnativen Zustände ist für mehrere Proteine detailliert erforscht worden (Abb. 10), z.B. für das gut untersuchte α -Lactalbumin.^[46, 73]

Abb. 9. Der Konformationsraum für das blockierte Alanin-Tetrapeptid. a) Das Peptid mit den Diederwinkeln ϕ und ψ des Rückgrats. b) Baumstruktur-Darstellung der 139 Minima der Energiefäche des Tetrapeptids. Die den Minima entsprechenden Energien (jeweils durch die unteren Enden der Linien angezeigt) sind auf der Energieskala [kcal mol^{-1}] relativ zur geringsten Energie angegeben. c) Ein vereinfachtes dreidimensionales Diagramm der Energiefäche des Hauptbeckens A, welches das Konformer mit der geringsten Energie beinhaltet (Becken D), sowie die Projektion auf zwei Hauptkomponenten (eine reduzierte Darstellung im Vergleich mit den sieben veränderlichen Diederwinkeln, welche die Konformation festlegen, siehe a)). Interessanterweise ist der zugängliche Bereich sehr weit ausgedehnt, bis hinunter zu einem Energieniveau von ca. 4 kcal mol^{-1} , bei dem sich die einzelnen Becken voneinander abgrenzen. Es gibt offensichtlich viele „Energiefallen“ auf dieser Oberfläche. (Mit Veränderungen aus Lit.^[63, 64] übernommen.)

Beim α -Lactalbumin findet man, daß sich die Signale für den Fern- und den Nah-UV-Circulardichroismus in Stopped-flow-Experimenten mit sehr verschiedenen Kinetiken entwickeln (Abb. 10b), was darauf hinweist, daß bereits vor der Bildung der dichtgepackten Tertiärstruktur und des nativen Zustands viele Sekundärstrukturelemente im kollabierten Zustand vorliegen.^[74] Untersuchungen von stabilen Analoga dieses unter kinetischer Kontrolle gebildeten kollabierten Zustands durch mehrere Methoden, einschließlich der Bestimmung, in welchem Ausmaß die Wasserstoffatome vor einem Austausch ge-

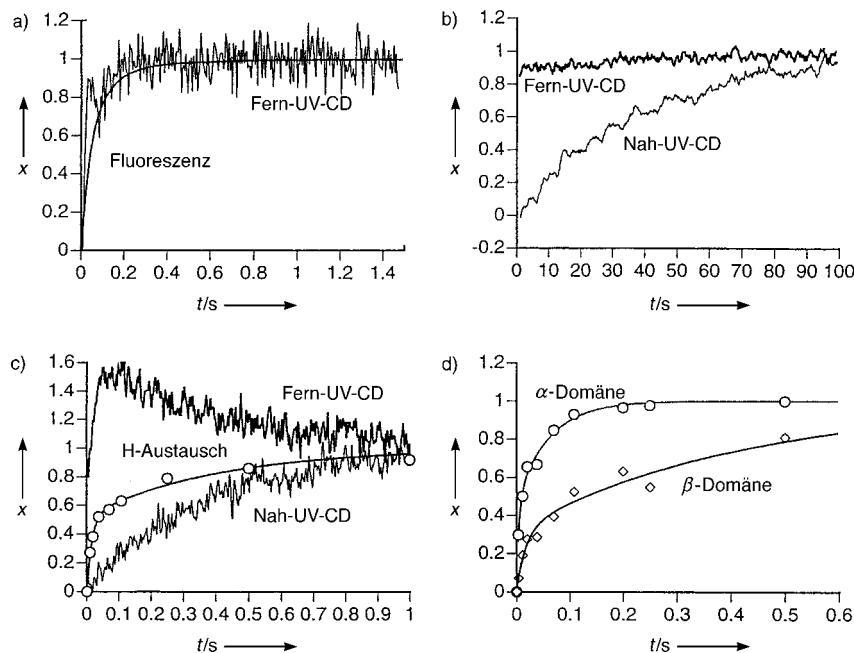


Abb. 10. Experimentelle Ergebnisse aus Untersuchungen der Kinetik der Proteinfaltung. a) Fern-UV-CD- und Fluoreszenz-Daten für das Protein ^{10}FN III ($x = \text{Anteil der nativen Form}$). Alle experimentellen Daten können durch einfache exponentielle Kurven beschrieben werden.^[165] b) Nah- und Fern-UV-CD-Daten für die Apo-Form des Rinder- α -Lactalbumins.^[84] Das Nah-UV-CD-Signal (ein Maß für die Tertiärstruktur) entwickelt sich viel langsamer als das Fern-UV-CD-Signal (ein Maß für die Sekundärstruktur). c) Nah- und Fern-UV-CD-Daten sowie Anteil des Wasserstoffs austauschs bei Hühnereiweiß-Lysozym.^[98] Daß das Fern-UV-CD-Signal zunächst über den Endwert hinausschießt, weist auf die Entwicklung nichtnativer Wechselwirkungen hin, und daß ein Schutz vor Wasserstoffs austausch vor der Bildung der Tertiärstruktur besteht, spricht für das Auftreten von Zwischenstufen mit beständiger Struktur. d) Einzelheiten des in c) gezeigten Schutzes vor Wasserstoffs austausch. Die beiden Domänen des Lysozyms bilden die vor Wasserstoffs austausch schützenden Strukturelemente mit unterschiedlichen Geschwindigkeiten, und die schnelle erste Phase der Schutzwirkung in beiden Fällen deutet darauf hin, daß sich einige Moleküle – verglichen mit dem Rest der Population – auf einer „Schnellspur“ falten.

schützt werden, deuten an, daß die Sekundärstruktur zumindest einen teilweisen, wenn nicht gar einen starken nativen Charakter hat.^[75, 76] Diese Sekundärstruktur ist nicht wichtig für die Faltung des 27mers,^[28, 55a] aber Simulationen eines 125meren Gitterproteins (kubischer $5 \times 5 \times 5$ -Grundzustand) haben ergeben, daß die Sekundärstruktur die Suche des kollabierten Zustands nach dem nativ-ähnlichen Faltungs-„Kern“ beschleunigt.^[54a] Allgemeiner formuliert scheint eine geeignete Mischung aus nativen Kontakt kurzer und langer Reichweite für eine schnelle Faltung wichtig zu sein. Experimentelle Untersuchungen mit α -Lactalbumin haben auch Hinweise auf ein insgesamt nativ-ähnliches Faltungsmuster

und auf spezifische tertäre Kontakte zumindest für einen Kern innerhalb dieser kollabierten Strukturen (Molten globules) ergeben^[77, 78] (Abb. 11), deren wichtige Rolle bei der Faltung erstmals von Ptitsyn et al.^[46, 79] vorgeschlagen wurde. Obwohl nach neueren Experimenten sowohl native als auch nichtnative Kontakte im Molten-globule-Zustand des α -Lactalbumins existieren können,^[80] scheinen die nativen für die Festlegung des Gesamtfaltungsmusters ausschlaggebend zu sein.^[81] Die Faltung der Polypeptidkette zum Molten globule kann eine geringe Kooperativität aufweisen,^[80, 82] während NMR-spektroskopische Untersuchungen zur Kinetik, bei denen die Bewegung einzelner Aminosäuren verfolgt wurde, darauf hinweisen, daß die für den nativen Zustand charakteristische dichte Packung in hochkooperativer Weise entsteht.^[83, 84] Eine Faltung dieses Typs mit einem späten Übergangszustand korreliert mit dem Simulationsergebnis aus Abbildung 6.

Bei einigen Proteinen hat man, im Unterschied zu den oben diskutierten, festgestellt, daß die mit mehreren physikalischen Methoden bestimmten Geschwindigkeiten, mit denen sich die verschiedenen Aspekte der Struktur entwickeln, sehr ähnlich sind^[3] (Abb. 10b). Dies legt nahe, daß es in diesen Fällen keine nennenswerte Anreicherung partiell gefalteter Zwischenprodukte gibt; d.h., Molekülzustände mit einer teilweise gebildeten Sekundärstruktur oder mit einer Sekundärstruktur ohne vollständig gebildete tertäre Wechselwirkungen sind nicht nennenswert populiert. Eine Reihe von Proteinen mit unterschiedlichen Faltungstopologien weist dieses Verhalten auf.^[3] Das wichtigste gemeinsame Merkmal scheint zu sein, daß alle diese Proteine klein sind (≤ 100 Aminosäuren) und in der nativen Struktur nicht mehr als eine Domäne enthalten. Die einfachste Interpretation ist, daß die erhaltenen Ergebnisse auf einer kinetischen „Zwei-Zustands“-Reaktion beruhen; d.h., daß es zu jeder gegebenen Zeit eine Molekülverteilung gibt, die hauptsächlich

aus einer Mischung von „entfalteten“ und nativen Polypeptidketten besteht. Ein Verhalten dieser Art könnte aus einer trichterförmigen Potentialenergiefläche resultieren (siehe Abb. 7), vorausgesetzt, es gibt eine hinreichend hohe Entropiebarriere, die zu einer einfach-exponentiellen Faltungskinetik führt. Die in Abbildung 7 gezeigten Ergebnisse aus Gittersimulationen ergaben eine Entropiebarriere, die spät (bei hohen Werten von Q) auftritt. Dennoch ist, wie bereits erwähnt, eine frühe Barriere (bei niedrigen Werten von Q) ebenfalls möglich; das würde bedeuten, daß nur ein kleiner Bruchteil der nativen Kontakte im Übergangsbereich existieren würde. Dies würde zu einem „Nucleations“-Mechanismus führen, bei

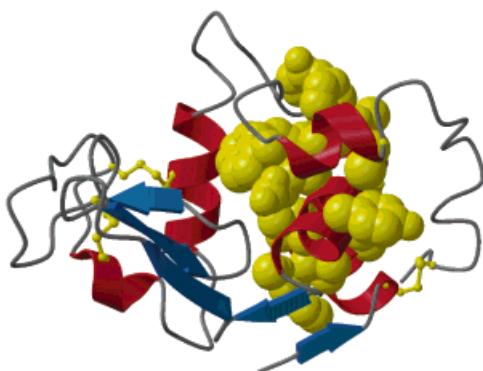


Abb. 11. Ergebnis der NMR-spektroskopischen Untersuchung des Schutzes vor Wasserstoffaustausch im hydrophoben Kern des Molten-globule-Zustands von Pferde-Lysozym. Dieser Kern besteht vor allem aus Aminosäuren, die im nativen Zustand drei der vier α -Helices des Proteins bilden (Helices A, B und D). Verändert übernommen aus Lit.^[78] α -Helices sind rot, β -Falthalbstränge blau gezeichnet, Disulfidbrücken sind als Kugel-Stab-Modelle und die Seitenketten des hydrophoben Kerns als Kalottenmodelle dargestellt; bei diesen wurden statt der van-der-Waals-Radien etwas kleinere Werte für die Radianen verwendet.

dem die Bildung eines Nucleus unter Beteiligung spezifischer Kontakte eines kleinen Teils der Aminosäuren der geschwindigkeitsbestimmende Schritt der Faltungsreaktion ist.^[85, 86] Die Freie Energie des Systems beginnt erst abzunehmen, wenn die energetische Stabilisierung durch nativ-ähnliche Kontakte groß genug ist, um den bei der Nucleation eintretenden Entropieverlust durch die Einschränkung der zugänglichen Konformationen überzukompensieren; eine detaillierte Beschreibung der Nucleation bei der Bildung von Sichelzell-Fasern durch Hämoglobin ist diesbezüglich ebenfalls von Interesse.^[87] Man hat bei Gittermodellsimulationen festgestellt, daß Sequenzen, die die Energielücke zwischen dem Grundzustand und der mittleren Energie eines Random globule optimieren, sich schnell falten, und zwar über einen Nucleationsmechanismus, dessen Übergangsbereich einem Q -Wert von ca. 0.2 entspricht.^[16, 86] Ein solcher Nucleus kann immer noch eine hohe Entropie haben, da viele Konformationen mit ihm vereinbar sind. Ein Nucleationsmechanismus ist auf der Grundlage experimenteller Studien auch für die Faltung kleiner Proteine vorgeschlagen worden^[88] (siehe unten). Es scheint keinen Fall zu geben, bei dem die experimentellen Daten auf der Millisekunden-Zeitskala auf eine graduelle „Wachsen“ von strukturierten Bereichen in einem kleinen Protein hindeuten, wie man es für einen Trichter ohne eine Barriere der Freien Energie erwarten würde. Bei dem auf einer Nucleation basierenden Faltungsmechanismus ist die Struktur des Nucleus, die sich stark von der des nativen Zustands unterscheidet, von besonderem Interesse. Eine besonders geeignete Methode für die Erforschung des Übergangsbereichs ist das Protein-Engineering.^[89] Bei diesem Verfahren tauscht man gezielt Aminosäuren gegeneinander aus und mißt die Änderung der Faltungs- und Entfaltungsgeschwindigkeiten im Vergleich zur Auswirkung des Austauschs auf die Gesamtstabilität des Proteins. Wenn beispielsweise eine Aminosäure im Übergangsbereich nativ-ähnliche Kontakte hat, würde man erwarten, daß eine Mutation ungefähr den gleichen Effekt auf die Stabilität des Übergangszustands wie auf die des nativen Zustands ausübt. Demnach würde die Mu-

tation zu einer Änderung der Faltungs geschwindigkeit, nicht aber der Entfaltungsgeschwindigkeit führen; wenn eine Aminosäure keine Kontakte im Übergangsbereich hat, dann sollte das Umgekehrte zutreffen. Durch Untersuchung einer großen Zahl von Mutanten ist eine detaillierte Charakterisierung des Übergangsbereichs möglich.^[89] Die Methode ist auf mehrere Proteine angewendet worden, einschließlich des Chymotrypsin-Inhibitors CI2^[90, 91] aus Gerste, eines schnell faltenden Proteins mit einer „Zwei-Zustands“-Faltungskinetik der oben diskutierten Art. Man fand, daß eine kleine Gruppe von Aminosäuren im Übergangsbereich nativ-ähnliche Kontakte hat. Diese sind in der durch eine Simulation optimierten Sequenz von CI2 sowie in einer Reihe verwandter Sequenzen hochgradig konserviert.^[92] Dies deutet auf einen breiten Übergangsbereich hin – anstelle eines wohldefinierten Übergangszustands –, weil ein Ensemble von Zuständen mit sehr verschiedenen Strukturen denselben oder einen ähnlichen Nucleus haben kann.^[93] Eine neuere Untersuchung der Faltungskinetiken von zwei Formen von Cytochrom c, die sehr ähnliche native Strukturen, aber nur 60 % Sequenzidentität aufweisen, deutet auch auf einen frühen Übergangszustand hin, in dem nur die konservierten Bereiche eine wohldefinierte Struktur aufweisen.^[94]

Obwohl die beiden in diesem Abschnitt beschriebenen experimentellen Szenarien der Faltung sehr verschieden erscheinen, ist ihnen gemeinsam, daß am Faltungsprozeß ein Absuchen des Konformationsraums nach nativ-ähnlichen Kontakten beteiligt ist, das die Aminosäuren einander räumlich nahebringt. Der Unterschied in den Flächen der mittleren effektiven Energie und die resultierenden Kinetiken spiegeln die Wichtigkeit der Kontakte zwischen bestimmten Aminosäuren wider. Für einige Sequenzen ist ein kollabierter Zustand gegenüber den langgestreckten Zuständen selbst dann energetisch begünstigt, wenn die meisten Kontakte nur Aminosäuren zusammenbringen, die vom richtigen Typ sind (z.B. eine hydrophobe Seitenkette enthalten). Bei der Faltung auf einer solchen Fläche findet die Konformationssuche im kollabierten Zustand statt, und die Umordnung der Kontakte bei der Bildung des späten Übergangsbereichs erfordert möglicherweise die Überwindung nennenswerter Barrieren, die im entfalteten Zustand nicht vorhanden sind. Bei anderen Proteinen ist eine kollabierte Struktur nur beim Vorhandensein einer Gruppe hochspezifischer und energetisch günstiger (nativ-ähnlicher) Kontakte stabil. In diesem Fall kann die Suche nach einem relativ kleinen Nucleus hauptsächlich im entfalteten Zustand stattfinden, und der Übergangsbereich wird frühzeitig während des Faltungsprozesses durchlaufen. Mutationen und/oder Veränderungen der äußeren Bedingungen können die Energiefläche umgestalten und die Details der Faltungsreaktion verändern, indem sie die Balance zwischen der effektiven Energie und der Konformationsentropie so verändern, daß der Übergangsbereich verschoben wird. Dieser Effekt ist experimentell bei mehreren Proteinen festgestellt worden.^[95, 96] Wesentliche Veränderungen in der Position des Übergangsbereichs, die auf dessen labilen Charakter und seine kritische Abhängigkeit von der Balance zwischen effektiver Energie und Entropie hinweisen, sind für einige Ala → Gly-Mutanten (Austausch von Alanin gegen Glycin) der monomeren Form des sich schnell faltenden λ -Repressorproteins nachgewiesen worden.^[97]

7. Komplexität der Faltung

Viele Reaktionen kleiner Moleküle haben komplexere Kinetiken als die ($H + H_2$)-Reaktion. Es kann eine Serie von Zwischenstufen und mehr als ein Produkt geben. Man kann die Theorie der Reaktionsgeschwindigkeit ohne weiteres ergänzen, um solche Merkmale zu berücksichtigen. Dementsprechend ist das bei Proteinen, insbesondere bei größeren, experimentell beobachtbare Verhalten komplexer als bei den im vorhergehenden Abschnitt behandelten Fällen. Um die Proteinfaltung besser zu verstehen, kann man die Erkenntnisse verwenden, die man aus Simulationen einfacherer Systeme erhalten hat, und aus ihnen Energieflächen ableiten, die mit den experimentell beobachteten Eigenschaften vereinbar sind.

Eine Fläche für ein komplexes Faltungsszenario ist in Abbildung 12 schematisch dargestellt. Sie beruht hauptsächlich auf Experimenten mit Lysozym und wurde so gestaltet, daß sie den in den Abbildungen 10c und d dargestellten Ergebnissen entspricht. Insbesondere enthält die Fläche Merkmale, die zwei der im Faltungsprozeß auftretenden komplexen Eigenschaften erklären.^[69, 98–101] Das erste ist, daß sich unterscheid-

bare Zwischenstufen während der Faltung anreichern. Dies deutet darauf hin, daß es in der Energiefläche vor dem Übergangsbereich für die Bildung des nativen Zustands signifikante Minima gibt. Das zweite ist, daß der native Zustand für eine Teilpopulation der rückfaltenden Moleküle schneller erreicht wird.^[100, 101] Dies deutet an, daß es Trajektorien für schnelle und langsame Faltung gibt und daß ein „Grat“ in der Energiefläche den Übergang zwischen ihnen behindert. Diese Ergebnisse kann man mit dem von Guo und Thirumalai diskutierten Mechanismus^[102] des „kinetic partitioning“ (etwa: kinetische Trennung) in Beziehung setzen. Um ein solches Verhalten in die Energiefläche einbauen zu können, muß man den Fortschritt der Reaktion durch mehr als nur die einfache, die nativen Kontakte betreffende Variable Q kennzeichnen, die wir für die oben beschriebenen einfacheren Flächen benutzt haben. Man benötigt mindestens eine weitere Variable. Es ist bezeichnend für unser relativ großes Unwissen über die detaillierten Mechanismen der Faltung, daß es nicht genügend Informationen gibt, um der zusätzlichen Variablen eine genau definierte physikalische Bedeutung zuzuordnen. Dabei könnte beispielsweise die Zahl der nichtnativen Kontakte in einem teilweise gefalteten Zustand beteiligt sein. Diese wäre, ebenso wie der Anteil der exponierten Oberfläche des Proteins, ein Maß für dessen Kompaktheit.

Bei der Faltung des Lysozyms gibt es Zwischenstufen, weil die beständige Struktur sich nicht für das ganze Protein in kooperativer Weise aus dem kollabierten Zustand bildet, der frühzeitig in der Faltungsreaktion entsteht; dieser Kollaps entspricht dem steilen Gefälle am oberen Rand von Abbildung 12b. Statt dessen entwickeln sich nativ-ähnliche Strukturen unabhängig voneinander in den beiden Domänen, die das native Faltungsmuster bilden (siehe Abb. 1) und von denen die eine überwiegend α -Helices, die andere hingegen einen beträchtlichen Anteil an β -Faltblatt enthält.^[103] Da die α -Domäne sich schneller faltet als die andere, ergibt sich ein teilweise gefalteter Zustand, in dem die sich schneller faltende Domäne bereits strukturiert ist. In der Fläche in Abbildung 12b entspricht dies dem tiefen Minimum auf der rechten Seite. Innerhalb jeder Domäne ist die Bildung beständiger Strukturen im wesentlichen kooperativ und kann – muß aber nicht (siehe unten) – ähnlich schnell stattfinden wie die vollständige Faltung eines kleinen Ein-Domänen-Proteins.^[99, 101] Die beiden Domänen vereinigen sich in einem langsameren Schritt, der mit der Bildung des vollständig nativ-ähnlichen Zustands innerhalb der Domänen gekoppelt ist; dies wird durch die Barriere auf der linken Seite von Abbildung 12b angezeigt. In Übereinstimmung mit diesen Experimenten hat man in Gittersimulationen längerer Ketten mehr als eine Faltungsdomäne gefunden.^[54, 104] Die Ergebnisse deuten darauf hin, daß Domänen^[105] ein wichtiges Merkmal bei der Faltung größerer Proteine sind, im Unterschied zu der von kleinen Proteinen (und kurzen Ketten in Gittersimulationen), bei denen oft die gesamte Polypeptidkette die einzige kooperative Einheit ist.^[3, 89]

Es gibt viele Beweise dafür, daß das Auftreten von Zwischenstufen bei der Faltung aller Proteine außer den kleinsten die Regel ist. Wichtige Informationen darüber kann man aus den Ergebnissen von Protein-Engineering-Experimenten erhalten, wie sie in Abschnitt 6 für die Charakterisierung des

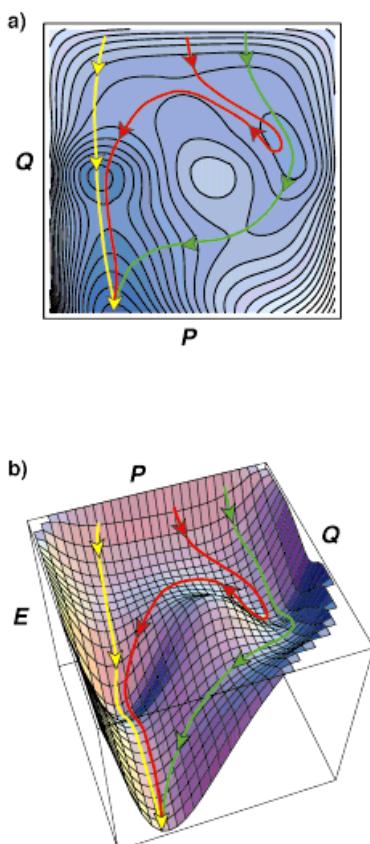


Abb. 12. Aus experimentellen Daten zur Faltung des Lysozyms abgeleitete Fläche der gemittelten effektiven Energie (siehe Text, Abschnitt 10.2 und Abb. 10), dargestellt a) als Konturdiagramm und b) als dreidimensionale Oberfläche. Mehrere mögliche Faltungstrajektorien sind eingezzeichnet: eine Trajektorie für die schnelle Faltung (gelb), eine Trajektorie für die lange Faltung (grün), die über eine hohe Energiebarriere verläuft, und eine Trajektorie für die lange Faltung (rot), die zunächst zu einem weniger gefalteten Zustand zurückkehrt und dann dem der schnell faltenden Variante entsprechenden Tal folgt.

Übergangsbereichs beschrieben worden sind. Eines der größeren Proteine, für die solche Untersuchungen gemacht wurden, ist Barnase, eine Ribonuclease aus *Bacillus amyloliquefaciens*^[89, 106] aus 110 Aminosäuren. Wie Lysozym bildet dieses Protein eine wohldefinierte Zwischenstufe, die in erheblichem Umfang nativ-ähnlich strukturiert ist. Die Protein-Engineering-Experimente ergeben, daß zusätzliche nativ-ähnliche Strukturelemente im Übergangsbereich der Faltung existieren, der bei wesentlich höheren Q -Werten auftritt als bei kleinen Proteinen wie C12. Wie bei Lysozym ist die Zusammenlagerung teilweise vorgeformter Domänen möglicherweise am geschwindigkeitsbestimmenden Schritt beteiligt.^[89]

Der Nachweis kinetischer Zwischenstufen bei der Faltung größerer Proteine wie Lysozym und Barnase hat eine lebhafte Diskussion über ihre Bedeutung ausgelöst.^[76, 107–109] Es gibt hierzu zwei sehr unterschiedliche Ansichten. Eine ist, daß diese Zwischenstufen produktiv sind und dazu beitragen, den Faltungsprozeß in Richtung des nativen Zustands zu lenken, wobei sie als Schablonen für die Bildung nativ-ähnlicher Strukturen dienen und die zu durchlaufenden Konformationen begrenzen. Dies kann die Suche beschleunigen, was die Proteinfaltung wiederum von Reaktionen kleiner Moleküle unterscheidet, bei denen das Auftreten von Zwischenstufen im allgemeinen die Reaktionsgeschwindigkeit verringert. Alternativ könnten die Zwischenstufen kinetische Fallen sein, an denen z.B. falsch gefaltete Konformationen beteiligt sind, die den Faltungsvorgang behindern und die Bildung des nativen Zustands verlangsamen; solche Zwischenstufen tauchen in der Gittersimulation für das 125mer auf.^[54a] Dies entspricht der Situation bei den Reaktionen kleiner Moleküle, bei denen eine Zwischenstufe die Reaktionsgeschwindigkeit verringert, obwohl sie vielleicht eine entscheidende Rolle für die Bildung des Produkts spielt. Wahrscheinlich treten bei der Faltung beide Arten von Zwischenstufen auf, wobei ihre Rolle von der Aminosäuresequenz und den Umgebungsbedingungen abhängt, wie Experimente mit Ubiquitin^[96] und Cytochrom c^[110] ergeben haben.

Das Auftreten von sich langsam faltenden neben sich schnell faltenden Molekülen bei Lysozym erhellt die Bedeutung der Zwischenstufen für die Faltung. Bei manchen Systemen sind unterschiedliche Populationen, die sich mit verschiedenen Geschwindigkeiten falten, an der Isomerisierung des Protein-Rückgrats beteiligt.^[111] Insbesondere sind in einer entfalteten Polypeptidkette das *cis*- und das *trans*-Isomer des Proline ähnlich stabil, und die Barriere für ihre Umwandlung ineinander ist hoch; die Faltungsreaktion schließt deshalb für die Moleküle, die zu Beginn der Faltung als nichtnative Isoptere vorliegen, einen zusätzlichen langsamem Schritt ein. Im Fall von Lysozym scheint jedoch der Ursprung der kinetischen Heterogenität nicht mit Prolin-Resten zusammenzuhängen.^[70, 100] Das Auftreten paralleler Faltungswege weist darauf hin, daß relativ früh in der Faltungsreaktion Zustände mit unterschiedlichen Konformationen entstehen und daß erhebliche Barrieren für die gegenseitige Umwandlung zwischen den sich langsam und den sich schnell faltenden Populationen existieren; die sich langsam faltenden Zustände müssen eine höhere Barriere überwinden, um den nativen Zustand zu erreichen oder um sich zu entfalten und zum

schnellen Faltungsweg zurückzugelangen (siehe Abb. 12). Die experimentellen Studien zeigen, daß die auf den Trajektorien der langsamen Faltung befindlichen Zwischenstufen Elemente der nativ-ähnlichen Struktur in einer der beiden Domänen des Faltungsmusters aufweisen, während jene auf der Trajektorie der schnellen Faltung solche Strukturen in beiden Domänen enthalten.^[99, 101]

Diese Befunde implizieren, daß am geschwindigkeitsbestimmenden Schritt des Faltungsprozesses bei den sich langsam faltenden Lysozymmolekülen die Umordnung falsch gefalteter Zustände beteiligt ist.^[70] Da die kollabierten Zustände des Lysozyms beträchtliche Sekundärstrukturanteile und vier Disulfidbrücken aufweisen, ist es naheliegend, daß für diese Umordnungsprozesse erhebliche Barrieren bestehen, weshalb sie experimentell nachweisbar sind. Diese Schlußfolgerungen werden durch die Ergebnisse aus Gittersimulationen gestützt. Im kollabierten Zustand, der bei den Simulationen für das 27mer auftritt, gibt es viele nichtnative Kontakte; d.h., die Zahl N der Kontakte ist viel größer als Q (siehe Abschnitt 10.1). Obwohl die resultierenden Zwischenstufen nicht stabil genug sind, machen sie sich als Unebenheiten in der Energiefläche bemerkbar (siehe Abb. 6c). Die Energiefläche ist so beschaffen, daß die Polypeptidkette sich nicht wesentlich strecken muß, um aus den Minima zu entkommen und sich dem nativen Zustand zu nähern. Trajektorien, für die falsch gefaltete Zwischenstufen eine Rolle spielen, traten in Gittersimulationen für 125mere auf.^[54a] In diesen Simulationen wird in einem frühen Stadium der Reaktion ein Faltungskern (Nucleus) gebildet (siehe Abschnitt 6). Er besteht aus ca. 30 Aminosäuren und weist einen beträchtlichen Sekundärstrukturanteil auf. Die Faltungstrajektorien umfassen sowohl Populationen, in denen sich der Rest der Polypeptidkette um diesen Kern zusammenlagert und direkt den nativen Zustand bildet, als auch solche, in denen falsch gefaltete, aber relativ stabile Strukturen gebildet werden, die die Faltungsreaktion verlangsamen. Bei diesen ist ein langsamer Schritt unter beträchtlicher Entfaltung erforderlich, bevor der native Zustand erreicht werden kann; d.h., der Wert von Q wird erheblich verringert, bevor er zunehmen kann. Dieser Fall tritt beispielsweise für Trajektorien ein, bei denen zwei Teile des Systems nativ-ähnliche Strukturen entwickelt haben, aber falsch zueinander orientiert sind. Die Ergebnisse der Simulationen für 125mere sind weitgehend mit den experimentellen Daten für Lysozym und auch mit neueren Ergebnissen für Cytochrome c in Einklang.^[96, 112]

8. Eine vereinheitlichte Betrachtung der Proteinfaltung

Es ist in diesem Stadium nützlich, das allgemeine Bild der Proteinfaltung darzustellen, das sich aus der Theorie der Heteropolymere, aus Gittersimulationen vereinfachter Modelle und aus Experimenten zu entwickeln beginnt. (Wir werden in diesem Abschnitt, in dem die bislang vorgestellten Ideen und Ergebnisse zusammengeführt werden, die Literaturzitate nicht wiederholen.) Unter für die Rückfaltung der Polypeptidkette günstigen Bedingungen stabilisieren die van-der-Waals- und die elektrostatischen Wechselwirkungen inner-

halb des Proteins sowie die zwischen Protein und Lösungsmittel den nativen Zustand. Doch die höhere Stabilität des nativen im Vergleich zu der des denaturierten Zustands erklärt allein noch nicht, wie die Polypeptidkette vom letzteren ausgehend den ersten findet (also einen einzigen Zustand aus einer astronomisch hohen Zahl denaturierter Konformationen). Wie die Gittersimulationen ergeben, besteht der entscheidende Punkt darin, daß die Suche nicht zufällig verläuft, sondern durch die Tatsache erheblich eingeschränkt wird, daß nur ein Teil der Fläche der gemittelten effektiven Energie zugänglich ist (siehe Abb. 6 und 7). Das heißt, gewisse Bereiche haben eine zu hohe Energie, als daß sie mit nennenswerter Wahrscheinlichkeit von Faltungstrajektorien erreicht werden könnten. Diese Einschränkungen für den abzusuchenden Raum und die generelle Neigung des Energiegefälles in Richtung des nativen Zustands bedeuten, daß bei der Polypeptidkette nie das im Levinthal-Paradoxon formulierte Problem besteht.

Im einen Szenario der Proteinfaltung kollabiert die Polypeptidkette schnell zu einem kompakten Knäuel. Der wichtigste Faktor, der zu einem solch frühen Kollaps führt, ist das Verbergen der hydrophoben Gruppen, doch die Art des Kollapses hängt von der Heterogenität der stabilisierenden Wechselwirkungen ab. Der Kollaps kann zu einem unorganisierten Knäuel ohne spezifische Kontakte und mit einer relativ großen Konformationsentropie führen oder zu einer kompakten Struktur mit einigen sehr wahrscheinlichen Kontakten und einer geringeren Konformationsentropie. Im ersten Fall erfordert der Kollaps keinen Suchvorgang (d. h., Kontakte können in beliebiger Reihenfolge gebildet werden) und kann deshalb schnell erfolgen. Fluktuationen innerhalb des kollabierte Zustands bringen verschiedene Gruppen miteinander in Berührung, und die Suche nach den kritischen, zum Erreichen des Übergangsbereichs erforderlichen Kontakten ist der langsame Schritt. Diese Art der Faltung hat viele Merkmale, die den in Abbildung 6 gezeigten Flächen von effektiver und von Freier Energie entsprechen. Unter anderen Umständen entspricht der Kollaps der Kette nicht einer Abnahme der Freien Energie ohne das Auftreten spezifischer Wechselwirkungen. Der Suchvorgang, der die erforderlichen Aminosäuren zusammenbringt, kann durch intrinsische Strukturpräferenzen und die Bildung hydrophober Cluster, die selbst im hochgradig entfalteten Zustand vorliegen, unterstützt werden. Diese rudimentären Strukturelemente können weiterhin durch die Bildung von Wasserstoffbrückenbindungen innerhalb der Hauptkette (z. B. zwischen benachbarten Strängen eines β -Faltblatts oder in α -Helices) stabilisiert werden, so daß größere Strukturbereiche in synergistischer Weise entstehen. Vorübergehende Fluktuationen solcher sich bildenden Strukturen innerhalb eines nichtkompakten Zustands führen nur dann zum Kollaps, wenn die erforderlichen stabilisierenden nativ-ähnlichen Kontakte gebildet werden. Sie bilden einen Faltungskern (Nucleus), um den herum der Rest der Struktur entstehen kann. Der langsame Schritt entspricht der Suche nach der Kernstruktur, von der aus die Faltung rasch zum nativen Zustand fortschreiten kann. Es gibt einen relativ frühen Übergangsbereich, und Proteine mit einer solchen Fläche der Freien Energie werden sich wahrscheinlich schnell und in einer scheinbaren „Zwei-Zustands“-Reaktion

falten. Je instabiler der native Zustand ist, um so mehr nativ-ähnliche Kontakte werden für die Bildung des kompakten Molten-globule-Zustands benötigt und um so langsamer ist der Faltungsprozeß.

Mit zunehmender Größe der Proteine werden die Szenarien komplexer, obwohl die Wechselwirkungen die gleichen bleiben und die Faltungsmechanismen auf den für kleinere Systeme beschriebenen Elementen aufbauen. Bei der stochastischen Suche oder beim Nucleationsmechanismus kann ein begrenzter Faltungskern von der Größe eines kleinen Proteins gefunden werden, und zusätzliche Beschränkungen der Suche sind erforderlich, um die Faltung in einer annehmbaren Zeit zu ermöglichen. Diese können zum Teil auf den Präferenzen beruhen, welche die Kette bereits im Random-coil-Zustand hat.^[113, 114] Unter nativen Bedingungen können Wechselwirkungen kurzer Reichweite zur Bildung von Strukturelementen (z. B. β -Turns oder gerade entstehende Helices) führen, die als Initiationsstellen die Wahrscheinlichkeit einer Zusammenlagerung zu Strukturen mit Kontakten längerer Reichweite erhöhen. In vielen Fällen werden sich falsche Strukturen bilden, die aufgelöst werden müssen. In den frühen Stadien der Faltung wird dies wahrscheinlich schnell stattfinden, doch später können signifikante Barrieren gegenüber der Umordnung falsch gefalteter Elemente zu einer Verlangsamung von Faltungsschritten führen. Der Kern muß hinreichend stabil sein, damit er lange genug für die Zusammenlagerung der restlichen Polypeptidkette existiert. Zuviel Stabilität in Nicht-Kern-Bereichen kann zu Zwischenstufen führen, von denen manche falsch gefaltet sein können. Populationen sich faltender Ketten können verschiedene Bereiche des Konformationsraums durchlaufen, so daß sie unterschiedliche Barrieren überwinden müssen, was zu einer komplexen Faltungskinetik führt. Diese Art der Faltung mit komplexer Kinetik tritt sowohl in Gittersimulationen für größere Proteine als auch im Experiment auf. Sie erfolgt dann mit hoher Wahrscheinlichkeit, wenn es vergleichsweise strukturierte Teile des Proteins gibt, die in der Sequenz weit voneinander entfernt sind und für die Bildung des nativen Zustands zusammengebracht werden müssen. Daraus folgt, daß die grundlegenden Merkmale des Faltungsmechanismus dieselben bleiben, auch wenn sich die Ereignisse bei der Faltung unterschiedlicher Proteine oder bei der derselben Proteins unter verschiedenen Bedingungen sehr unterscheiden mögen. Die Faltung ist ein fortschreitender Prozeß, in dem nichtnative und native Kontakte, von denen einige besonders wichtig sein können, nativ-ähnliche Strukturelemente stabilisieren. Die Energie-„Gebühr“ für nichtnative Wechselwirkungen oder eine falsche Faltung nimmt während des Faltungsprozesses zu. Die Polypeptidkette wird von stabilisierenden Wechselwirkungen in für die Faltung produktive Bereiche des Konformationsraums gelenkt, die eine wirkungsvolle Suche nach dem nativen Zustand ermöglichen. Obwohl Energiegefälle, die den Konformationsraum einschränken, für die Faltung in einer annehmbaren Zeit essentiell sind, ist ein kontinuierlicher trichterartiger Gradient der effektiven Energie nicht erforderlich, und es ist unwahrscheinlich, daß er in realen Proteinen existiert. Ein eher „stufenweises“ Verhalten ist wahrscheinlich, bei dem Bereiche der Energiefläche aufgesucht werden, in denen der Energiegradient abwärts verläuft, aber auch solche, in

denen er relativ konstant ist oder sogar aufwärts verläuft. Ein analoges Verhalten ist in Simulationen für Edelgas-Cluster festgestellt worden.^[23] Anders als bei Reaktionen kleiner Moleküle sind die Trajektorien, denen die einzelnen Polypeptidketten folgen, wahrscheinlich – zumindest zu Beginn der Faltung – sehr verschieden, wobei die Zahl gemeinsamer Merkmale zunimmt, wenn das System sich dem nativen Zustand nähert.

Um die vorstehende Beschreibung des Faltungsvorgangs zu vervollständigen, muß man verstehen, wie dieser in der Sequenz codiert ist, denn es ist die Sequenz, welche die Energiefläche jedes einzelnen Proteins bestimmt. Diese Information kann man nicht direkt aus den verfügbaren Gittersimulationen erhalten, die verwendet wurden, um den vollständigen Konformationsraum zu erkunden. Diese sind übermäßig vereinfacht und enthalten zu wenig Detailinformation, als daß man ihre jeweiligen Ergebnisse mit denen für einzelne Proteine in Beziehung setzen könnte. Dennoch gibt die gegenwärtige Analyse des Faltungsmechanismus Hinweise darauf, wie die Faltung ablaufen könnte. Ein wichtiger Gesichtspunkt des Faltungsvorgangs ist der, daß die spezifischen Eigenschaften der einzelnen Aminosäurereste wichtiger werden, wenn die Polypeptidkette sich dem nativen Zustand nähert. Die lokalen Wechselwirkungen im entfalteten Zustand führen zu einer signifikanten Wahrscheinlichkeit für die Bildung von nativ-ähnlichen Helices und von Faltblattsträngen innerhalb des Ensembles von sich schnell ineinander um wandelnden Konformeren.^[113, 114] Während eine sich faltende Polypeptidkette einen kollapsartigen Übergang durchläuft, führt die nichtzufällige Verteilung von Aminosäureresten unterschiedlichen Charakters (z.B. von solchen mit hydrophoben oder hydrophilen Seitenketten) zu einer Architektur (einem globalen Faltungsmuster), welche die Eigenschaften des nativen Zustands aufweist. In diesem Stadium der Faltung, das oft zu einem Molten globule führt, sind jedoch die spezifischen Eigenschaften der Reste von geringer Bedeutung, und sehr unterschiedliche Sequenzen mit ähnlichen globalen Eigenschaften können dasselbe Faltungsmuster bilden. Für die Festlegung dieser Architektur sind Kontakte zwischen einigen bestimmten Aminosäureresten wichtiger als andere, da sie Bereiche der Polypeptidkette zusammenbringen, die in der Sequenz nicht benachbart sind. Ein weiterer essentieller Gesichtspunkt der Suche nach dem korrekten Faltungsmuster ist die Notwendigkeit, hochgradig destabilisierende Situationen zu vermeiden, z.B. eine nicht kompensierte Ladung im Inneren des Proteins. Bei der endgültigen Zusammenlagerung zum nativen Zustand werden die Wechselwirkungen in allen Bereichen der Sequenz wichtig. Die Bildung der dichtgepackten Struktur optimiert die van-der-Waals- und andere stabilisierende Wechselwirkungen. Selbst eine Aminosäure in einer nichtnativen Umgebung (beispielsweise eine mit einer Seitenkette, die zu einem großen Hohlraum im Proteininneren führt) kann die Struktur erheblich destabilisieren, obwohl durchaus eine Toleranz gegenüber dem Ersatz bestimmter Reste durch andere besteht.^[115, 116] Da der native Zustand in einer tiefen Energiesenke liegt, erreichen die meisten Polypeptidketten diesen Zustand und bleiben dort, wenn die Bedingungen die Faltung begünstigen. Diese physikalischen Aspekte der Thermodynamik und Kinetik der Proteinfaltung ha-

ben ihre Entsprechung im theoretischen Modell in der Verwendung von Threading(„Einfädelungs“)-Algorithmen zur Identifizierung der aus einer gegebenen Sequenz resultierenden Struktur.^[9, 117–122]

Es ist unwahrscheinlich, daß die Packung der Seitenketten an der Lösung des Levinthal-Paradoxons beteiligt ist, und daher sind die Seitenketten auch nicht in den Gittermodellen enthalten, die für die Entwicklung der hier vorgestellten Konzepte verwendet wurden. Ihre Packung könnte jedoch wichtig sein für einen Übergang erster Ordnung zwischen dem tiefen Energieminimum der nativen Struktur und der großen Population locker gepackter Molten globules mit nativ-ähnlichen Rückgratstrukturen. Tatsächlich scheint es beim Design kleiner Proteine relativ einfach zu sein, ein kompaktes Molten globule mit der korrekten Topologie zu erhalten (z.B. ein Vier-Helix-Bündel).^[123, 124] Die Schwierigkeit dabei ist, dieses dann so zu modifizieren, daß es die nativen Eigenschaften erhält, die aus der feststoffartigen Packung der Seitenketten resultieren. Die Sequenz eines realen Proteins erreicht das Minimum des nativen Zustands durch eine Spezifität, die über die einfachen Muster von hydrophoben Seitenketten einerseits und hydrophilen andererseits hinausgeht; die Sequenz muß so aufgebaut sein, daß bestimmte Reste sich korrekt pakken können, um so die einzigartige native Struktur zu bilden. Eine solche Packung muß so stabil sein, daß sie den zusätzlichen Entropieverlust überwinden kann, den der Übergang von den vielen Konformationen des Molten globule zum einzigartigen nativen Zustand mit sich bringt. Das letzte Stadium der Faltung besteht demnach in der kooperativen Bildung einer hochgradig kompakten Struktur, in der die spezifische Packung der Seitenketten eine Schlüsselrolle spielt.

9. Zusammenfassung und Ausblick

Detaillierte Rechnungen für die Reaktion $H + H_2$ sowie das Konzept der Reaktionsflächen haben das grundlegende Verständnis chemischer Reaktionen revolutioniert. Diese „neue Sichtweise“,^[107, 125, 126] die auf den Flächen der gemittelten effektiven Energie oder Energie-„Landschaften“^[14, 17, 127] beruht, hat derzeit eine ähnliche Bedeutung für unser Verständnis der Proteinfaltung. Die neue Sichtweise ist aus theoretischen Modellen und aus Gittersimulationen entstanden. Letztere haben eine unverzichtbare Rolle gespielt, weil sie viele komplexe Eigenschaften der Proteinfaltung berücksichtigen, aber noch einfach genug sind, um detaillierte Untersuchungen des Faltungsvorgangs zu ermöglichen.^[6, 52] Es gibt einen wesentlichen Unterschied zur Sichtweise Levinthals, die wir in Abschnitt 1 dargestellt haben. Die astronomisch hohe Zahl von Zuständen für das denaturierte Protein muß nur dann in einer zufälligen Weise nach dem nativen Zustand abgesucht werden, wenn alle außer diesem im wesentlichen die gleiche Energie aufweisen. Diese Version ist als das „Golfplatz-Modell“ der Energiefläche bezeichnet worden.^[25, 55a] Wie die Gittersimulationen für das 27mer ergeben, enthält sogar ein Satz von Zufallssequenzen einen hohen Anteil an Sequenzen mit einer schnellen Faltung.^[55a] Die Suche unter den denaturierten Konformationen ist stark eingeschränkt, weil nur ein kleiner Teil der Zustände hinreichend niedrige Ener-

gien hat, um bei einer gegebenen Temperatur erreicht werden zu können. Möglicherweise wird die Suche auch durch eine Energieabnahme in Richtung des nativen Zustands geleitet. Dies führt zu einer Reihe von Faltungsszenarien, bei denen es viele Wege gibt, um vom ausgestreckten Knäuel zum nativen Zustand zu gelangen. Die resultierende Vielfalt von Faltungspopulationen erfordert die Verwendung reduzierter Fortschrittsvariablen auf der Fläche der mittleren effektiven Energie für die Beschreibung der Proteinfaltung (anstelle der Reaktionskoordinate auf einer Potentialenergieläche bei kleinen Molekülen). Die geeignete(n) Fortschrittsvariable(n) muß (müssen) die wesentlichen Merkmale der Faltungsreaktion erfassen und experimentell untersucht werden können. Eine Art von Fortschrittsvariablen, der Anteil nativ-ähnlicher Kontakte, ist für einfache Flächen hinreichend, doch wie wir bei der Analyse der Faltung von Lysozym gesehen haben, sind bei komplexeren Systemen weitere Variablen erforderlich. Darüber hinaus ist deutlich geworden, daß man für das Verständnis der grundlegenden Prinzipien der Proteinfaltung außer der gesamten Freien Energie auch die mittlere effektive Energie und die Konformationsentropie berücksichtigen muß. Dies ist beispielsweise essentiell für die Analyse des Verhaltens der Temperaturabhängigkeit der Proteinfaltung, das nicht durch eine Arrhenius-Beziehung beschrieben werden kann.

Infolge der großen Zahl signifikanter Freiheitsgrade ist der Übergangs-„Zustand“ viel komplexer als bei Reaktionen kleiner Moleküle. Gittersimulationen deuten darauf hin, daß es wahrscheinlich einen oder auch mehrere Übergangsbereiche mit mehreren Konformationen gibt, die schnell zum nativen Zustand führen. Falls es sich um einen späten Übergangsbereich handelt, besteht er aus einem Ensemble von Strukturen, die alle dem nativen Zustand ähneln, und der interessante, an der Lösung des Levinthal-Paradoxons beteiligte Schritt betrifft den Mechanismus, durch welchen diese Lösung gefunden wird. Ein früher Übergangsbereich kann hingegen viele unterschiedliche Strukturen beinhalten, wobei jeweils nur ein kleiner Teil der Struktur wohldefiniert ist. Die Kenntnis dieses Teils ist dann ausschlaggebend für das Verständnis des Faltungsvorgangs.

Ein wichtiges Ergebnis der Gittersimulationen ist, daß das im Experiment auftretende unterschiedliche Verhalten von Proteinen, die sich aus dem Zustand eines ausgedehnten Knäuels schnell und ohne Zwischenstufen zu durchlaufen falten können, und jenen, die kollabierte Zustände und kinetische Fallen durchlaufen, keinen qualitativen Unterschied in den Flächen der mittleren effektiven Energie erfordert. Statt dessen können die Unterschiede im Faltungsverhalten aus einer Veränderung der Balance zwischen der Konformationsentropie und der mittleren effektiven Energie als Funktion der Fortschrittsvariablen resultieren. Es ist möglich, experimentell von einem Faltungsverhalten zum anderen zu wechseln, indem man z.B. durch Mutagenese zusätzliche hydrophobe Gruppen in den Kern des Proteins einbaut (siehe Abschnitt 6).^[96] In Simulationen kann man die Umstellung des Verhaltens erreichen, indem man die Temperatur senkt, um den entropischen Beitrag zu vermindern (siehe Abb. 6 und 7), und indem man die anziehende hydrophobe Wechselwirkung zwischen den Aminosäuren erhöht.^[128]

Die Vielfalt der Wege, auf denen der native Zustand erreicht werden kann, deutet darauf hin, daß die Verwendung einfacher kinetischer Schemata für die Analyse der Proteinfaltungsreaktion neu bewertet werden muß. Solange Messungen auf einer Zeitskala ($\leq \text{ms}$) ein einfach-exponentielles Verhalten ergeben oder die Ergebnisse durch einige wenige Exponentialterme beschrieben werden können, ist die Annahme „chemischer“ Spezies eine geeignete Modellvorstellung. Diese Spezies entsprechen jedoch möglicherweise nicht den konformativ gut definierten Einheiten bei den Reaktionen kleiner Moleküle, sondern eher einer sehr breiten Verteilung von Strukturen. Dieses einfache Schema ist angemessen, wenn die Prozesse auf unterschiedlichen Zeitskalen ablaufen, die Einstellung des Gleichgewichts zwischen den Konformeren einer chemischen Spezies im Vergleich zur Umwandlung zwischen unterschiedlichen Spezies also schnell abläuft. Beispielsweise treten in den Faltungssimulationen, die den Abbildungen 6 und 7 entsprechen, einfach-exponentielle Kinetiken auf. Man kann diese nach dem Standardschema U \rightarrow N beschreiben, in dem die entfaltete Spezies U die Bereiche des Random coil und des kollabierten Knäuels einschließt (10^{16} Zustände für das 27mer) und N dem einzigartigen nativen Zustand entspricht. Der Kollaps zu einem Knäuel, der ein wichtiger Teil des Faltungsvorgangs ist, erscheint in diesem einfachsten kinetischen Schema nicht, weil er schnell stattfindet und die Gesamtreaktionsgeschwindigkeit nicht beeinflußt. Selbst wenn die verfügbaren Daten mit einem einfachen kinetischen Modell beschrieben werden können, heißt das noch nicht, daß die Faltungsreaktion verstanden ist. Die Simulationsergebnisse und deren Analyse weisen darauf hin, daß es „versteckte“ komplexe Eigenschaften des Faltungsprozesses gibt. Die Prozesse bei Proteinen, die sich unter Beteiligung von „Zwei-Zustands“-Übergängen falten, werden auf einer kurzen Zeitskala wahrscheinlich vielstufig erscheinen. Dies ist analog zur Zunahme der beobachteten komplexen Eigenschaften mit jeder Verbesserung der experimentellen Techniken, wie man sie bei anderen scheinbar einfachen Reaktionen von Proteinen – z.B. bei der erneuten Ligandenbindung nach einer Photodissociation bei Myoglobin^[129] – findet.

Es besteht ein erheblicher Bedarf an Experimenten, die eine Untersuchung der Proteinfaltung in einem kürzeren Zeitrahmen und mit höherer Strukturauflösung ermöglichen.^[130] Genau diese Techniken werden derzeit etabliert, und mit einigen davon sind bereits erste Ergebnisse erhalten worden. Von besonderer Bedeutung ist die NMR-Spektroskopie, mit der es neuerdings möglich ist, den Faltungsprozeß in Echtzeit zu verfolgen.^[84, 131, 132] Zu den neueren Entwicklungen zählen die Einführung neuartiger zweidimensionaler Methoden zur Untersuchung der Faltung auf der Ebene einzelner Aminosäurereste^[83] sowie die Möglichkeit, zumindest einige Aspekte des Faltungsprozesses auf einer Zeitskala von Hundertstelsekunden zu verfolgen.^[133] Durch Messungen des Kern-Overhauser-Effekts (NOE) zur Bestimmung von intramolekularen Abständen kann die Entwicklung nativer und nichtnativer Kontakt während der Faltung untersucht werden – ein Schlüsselaspekt bei detaillierten Vergleichen mit Simulationen.^[80] Andere spektroskopische Methoden, die spezifische Strukturinformationen liefern, wie die IR-Spektroskopie^[134] und die Raman-Spektroskopie,^[135] werden neuerdings auch zum Studium

der Proteinfaltung verwendet. Ergebnisse aus Untersuchungen der größeren Fluktuationen in Proteinen im Gleichgewichtszustand durch Amidprotonenaustausch und andere Techniken werden mit den Energieflächen der Faltung und Entfaltung in Beziehung gesetzt.^[136, 137]

Zur Zeit werden Methoden entwickelt, um die sehr frühen Ereignisse bei der Faltung durch Verwendung schneller Mischungsapparaturen oder Anwendung von Initiationsmethoden auf der Nanosekunden-Zeitskala zu erforschen.^[138–142] Untersuchungen der Ereignisse bei der Entfaltung im Sub-Mikrosekunden-Bereich^[134] sind als Ergänzung des Studiums des Faltungsvorgangs von Interesse. Des weiteren können solche Entfaltungsexperimente direkt mit detaillierten, alle Atome enthaltenden Simulationen der durch hohe Temperaturen oder niedrige pH-Wert eingeleiteten Entfaltung^[2, 34, 35, 143–145] verglichen werden. Gleichzeitig mit den experimentellen Entwicklungen werden auch theoretische Ansätze verbessert, um zu immer realistischeren Faltungssimulationen zu gelangen. Dazu gehören beispielsweise Rechnungen für alle Atome unter Einbeziehung eines Modells für das Lösungsmittel,^[146a] Gittermodelle mit expliziter Einbeziehung der Seitenketten,^[146b] vereinfachte, gitterlose („off-lattice“-)Modelle poly-peptidartiger Ketten^[147] und die Ausweitung der Methoden, die zur Bestimmung der Flächen der mittleren effektiven Energie und der Konformationsentropie des Faltungsprozesses verwendet werden, auf Modelle, die alle Atome des Proteins und des Lösungsmittels enthalten.^[148] Die ständige Verbesserung der Computerleistung ist für den Fortschritt auf diesem Gebiet von äußerster Wichtigkeit.

Mit ihrer gegenwärtigen Detailgenauigkeit können Gittersimulationen für kein einziges Protein den Faltungsmechanismus bestimmen. Die Nützlichkeit der Simulationen liegt darin, daß sie allgemeine Konzepte liefern und mögliche Mechanismen andeuten können. Einige Überprüfungen der Schlußfolgerungen aus Gittersimulationen sind bereits veröffentlicht worden.^[6, 88, 149] Dennoch ist unser Wissen derzeit nicht ausreichend um nachzuweisen, daß ein Protein eine Energiefläche aufweist, die einem einfachen „Trichter“^[57] entspricht, und wir haben bereits angedeutet, daß eine solche Fläche wahrscheinlich gar nicht existiert, und sei es nur aus dem Grund, daß es keinen Evolutionsdruck für die Entstehung einer für Theoretiker idealisierten Welt gibt. Allerdings sind die verfügbaren experimentellen Daten derart beschränkt, daß die auf Trichtern beruhenden qualitativen Interpretationen einige der verfügbaren Faltungsstudien adäquat beschreiben.^[93] In diesem Beitrag haben wir vorgeschlagen, wie man durch das Zusammenspiel von Simulationen und Experimenten über solche Beschreibungen hinausgehen kann. Man muß zunächst eine angenäherte Fläche der effektiven Energie mit den zugehörigen Strukturen kartieren und dann die Dynamik der Faltung auf dieser Fläche beschreiben und ermitteln, wie die Struktur sich mit der Zeit entwickelt. Bei der Bestimmung detaillierter Strukturen auf kürzeren Zeitskalen und bei der Simulation realistischer Modelle werden derzeit erhebliche Fortschritte gemacht. Dies macht es wahrscheinlich, daß man die Grundprinzipien der Faltung realer Proteine in den nächsten Jahren verstanden haben wird.

Diese Prinzipien haben praktische Bedeutung für die Nutzung der Fortschritte bei der Genomforschung^[150] und für das

Design neuartiger Proteine mit spezifischen Eigenschaften.^[151] Darüber hinaus muß betont werden, daß ein Prozeß, der für ein Lebewesen so wichtig ist wie die Proteinfaltung, durch eine ausgeklügelte Reihe von Katalysatoren gesteuert werden muß. Enzyme beschleunigen potentiell langsame Faltungseignisse wie die Isomerisierung des Prolins und die Bildung von Disulfidbrücken.^[4, 111] Außerdem gibt es in der Zelle molekulare Chaperons, d. h. Proteine (Faltungshelfer), die die intermolekularen Wechselwirkungen verhindern, die sonst unter den Bedingungen im Zellinneren zur Aggregation führen würden.^[4, 152] Einige von ihnen weisen möglicherweise eine Entfaltungsaktivität (unfoldase activity) auf,^[153] die es falsch gefalteten Proteinen ermöglicht, aus den kinetischen Fallen zu entkommen, die wir in Abschnitt 7 diskutiert haben. Ihre Wirkungsweise kann auch mit Energieflächen (ähnlich jenen, die wir hier für die Faltung isolierter Systeme vorgestellt haben) beschrieben werden,^[154] obwohl beim gegenwärtigen Wissensstand bereits einfache kinetische Modelle recht zufriedenstellende Beschreibungen liefern. Ein weiteres faszinierendes Problem hängt mit der Faltung von Proteinen in Membranen oder an deren Oberflächen zusammen. Die experimentelle Erforschung dieses Vorgangs steht erst am Anfang.^[155] Angesichts der großen Zahl membranassozierter Proteine wirft die Art des Faltungsprozesses in einer überwiegend hydrophoben Umgebung wichtige Fragen auf.

Parallel zu unserem sich vergrößernden Wissen über die Prinzipien der Proteinfaltung wächst auch das Bewußtsein, daß Faltung und Entfaltung eine wichtige Rolle bei den Mechanismen und der Steuerung einer Vielfalt von zellulären Prozessen spielen. Dazu gehören die Translokation von Proteinen durch Membranen in die richtigen Zellkompartimente^[155] sowie die Regulierung der komplexen Ereignisse, die während des Zelencyclus stattfinden.^[156] Darüber hinaus werden Proteine, die sich nicht korrekt und effizient falten, zunehmend mit Fehlfunktionen biologischer Systeme in Zusammenhang gebracht, und man kennt inzwischen eine beachtliche Reihe von Krankheiten, die mit der falschen Faltung von Proteinen zusammenhängen.^[157] Einige von ihnen, z.B. Multiple Sklerose, gehen auf Mutationen zurück, welche die normale Faltung und Sekretion bestimmter Proteine beeinträchtigen.^[158] Andere, beispielsweise die Alzheimer- oder die Creutzfeld-Jakob-Krankheit, werden mit der nachträglichen Umwandlung normalöslicher Proteine in unlösliche Amyloid-Plaques und -Fasern in Verbindung gebracht.^[159] Solche intermolekularen Wechselwirkungen betreffen Bereiche des Konformationsraums außerhalb derer, die in diesem Aufsatz diskutiert werden, doch darf man annehmen, daß die zugrundeliegenden Kräfte sowie die Faltungsprozesse ähnlich sind, abgesehen von einer Konzentrationsabhängigkeit als zusätzlicher Variable. Eines der Proteine, von denen man weiß, daß Mutationen zu einer durch Amyloid ausgelösten Erkrankung führen können, ist Lysozym,^[160] dessen Faltungsverhalten ausführlich beschrieben wurde (siehe Abschnitt 4). Nach neueren Untersuchungen sind die zur Aggregation neigenden Spezies, die zur Bildung von Amyloid-Fasern führen, den normalen Faltungsintermediaten sehr ähnlich. Es gibt auch Proteine, bei denen sich eine monomere Untereinheit nur im dimeren oder oligomeren Assoziat effizient faltet.^[161, 162] Ein detaillierteres Verständnis sowohl des nor-

malen als auch des abweichenden Faltungsverhaltens könnte demnach für die Entwicklung neuartiger therapeutischer Strategien zur Verhinderung des Ausbruchs oder der Abschwächung der Auswirkungen solcher Krankheiten wichtig sein.

Wir hoffen, wir konnten zeigen, daß der Mechanismus der Proteinfaltung eines der faszinierendsten Probleme auf dem Gebiet chemischer Reaktionen ist. Wenn dieser Aufsatz das Interesse von Chemikern an diesem sich schnell entwickelnden Thema weckt und einige von ihnen veranlaßt, zu dessen weiterer Erforschung beizutragen, hat er seinen Zweck nicht verfehlt.

10. Anhang: Theoretische und experimentelle Methoden zur Untersuchung der Proteinfaltung

10.1. Gittermodelle der Proteinfaltung

In einem Gittermodell werden die Aminosäuren auf Gitterpunkten positioniert und gehen je nach Art der beteiligten Seitenketten spezifische Wechselwirkungen mit ihren nächsten Nachbarn ein. Hier wird ein kubisches Gitter benutzt,^[30, 54a] obwohl auch komplexere Gerüste verwendet wurden.^[128] Der Faltungsvorgang im Gitter verläuft über Monte-Carlo-Bewegungen der Aminosäuren, bis die native Konformation (die mit der niedrigsten Energie) erreicht ist. Um die in den Abbildung 6 und 7 verwendeten Gittermodelle zu veranschaulichen, betrachten wir eine einfache Polypeptidkette aus 27 Aminosäuren (27mer) und ihre Faltungstrajektorie (Abb. 13). In einem unendlichen kubischen Gitter gibt es ca. 10^{16} Konformationen für diese Kette. Setzt man eine in der Summe hydrophobe Anziehung zwischen den Aminosäuren voraus, so wird der native Zustand des Modellproteins einem $3 \times 3 \times 3$ -Würfel entsprechen (die Struktur ganz rechts in Abbildung 13 d), der von der Polypeptidkette völlig ausgefüllt wird. Es gibt 103 346 Möglichkeiten, ein 27mer in diesen Würfel einzupassen. Welche davon dem nativen Zustand entspricht, wird durch die Sequenz festgelegt, also von der Art der Wechselwirkungen zwischen benachbarten Paaren von Aminosäuren bei gegebener Konformation. Da es im Würfel 28 Kontakte zwischen nächsten Nachbarn gibt, kann der native Zustand anhand dieser Kontakte bestimmt werden. In diesem Modell werden einfache Wechselwirkungspotentiale verwendet, um die effektive Energie einer gegebenen Konformation darzustellen. Dies ermöglicht es, die Thermodynamik und die Dynamik des Systems in einer vertretbaren Computerrechenzeit zu bestimmen.

Obwohl die in den Gittersimulationen verwendeten Potentiale vereinfacht sind, liefern sie allgemeine Informationen über mögliche Szenarien der Proteinfaltung, um die es hier vor allem geht. Um das Verhalten eines Proteins im Detail zu ermitteln, benötigt man kompliziertere Modelle und genauere Wechselwirkungspotentiale. Die aus den Faltungssimulationen mit Gittermodellen und aus den ergänzenden Ergebnissen der detaillierten, alle Atome umfassenden Simulationen der Entfaltung gewonnenen Einblicke können mit experimentellen Daten kombiniert werden, um so zu einem

Verständnis der Proteinfaltung zu gelangen. Die Simulationen tragen hierzu insofern wesentlich bei, als daß sie es ermöglichen, die Strukturen, die Energetik und die Dynamik der Reaktionen mit einer Detailgenauigkeit zu analysieren, die experimentell nicht erreichbar wäre.

Um den zeitlichen Verlauf einer Trajektorie bei einer Temperatur nahe der Übergangstemperatur T_m der Faltung zu veranschaulichen, ist in Abbildung 13 e die Energie E , die Zahl N der Kontakte, die als Maß der Kompaktheit dem Gyrationsradius ähnelt, und der Anteil Q der nativen Kontakte aufgetragen. Für ein 27mer steigt N von 0 für die völlig ausgestreckte Kette (keine Kontakte) bis auf 28 für den vollständig kompakten Würfel an, ganz gleich, ob dies der native Zustand ist oder nicht. Man sieht, daß N bereits früher ansteigt als Q . Dies entspricht einem schnellen Kollaps (nach etwa 10^4 Monte-Carlo-Schritten) und der Bildung eines ungeordneten, relativ kompakten Knäuels. Dem Kollaps folgt ein langsamer Suchprozeß (10^7 Schritte) innerhalb der 10^{10} halbkompakten Zustände (die Struktur ändert sich vollständig in etwa 10^3 Schritten, doch die Zahl der Kontakte bleibt groß), bis einer von 10^3 Zuständen des Übergangsbereichs gefunden wird. Da Q in diesem Übergangsbereich zwischen 0.8 und 0.9 liegt (entsprechend 22–26 nativen Kontakten), ähneln die Übergangszustände einander und auch dem nativen Zustand. Von dort aus gibt es, wie man es für einen Übergangszustand erwartet, eine rasche Umordnung (10^4 Schritte) zum nativen Zustand.

Man kann eine Beziehung zwischen der Zahl der Monte-Carlo-Schritte und der Zeit herstellen, da die Pseudo-Diederwinkel im Gittermodell (z. B. der Diederwinkel, der von den Aminosäuren 2–5 in Abbildung 13 d definiert wird) als die $C_\alpha\text{-}C_\alpha\text{-}C_\alpha\text{-}C_\alpha$ -Pseudo-Diederwinkel eines Proteirückgrats angesehen werden können. Die Geschwindigkeit für die Änderung des C_α -Pseudo-Diederwinkels liegt in der Größenordnung von 10^9 s⁻¹,^[166, 167] so daß 10^6 Monte-Carlo-Schritte etwa 10^{-3} s entsprechen. In einer typischen Faltungssimulation des Modellproteins werden ca. 10^7 Schritte benötigt; dies entspricht einer Zeit von 10 ms, die der für die Faltung kleiner Proteine experimentell ermittelten ähnelt.

10.2. Strukturuntersuchungen zur Proteinfaltung

Die Faltung eines Proteins kann im Labor auf mehrere Arten ausgelöst werden. Eine der einfachsten ist, das Protein durch ein chemisches Denaturierungsmittel, wie Guanidiniumchlorid, in hoher Konzentration zu entfalten und die Lösung dann schnell zu verdünnen, so daß die Konzentration des Denaturierungsmittels bis in einen Bereich absinkt, in dem der native Zustand thermodynamisch stabil ist. Dabei stellt sich die Frage, wie die Strukturänderungen verfolgt werden können, da die Faltung oft in weniger als einer Sekunde beendet ist und (wie die Energiefläche zeigt) zumindest bis auf die letzten Schritte extrem heterogen verläuft.

Eine hierfür entwickelte Strategie ist die Verwendung einer Serie von komplementären Stopped-flow-Techniken, mit denen jeweils ein ganz bestimmter Aspekt der Bildung einer nativ-ähnlichen Struktur untersucht werden kann.^[65, 98] Dies ist

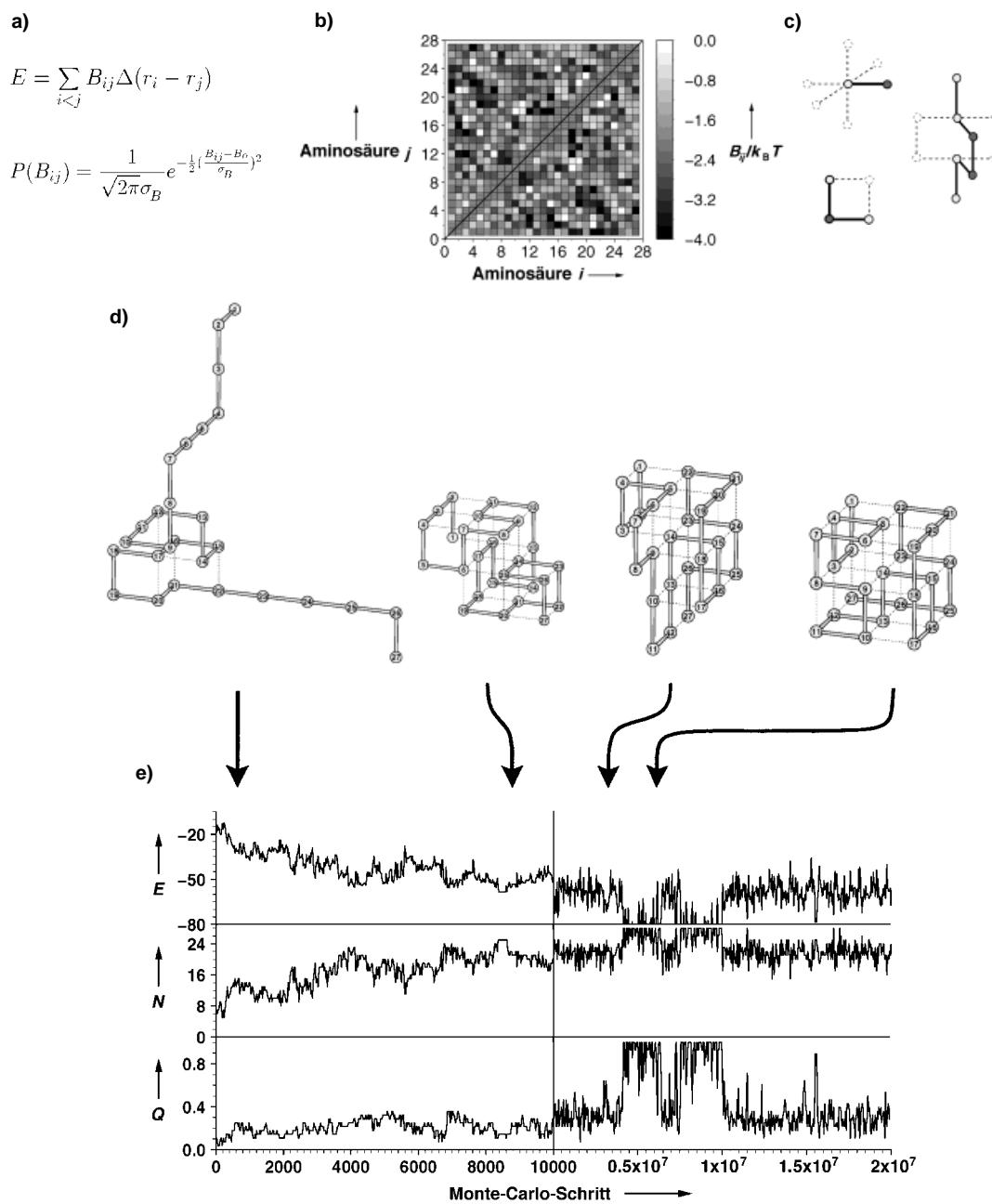


Abb. 13. a) Die Gesamtenergie E einer beliebigen Kettenkonformation entspricht der Summe der Kontaktenergien B zwischen den Aminosäurebausteinen, die sich aus einer Gauß-Verteilung mit dem Mittelwert B_0 und der Standardabweichung σ_B ableiten lassen. Dabei sind r_i und r_j die Ortsvektoren der Aminosäuren mit den Sequenznummern i bzw. j . Der Wert von B_0 beträgt $-2k_B T$ und entspricht der mittleren Anziehung zwischen den Aminosäurebausteinen, womit dem hydrophoben Effekt Rechnung getragen wird. Der Wert von σ_B beträgt $k_B T$ und entspricht der Variationsbreite der Kontaktenergien der Aminosäuren. b) Die Kontaktenergiematrix für die Sequenz, die in e) und in Abbildung 6 verwendet wurde. c) Die drei erlaubten Monte-Carlo-Bewegungen: Umlappen eines freien Endes (oben links), Umlappen einer Ecke (unten links) und die kurbelartige Drehung zweier Aminosäurebausteine (rechts). d) Beispielstrukturen für unterschiedliche Bereiche der in e) gezeigten Faltungstrajektorie (die Zugehörigkeiten sind durch Pfeile angezeigt): Random coil, ungeordnetes halbkompaktes Knäuel, Übergangszustand und nativer Zustand (von links nach rechts). e) Die Faltungstrajektorie mit E , N und Q als Funktion der Zahl der Monte-Carlo-Schritte. (Mit Veränderungen aus Lit.^[28, 55a] übernommen.)

schematisch in Abbildung 14 dargestellt. Anhand des zeitlichen Verlaufs des Fern-UV-CD-Signals kann z. B. die Bildung der Sekundärstruktur (insbesondere von α -Helices) verfolgt werden, und der zeitliche Verlauf des Nah-UV-CD-Signals kann Hinweise auf die nativ-ähnliche Packung der aromatischen Seitenketten geben. Der Untersuchungsbereich der so verwendeten Techniken nimmt rasch zu. Eine der bedeutenderen Methoden ist die Möglichkeit, die äußeren Abmessun-

gen des Proteins (R_g) durch zeitaufgelöste Röntgenkleinwinkelstreuung zu bestimmen.^[69]

Selbst bei den besten konventionellen Mischungsverfahren beträgt die Totzeit typischerweise einige Millisekunden, doch es gibt Fortschritte bei der Entwicklung von Geräten, mit denen sie erheblich verringert werden kann. Darüber hinaus werden neue Methoden zur Auslösung der Rückfaltungsreaktion eingeführt.^[138–142] Dazu zählt die Verwendung von Tem-

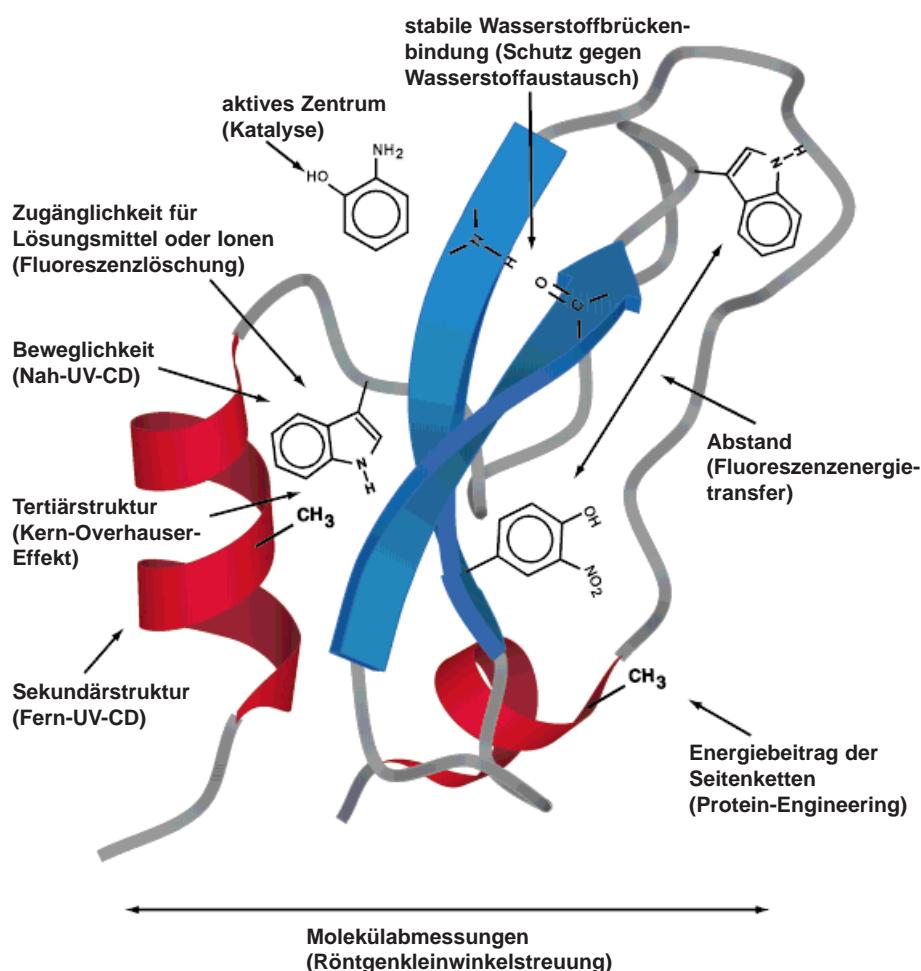


Abb. 14. Schematische Darstellung einiger charakteristischer Eigenschaften eines globulären Proteins, die während der Faltung durch biophysikalische Methoden (stopped oder quenched flow) mit einer Zeitauflösung im Bereich von Sekunden bis Millisekunden verfolgt werden können (mit Veränderungen aus Lit.^[130] übernommen). Beispiele für spezifische Methoden, mit denen unterschiedliche Merkmale untersucht werden können, sind angegeben. Zu den anderen Eigenschaften, die untersucht werden können, aber hier nicht aufgeführt sind, gehören das Vorhandensein hydrophober Hohlräume sowie die thermodynamische Gesamtstabilität. Obwohl keine Methode allein alle während der Faltung auftretenden Strukturänderungen erfassen kann, liefert die Anwendung mehrerer komplementärer Methoden ein detaillierteres Bild von der Verteilung der Konformationen, die in verschiedenen Stadien der Faltung eingenommen werden.

peratursprünge unter Bedingungen, bei denen eine Kälte-denaturierung stattgefunden hat und ein Temperaturanstieg die Rückfaltung auslöst.^[139, 141] In manchen Fällen löst die schnelle Änderung der Oxidationsstufe eines Metalloproteins die Faltungsreaktion aus.^[140] Mit solchen Methoden können Erkenntnisse über Faltungereignisse auf der Sub-Mikrosekunden-Zeitskala erhalten werden.

Ein besonderes Ziel der Strukturuntersuchungen ist es, die Bildung von Kontakten auf der Ebene einzelner Aminosäurereste zu verfolgen. Diese Daten können direkt mit den Ergebnissen der Simulationen verglichen werden. Die NMR-Spektroskopie ist bekanntermaßen das Verfahren, mit dem dieses Ziel – zumindest im Prinzip – erreicht werden kann, und in der Praxis werden hier derzeit beträchtliche Fortschritte gemacht. Wenn die Faltung relativ langsam erfolgt, auf der Zeitskala von Sekunden oder darüber, kann man eindimensionale Spektren sequentiell aufzeichnen und deren Veränderung verfolgen.^[131, 132] Ein Beispiel hierfür ist in Abbildung 15 wie-

dergegeben. Darüber hinaus wurde gezeigt, daß es mit nur einem zweidimensionalen NMR-Experiment möglich ist, die Faltung an einzelnen Aminosäureresten eines isotopenmarkierten Proteins zu untersuchen, was den Weg zu hochauflösenden Experimenten bahnt.^[83] Ebenso wichtig war die Demonstration, daß man durch das Anregen einer Kernpolarisation mit Lasern (Photo-CIDNP-Experimente; CIDNP = chemically induced dynamic nuclear polarization) die Zeitspanne, innerhalb derer die Ereignisse verfolgt werden können, bis auf ca. 10 ms verringern kann, was die Zeitauflösung der NMR-Techniken in die Nähe derer bei der optischen Spektroskopie bringt.

Ein weiterer Ansatz, die Faltung mit NMR-Methoden zu untersuchen, sind die Methoden der Wasserstoffatom-Pulsmarkierung.^[150, 168, 169] Diese beruhen darauf, daß labile Wasserstoffatome (zum Beispiel jene, die an das Amid-Stickstoffatom der Peptidbindung gebunden sind) in entfalteten Zuständen schnell gegen die Wasserstoffatome des Lösungsmittels ausgetauscht werden. Sie sind aber während der Faltung des Proteins oft gegen diesen Austausch geschützt, weil die Amidgruppen dann an Wasserstoffbrückenbindungen beteiligt oder im Inneren des Proteins verborgen sind. Zu verschiedenen

Zeitpunkten nach Beginn der Faltung des Proteins kann man das Ausmaß des Schutzes einzelner Wasserstoffatome ermitteln, indem man einen durch hohen pH-Wert verursachten „Markierungs“-Puls anwendet, bei dem frei zugängliche Amidprotonen durch Deuterium ersetzt werden. Der Austausch wird dann bei niedrigem pH-Wert unterdrückt („quenching“), und das Protein kann sich zurückfalten. Das NMR-spektroskopisch ermittelte Maß der Markierung an unterschiedlichen Stellen spiegelt dann die jeweilige Proteinstruktur im Verlauf des Faltungsprozesses wider. Die Zahl der markierten Stellen kann auch durch Massenspektrometrie ermittelt werden, was wichtige Informationen über die Verteilung von Strukturtypen während des Faltungsprozesses liefert.^[170]

Wir danken Azat Badretdinov, Aaron Dinner, Alan Fersht, Andrew Miranker, Lorna Smith und Eugene Shakhnovich für hilfreiche Diskussionen, Oren Becker, der Abbildung 9 zur Verfügung stellte, und Eric Feyont für die Hilfe bei der Erstellung

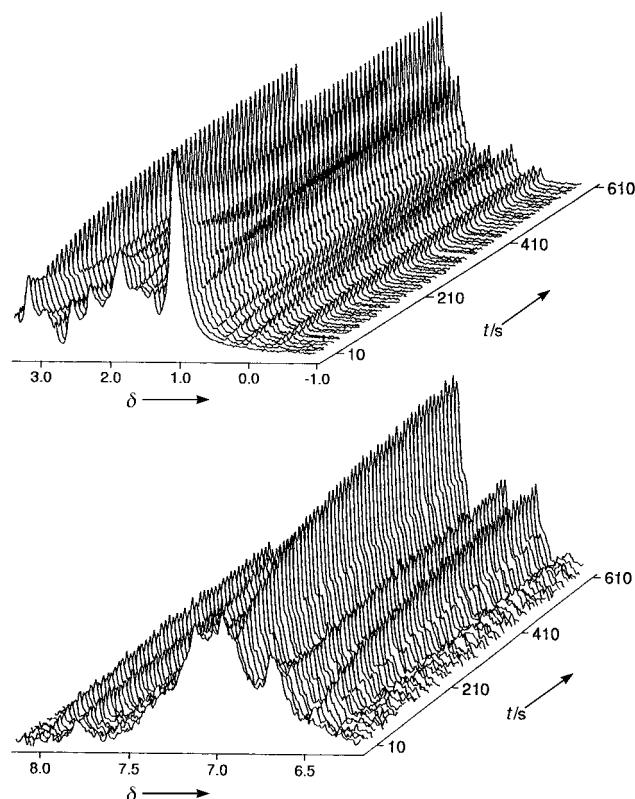


Abb. 15. Untersuchung der Proteinfaltung in Echtzeit durch ^1H -NMR-Spektroskopie. In 6 M Guanidiniumchlorid gelöstes Rinder- α -Lactalbumin wurde direkt im NMR-Gerät in einen Rückfaltungspuffer verdünnt. Die Spektren wurden zu bestimmten Zeitpunkten zwischen 1.2 s und 10.3 min nach der Auslösung der Rückfaltung aufgezeichnet, und ihre gestaffelte Darstellung zeigt die Veränderungen im Bereich der Signale von Methylen- und Methylenegruppen (oben) sowie von aromatischen Seitenketten (unten). Das Verschwinden der breiten und einander überlappenden Signale des anfänglich gebildeten Molten-globule-Zustands und das Auftreten der gut aufgelösten Signale, die für den nativen Zustand charakteristisch sind, ist klar ersichtlich. (Übernommen aus Lit.^[132])

der \$^1\text{H}\$-NMR-Spektroskopie. In 6 M Guanidiniumchlorid gelöstes Rinder- α -Lactalbumin wurde direkt im NMR-Gerät in einen Rückfaltungspuffer verdünnt. Die Spektren wurden zu bestimmten Zeitpunkten zwischen 1.2 s und 10.3 min nach der Auslösung der Rückfaltung aufgezeichnet, und ihre gestaffelte Darstellung zeigt die Veränderungen im Bereich der Signale von Methylen- und Methylenegruppen (oben) sowie von aromatischen Seitenketten (unten). Das Verschwinden der breiten und einander überlappenden Signale des anfänglich gebildeten Molten-globule-Zustands und das Auftreten der gut aufgelösten Signale, die für den nativen Zustand charakteristisch sind, ist klar ersichtlich. (Übernommen aus Lit.^[132])

Eingegangen am 14. Juli 1997 [A 242]

Übersetzt von Dr. Michael Groß, Oxford (Großbritannien)

- [1] R. E. Levine, R. B. Bernstein, *Molecular Reaction Dynamics and Chemical Reactivity*, Oxford University Press, Oxford, 1987.
- [2] „Protein Folding: Theoretical Studies of Thermodynamics and Dynamics“: M. Karplus, E. Shakhnovich in *Protein Folding* (Hrsg.: T. E. Creighton), Freeman, San Francisco, 1992, S. 127–195.
- [3] A. D. Miranker, C. M. Dobson, *Curr. Opin. Struct. Biol.* **1996**, 6, 31–42.
- [4] M.-J. Getting, J. Sambrook, *Nature* **1992**, 360, 533–45.
- [5] C. B. Anfinsen, *Science* **1973**, 181, 223–230.
- [6] M. Karplus, A. Šali, *Curr. Opin. Struct. Biol.* **1995**, 5, 58–73.

- [7] C. Levinthal in *Mossbauer Spectroscopy in Biological Systems, Proceedings of a Meeting held at Allerton House, Monticello, Illinois* (Hrsg.: P. Debrunner, J. C. M. Tsibris, E. Münck), University of Illinois Press, Urbana, 1969, S. 22.
- [8] C. Levinthal, *J. Chim. Phys.* **1968**, 65, 44–45.
- [9] J. Moult, R. Judson, K. Fidelis, J. Pedersen, *Proteins: Struct. Funct. Genet.* **1995**, 23, ii–iv.
- [10] H. Wu, *Am. J. Physiol.* **1929**, 90, 562.
- [11] J. S. Fruton, *Molecules and Life*, Wiley, New York, 1972.
- [12] „Molekulare Dynamik chemischer Elementarreaktionen (Nobel-Vortrag)“: D. R. Herschbach, *Angew. Chem.* **1987**, 99, 1251–1275; *Angew. Chem. Int. Ed. Engl.* **1987**, 26, 1221–1243.
- [13] M. Karplus, R. N. Porter, R. D. Sharma, *J. Chem. Phys.* **1965**, 43, 3259–3287.
- [14] J. D. Bryngelson, J. N. Onuchic, N. D. Soccia, P. G. Wolynes, *Proteins: Struct. Funct. Genet.* **1995**, 21, 167–195.
- [15] K. A. Dill, S. Bromberg, K. Yue, K. M. Fiebig, D. P. Yee, P. D. Thomas, H. S. Chan, *Protein Sci.* **1995**, 4, 561–602.
- [16] E. I. Shakhnovich, *Curr. Opin. Struct. Biol.* **1997**, 7, 29–40.
- [17] K. A. Dill, H. S. Chan, *Nature Struct. Biol.* **1997**, 4, 10–19.
- [18] A. H. Zewail, J. Polanyi, *Acc. Chem. Res.* **1995**, 28, 119–124.
- [19] R. N. Porter, M. Karplus, *J. Chem. Phys.* **1964**, 40, 1105–1115.
- [20] G. C. Schatz, *J. Phys. Chem.* **1996**, 100, 12839–12847.
- [21] M. Oliveberg, A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **1995**, 92, 8926–8929.
- [22] D. J. Wales, *Science* **1996**, 271, 925–929.
- [23] K. D. Ball, R. S. Berry, R. E. Kunz, F.-Y. Li, A. Prokova, D. J. Wales, *Science* **1996**, 271, 963–965.
- [24] D. L. Freeman, J. D. Doll, *Annu. Rev. Phys. Chem.* **1996**, 47, 43–80.
- [25] J. D. Bryngelson, P. G. Wolynes, *J. Phys. Chem.* **1989**, 93, 6902–6915.
- [26] E. I. Shakhnovich, A. M. Gutin, *Biophys. Chem.* **1989**, 34, 187–199.
- [27] V. I. Abkevich, A. M. Gutin, E. I. Shakhnovich, *J. Chem. Phys.* **1994**, 101, 6052–6062.
- [28] A. Šali, E. Shakhnovich, M. Karplus, *Nature* **1994**, 369, 248–251.
- [29] H. A. Kramers, *Physica* **1940**, 7, 284–293.
- [30] R. W. Pastor, M. Karplus, *J. Chem. Phys.* **1989**, 91, 211–218.
- [31] S. H. Northrup, M. R. Pear, C.-Y. Lee, J. A. McCammon, M. Karplus, *Proc. Natl. Acad. Sci. USA* **1982**, 79, 4035–4039.
- [32] T. Lazaridis, G. Archontis, M. Karplus, *Adv. Protein Chem.* **1995**, 47, 231–306.
- [33] D. A. McQuarrie, *Statistical Mechanics*, Harper & Row, New York, 1976, S. 641.
- [34] A. Caflisch, M. Karplus, *Proc. Natl. Acad. Sci. USA* **1994**, 91, 1746–1750.
- [35] A. Caflisch, M. Karplus, *J. Mol. Biol.* **1995**, 252, 672–708.
- [36] G. I. Makhatadze, P. L. Privalov, *Adv. Protein Chem.* **1995**, 47, 308–417.
- [37] P. L. Privalov, *Adv. Protein Chem.* **1979**, 33, 167–241.
- [38] F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, *J. Mol. Biol.* **1977**, 112, 535–542.
- [39] L. Holm, C. Sander, *Science* **1996**, 273, 595–602.
- [40] C. Chothia, *Nature* **1992**, 360, 543–544.
- [41] C. A. Orengo, D. T. Jones, J. M. Thornton, *Nature* **1994**, 372, 631–634.
- [42] A. V. Finkelstein, A. Y. Badretdinov, A. M. Gutin, *Proteins* **1995**, 23, 142–150.
- [43] A. R. Fersht, *Curr. Opin. Struct. Biol.* **1994**, 5, 79–84.
- [44] R. Elber, M. Karplus, *Science* **1987**, 235, 318–321.
- [45] C. L. Brooks III, M. Karplus, B. M. Pettitt, *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics (Advances in Chemical Physics, Vol. 71)*, Wiley, New York, 1988.
- [46] O. B. Ptitsyn, *Adv. Protein Chem.* **1995**, 47, 83–229.
- [47] A. Dinner, A. Šali, M. Karplus, E. Shakhnovich, *J. Chem. Phys.* **1994**, 101, 1444–1451.
- [48] D. Stigter, D. O. V. Alonso, K. A. Dill, *Proc. Natl. Acad. Sci. USA* **1991**, 88, 4176–4180.
- [49] C. Tanford, *Adv. Protein Chem.* **1970**, 24, 1–95.
- [50] W. Englander, P. Mayne, *Annu. Rev. Biophys. Biomol. Struct.* **1992**, 21, 243–265.
- [51] M. Karplus, D. L. Weaver, *Protein Sci.* **1994**, 3, 650–668.
- [52] E. Shakhnovich, *Folding Des.* **1996**, 1, R50–54.

- [53] E. P. Wigner, *Trans. Faraday Soc.* **1938**, 34, 29.
- [54] a) A. Dinner, A. Šali, M. Karplus, *Proc. Natl. Acad. Sci. USA* **1996**, 93, 8356–8361; b) T. Lazaridis, M. Karplus, *Science* **1997**, 278, 1928–1931.
- [55] a) A. Šali, E. Shakhnovich, M. Karplus, *J. Mol. Biol.* **1994**, 235, 1614–1636; b) A. Šali, E. Shakhnovich, M. Karplus, unveröffentlichte Ergebnisse.
- [56] M. Karplus, A. Caflisch, A. Šali, E. Shakhnovich in *Modelling of Biomolecular Structures and Mechanisms* (Hrsg.: A. Pullman), Kluwer, Dordrecht, **1995**, S. 69–84.
- [57] J. N. Onuchic, P. G. Wolynes, Z. Luthey-Schulten, N. D. Soccia, *Proc. Natl. Acad. Sci. USA* **1996**, 92, 3626–3630.
- [58] D. L. Bunker, *Theory of Elementary Gas Phase Reaction Rates*, Pergamon, Oxford, **1966**.
- [59] R. Zwanzig, *Proc. Natl. Acad. Sci. USA* **1995**, 92, 9801–9804.
- [60] N. D. Soccia, J. N. Onuchic, P. G. Wolynes, *J. Chem. Phys.* **1996**, 104, 5860–5868.
- [61] S. J. Segawa, M. Sugihara, *Biopolymers* **1984**, 23, 2473–1488.
- [62] M. Oliveberg, A. R. Fersht, *Biochemistry* **1996**, 35, 2726–2737.
- [63] O. M. Becker, M. Karplus, *J. Phys. Chem.* **1997**, 101, 1495–1517.
- [64] O. M. Becker, *J. Mol. Struct. (Theochem.)*, im Druck.
- [65] P. A. Evans, S. E. Radford, *Curr. Opin. Struct. Biol.* **1994**, 4, 100–106.
- [66] K. B. Wiberg, *Physical Organic Chemistry*, Wiley, New York, **1964**.
- [67] D. R. Herschbach, M. S. Johnson, D. Rapp, *J. Chem. Phys.* **1959**, 31, 1652–1661.
- [68] O. B. Ptitsyn, R. H. Pain, G. V. Semisotnov, E. Zerovnik, O. I. Razgulysev, *FEBS Lett.* **1990**, 262, 2024.
- [69] D. Eliezer, P. A. Jennings, P. E. Wright, S. Doniach, K. O. Hodgson, H. Tsuru, *Science* **1995**, 270, 487–488.
- [70] S. E. Radford, C. M. Dobson, P. A. Evans, *Nature* **1992**, 358, 302–307.
- [71] J. I. Guijarro, M. Jackson, A. F. Chaffotte, M. Delepine, H. H. Mantsch, M. E. Goldberg, *Biochemistry* **1995**, 34, 2998–3008.
- [72] W. Kauzmann, *Adv. Protein Chem.* **1959**, 14, 1–63.
- [73] K. Kuwajima, *FASEB J.* **1996**, 10, 102–109.
- [74] M. Ikeguchi, M. Kuwajima, M. Mitani, S. Sugai, *Biochemistry* **1986**, 25, 6965–6972.
- [75] J. Baum, C. M. Dobson, P. A. Evans, C. Hanley, *Biochemistry* **1989**, 28, 7–13.
- [76] B. A. Schulman, C. Redfield, Z.-Y. Peng, C. M. Dobson, P. S. Kim, *J. Mol. Biol.* **1995**, 253, 651–657.
- [77] L. C. Wu, Z.-Y. Peng, P. S. Kim, *Nature Struct. Biol.* **1995**, 2, 281–286.
- [78] L. A. Morozova, D. T. Haynie, C. Arico-Muendel, H. V. Van Dael, C. M. Dobson, *Nature Struct. Biol.* **1995**, 2, 871–875.
- [79] D. A. Dolgikh, R. I. Gilmanshin, E. V. Brazhnikov, V. E. Bychkova, G. V. Semisotnov, S. Y. Venyaminov, O. B. Ptitsyn, *FEBS Lett.* **1981**, 136, 311–315.
- [80] J. Balbach, V. Forge, W. S. Lau, J. Jones, N. A. J. van Nuland, C. M. Dobson, *Proc. Natl. Acad. Sci. USA*, im Druck.
- [81] B. Schulman, P. S. Kim, C. M. Dobson, C. Redfield, *Nature Struct. Biol.*, im Druck.
- [82] B. A. Schulman, P. S. Kim, *Nature Struct. Biol.* **1996**, 3, 682–687.
- [83] J. Balbach, V. Forge, W. S. Lau, N. A. J. van Nuland, K. Brew, C. M. Dobson, *Science* **1996**, 274, 1161–1163.
- [84] J. Balbach, V. Forge, N. J. van Nuland, S. Winder, P. J. Hore, C. M. Dobson, *Nature Struct. Biol.* **1995**, 2, 865–870.
- [85] D. B. Weiflaufer, *Proc. Natl. Acad. Sci. USA* **1973**, 70, 697–701.
- [86] V. E. Abkevich, A. M. Gutin, E. Shakhnovich, *Biochemistry* **1994**, 33, 10026–10036.
- [87] W. Eaton, *Adv. Protein Chem.* **1990**, 40, 63–279.
- [88] A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **1995**, 92, 10869–10873.
- [89] A. R. Fersht, *FEBS Lett.* **1993**, 325, 5–16.
- [90] A. R. Fersht, L. S. Itzhaki, N. F. ElMasry, J. M. Matthews, D. E. Otzen, *Proc. Natl. Acad. Sci. USA* **1994**, 91, 10426–10429.
- [91] L. S. Itzhaki, D. E. Otzen, A. R. Fersht, *J. Mol. Biol.* **1995**, 254, 260–288.
- [92] E. Shakhnovich, V. Abkevich, O. B. Ptitsyn, *Nature* **1996**, 379, 96–98.
- [93] J. N. Onuchic, N. D. Soccia, Z. Luthey-Schulten, P. G. Wolynes, *Folding Des.* **1996**, 1, 441–450.
- [94] G. A. Mines, T. Pascher, S. C. Lee, J. R. Winkler, H. B. Gray, *Chem. Biol.* **1996**, 3, 491–496.
- [95] M. S. Kay, R. L. Baldwin, *Nature Struct. Biol.* **1996**, 3, 439–445.
- [96] S. Khorasanizadeh, I. D. Peters, H. Roder, *Nature Struct. Biol.* **1996**, 3, 193–205.
- [97] R. E. Burton, C. S. Huang, M. A. Daugherty, T. L. Calderone, T. G. Oas, *Nature Struct. Biol.* **1997**, 4, 305–310.
- [98] C. M. Dobson, P. A. Evans, S. E. Radford, *Trends Biochem. Sci.* **1994**, 19, 71–72.
- [99] S. E. Radford, C. M. Dobson, *Philos. Trans. R. Soc. London B* **1995**, 348, 17–25.
- [100] T. Kiefhaber, *Proc. Natl. Acad. Sci. USA* **1995**, 92, 9029–9033.
- [101] A. Matagne, S. E. Radford, C. M. Dobson, *J. Mol. Biol.* **1997**, 267, 1068–1074.
- [102] Z. Guo, D. Thirumalai, *Biopolymers* **1995**, 36, 83–102.
- [103] A. Miranker, S. E. Radford, M. Karplus, C. M. Dobson, *Nature* **1991**, 349, 633–636.
- [104] V. I. Abkevich, A. M. Gutin, E. I. Shakhnovich, *Protein Sci.* **1995**, 4, 1167–1177.
- [105] A. R. Panchenko, Z. Luthey-Schulten, P. G. Wolynes, *Proc. Natl. Acad. Sci. USA* **1996**, 93, 2008–2113.
- [106] A. R. Fersht, *Biochem. Soc. Trans.* **1994**, 22, 267–273.
- [107] R. L. Baldwin, *J. Biomol. NMR* **1995**, 5, 103–109.
- [108] T. E. Creighton, N. J. Darby, J. Kemmink, *FASEB J.* **1996**, 10, 110–118.
- [109] H. Roder, W. Colón, *Curr. Opin. Struct. Biol.* **1997**, 7, 15–28.
- [110] T. R. Sosnick, L. Mayne, S. W. Englander, *Proteins* **1996**, 24, 413–426.
- [111] F. X. Schmid in *Protein Folding* (Hrsg.: T. E. Creighton), Freeman, San Francisco, **1992**, S. 197–241.
- [112] T. R. Sosnick, L. Mayne, R. Hiller, S. W. Englander, *Nature Struct. Biol.* **1994**, 1, 149–156.
- [113] V. Munoz, L. Serrano, *Proteins* **1994**, 20, 301–311.
- [114] L. J. Smith, K. Fiebig, H. Schwalbe, C. M. Dobson, *Folding Des.* **1996**, 1, 95–106.
- [115] M. Blaber, W. A. Baase, N. Gassuer, B. W. Matthews, *J. Mol. Biol.* **1993**, 246, 317–330.
- [116] M. H. J. Cordes, A. R. Davidson, R. T. Sauer, *Curr. Opin. Struct. Biol.* **1996**, 6, 3–10.
- [117] J. Novotny, R. E. Bruccoleri, M. Karplus, *J. Mol. Biol.* **1984**, 177, 787–818.
- [118] J. U. Bowie, R. Luthy, D. Eisenberg, *Science* **1991**, 253, 164–170.
- [119] A. V. Finkelstein, B. A. Reva, *Nature* **1991**, 351, 497–499.
- [120] M. J. Sippl, *J. Comput. Aided Mol. Des.* **1993**, 7, 473–501.
- [121] C. M. R. Lemer, M. J. Rooman, S. J. Wodak, *Proteins* **1995**, 23, 337–355.
- [122] D. T. Jones, J. M. Thornton, *Curr. Opin. Struct. Biol.* **1996**, 6, 210–216.
- [123] S. Kamtekar, J. M. Schiffer, H. Xiong, J. M. Babik, M. H. Hecht, *Science* **1993**, 262, 1680–1685.
- [124] M. H. Hecht, J. S. Richardson, D. C. Richardson, R. C. Ogden, *Science* **1990**, 249, 884–891.
- [125] S. C. Harrison, R. Durbin, *Proc. Natl. Acad. Sci. USA* **1985**, 82, 4028–4030.
- [126] R. L. Baldwin, *Nature* **1994**, 369, 183–184.
- [127] „Physics & Biology“: H. Frauenfelder, *Physics in a Technological World*, AIP, New York, **1988**.
- [128] A. M. Gutin, V. I. Abkevich, E. I. Shakhnovich, *Biochemistry* **1995**, 34, 3066–3076.
- [129] P. J. Steinbach, A. Ansari, J. Berendzen, D. Braunstein, K. Chu, B. R. Cowen, D. Ehrenstein, H. Frauenfelder, J. B. Johnson, D. C. Lamb, *Biochemistry* **1991**, 30, 3988–4001.
- [130] K. W. Plaxco, C. M. Dobson, *Curr. Opin. Struct. Biol.* **1996**, 6, 630–636.
- [131] C. Freiden, S. D. Hoeltzli, I. J. Ropson, *Protein Sci.* **1993**, 2, 2007–2014.
- [132] N. A. J. van Nuland, J. Balbach, V. Forge, C. M. Dobson, *Nature Struct. Biol.* **1995**, 2, 865–870.
- [133] P. J. Hore, S. Winder, C. H. Roberts, C. M. Dobson, *J. Am. Chem. Soc.* **1997**, 119, 5049–5050.
- [134] C. M. Phillips, Y. Mizutani, R. M. Hochstrasser, *Proc. Natl. Acad. Sci. USA* **1995**, 92, 7292–7296.
- [135] G. Wilson, L. Hecht, L. D. Barron, *J. Mol. Biol.* **1996**, 261, 341–347.
- [136] Y. Bai, S. W. Englander, *Proteins* **1996**, 24, 145–151.
- [137] D. Shortle, Y. Wang, J. R. Gillespie, J. O. Wrable, *Protein Sci.* **1996**, 5, 991–1000.

- [138] C. M. Jones, E. R. Henry, U. Hu, C.-K. Chan, S. D. Luck, A. Bhuyan, H. Roder, J. Hofrichter, W. A. Eaton, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11860–11864.
- [139] B. Nolting, R. Goblik, A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 10668–10672.
- [140] T. Pascher, J. P. Chesick, J. R. Winkler, H. B. Gray, *Science* **1996**, *271*, 1558–1560.
- [141] R. M. Ballew, J. Sabelko, M. Gruebele, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 5759–5764.
- [142] W. A. Eaton, P. A. Thompson, C.-K. Chan, S. J. Hagen, J. Hofrichter, *Structure* **1996**, *4*, 113–1139.
- [143] A. E. Mark, W. F. van Gunsteren, *Biochemistry* **1992**, *31*, 7745–7788.
- [144] V. Daggett, M. Levitt, *Curr. Opin. Struct. Biol.* **1994**, *4*, 291–295.
- [145] A. J. Li, V. Daggett, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10430–10434.
- [146] a) T. Lazaridis, M. Karplus, *Science* **1997**, *278*, 1928–1931; b) A. Kolinski, J. Skolnick, *Proteins* **1994**, *18*, 338–352.
- [147] A. Guo, D. Thirumalai, J. D. Honeycutt, *J. Chem. Phys.* **1992**, *97*, 525–535.
- [148] E. M. Boczek, C. L. Brooks III, *Science* **1995**, *269*, 393–396.
- [149] A. R. Fersht, *Curr. Opin. Struct. Biol.* **1997**, *7*, 3–9.
- [150] G. L. G. Miklos, G. M. Rubin, *Cell* **1996**, *86*, 521–529.
- [151] J. S. Marvin, E. E. Corcoran, N. A. Hattangadi, J. V. Zhang, S. A. Greene, H. W. Hellinga, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 4366–4371.
- [152] F. U. Hartl, *Nature* **1996**, *381*, 571–580.
- [153] R. Zahn, S. Perrett, G. Stenberg, A. R. Fersht, *Science* **1996**, *271*, 642–645.
- [154] M. Todd, G. Lorimer, D. Thirumalai, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 4030–4034.
- [155] M. L. Riley, B. A. Wallace, S. L. Flitsch, P. J. Booth, *Biochemistry* **1997**, *36*, 192–196.
- [156] R. W. King, R. J. Deshaies, J. M. Peters, M. W. Kirchner, *Science* **1996**, *274*, 1652–1659.
- [157] P. J. Thomas, B.-H. Qu, P. L. Pederson, *Trends Biochem. Sci.* **1995**, *20*, 456–459.
- [158] S. C. Hyde, P. Emsley, M. J. Hartshorn, M. M. Mimmack, U. Gileadi, S. R. Pearce, M. P. Gallagher, D. R. Gill, R. E. Hubbard, C. F. Higgins, *Nature* **1990**, *346*, 362–365.
- [159] S. B. Prusiner, *Biochemistry* **1992**, *31*, 12277–12288.
- [160] D. R. Booth, M. Sunde, V. Bellotti, C. V. Robinson, W. L. Hutchinson, P. E. Fraser, P. N. Hawkins, C. M. Dobson, S. E. Radford, C. C. F. Blake, M. B. Pepys, *Nature* **1997**, *385*, 787–793.
- [161] C. D. Waldburger, T. Jonsson, R. T. Sauer, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 2629–2634.
- [162] J. King, C. Haase-Pettingell, A. S. Robinson, M. Speed, A. Mitraki, *FASEB J.* **1996**, *10*, 57–66.
- [163] P. J. Kraulis, *J. Appl. Crystallogr.* **1991**, *24*, 946.
- [164] E. A. Merritt, M. E. Murphy, *Acta Crystallogr. Sect. D* **1994**, *50*, 869–873.
- [165] K. W. Plaxco, C. Spitzfaden, I. D. Campbell, C. M. Dobson, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10703–10706.
- [166] J. A. McCammon, S. H. Northrup, M. Karplus, R. M. Levy, *Biopolymers* **1980**, *19*, 2033–2045.
- [167] K. Yapa, D. L. Weaver, M. Karplus, *Proteins: Struct. Funct. Genet.* **1992**, *12*, 237–265.
- [168] J. B. Udgaoonkar, R. L. Baldwin, *Nature* **1988**, *335*, 694–699.
- [169] H. Roder, G. A. Elove, S. W. Englander, *Nature* **1988**, *335*, 700–704.
- [170] A. Miranker, C. V. Robinson, S. E. Radford, R. T. Aplin, C. M. Dobson, *Science* **1993**, *262*, 896–900.