
Mechanik des Grün Fluoreszierenden Proteins

Hendrik Dietz



München, 2004

Mechanik des Grün Fluoreszierenden Proteins

Diplomarbeit
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Hendrik Dietz
aus Dresden

München, den 22. März 2004

Erstgutachter: Prof. Dr. Matthias Rief
Zweitgutachter: Prof. Dr. Roland R. Netz

Inhaltsverzeichnis

Zusammenfassung	xi
1 Einleitung	1
2 Das Grün Fluoreszierende Protein	3
2.1 Biolumineszenz	3
2.2 Historisches	4
2.3 <i>Wildtype</i> -GFP und <i>Cycle3</i> -GFP	5
2.4 Fluoreszenz als Indikator des nativen GFP	6
3 Theoretische Grundlagen	9
3.1 Elastizität von Polymeren	9
3.2 Das Wormlike-Chain Modell (WLC)	10
3.3 Gültigkeit des WLC-Modells	11
3.4 Verwendung des WLC-Modells	12
3.5 Konformationskinetik von Proteinen	12
3.6 Übergangsraten unter konstanter Kraft	15
3.7 Übergangsraten unter zeitabhängiger Kraft	16
4 Methoden	19
4.1 Das Kraftspektrometer	19
4.2 Eigenschaften der Blattfedern	20
4.3 Chimären-Proteine	21
4.4 Durchführung der Messungen	22
4.5 Bestimmung von Entfaltungskräften	23
4.6 Monte-Carlo Simulationen	24
5 Ig8-GFP und DdFLN(1-5)-GFP	27
5.1 Einleitung	27
5.2 Ig8-GFP	27
5.3 DdFLN(1-5)-GFP	30
5.4 Diskussion	33

6	Natives GFP und GFP_{11-230}	35
6.1	Einleitung	35
6.2	Eichung von WLC-Konturlängen	35
6.3	Längenzunahme nach GFP-Entfaltung	36
6.4	Interpretation der Längenzunahme	37
6.5	Kraftplateaus vor der GFP_{11-230} -Entfaltung	40
6.6	Übergang native Struktur $S_n \rightarrow GFP_{11-230}$	42
6.7	Stabilität von GFP_{11-230}	44
6.8	Energielandschaft - Teil 1	47
6.9	Zusammenfassung	49
7	Weitere Zwischenzustände der GFP-Entfaltung: S_1 und S_2	51
7.1	Einleitung	51
7.2	Unterstrukturen der GFP-Entfaltung	51
7.3	Messung von Lebensdauern	53
7.4	Kraftabhängigkeit von Lebensdauern	55
7.5	Längenzuwachs beim Übergang $GFP_{11-230} \rightarrow S_1$	59
7.6	Interpretation der Lebensdauern	60
7.7	Ein weiterer Zwischenzustand S_2	67
7.8	Energielandschaft - Teil 2	70
7.9	Zusammenfassung	70
8	Energielandschaft des GFP	71
8.1	Mechanischer Entfaltungspfad des GFP	71
8.2	Energielandschaft	73
9	Ausblick	75
9.1	GFP als molekularer Kraftsensor	75
9.2	Temperaturabhängigkeit der GFP-Stabilität	76
9.3	Rückfaltung von GFP	77
9.4	Weitere Stichpunkte	77
A	Kommunizierende Entfaltungseignisse	79
B	Mustererkennung	83
B.1	Einleitung	83
B.2	Erkennen eines Musters in Kraftkurven	84
B.3	Bestimmung der Übereinstimmungsgüte	85
B.4	Prinzipielle Tauglichkeit der Mustererkennung	86
B.5	Statistische Analyse von Datensätzen	88
B.6	Zusammenfassung	89
C	Eichung der Blattfedern	91

D	Molekularbiologische Methoden	93
D.1	Gentechnische Methoden	93
D.2	Proteinexpression	94
E	Sequenzinformationen	97
E.1	GFP	97
	Danksagung	102
	Erklärung	104

Abbildungsverzeichnis

2.1	Schema Biolumineszenz	3
2.2	Schema Energietransfer P^* nach F	4
2.3	Strukturbild des GFP	6
2.4	Anteil natives GFP als Funktion der GdnHCl-Konzentration	7
2.5	Fluoreszenz-Spektrum DdFLN(1-5)-GFP	8
3.1	Anpassung WLC-Modell an Polypeptid-Kraftkurve	12
3.2	Schema Energielandschaft	14
3.3	Wahrscheinlichkeitsdichten dP/dF , Variation η	18
3.4	Wahrscheinlichkeitsdichten dP/dF , Variation d_{S_n-T}	18
4.1	Schema Kraftspektrometer	19
4.2	Olympus Bio-Lever	21
4.3	Bestimmung von Entfaltungskräften	24
5.1	Schema Ig8-GFP-Chimäre	28
5.2	Ig8-GFP Kraftkurven	28
5.3	Ig8-GFP Kraftkurve, WLC Anpassung	29
5.4	Schema DdFLN(1-5)-GFP-Chimäre	31
5.5	DdFLN(1-5)-GFP Kraftkurven	31
5.6	Erläuterung DdFLN(1-5)-GFP Kraftkurve	32
5.7	DdFLN-GFP Kraftkurve, WLC Anpassung	32
5.8	Überlagerung Ig8-GFP Kraftkurve mit DdFLN(1-5)-GFP Kraftkurve	33
6.1	Verteilung Zunahme der Konturlänge bei GFP-Entfaltung	36
6.2	Schema Konturlängen in verschiedenen Zuständen	38
6.3	Erwartete Längenzunahme bei Entfaltung einzelner Aminosäuren von der Struktur S_n	39
6.4	Strukturvorschläge GFP_{11-230} und GFP_{4-222}	40
6.5	Kraftplateaus in Ig8-GFP Kraftkurven	41
6.6	Kraftplateaus in DdFLN(1-5)-GFP Kraftkurven	41
6.7	Anpassung Gleichgewichts-Modell an Ig8-GFP Kraftkurven	43
6.8	Anpassung Gleichgewichts-Modell an DdFLN(1-5)-GFP Kraftkurven	43
6.9	Verteilung von GFP_{11-230} -Entfaltungskräften (Experiment)	45

6.10	Simulierte Kraftkurve mit GFP_{11-230} -Entfaltung	45
6.11	Verteilung von Entfaltungskräften für GFP_{11-230} (simuliert und analytisch)	47
6.12	Schema der Energielandschaft des GFP - Teil 1	49
7.1	Unterstrukturen in der GFP-Entfaltung in Ig8-GFP Kraftkurven	52
7.2	Unterstrukturen in der GFP-Entfaltung in DdFLN(1-5)-GFP Kraftkurven	52
7.3	Überlagerung der GFP-Entfaltung in DdFLN(1-5)-GFP Kraftkurven	53
7.4	Auschnitte von drei DdFLN(1-5)-GFP Kraft-Zeitkurven	54
7.5	Verteilung von Lebensdauern	54
7.6	Verteilung von Lebensdauern, einfach-exponentiell gefittet	55
7.7	Auswirkung einer Kraftverteilung auf Lebensdauern: Breite	57
7.8	Auswirkung einer Kraftverteilung auf Lebensdauern: Position	57
7.9	Kraftwirkung auf den Zwischenzustand ist determiniert	58
7.10	Längenzuwachs beim Übergang $GFP_{11-230} \rightarrow S_1$	60
7.11	Erwartete Längenzunahme bei Entfaltung einzelner Aminosäuren von der Struktur GFP_{11-230}	61
7.12	Strukturvorschläge für den Zustand S_1	61
7.13	Verteilung der auf S_1 wirkenden Kräfte	62
7.14	Anpassung $P(t)$ mit Kraftverteilung $g(F)$ an die gemessene Verteilung von Lebensdauern	63
7.15	Anpassung $P(t)$ mit Kraftverteilung $g(F)$ - niedrige und hohe Kräfte . . .	63
7.16	Simulierte Entfaltung $GFP_{11-230} \rightarrow S_1 \rightarrow S_u$	64
7.17	Simulierte Lebensdauern des Zustandes S_1	65
7.18	Simulierte Verteilungen der Kräfte $F_{intersect}$ und F_S	65
7.19	Simulierte Entfaltungskräfte, einzelner Zustand S_1	66
7.20	Hochaufgelöste GFP_{11-230} -Entfaltungen	67
7.21	Anpassung des WLC-Modells an Zwischenzustände	68
7.22	Vergrößerung zweier GFP_{11-230} -Entfaltungen	69
8.1	Hochauflösende Kraftkurve mit den Übergängen $S_n \rightarrow GFP_{11-230} \rightarrow S_1 \rightarrow$ $S_2 \rightarrow S_u$	71
8.2	Schema des Übergangs $S_n \rightarrow GFP_{11-230}$	72
8.3	Schema des Übergangs $GFP_{11-230} \rightarrow S_1$	72
8.4	Schema der Energielandschaft des GFP	74
9.1	Entfaltungskräfte GFP_{11-230} in zwei Temperaturbereichen	77
9.2	Lebensdauern S_1 in zwei Temperaturbereichen	78
A.1	Simulierte Kraftkurve mit zwei Entfaltungen	79
A.2	Simulierte Kraftverteilungen: Variation $\Delta L/L$	80
A.3	Simulierte Kraftverteilungen: Variation k_c	80
A.4	Simulierte Kraftverteilungen: $S_a = S_b$, Variation $\Delta L/L$	82
B.1	Muster $g(x)$ und γ - Histogramm	86

B.2	Kraftkurven aus bestimmten γ -Intervallen	87
B.3	Vergleich Ig8-DdFLN(4) Daten mit DdFLN(1-5)-GFP Daten	89
C.1	Rauschspektrum Bio-Lever Typ A	92

Tabellenverzeichnis

4.1	Experimentelle Parameter	23
E.1	DNA-Sequenz <i>Cycle3</i> -GFP	97
E.2	Aminosäuren-Sequenz <i>Cycle3</i> -GFP	98

Zusammenfassung

Im Rahmen dieser Diplomarbeit wurde die Mechanik einzelner *Grün Fluoreszierender Proteine* (GFP) mit einem Kraftspektrometer untersucht. Das GFP ist ein topologisch komplexes Protein, das eine Vielzahl von Anwendungen in der Zellbiologie gefunden hat.

Für die Untersuchungen wurden einzelne GFP-Moleküle mittels gentechnischer Methoden mit kraftspektroskopisch bereits untersuchten Proteinen fusioniert. Die Messungen an diesen Chimären-Proteinen zeigten, dass die Entfaltung des GFP mechanisch induziert werden kann und zu einem komplexen Muster in den Kraftkurven führt.

Die mechanische Entfaltung des GFP erfolgt sukzessive über drei beobachtete Zwischenzustände. Der Übergang des nativen GFP in den ersten Zwischenzustand äußert sich in einem Kraftplateau bei 35 pN und findet nahe am thermischen Gleichgewicht statt. Es konnte gezeigt werden, dass dabei eine am Aminoterminus des GFP befindliche α -Helix von der nativen Struktur abgelöst wird. Die Analyse der Kraftkurven mit einem elastisch gekoppelten Zwei-Zustandsmodell ergab, dass für diesen Übergang eine Arbeit von ca. $20 k_B T$ aufgewendet werden muss und der erste Zwischenzustand somit energetisch ca. $4 k_B T$ über dem entfalteten Zustand des GFP liegt. Der erste Zwischenzustand zerfällt in einem Nichtgleichgewichtsprozess bei Kräften von 102 ± 40 pN. Eine Analyse der Verteilung gemessener Entfaltungskräfte ergab eine natürliche Lebensdauer dieses Zwischenzustandes von 14 ± 6 s und eine Potentialbreite von 0.28 ± 0.03 nm. Durch die Messung von Längenzunahmen beim Zerfall des ersten Zwischenzustandes konnte gezeigt werden, dass bei diesem Übergang ein kompletter β -Strang aus der GFP-Struktur herausbricht. Die Entfaltung des ersten Zwischenzustandes erzeugt Bedingungen konstanter Kraftbelastung für den zweiten Zwischenzustand. Zur Charakterisierung der Lebensdauer des zweiten Zwischenzustandes wurden hochauflösende Messungen mit schnellen Blattfedern durchgeführt. Eine Analyse der gemessenen Lebensdauerverteilung ergab eine natürliche Lebensdauer dieses Zustandes von etwa 10 s und eine Potentialbreite von 0.55 ± 0.1 nm. Durch hochauflösende Messungen mit rauscharmen Blattfedern konnte ein dritter Zwischenzustand beobachtet werden. Der Zerfall dieses Zustandes führt das GFP in den komplett entfalteten Zustand.

Aus den Ergebnissen dieser Untersuchungen wurde ein detailliertes Bild der Energielandschaft entlang des Pfades der mechanischen Entfaltung des GFP rekonstruiert. Die Experimente sind zudem grundlegend für die Konstruktion eines GFP-basierten molekularen Kraftsensors.

Kapitel 1

Einleitung

Proteine machen mehr als 50% der Trockenmasse einer Zelle aus. Sie sind an nahezu allen zellulären Prozessen maßgeblich beteiligt. Proteine katalysieren chemische Reaktionen, sie verdauen, regulieren und formen Kanäle und Pumpen. Proteine transportieren ganze Organellen entlang von Schienen, die aus Proteinen gebildet werden. Sie sorgen für Kommunikation zwischen Zellen und erzeugen Lumineszenz. Dies sind nur Beispiele für die vielfältigen Aufgaben, die von Proteinen wahrgenommen werden. Trotz ihrer unterschiedlichen Aufgaben haben die meisten Proteine jedoch eines gemeinsam: ihre biologische Funktion erreichen sie nur über die korrekte Faltung in eine hochspezifische, dreidimensionale Struktur.

Wie sich Proteine in ihre funktionale Form falten, ist Thema der aktuellen Forschung. Maßgeblich für diesen Prozess ist eine komplexe Energielandschaft. In der Vergangenheit konnten Details dieser Energielandschaft experimentell nur sehr schwer bestimmt werden.

Mit der Entwicklung der Einzelmolekül-Kraftspektroskopie ist die Untersuchung der mechanischen Eigenschaften einzelner Proteinmoleküle möglich geworden. Mit dieser Technik kann die Energielandschaft von Proteinen entlang kontrollierter Pfade erforscht werden. Es können Informationen über energetische Barrieren und Potentialbreiten entlang dieser mechanischen Faltungs- und Entfaltungspfade gewonnen werden.

Bei den meisten der bisher kraftspektroskopisch untersuchten Proteine wurde eine Faltung und Entfaltung über einen Zwei-Zustandsprozess beobachtet. Die Energielandschaften weisen dabei nur eine energetische Barriere zwischen dem gefalteten und entfalteten Zustand auf. In der Natur gibt es jedoch eine große Zahl topologisch ungleich komplexerer und größerer Proteine, als die, die bisher untersucht worden sind. Daher stellen sich folgende Fragen:

- Äußert sich Größe und Komplexität eines Proteins auch in der Energielandschaft?
- Gibt es Zwischenzustände des Proteins entlang bestimmter Pfade in der Energielandschaft?
- Können Unterschiede in der Stabilität verschiedener Teile der gefalteten Struktur identifiziert werden?

Mit der vorliegenden Diplomarbeit wird anhand kraftspektroskopischer Untersuchungen des *Grün Fluoreszierenden Proteins* (GFP) aus der Qualle *Aequorea Victoria* ein Beitrag zur Beantwortung dieser Fragen geleistet.

Das GFP als Beispiel für ein großes und topologisch komplexes Protein hat aufgrund seiner günstigen Eigenschaften in der Zellbiologie eine Vielzahl von Anwendungen gefunden. Das GFP kodierende Gen enthält alle für die Fluoreszenz nötigen Informationen und das Protein ist zudem ungiftig. Aus diesem Grund ist es zum wichtigsten Fluorophor für die *in vivo* Fluoreszenzmarkierung geworden.

Über die mechanischen Eigenschaften des GFP ist bisher nichts bekannt. Faltungsstudien am GFP sind rar und werden durch die sehr langsame Kinetik des Proteins erschwert. Da die Fluoreszenz des GFP einen Indikator für die native GFP-Struktur bildet, kann eine Untersuchung der mechanischen Stabilität einzelner GFP-Moleküle zudem die Grundlage für die Konstruktion GFP-basierter molekularer Kraftsensoren schaffen.

Kapitel 2

Das Grün Fluoreszierende Protein

2.1 Biolumineszenz

Es gibt chemische Reaktionen, bei denen überschüssige Reaktionsenergie in Form von Licht emittiert wird. Derart verursachte Lichtemission wird als Chemolumineszenz bezeichnet. Ist für diese Reaktion ein katalysierendes Enzym notwendig, so nennt man die Erscheinung auch Biolumineszenz.

Biolumineszenz ist in der Evolution des Lebens auf der Erde in vielen verschiedenen Organismen wie Bakterien, einzelligen Algen, Polypen, Medusen, Käfern und Fischen entstanden.

Chemisch gesehen handelt es sich bei den Prozessen, die zur Biolumineszenz in diesen Organismen führen, immer um durch bestimmte Enzyme (*Luciferasen*) katalysierte Reaktionen von molekularem Sauerstoff mit verschiedensten Substraten (*Luciferine*). Dieser Prozess ist in Abbildung 2.1 schematisch dargestellt.

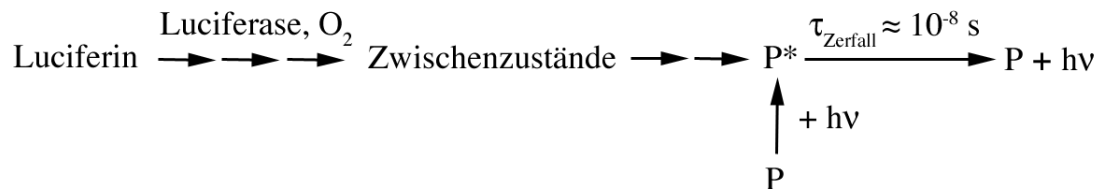


Abbildung 2.1: Luciferin wird mittels Luciferase und molekularem Sauerstoff in Zwischenzustände umgewandelt, bis schließlich ein Molekül P*, in einem angeregten Zustand erzeugt wird. Der angeregte Zustand zerfällt schnell unter Emission von sichtbarem Licht. Graphik adaptiert aus [43].

Mehrere Faktoren beeinflussen die Farbe der Biolumineszenz. Im einfachsten Fall stimmt das Emissionsspektrum mit dem Fluoreszenzspektrum des angeregten Reaktionsproduktes (P* in Abb. 2.1) überein. Die Lichtemission kann jedoch auch von einem an die Luciferase

gebundenem Reaktionsprodukt ausgehen, dessen Intensität und Spektrum sich vom freien, emittierenden Molekül unterscheidet.

Eine weitere Möglichkeit ist, dass die Energie des angeregten Reaktionsproduktes P^* zu einem zweiten Fluorophor (F in Abb. 2.2) übertragen wird, in dem z.B. das zweite Fluorophor F das von P^* emittierte Licht absorbiert und anschließend Licht mit dem für F charakteristischen Spektrum emittiert. Es existieren jedoch auch Mechanismen, die strahlungslose Energieübertragung ermöglichen. Befinden sich die zwei Fluorophore in atomaren Abständen zueinander, so kann es zum Energietransfer über einen Elektronenaustausch kommen (Dexter Energie Transfer). Eine zweite strahlungslose Energieübertragung ist der sogenannte Förster Resonanz Energie Transfer (FRET). Eine Voraussetzung dafür ist, dass das Emissionsspektrum von P^* teilweise mit dem Absorptionsspektrum der zweiten Spezies F überlappt. Die Effizienz dieser Form der Energieübertragung ist eine Funktion von r^{-6} , wobei r der Abstand zwischen den zwei Spezies P^* und F ist.

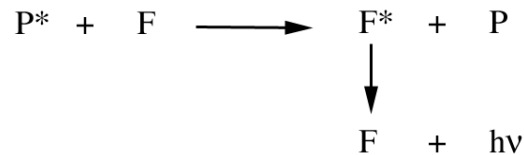


Abbildung 2.2: Ein zweites Fluorophor F kann mit der Energie von P^* angeregt werden und der eigenen Charakteristik Licht emittieren. Graphik adaptiert aus [43].

Ein Protein, das ein solches als F bezeichnetes zweites Fluorophor darstellt, ist das *Grün Fluoreszierende Protein*.

2.2 Historisches

Die Qualle *Aequorea Victoria* ist ein biolumineszenter Organismus. 1962 wurde durch Shimomura *et al.* das für die Biolumineszenz verantwortliche Enzym *Aequorin* isoliert [41]. Dabei wurde gleichzeitig ein Protein isoliert, das zu leicht grünlichen Lösungen führte und unter UV-Beleuchtung eine kräftige Fluoreszenz bei Wellenlängen um 508 nm aufwies. *Aequorin* und das seiner Erscheinung nach benannte *Grün Fluoreszierende Protein* bilden in *Aequorea Victoria* ein System wie in Abb. 2.2 beschrieben. Über strahlungslosen Energietransfer nach Förster wird die eigentlich bläuliche Lumineszenz des reinen *Aequorin*s über das GFP in grünliche Lumineszenz verwandelt.

1992 wurde durch Prasher *et al.* erstmalig das GFP kodierende Gen aus *Aequorea Victoria* kloniert [30]. Die Expression dieses Gens in anderen Organismen führte zur Fluoreszenz. Das Gen allein enthält demnach alle für die Entwicklung der Fluoreszenz nötigen Informationen. Dieser Umstand hat nicht zuletzt zu der Vielfalt der heutigen Anwendungen des GFP geführt. So wird das GFP z.B. intensiv als Genexpressions-Marker für *in vivo* Untersuchungen verwendet. Fusioniert man das GFP-Gen mit dem Gen eines anderen Proteins, können dank der Fluoreszenzeigenschaften des GFP Informationen über Zeit und Ort der Expression des Proteins gewonnen werden.

Grün fluoreszierende Proteine finden sich in zahlreichen Polypen und Medusen der Weltmeere. Sie treten wie in *Aequorea Victoria* als Begleitproteine zu anderen bläulich chemolumineszenten Proteinen wie Aequorin auf und bestimmen die Farbe der Biolumineszenz zu grün. Es ist unklar, warum diese Organismen überhaupt biolumineszent sind, warum eine grüne Emission zu einem evolutionären Vorteil gegenüber der bläulichen Emission führt und warum diese Organismen den Energieaufwand betreiben, ein separates, grün leuchtendes Protein zu synthetisieren anstatt die bereits vorhandene Aequorin-Entsprechung zu einer grünen Emission hin zu mutieren.

Der Begriff *Grün Fluoreszierendes Protein* bezieht sich im folgenden ausschließlich auf das GFP aus der Qualle *Aequorea Victoria*.

2.3 *Wildtype*-GFP und *Cycle3*-GFP

Das GFP ist ein aus 238 Aminosäuren bestehendes wasserlösliches, globuläres Protein. Die Kristallstruktur des GFP ist in Abb. 2.3 abgebildet (pdb-Datei: 1EMB [4]). Das GFP bildet ein sogenanntes β -Barrel: 11 β -Faltblätter sind derart aneinander gelagert, dass eine einem Fass ähnelnde Struktur entsteht. α -helikale Motive bilden zu beiden Seiten des β -Barrels eine Art Deckel für das Fass. Diagonal durch das Fass verläuft eine α -Helix, die das für die Fluoreszenz verantwortliche Chromophor im Zentrum der Struktur positioniert. Das Chromophor selbst wird aus drei Aminosäuren über einen komplexen, autokatalytischen Ringschluss unter Zufuhr von Sauerstoff gebildet (für die Details sei auf [46] verwiesen). Diese drei Aminosäuren sind die Residuen 65 bis 67 (Serin, Tyrosin, Glycin).

Durch Einschluß des Chromophores in das Fass ist es damit vor äußeren Einflüssen aus dem Lösungsmittel, die die Fluoreszenz beeinträchtigen könnten, geschützt.

Die Qualle *Aequorea Victoria* hat ihren Lebensbereich hauptsächlich im kalten Nordwest-Pazifik und GFP formt sich am besten bei den dort herrschenden Temperaturen. Für die meisten Anwendungen sind jedoch physiologische Temperaturen im Bereich von 37°C interessant. Bei diesen Temperaturen bildet sich die native Struktur des GFPs jedoch äußerst schlecht, was auf eine starke Tendenz zur Aggregation zurückzuführen ist [17]. Allerdings beschränkt sich die Temperaturempfindlichkeit des Proteins scheinbar nur auf die Faltung. Richtig gefaltetes GFP ist stabil und zeigt Fluoreszenz bis hin zu etwa 65°C [41].

Über DNA-Shuffling wurde eine Mutante des GFP erzeugt, die sich auch bei 37°C effizient faltet [11]. Diese Mutante ist unter dem Namen *Cycle3*-GFP bekannt. Sie unterscheidet sich vom *Wildtype*-GFP nur durch drei Punktmutationen: Phe→Serin an Position 99, Met→Thr an Position 153 und Val→Ala an Position 163. *Cycle3*-GFP ist bei 37°C besser löslich als *Wildtype*-GFP und zeigt eine niedrigere Tendenz zur Aggregation während des Faltungsprozesses. Die *Cycle3*-Mutante zeigt identische spektrale Eigenschaften wie *Wildtype*-GFP und wird weithin verwendet. Die Kristallstruktur von *Cycle3*-GFP konnte aufgelöst werden (pdb-Datei: 1B9C [4]) und zeigt kaum Unterschiede zur *Wildtype*-Struktur.

In der vorliegenden Arbeit wird *Cycle3*-GFP¹ hinsichtlich der mechanischen Eigen-

¹Das verwendete GFP zeigte noch eine weitere Punktmutation: Ser→Gly an Position 2

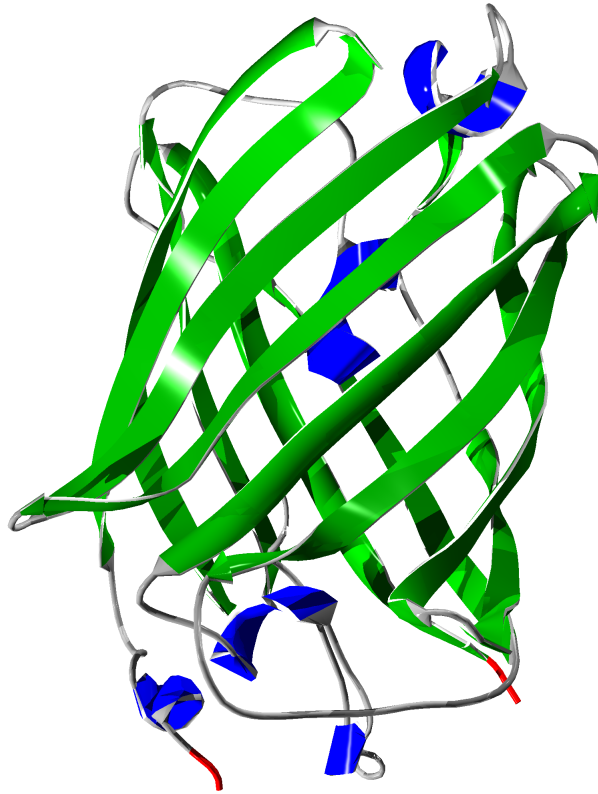


Abbildung 2.3: *Wildtype*-GFP Struktur, pdb-Datei 1EMB [4]. Rot markiert: Amino- und Carboxyterminus des Moleküls.

schaften untersucht. Die für die Expression verwendete DNA-Sequenz ist in Tabelle E.1 angegeben. Die zugehörige Aminosäuren-Sequenz findet sich in Tabelle E.2.

2.4 Fluoreszenz als Indikator des nativen GFP

Die native Struktur des GFP bildet sich auf einer Zeitskala von einigen Minuten [46][41][17]. Die erstmalige, korrekte Faltung bringt die Aminosäuren 65 bis 67 (Serin, Tyrosin, Glycin) in eine Position, die den Reifungsprozess des Chromophores aus diesen Aminosäuren ermöglicht. Dieser Prozess dauert bis zu einigen Stunden [46].

Hat sich das Chromophor einmal gebildet, kann GFP z.B. durch Zugabe von GdnHCl denaturiert werden. Von Fukuda *et al.* wurde über Messungen des zirkulären Dichroismus einer GFP-Lösung der Anteil von nativem GFP in der Lösung als Funktion der GdnHCl-Konzentration bestimmt [17]. Gleichzeitig wurde die Intensität der Fluoreszenz bei 508 nm Wellenlänge als Funktion der GdnHCl-Konzentration bestimmt.

Beide Kurven in Abb. 2.4 stimmen sowohl für *Wildtype*-GFP und *Cycle3*-GFP jeweils überein. Das Verlöschen der Fluoreszenz des GFP bildet daher einen Indikator für den Verlust der nativen Struktur. Dies ist das wesentliche Argument, das eine mögliche Anwendung

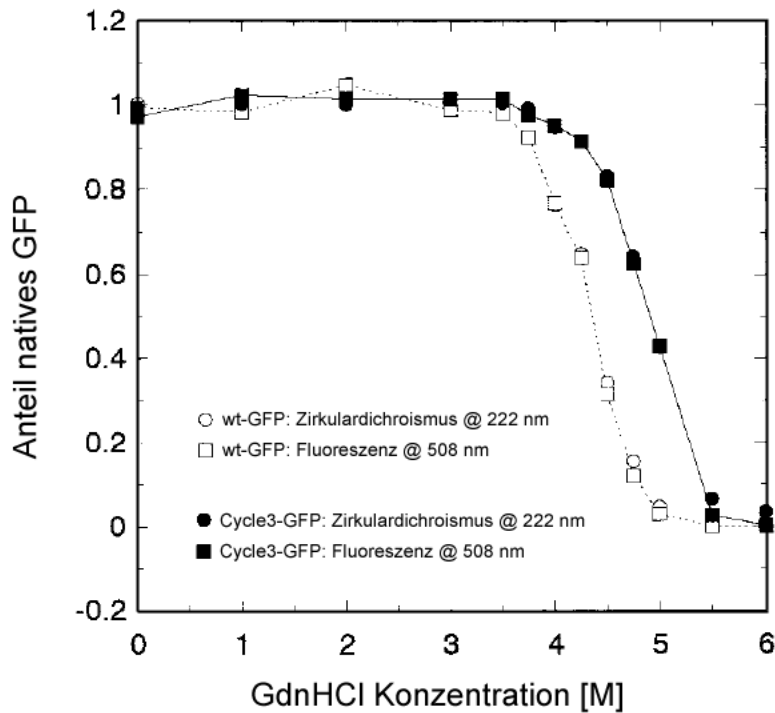


Abbildung 2.4: Fluoreszenzintensität bei 508 nm und Zirkulardichroismus bei 222 nm als Funktion der GdnHCl-Konzentration von *Wildtype*-GFP und *Cycle3*-GFP. Graphik adaptiert aus [17].

des GFP als molekularen Kraftsensor begründet.

In der genannten Studie wurde ebenfalls die Rückfaltung des GFP nach Denaturierung mit GdnHCl untersucht. Die grüne Fluoreszenz des GFP tritt dabei im selben Maße wieder auf, wie die native Struktur gebildet wird. Die Fluoreszenz des GFP bildet somit auch einen Indikator für die Bildung der nativen Struktur [17], sofern das Chromophor vormals bereits gereift war.

Nach den Erkenntnissen aus dieser Studie bildet sich die native Struktur bei natürlichen Bedingungen mit einer Rate von $k_{on}^0 \simeq 10^{-2} \text{ s}^{-1}$, also auf einer Zeitskala von Minuten (andere Literaturwerte liegen bei 10 Minuten [41]). Die natürliche Übergangsrate des GFP in den entfaltenen Zustand k_{off}^0 liegt hingegen bei etwa 10^{-9} s^{-1} [17]. Das GFP ist somit ein sehr stabiles Protein, dessen langsame Faltungs- und Entfaltungskinetik entsprechende Untersuchungen erschwert. Es findet sich daher auch wenig Literatur zu diesem Thema, obwohl das GFP zu einem vielgenutzten Werkzeug in der Wissenschaft geworden ist.

In Abb. 2.5 ist ein Fluoreszenzspektrum der in dieser Arbeit untersuchten DdFLN(1-5)-GFP-Chimäre abgebildet, um die grundlegende Eigenschaft des GFP darzustellen. Im folgenden wird jedoch ausschließlich über die mechanischen Eigenschaften des GFP gesprochen werden.

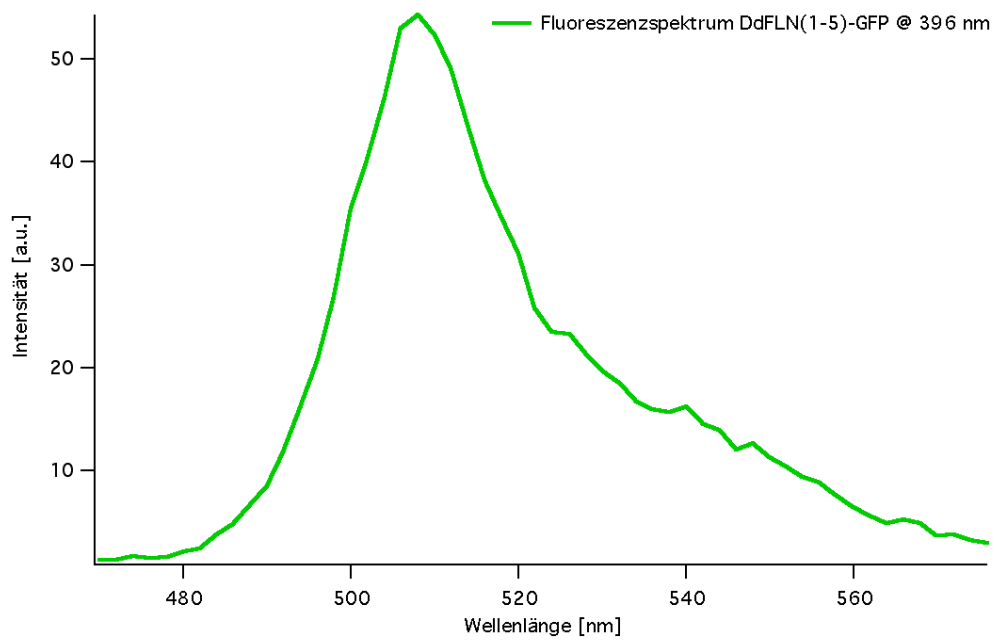


Abbildung 2.5: Fluoreszenzspektrum von DdFLN(1-5)-GFP bei einer Anregungswellenlänge von 396 nm.

Kapitel 3

Theoretische Grundlagen

3.1 Elastizität von Polymeren

Proteine sind Polymere (Polypeptide). Während klassische Polymere aber Ketten aus immer gleichen Gliedern sind, bestehen Proteine aus einer Abfolge von Gliedern, deren Rückgrate zwar identisch sind, jedoch Restgruppen aufweisen, die sich erheblich unterscheiden können. Trotzdem ist das elastische Verhalten denaturierter Proteine prinzipiell sehr ähnlich zu dem klassischer Polymere: Um sie zu strecken, benötigt man eine Kraft.

Die für eine bestimmte Streckung aufzuwendende Arbeit begründet sich zum Teil dadurch, dass mit der Streckung die Anzahl möglicher Konformationen, die das Polymer einnehmen kann, erheblich eingeschränkt wird und somit die Entropie des Polymers reduziert wird. Diese *entropische Elastizität* dominiert vor allem bei niedrigen Kräften bzw. Streckungen. Bei hohen Kräften und nahezu ausgestrecktem Polymer dominieren Beiträge, die von Bindungspotentialen innerhalb des Polymers stammen (Strecken von Bindungswinkeln, Dehnen von Bindungen). Dies führt zu einer *enthalpischen Elastizität*.

Es gibt mehrere Modelle, die zur Beschreibung der entropischen Elastizität von Polymeren entwickelt worden sind. Die bekanntesten darunter sind das FJC (*Freely-Jointed-Chain*)-Modell und das FRC (*Freely-Rotating-Chain*)-Modell [15]. Kürzlich wurde das EJC (*Elastically-Jointed-Chain*)-Modell vorgeschlagen, das zur Beschreibung von Objekten aus globulären Unterstrukturen (wie z.B Aktin oder Tubulin) geeignet ist [26]. Ein bekanntes Modell, das im Prinzip eine Mischform aus enthalpischen und entropischen Beiträgen darstellt, ist das WLC (*Wormlike-Chain*)-Modell (auch *Kratky-Porod*-Modell) [22][23].

Während die drei erstgenannten Modelle die möglichen Konfigurationen einer Kette aus diskreten Gliedern betrachten, ist das WLC-Modell ein Kontinuumsmodell. Es hat sich gezeigt, dass das WLC-Modell teilweise das Elastizitätsverhalten von Polypeptiden gut beschreiben kann. Das WLC-Modell kommt in der vorliegenden Arbeit zur Anwendung. Die zugrunde liegenden Ideen sollen daher im folgenden Abschnitt kurz erläutert werden.

3.2 Das Wormlike-Chain Modell (WLC)

Die folgende Ableitung ist im wesentlichen aus [7][22] adaptiert worden.

Das WLC-Modell abstrahiert ein Polymer als kontinuierliche Raumkurve $\vec{r}(s)$ (Funktion der Bogenlänge s). Die Bogenlänge s kann dabei Werte zwischen Null und L annehmen, wobei L die Länge des Polymers im vollständig ausgestreckten Zustand bezeichnet. L wird auch *Konturlänge* des Moleküls genannt. Als Tangentenvektor bezeichnet man $\vec{t}(s) = d\vec{r}(s)/ds$. Die mit einer bestimmten Konformation verbundene Energiedichte e_b des Polymers wird im WLC-Modell als

$$e_b(s) = \frac{p \cdot k_B T}{2} \left| \frac{d^2 \vec{r}(s)}{ds^2} \right|^2 \quad (3.1)$$

angenommen. Dabei ist $|d^2 \vec{r}(s)/ds^2|^2 = \kappa^2$ das Quadrat der Krümmung der Raumkurve in einem durch s definierten Punkt. Der Parameter p heißt Persistenzlänge und beschreibt die Länge, ab der die Richtungen der Tangentenvektoren $\vec{t}(s)$ und $\vec{t}(s+p)$ nicht mehr korreliert sind. Je größer p , umso steifer ist das Polymer. Die Energie eines durch eine entlang der z -Achse wirkenden Kraft $\vec{F} = F \cdot \vec{e}_z$ gestreckten Polymers beträgt

$$E_{WLC} = \int_0^L \frac{p \cdot k_B T}{2} \cdot \left| \frac{d^2 \vec{r}(s)}{ds^2} \right|^2 - F \cdot \vec{t}(s) \cdot \vec{e}_z \, ds . \quad (3.2)$$

Mit Gl. 3.1 modelliert man einen enthalpischen Beitrag, mit dem über den Boltzmann-Faktor die möglichen Konformationen der Raumkurve $\vec{r}(s)$ gewichtet werden können. Die Zustandssumme des Systems ergibt sich daher aus:

$$Z_{WLC}(L, F, \vec{t}_{s=0}, \vec{t}_{s=L}) = \int \exp\left(-\frac{E_{WLC}}{k_B T}\right) \Gamma(\vec{t}) \quad (3.3)$$

Dabei bedeutet $\Gamma(\vec{t})$, dass über alle möglichen Pfade, die der Tangentenvektor \vec{t} für eine bestimmte Form der Raumkurve $\vec{r}(s)$ auf der Einheitskugel zeichnen kann (welche jeweils bei $\vec{t}(s=0)$ beginnen und bei $\vec{t}(s=L)$ aufhören) integriert werden muß. Gl. 3.3 hat die Form eines Feynmanschen Pfadintegrals und kann mit Methoden aus der Quantenmechanik gelöst werden [7]. Aus der Zustandssumme kann über

$$\langle d(F) \rangle = k_B T \cdot \frac{\partial \ln Z_{WLC}}{\partial F} \quad (3.4)$$

der mittlere End-zu-End Abstand $\langle d \rangle$ des Polymers als Funktion der wirkenden Kraft F bestimmt werden. d wird im folgenden auch als *Ausdehnung* des Polymers bezeichnet. In Bezug auf die durchgeführten Experimente ist die Angabe der vom Polymer ausgeübten Rückstellkraft $F(d)$ als Funktion des End-zu-End Abstandes d interessant. Diese ergibt sich aus

$$\langle F(d) \rangle = -k_B T \cdot \frac{\partial \ln Z_{WLC}}{\partial d} \quad (3.5)$$

Die Lösung von Gl. 3.4 bzw. Gl. 3.5 ist nur für Grenzfälle (kleine, extrem steife Polymere und sehr lange, flexible Polymere) analytisch angebar [22]. Es existieren jedoch Methoden, die eine numerische Lösung von Gl. 3.4 und Gl. 3.5 ermöglichen [7][22].

3.3 Gültigkeit des WLC-Modells

Für kleine Kräfte sagt das WLC-Modell eine relative Ausdehnung entlang der Kraftrichtung von

$$\frac{d_{F \rightarrow 0}}{L} \simeq \frac{2}{3} \frac{p}{k_B T} \cdot F \quad (3.6)$$

voraus. Dieser Kraftbereich zeigt einen linearen Zusammenhang zwischen angelegter Kraft und Ausdehnung. Experimentell ermittelt man für Polypeptide Persistenzlängen im Bereich von $p = 0.8$ nm im niedrigen Kraftbereich (0 bis 30 pN), wenn man das WLC-Modell in diesem Kraftbereich an die Daten anpasst [35]. Für große Kräfte hingegen gilt die folgende Beziehung für das WLC-Modell:

$$\frac{d_{F \rightarrow \infty}}{L} = 1 - \sqrt{\frac{k_B T}{4pF}} \quad (3.7)$$

Dieses Verhalten wird an DNA tatsächlich beobachtet [8].

Für Kräfte im Bereich von $F \sim k_B T/p$ kann eine von Marco und Siggia entwickelte Interpolationsformel für das Kraft-Ausdehnungs-Verhalten $F(d)$ eines WLC-Polymers verwendet werden [8][27]. Sie lautet:

$$F = \frac{k_B T}{p} \cdot \left(\frac{d}{L} + \frac{1}{4(1 - \frac{d}{L})^2} - \frac{1}{4} \right) \quad (3.8)$$

Gl. 3.8 reproduziert exakt die numerische Lösung von Gl. 3.5 in den Grenzfällen sehr hoher und sehr niedriger Kräfte. In den Zwischenbereichen führt sie zu Abweichungen von bis zu 10% von der (numerisch) exakten Lösung des WLC-Modelles. Es existieren jedoch Methoden, die die Voraussagen der Interpolationsformel im Vergleich zur exakten Lösung stark verbessern [7].

Vergleicht man im mittleren Kraftbereich (30-200 pN) die Voraussagen des WLC-Modells mit an Polypeptiden gemessenen Kraft-Ausdehnungskurven, so zeigt sich, daß der Parameter p im Bereich von 0.3-0.5 nm gewählt werden muss, um eine gute Anpassung des Modells an experimentelle Daten zu erreichen.

Der Umstand, dass die Persistenzlänge p in verschiedenen Kraftbereichen unterschiedlich gewählt werden muss, um das Elastizitätsverhalten eines Polypeptids zu reproduzieren, zeigt einen Mangel des WLC-Modells auf. Die verschiedenen Restgruppen der Aminosäuren verursachen verschiedene Wechselwirkungen, die das elastische Verhalten von Polypeptiden im niederen Kraftbereich beeinflussen. Diese Einflüsse bleiben im WLC-Modell unberücksichtigt und begründen die genannte Unzulänglichkeit des Modells in Bezug auf Polypeptide.

Bei höheren Kräften (> 500 pN) bekommt die Bindungsstruktur des Peptid-Rückgrates einen deutlichen Einfluss. Das WLC-Modell ist in diesem Kraftbereich soweit verbessert worden, dass nur noch die Konturlänge L als Fit-Parameter benötigt wird [20].

3.4 Verwendung des WLC-Modells

Das WLC-Modell hat eine Eigenschaft (insbesondere die Interpolationsformel Gl. 3.8), die für die quantitative Auswertung gemessener Kraftkurven sehr nützlich ist. Die vorausgesagte Kraft ist eine Funktion der relativen Ausdehnung, also $F = F(d/L)$. Das Elastizitätsverhalten realer Polypeptide ist auch hauptsächlich eine Funktion von d/L . Das WLC-Modell kann daher zum Messen von Gesamtlängen (bzw. Konturlängen) verwendet werden, vorausgesetzt, man kann erhaltene Längenangaben L eichen. Die ist z.B. über eine Kraftkurve an einem Polypeptid, dessen Gesamtlänge bzw. Anzahl der Aminosäuren bekannt ist, möglich. Dieser Weg wurde in dieser Diplomarbeit besprochen.

Abb. 3.1 zeigt ein Beispiel für eine Anpassung der WLC-Interpolationsformel Gl. 3.8 an eine gemessene Kraftkurve. Ein Wert von $p = 0.5$ nm für die Persistenzlänge führt zu guter Übereinstimmung zwischen beiden Kurven im Kraftbereich 20 bis 150 pN. Dieser Parameter und Gl. 3.8 ist für alle WLC-Anpassungen in dieser Arbeit verwendet worden.

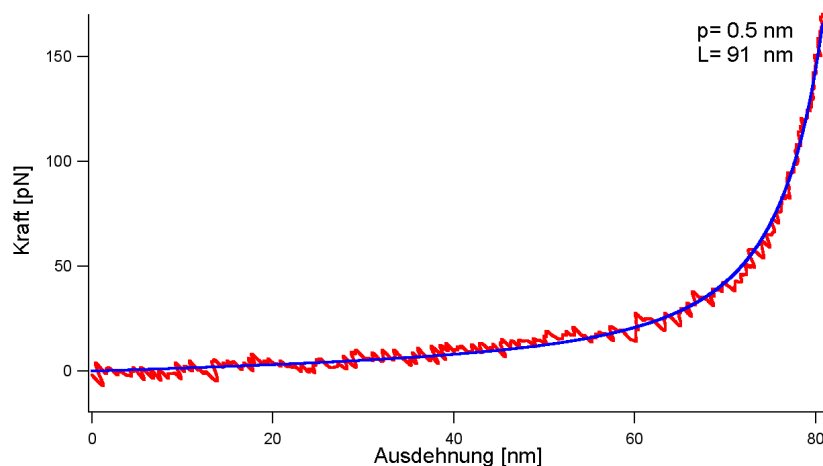


Abbildung 3.1: Anpassung des WLC-Modelles über Gl. 3.8 mit den Parametern $p = 0.5$ nm und $L = 91$ nm an eine Kraftkurve, die an einem Polypeptid gemessen wurde.

3.5 Konformationskinetik von Proteinen

Die Konformationskinetik von Proteinen spielt in dieser Arbeit eine wichtige Rolle. Dieser Abschnitt bietet daher eine kurze Einführung in das Thema. Die Inhalte dieses Abschnittes sind im wesentlichen aus [19] adaptiert worden.

Viele Proteine treten entweder in einer nativen (auch als gefaltet bezeichneten) Struktur (S_n) oder in einer ungefalteten, denaturierten Form (S_u) auf. Native Strukturen können über Methoden wie Röntgenstrukturaufklärung oder Kernmagnetische Resonanz bestimmt werden. Die native Struktur S_n stellt ein Ensemble einer großen Zahl von Konformationen

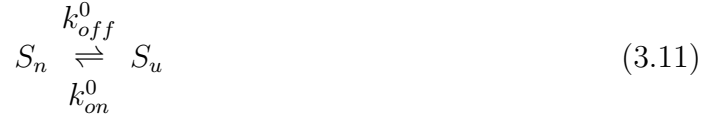
dar, die sich nicht stark von einem gemeinsamen Mittel unterscheiden. Ein ungefaltetes Protein S_u hingegen ist ein Ensemble von Konformationen, die so starke Fluktuationen aufweisen, dass eine gemeinsame Struktur nicht sinnvoll definierbar ist. Methoden der Strukturaufklärung führen hier zu keinen Ergebnissen. Die freie Enthalpie für einen solchen Ensemble-Zustand (S_n oder S_u) ist

$$G = U - TS \quad (3.9)$$

wobei U die mittlere innere Energie des Systems bezeichnet, T die Temperatur und S die Entropie ist. Die Wahrscheinlichkeit, ein einzelnes Protein in einem Zustand S_n oder S_u anzutreffen ist gleich zu dem Verhältnis der Konzentrationen von Proteinen in diesen Zuständen und ist durch den Boltzmann-Faktor gegeben.

$$\frac{[S_u]}{[S_n]} = K_{eq}^0 = \exp\left(-\frac{G(S_u) - G(S_n)}{k_B T}\right) \quad (3.10)$$

Betrachtet man ein einzelnes Protein, so kann es mit den Raten k_{off}^0 bzw. k_{on}^0 zwischen den beiden Zustands-Ensembles S_n und S_u wechseln.



Die zeitliche Veränderung der Konzentrationen kann im einfachsten Fall als linearer Zusammenhang modelliert werden:

$$\frac{d[S_n]}{dt} = -k_{off}^0[S_n] + k_{on}^0[S_u] \quad (3.12)$$

Im thermischen Gleichgewicht ist $d[S_n]/dt = 0$. Damit läßt sich ein Zusammenhang zu Gl. 3.10 herstellen.

$$\frac{k_{off}^0}{k_{on}^0} = K_{eq}^0 \quad (3.13)$$

Das bedeutet, dass die Wahrscheinlichkeit dafür, ein einzelnes Protein im Zustand S_u anzutreffen, durch die Gleichgewichtskonstante K_{eq}^0 angegeben wird. Die mittlere Lebensdauer des Zustandes S_n ist damit $\tau_n^0 = 1/k_{off}^0$ und die des Zustandes S_u ist $\tau_u^0 = 1/k_{on}^0$. Ein solches Verhalten lässt sich durch Einführung eines sogenannten Übergangszustandes T mit einer sehr viel höheren freien Enthalpie $G(T)$ zwischen den beiden Zuständen S_n und S_u erklären. Diese Situation ist in Abb. 3.2 dargestellt. Dabei ist als Reaktionskoordinate der End-zu-End Abstand (Ausdehnung) d des betrachteten Proteins eingeführt worden. Die Übergangsraten k_T vom Zustand T nach S_n oder S_u sei jeweils gleich. Die Wahrscheinlichkeit, das Protein im Zustand T anzutreffen, beträgt:

$$\frac{[T]}{[S_n]} = \exp\left(-\frac{G(T) - G(S_n)}{k_B T}\right) \quad (3.14)$$

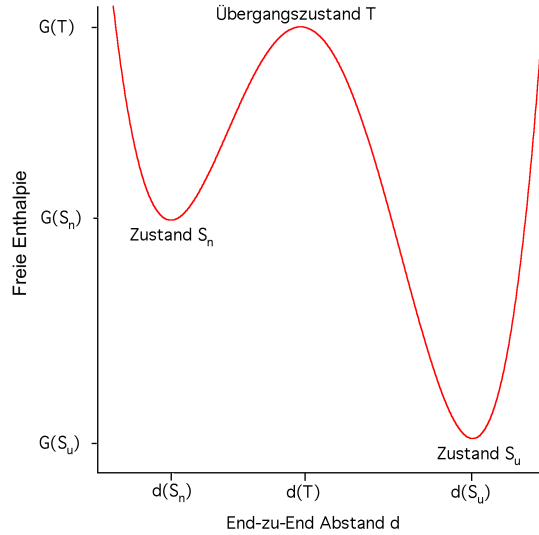


Abbildung 3.2: Schema der energetischen Verhältnisse entlang des End-zu-End Abstandes als Reaktionskoordinate. Das Gleichgewicht ist in diesem Fall stark zum Zustand S_u verschoben.

bzw.

$$\frac{[T]}{[S_u]} = \exp\left(-\frac{G(T) - G(S_u)}{k_B T}\right) \quad (3.15)$$

Schreibt man Gl. 3.12 für die Übergänge zwischen den Zuständen T und S_n sowie T und S_u und verlangt thermisches Gleichgewicht, so folgt:

$$\frac{k_{S_n \rightarrow T}}{k_T} = \frac{[T]}{[S_n]} \quad \text{bzw.} \quad \frac{k_{S_u \rightarrow T}}{k_T} = \frac{[T]}{[S_u]} \quad (3.16)$$

Thermisches Gleichgewicht soll aber auch zwischen den Zuständen S_n und S_u gelten. Damit muß Gl. 3.13 auch gelten. Dies ist erfüllt, wenn $k_{S_n \rightarrow T} = k_{off}^0$ und $k_{S_u \rightarrow T} = k_{on}^0$ ist. Es ergibt sich die sogenannte *Arrhenius-Gleichung* für die Übergangsraten:

$$k_{off}^0 = k_T \cdot \exp\left(-\frac{G(T) - G(S_n)}{k_B T}\right) \quad \text{und} \quad k_{on}^0 = k_T \cdot \exp\left(-\frac{G(T) - G(S_u)}{k_B T}\right) \quad (3.17)$$

Damit bestimmt die Höhe der Energiebarriere zwischen den beiden Zuständen S_n und S_u hauptsächlich die Übergangsraten. Wie weit das Gleichgewicht zu den Zuständen S_u bzw. S_n verschoben ist, hängt dagegen von der relativen Lage der zugehörigen freien Enthalpien $G(S_n)$ und $G(S_u)$ ab. Der Vorfaktor k_T in Gl. 3.17 (in der Literatur auch als *pre-exponential factor* und *frequency factor* bekannt) ergibt sich aus der genauen Form der Energielandschaft.

Die Bestimmung des Faktors k_T ist ein Ergebnis der *Kramers-Raten-Theorie* [19]. Dabei wird angenommen, das Protein befinde sich im Zustand S_n und diffundiere mit der

Federkonstante $\kappa_n = \partial_d G(S_n)$ um eine Ruheposition bei $d = d(S_n)$. Erreicht es die Position $d(T)$, kann es in den Zustand S_n überwechseln, in welchem das Protein nun mit der Federkonstante $\kappa_u = \partial_d G(S_u)$ um die Ruheposition bei $d = d(S_u)$ diffundiert. Mit diesen Annahmen kann der Vorfaktor k_T in Gl. 3.17 abgeschätzt werden. Es ergibt sich, dass der Faktor k_T in diesem auf Diffusion basierenden Modell abhängig von der Reaktionsrichtung $S_n \rightarrow S_u$ bzw. $S_u \rightarrow S_n$ bestimmt werden muß. Die Gleichgewichtsbedingung zwischen den Zuständen S_n und S_u bleibt jedoch erhalten.

$$k_T(S_n \rightarrow S_u) = \frac{1}{\pi} \frac{\kappa_u d_{T-S_u}}{\kappa_n d_{S_n-T} + \kappa_u d_{T-S_u}} \frac{\kappa_n}{\gamma_n} \sqrt{\frac{G(T) - G(S_n)}{k_B T}} \quad (3.18)$$

$$k_T(S_u \rightarrow S_n) = \frac{1}{\pi} \frac{\kappa_n d_{S_n-T}}{\kappa_n d_{S_n-T} + \kappa_u d_{T-S_u}} \frac{\kappa_u}{\gamma_u} \sqrt{\frac{G(T) - G(S_u)}{k_B T}} \quad (3.19)$$

Dabei ist $d_{S_n-T} = d(T) - d(S_n)$ und $d_{T-S_u} = d(S_u) - d(T)$. Die Faktoren γ_n und γ_u bezeichnen die Reibungskoeffizienten des Proteins in den entsprechenden Zuständen.

Experimentell hat es sich gezeigt, dass der Vorfaktor k_T im Bereich der Konformationskinetik von Proteinen einen Wert von etwa $k_T \simeq 10^9 \text{ s}^{-1}$ annimmt [5][24][44].

3.6 Übergangsraten unter konstanter Kraft

In dieser Arbeit werden einzelne Proteine in verschiedenen Zuständen mit Kraft belastet. Welche Auswirkungen eine zunächst zeitlich konstante Kraft auf die Konformationskinetik von Proteinen hat, soll daher anhand eines einfachen Modells in diesem Abschnitt kurz erläutert werden. Die Argumentation ist im wesentlichen aus [19] adaptiert worden.

Es werde nun eine äußere, zeitlich konstante Kraft F an das Protein mit den zwei Zuständen S_u und S_n angelegt. Der in Abb. 3.2 bezeichnete End-zu-End Abstand als Reaktionskoordinate liege parallel zur Richtung der wirkenden Kraft. Wenn sich durch den Übergang eines Proteins vom Zustand S_n nach S_u der End-zu-End Abstand des Moleküls um die Länge $d_{S_n-S_u} = d(S_u) - d(S_n)$ ändert, so ändert sich die freie Enthalpie in einem linearen Modell um den Betrag [3]:

$$\Delta G \simeq G(S_u) - G(S_n) - F d_{S_n-S_u} \quad (3.20)$$

Im thermischen Gleichgewicht gilt daher für die Gleichgewichtskonstante aus Gl. 3.10:

$$K_{eq}(F) = K_{eq}^0 \cdot \exp\left(\frac{F d_{S_n-S_u}}{k_B T}\right) \quad (3.21)$$

Je nach Kraftrichtung verschiebt sich das Gleichgewicht zum Zustand S_n oder S_u . Entsprechend ergibt sich für die Übergangsraten unter Kraftwirkung:

$$k_{off}(F) = k_{off}^0 \cdot \exp\left(\frac{F d_{S_n-T}}{k_B T}\right) \quad (3.22)$$

$$k_{on}(F) = k_{on}^0 \cdot \exp\left(-\frac{Fd_{T-S_u}}{k_B T}\right) \quad (3.23)$$

In diesem linearen Modell wird angenommen, dass

$$k_{on}^0(F) = k_{on}^0(0) \text{ und } k_{off}^0(F) = k_{off}^0(0) \quad (3.24)$$

gilt. Diese Raten werden in dieser Arbeit auch als *natürliche Übergangsraten* bezeichnet. Eine weitere Annahme dieses Modells bezieht sich auf die Position des Übergangszustandes T relativ zu den Zuständen S_n und S_u :

$$d_{S_n-T}(F) = d_{S_n-T}(0) \text{ und } d_{T-S_u}(F) = d_{T-S_u}(0) \quad (3.25)$$

Gilt für den mit dem Übergangszustand T verbundene End-zu-End Abstand $d(T) \simeq d(S_n)$, so hat die Kraft nur einen kleinen Einfluß auf die Übergangsraten k_{off} , während k_{on} sehr stark beeinflusst wird. Eine solche Situation ist z.B. bei einzelnen Titin-Domänen vorgefunden worden. Der Übergangszustand T ähnelt in Bezug auf den End-zu-End Abstand sehr stark dem nativen Zustand S_n [10].

3.7 Übergangsraten unter zeitabhängiger Kraft

Die Kräfte, mit denen in dieser Arbeit Proteine belastet werden, sind jedoch zeitabhängig, d.h. $F = F(t)$. Damit sind die Übergangsraten aus Gl. 3.22 und Gl. 3.23 auch zeitabhängig. Im einfachsten Fall ist dann die Wahrscheinlichkeit $dP_{S_n \rightarrow S_u}$ dafür, dass ein einzelnes Protein den Übergang $S_n \rightarrow S_u$ in einem Zeitintervall dt vollführt, durch die folgende Gleichung gegeben:

$$dP_{S_n \rightarrow S_u}(t, t + dt) = k_{off}(t) dt \cdot P_0 \quad (3.26)$$

P_0 ist die Wahrscheinlichkeit dafür, dass das Protein bis zum Zeitpunkt t den Übergang nicht vollzogen hat. Es gilt $P_0 = 1 - P_{S_n \rightarrow S_u}(t)$, wobei $P_{S_n \rightarrow S_u}(t)$ die Wahrscheinlichkeit dafür angibt, dass der Übergang $S_n \rightarrow S_u$ bis zum Zeitpunkt t schon vollzogen wurde. Nach Integration folgt:

$$P_{S_n \rightarrow S_u}(t) = 1 - \exp\left(-\int_0^t k_{off}(t') dt'\right) \quad (3.27)$$

Die Wahrscheinlichkeit dafür, das Protein bei einem Übergang von $S_u \rightarrow S_n$ beobachten zu können, ist hingegen nach einer ähnlichen Überlegung durch

$$P_{S_u \rightarrow S_n}(t) = 1 - \exp\left(-\int_0^t k_{on}(t') dt'\right) \quad (3.28)$$

gegeben. Hierbei ist die Anfangsbedingung jedoch, das Protein befinde sich im Zustand S_u .

Einsetzen von Gl. 3.22 in Gl. 3.27 und Variablensubstitution $t \rightarrow F$ liefert die Wahrscheinlichkeit für den Übergang $S_n \rightarrow S_u$ als Funktion der Kraft.

$$P_{S_n \rightarrow S_u}(F) = 1 - \exp \left(-k_{off}^0 \cdot \int_0^F \exp \left(\frac{F' d_{S_n-T}}{k_B T} \right) \frac{1}{\dot{F}'} dF' \right) \quad (3.29)$$

Der Faktor $\dot{F}' = dF'/dt$ in Gl. 3.29 wird als *Kraft-Ladungsrate* bezeichnet. Aus Gl. 3.29 lässt sich eine Wahrscheinlichkeitsdichte $dP_{S_n \rightarrow S_u}(F)/dF$ für den Übergang als Funktion der Kraft berechnen. Die Kraft, bei der $dP_{S_n \rightarrow S_u}(F)/dF$ maximal wird, ist dann die wahrscheinlichste Kraft, bei der es zu einem Übergang kommt. Diese Kraft definiert gleichzeitig die Stabilität eines bestimmten Zustandes bei einer bestimmten Kraft-Ladungsrate \dot{F} . Die Auswirkungen zeitabhängiger Kräfte werden ausführlich z.B. in [12] und [18] diskutiert.

Die zeitliche Abhängigkeit der in dieser Arbeit vorliegenden Kräfte wird teilweise als lineare Kraftrampe $F(t) = \eta \cdot t$ angenähert. Die zugehörige Kraft-Ladungsrate \dot{F} ist somit eine Konstante. In diesem Fall ergibt sich folgende, bereits in [31] diskutierte Wahrscheinlichkeitsdichte:

$$\frac{dP(F)}{dF} = \frac{k_{off}^0}{\eta} \exp \left(-\frac{k_{off}^0}{\eta} \frac{k_B T}{d_{S_n-T}} \left[\exp \left(\frac{F d_{S_n-T}}{k_B T} \right) - 1 \right] + \frac{F d_{S_n-T}}{k_B T} \right) \quad (3.30)$$

Die wahrscheinlichste Kraft \bar{F} , bei der ein Übergang $S_n \rightarrow S_u$ stattfinden wird, ist durch

$$\bar{F} = \frac{k_B T}{d_{S_n-T}} \cdot \ln \left(\frac{\eta d_{S_n-T}}{k_{off}^0 k_B T} \right) \quad (3.31)$$

gegeben. Erhöht man die Steigung η der Kraftrampe, so verschiebt sich die Position des Maximums der Wahrscheinlichkeitsdichte ($= \bar{F}$) logarithmisch zu höheren Kräften. Entsprechend verschiebt sich \bar{F} logarithmisch zu niederen Kräften, falls die natürliche Übergangsrates k_{off}^0 erhöht wird. Die Breite der Wahrscheinlichkeitsverteilung bleibt davon jedoch unberührt. Abb. 3.3 verdeutlicht dies.

\bar{F} und die Form der Verteilung werden jedoch stark durch den Abstand d_{S_n-T} entlang der Kraftrichtung beeinflusst. Verkleinert man diesen Parameter, so verbreitert sich die Verteilung stark und \bar{F} wird deutlich höher. Eine Vergrößerung des Abstandes d_{S_n-T} führt hingegen zu schmaleren Verteilungen bei niedrigeren \bar{F} . In Abb. 3.4 sind Wahrscheinlichkeitsdichten aufgetragen, bei denen der Parameter d_{S_n-T} lediglich um einen Faktor 2 geändert wurde.

Die genaue Zeitabhängigkeit der Kräfte $F(t)$, mit denen in der vorliegenden Arbeit einzelne Proteine belastet wurden, waren jedoch Überlagerungen aus einer linearen Kraftrampe und einem von der aktuell wirkenden Kraft abhängigen Term:

$$F(t) = \eta t - k_c d(F(t)) \quad (3.32)$$

$d(F(t))$ ist der aktuelle Abstand zwischen AFM-Spitze und Probenoberfläche. Er ist gleich zum End-zu-End Abstand des dazwischen eingespannten Polypeptids. Dieser Abstand ist

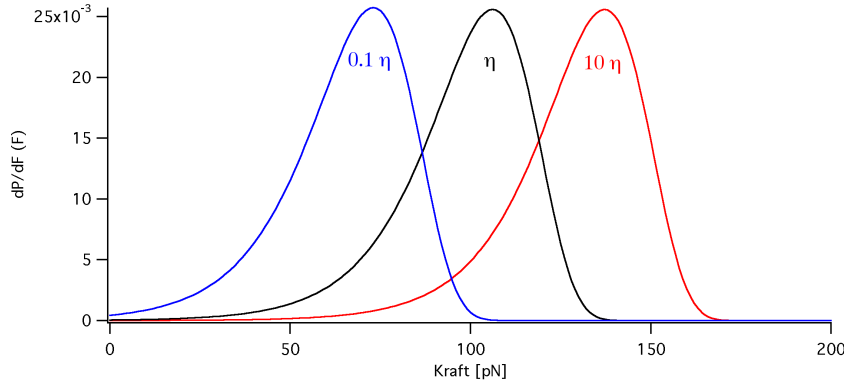


Abbildung 3.3: dP/dF als Funktion der Kraft F für drei verschieden steile Krafttrampen $F(t) = \eta t$.

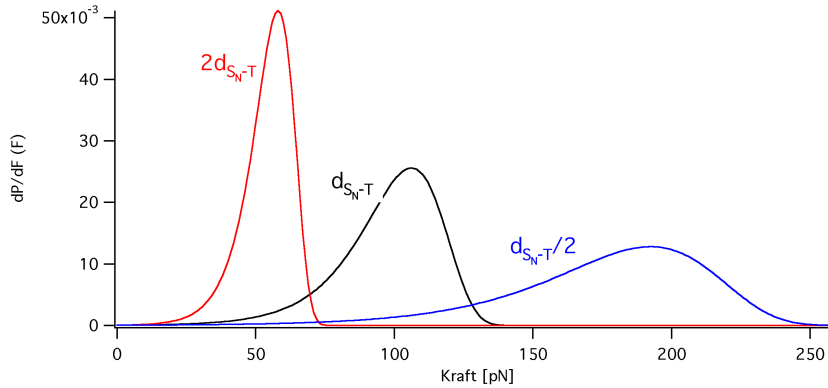


Abbildung 3.4: dP/dF als Funktion der Kraft F für drei verschiedene Abstände d_{S_n-T} zwischen nativem Zustand S_n und Übergangszustand T .

kraftabhängig, da er durch das Elastizitätsverhalten des Polypeptids bestimmt wird. Der Faktor η ist durch $\eta = k_c \cdot v_p$ gegeben, wobei v_p die Geschwindigkeit ist, mit der die Probenoberfläche bewegt wird und k_c die Federkonstante der verwendeten Blattfeder bezeichnet. Für kleine Kräfte gilt Gl. 3.6, und die Kraftladungsrate \dot{F} bleibt eine Konstante, womit die obigen Überlegungen gültig bleiben:

$$\dot{F} = \frac{k_c v_p}{1 + \frac{2}{3} \frac{k_c p L}{k_B T}} \quad (3.33)$$

Die experimentell erreichten Kräfte lagen jedoch im Gültigkeitsbereich von der in Gl. 3.8 angegebenen Interpolationsformel. In diesem Kraftbereich kann es zu deutlichen Abweichungen von einer durch eine lineare Krafttrampe hervorgerufenen Verteilung dP/dF kommen [13].

Kapitel 4

Methoden

4.1 Das Kraftspektrometer

Das Kraftspektrometer ist eine Abwandlung eines Atom-Kraftmikroskops, kurz AFM. Die Technologie wurde 1986 durch Gerd Binnig, Calvin Quate und Christoph Gerber [6] für die Abbildung von Festkörperoberflächen entwickelt. In einem AFM rastert eine sehr scharfe Spitze die zu untersuchende Oberfläche ab und folgt aufgrund verschiedener Wechselwirkungen der Topographie. Ein Kraftspektrometer hingegen befähigt den Experimentator, Untersuchungen zur Elastizität einzelner Moleküle durchzuführen [29]. Insbesondere können biologische Proben unter physiologischen Bedingungen untersucht werden.

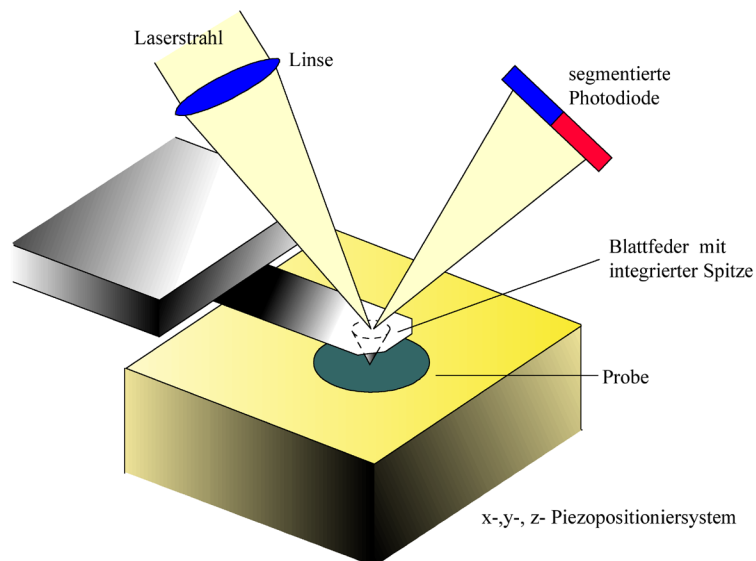


Abbildung 4.1: Schematischer Aufbau des in dieser Arbeit verwendeten Kraftspektrometers. Verbiegungen der Blattfeder führen zu verschiedenen Auslenkungen des reflektierten Laserstrahls. Dies wird durch eine segmentierte Photodiode detektiert.

Das in dieser Arbeit verwendete Kraftspektrometer ist schematisch in Abbildung 4.1

gezeigt. Ein in drei Dimensionen funktionierendes Positioniersystem auf Basis von Piezokristallen trägt die zu untersuchende Probe. Für kraftspektroskopische Messungen reicht die zur Probenfläche vertikale Dimension völlig aus, die zwei lateralen Bewegungsrichtungen sind nützlich für die Ansteuerung bestimmter Orte auf der Probenfläche. Kapazitive Abstandssensoren ermöglichen die nanometergenaue Angabe von Positionsänderungen. Für das Auslesen von Kräften wird ein Lichtzeiger auf eine Blattfeder mit bekannter Federkonstante k_c fokussiert. Der reflektierte Strahl wird auf eine segmentierte Photodiode geworfen und ermöglicht die Detektion von Verbiegungen der Blattfeder. Die Blattfeder trägt auf der der Probe zugewandten Seite eine scharfe Spitze (Krümmungsradius ca. 30 nm).

Eine kraftspektroskopische Messung besteht aus dem gleichzeitigen Detektieren von an der Blattfeder wirkenden Kräften $F(t)$ und absoluten Positionen $z(t)$ der Probenfläche während man die Probenfläche vertikal an die Spitze unter der Blattfeder annähert, eine kurze Kontaktzeit erlaubt, und dann die Probenfläche wieder entfernt. Besteht eine Wechselwirkung zwischen der Spitze und der Probenfläche, so wird die Blattfeder ausgelenkt. Hat man die Blattfeder geeicht¹, kann auf die wirkenden Kräfte geschlossen werden. Aus den so erhaltenen Datensätzen $F(t)$ und $z(t)$ kann über

$$d(t) = z(t) - \frac{F(t)}{k_c} \quad (4.1)$$

der jeweils gültige Abstand zwischen Probenfläche und Spitze der Blattfeder berechnet werden. Die Auftragung von $F(t)$ über $d(t)$ führt zu sogenannten *Kraft-Abstandskurven* $F(d)$. Hat man zwischen Oberfläche und Spitze z.B. ein Polypeptid gespannt, so gibt der Abstand d gleichzeitig die momentane Ausdehnung (bzw. den End-zu-End-Abstand) des Polypeptids an. Die Darstellungen von $F(d)$ werden daher auch als *Kraft-Ausdehnungskurven* bezeichnet.

Ein Kraftspektrometer ermöglicht auf diese Weise die Untersuchung des elastischen Verhaltens einzelner Moleküle. Die erreichbare Genauigkeit der Längenmessung liegt dabei im Nanometerbereich. Kräfte können mit einer Auflösung im Pikonewtonbereich gemessen werden.

4.2 Eigenschaften der Blattfedern

Es kamen verschiedene Typen von Blattfedern zum Einsatz. Es standen dreiecksförmige *Microlever* (Fa. Park Scientific Instruments, jetzt: Veeco Metrology LLC, Santa Barbara, USA) zur Verfügung, deren Federkonstanten im Bereich von 15 pN/nm lagen. Des Weiteren wurden *Bio-Lever* (Hersteller: Olympus, Bezug über Fa. Atomic Force, München, D) verwendet. Letztere zeichneten sich durch ein im Vergleich zu den Microlevern deutlich besseres Auflösungsvermögen in der Kraftmessung aus. Dies ist unter anderem auf die verringerte Größe der Bio-Lever zurückzuführen [42].

¹Die verwendete Methode zur Eichung der Blattfeder wird in Anhang C erläutert.

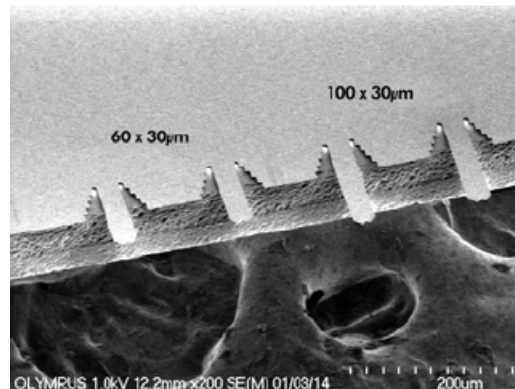


Abbildung 4.2: Elektronenmikroskopische Aufnahme der Olympus Bio-Lever Typ A und B. Graphik adaptiert aus Werbematerial Fa. Asylum Research, Santa Barbara, USA.

Zur Verfügung standen zwei verschiedene Bio-Lever Modelle aus SiN, die beide auf einem Chip geliefert wurden. Die Blattfedern mit Spitze sind werkseitig auf allen Seiten mit Gold bedampft worden. Die Bio-Lever haben Dimensionen von $30\ \mu\text{m}$ Breite und $60\ \mu\text{m}$ (Typ A) bzw. $100\ \mu\text{m}$ (Typ B) Länge. Die Dicke der Blattfedern liegt bei $0.16\ \mu\text{m}$. Die Federkonstanten k_c der Bio-Lever Typ A liegt im Bereich von $30\ \text{pN/nm}$, die der längeren Bio-Lever Typ B im Bereich von $6\ \text{pN/nm}$. Die Bio-Lever Typ A zeigen des weiteren in wässriger Lösung eine höhere Resonanzfrequenz (im Bereich von $8\ \text{kHz}$), ein Umstand, der sich für die Detektion schneller und kurzlebiger Ereignisse als nützlich erwies.

4.3 Chimären-Proteine

Im Zentrum dieser Arbeit stand die kraftspektroskopische Untersuchung von GFP.

Die Untersuchung einzeln vorliegender GFP-Domänen stünde aufgrund der Struktur des GFP vor großen Schwierigkeiten, da u.a. eine eventuell beobachtete GFP-Entfaltung von einer unspezifischen Wechselwirkung nicht zu unterscheiden wäre.

Es hat sich gezeigt, dass es für die Kraftspektroskopie sehr vorteilhaft ist, wenn die zu untersuchenden Proteine eine kettenartige Quartärstruktur besitzen. Die sequentielle Entfaltung einzelner Domänen (Kettenglieder) führt zu einem klar definierten Muster und bietet die Möglichkeit, Einzelmolekül-Kraftkurven zu identifizieren [28][34][38][45].

Eine mögliche Vorgehensweise zur Untersuchung des GFP ist daher das Einfügen einer GFP-Domäne zwischen Einzeldomänen solcher modularer Proteine. In kraftspektroskopischen Messungen an solchen Chimären-Proteinen können anhand des Entfaltungsmusters des Trägerproteins Einzelmolekül-Kraftkurven identifiziert werden. Das Trägerprotein kann des weiteren als Griff für die Anbindung an Oberfläche und Blattfederspitze dienen. Auf diese Weise ist sichergestellt, dass die am GFP wirkenden Kräfte immer am Amino- und Carboxyterminus des Moleküls angreifen. Wird die Stabilität der Trägerdomänen passend gewählt, so kann erreicht werden, dass die Entfaltung des GFP erst dann geschieht, wenn Wechselwirkungen mit der Probenoberfläche keine Rolle mehr spielen.

Dieser Weg ist in der vorliegenden Arbeit beschrieben worden. Dazu ist in das bereits intensiv studierte Ig8-Protein [34][28] (8 Immunoglobulin-ähnliche Domänen aus menschlichem Herzmuskeltitin, Domänen Ig27 bis Ig34) mit gentechnischen Methoden² eine *Cycle3*-GFP-Domäne zwischen den Titin-Domänen Ig30 und Ig31 eingefügt worden.

Ein weiteres modulares Protein, dessen Entfaltungsmuster durch Ingo Schwaiger am Lehrstuhl für Angewandte Physik aufgeklärt wurde, ist durch die Domänen 1 bis 5 des *Dictyostelium discoideum Filamins*, DdFLN(1-5), gegeben [16][38]. Zwischen die Domänen 3 und 4 des DdFLN(1-5) ist ebenfalls mit gentechnischen Methoden eine *Cycle3*-GFP-Domäne eingefügt worden (vgl. Anhang D für die Vorgehensweise). Dieses Chimären-Protein wird im folgenden als DdFLN(1-5)-GFP bezeichnet.

Einzelne Titin-Domänen entfalten typischerweise bei Zugkräften im Bereich von 150 bis 250 pN, während DdFLN-Domänen bei Belastungen von 50 bis 100 pN entfalten. Dieser Unterschied in der Stabilität hat sich für die Untersuchung des GFP als sehr vorteilhaft erwiesen.

4.4 Durchführung der Messungen

Die zu untersuchenden Proteine wurden mit molekularbiologischen Methoden (siehe Anhang D) hergestellt und bei 4°C in einer 70mM-Imidazolösung gelagert ($pH = 8.0$). Für die Messungen wurden sie in PBS (*Phosphate Buffered Saline*) - Pufferlösung mit einem pH -Wert von 7.4 überführt. Eine kleine Menge der Probenlösung (etwa 10 μl mit einer Konzentration $c \simeq 0.5\mu\text{g}/\mu\text{l}$) wurde dann auf eine frisch bedampfte bzw. frisch geätzte Goldoberfläche der Dicke ~ 50 nm aufgetragen und etwa fünf Minuten bei Zimmertemperatur zur Adsorption auf die Oberfläche inkubiert. Der Mechanismus der Anbindung der Moleküle an die Goldoberfläche ist unklar und wird als unspezifisch bezeichnet. Die für ein Experiment verwendeten, ebenfalls mit Gold bedampften Blattfedern waren jeweils noch unbenutzt.

Für die Messung einer einzelnen Kraftkurve wurde die Blattfeder an die Oberfläche angenähert und mit einer Kraft von bis zu 1000 pN zur Anbindung der Moleküle an die Spitze der Blattfeder eingedrückt. Die Anbindung geschah wiederum unspezifisch. Anschließend wurde die Blattfeder mit der Geschwindigkeit v_p wieder von der Oberfläche entfernt, wobei mit der Samplingrate f_s gleichzeitig die Position der Oberfläche $z(t)$ und die wirkenden Kräfte $F(t)$ detektiert wurden.

In einem Experiment wurden bis zu 2000 einzelne Kraftkurven gemessen.

Die im Kapitel 5 und in den Abschnitten 6.5 und 6.6 vorgestellten Daten wurden zum Teil mit Microlevern und zum Teil mit Bio-Levern Typ A und B gemessen. Die eingestellten Probengeschwindigkeiten v_p lagen dabei im Bereich von 50-1000 nm/sec.

Die Messungen, deren Ergebnisse in den Kapiteln 6 und 7 vorgestellt und analysiert werden, erforderten es, durch den Experimentator beeinflussbare Parameter möglichst konstant zu lassen. Es sind ausschließlich Bio-Lever verwendet worden. Die folgende Tabelle

²Das Ig8-GFP kodierende Gen wurde in Vorarbeiten am Lehrstuhl für Angewandte Physik durch Angelika Kardinal erzeugt.

listet die für die einzelnen Abschnitte gültigen Umgebungsbedingungen wie Federkonstante k_c , Resonanzfrequenz ω_0 , Probengeschwindigkeit v_p , sowie Samplingrate f_s und Bandbreite f_c der Signalfilterung auf.

Abschnitt	k_c [pNnm ⁻¹]	ω_0 [kHz]	v_p [nms ⁻¹]	f_s [kHz]	f_c [kHz]
6.3	30	8	300	20	10
6.7	30	8	300	20	10
7.3	30	8	300	20	10
7.5	30	8	300	20	10
7.6	30	8	300	20	10
7.7	6	1.6	90	20	10

Tabelle 4.1: Experimentelle Parameter in den genannten Abschnitten der vorliegenden Arbeit.

4.5 Bestimmung von Entfaltungskräften

Für diese Arbeit im besonderen Maße interessant waren die Kräfte, bei denen es zu Übergängen zwischen verschiedenen Konformationen der untersuchten Moleküle kam³. Mit einem solchen Übergang (Entfaltung) ist in den hier untersuchten Fällen eine Zunahme der Konturlänge ΔL der Moleküle verbunden. Diese Verlängerung geschieht sehr schnell auf einer Zeitskala ($< 10^{-6}$ s), die experimentell nicht aufgelöst werden kann. Nach einem solchen Übergang relaxiert die ausgelenkte Blattfeder ungestört bis auf ein Kraftniveau, das durch das Kraft-Ausdehnungsverhalten des um ΔL verlängerten Moleküls vorgegeben wird. Die Relaxation der Blattfeder ist in der Kraft-Abstandsdarstellung in guter Näherung durch eine Gerade mit Steigung $-k_c$ gegeben. Abb. 4.3 zeigt beispielhaft eine Kraft-Ausdehnungskurve an einem Polypeptid, das mehrere Konformationsübergänge vollzogen hat. Die Kraftkurve zeigt (wie alle weiteren gemessenen Kraftkurven) jedoch auch ein Krafrauschen, das im wesentlichen auf thermische Anregung zurückzuführen ist. Dieses Rauschen erschwert die genaue Bestimmung von Entfaltungskräften.

Um die Einflüsse des Rauschens bei der Bestimmung von Kräften zu minimieren, ist das WLC-Modell verwendet worden. Passt man über die WLC-Interpolationsformel in Gl. 3.8 berechnete Kraftkurven an Anstiege vor einer Entfaltung in gemessenen Kraftkurven an, so können über die Schnittpunkte der WLC-Kurven mit der Phase der Blattfeder-Relaxation Entfaltungskräfte genau bestimmt werden (vgl. Abb. 4.3). In der Praxis führt diese Vorgehensweise zu einer Ungenauigkeit in der Kraftmessung von ± 3 pN.

Der Nullpunkt der Kraftachse lässt sich über Anpassen einer konstanten Funktion in den Bereich des Rauschens nach Abriss eines Moleküls in jeder Kraftkurve bestimmen (vgl. Abb. 4.3, Bereich Ausdehnung > 240 nm). Dies führt auch zu einer Ungenauigkeit von ± 3 pN.

³Diese Kräfte werden im folgenden auch als Entfaltungskräfte bezeichnet.

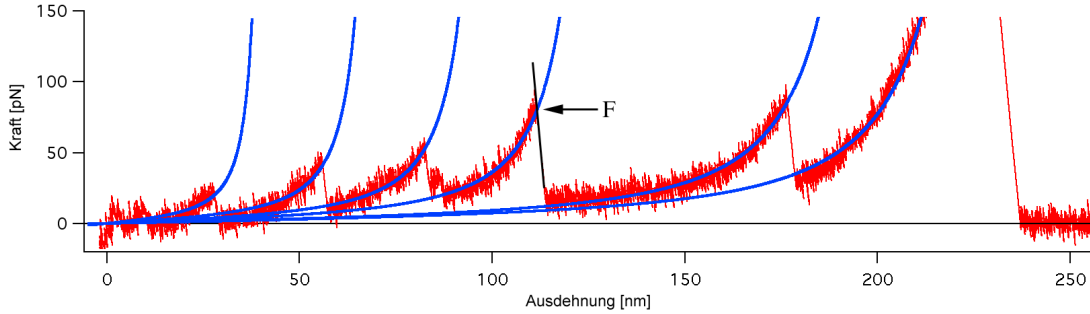


Abbildung 4.3: An eine gemessene Kraftkurve wurde das WLC-Modell angepasst. $F = 0$ kann über Anpassung einer Null-Linie in die Phase nach dem Abriss eines Moleküls (Bereich Ausdehnung > 240 nm) bestimmt werden. Die Kraft F , bei der es zu einem Konformationsübergang kam, ergibt sich durch den Schnittpunkt der WLC-Kurve mit der Blattfeder-Relaxation.

Zusammenfassend führt die beschriebene Vorgehensweise zur Bestimmung von Entfaltungskräften zu einer Ungenauigkeit von ± 6 pN. Hinzu kommt die relative Unsicherheit in der Bestimmung der Federkonstante, es sind Kräfte nur bis auf $\pm 10\%$ angebbar.

4.6 Monte-Carlo Simulationen

Die Ergebnisse aus den in dieser Arbeit durchgeführten Messungen wie Kraftverteilungen und Lebensdauererwartungen können mit dem linearen Zwei-Zustandsmodell aus Abschnitt 3.6 interpretiert werden. Eine Möglichkeit, die genauen experimentellen Bedingungen in eine Analyse der Verteilungen einzubeziehen, ist die Durchführung von Monte-Carlo Simulationen.

In einer solchen Simulation wird ein virtuelles Polypeptid der Konturlänge L mit der Geschwindigkeit v_p beginnend mit der Ausdehnung $d = 0$ gestreckt. Dies führt zu einer zusätzlichen Ausdehnung $\Delta d = v_p \cdot \Delta t$ in jedem Zeitschritt Δt . Nach jedem Zeitschritt wird die momentan wirkende Kraft F für die momentane Ausdehnung d über die WLC-Interpolationsformel

$$F(d) = \frac{k_B T}{p} \left(\frac{d}{L} + \frac{1}{4(1 - d/L)^2} - \frac{1}{4} \right) \quad (4.2)$$

berechnet. Zur Simulation einer in dem simulierten Polypeptid enthaltenen Struktur S_a , welche in eine andere Konformation mit einem zusätzlichen Beitrag ΔL zur Konturlänge übergehen kann, wird nach jedem Zeitschritt eine Übergangsrate $k_{off}(F)$ für die aktuell wirkende Kraft berechnet:

$$k_{off}(F) = k_{off}^0 \cdot \exp\left(\frac{F \Delta L}{k_B T}\right) \quad (4.3)$$

Die Wahrscheinlichkeit ΔP_u für einen Konformationsübergang in jedem Zeitschritt Δt ist dann durch

$$\Delta P_u = k_{off}(F) \cdot \Delta t \quad (4.4)$$

gegeben. In den durchgeführten Simulationen wurde für jedes Zeitintervall die Wahrscheinlichkeit ΔP_u für einen Übergang berechnet und mit einer durch einen Zufallsgenerator erzeugten Zahl γ im Bereich von $[0,1]$ verglichen. Lag ΔP_u über γ , so wurde für einen Übergang entschieden und das Polypeptid schlagartig um ΔL verlängert. Die Eigenschaften einer Blattfeder der Federkonstante k_c sind durch Einführung einer Relaxationsphase nach einem Übergang mit dem Kraft-Abstandsverhalten

$$F_{relax} = -k_c \cdot d \quad (4.5)$$

einbezogen worden (siehe Abschnitt 7.4). Wenn F_{relax} bis auf die durch Gl. 4.2 für ein Polypeptid der Länge $L + \Delta L$ vorhergesagte Kraft (im folgenden als $F_{intersect}$ bezeichnet) abgesunken war, ist das Kraft-Ausdehnungsverhalten wieder mit Gl. 4.2 berechnet worden. Es können nun nachfolgende Übergänge nach dem selben Schema in die Simulation einbezogen werden. Im allgemeinen müssten in jedem Zeitschritt auch Wahrscheinlichkeiten für Rückreaktionen (über $k_{on}(F)$ -Raten) berechnet werden. In der experimentellen Situation, die hier simuliert wurde, sind diese verschwindend gering und können daher vernachlässigt werden [33].

Kapitel 5

Ig8-GFP und DdFLN(1-5)-GFP

5.1 Einleitung

In diesem Kapitel werden die Kraftkurven vorgestellt, die an einzelnen Ig8-GFP und DdFLN(1-5)-GFP Molekülen gemessen wurden. Die Kraftkurven zeigen ein neues Entfaltungsereignis. Es wird gezeigt, dass dieses Ereignis durch die mechanisch induzierte Entfaltung des GFP verursacht wird.

5.2 Ig8-GFP

An der Ig8-GFP-Chimäre wurden kraftspektroskopische Messungen durchgeführt. Einzelheiten über Versuchsdurchführung und Anbindung der Proteine an Oberflächen und Blattfeder-Spitzen sind in Abschnitt 4.4 angegeben. Abbildung 5.1 zeigt schematisch die Struktur dieser Protein-Chimäre.

In Abb. 5.2 ist eine Auswahl verschiedener Ig8-GFP-Kraftkurven zusammengestellt. Ab einer Ausdehnung von etwa 80 nm ist in allen Kraftkurven ein sich wiederholendes, Sägezähnen ähnelndes Muster zu erkennen. Es wird durch die sequentielle Entfaltung einzelner Ig-Domänen verursacht [34]. Die Zahl dieser Entfaltungsereignisse schwankt von Kraftkurve zu Kraftkurve. Dies hängt mit der unspezifischen Anbindung der Proteine an Blattfederspitze und Oberfläche zusammen. Die gedehnten Moleküle können z.B. bereits vor der völligen Entfaltung aller im Molekül enthaltenen Domänen von der Spitze oder der Probenoberfläche abreißen. Es kann aber auch durchaus sein, dass z.B. die Domänen Ig27 und Ig28 an der Spitze kleben, während Ig34 an der Oberfläche klebt. Das führt bei Entfernung der Spitze von der Oberfläche dazu, dass nur die Domänen Ig29-30, GFP und Ig30-33 mit Kraft belastet werden. Deshalb sollte man bei einer solchen Anbindungssituation nur fünf Ig-Entfaltungen beobachten, gefolgt vom Abriss des Moleküls. Dies ist z.B. bei der Kurve Σ der Fall. In dieser Situation wurde auch die im Molekül enthaltene GFP-Domäne mit Kraft belastet. Wenn die mechanische Stabilität des GFP geringer als die Stabilität der Anbindung des ganzen Moleküls an Spitze und Oberfläche ist, so sollte die Entfaltung des GFP in den Kraftkurven beobachtbar sein.

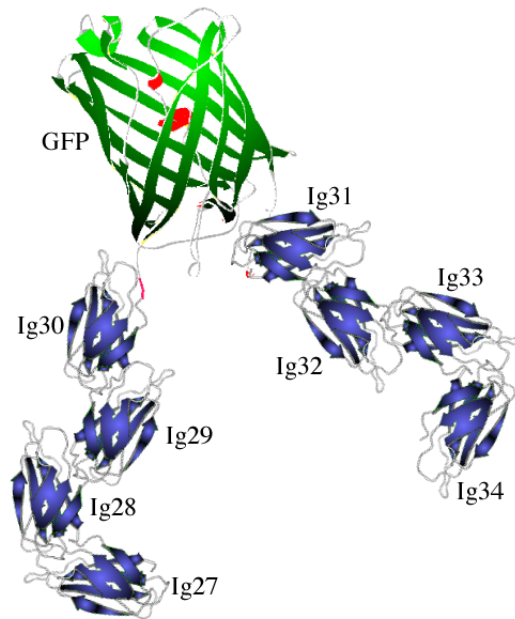


Abbildung 5.1: Schematische Darstellung der Struktur der Ig8-GFP-Chimäre. Blau dargestellt: Ig-Domänen 27-30 und Ig-Domänen 31-34. Grün dargestellt: GFP.

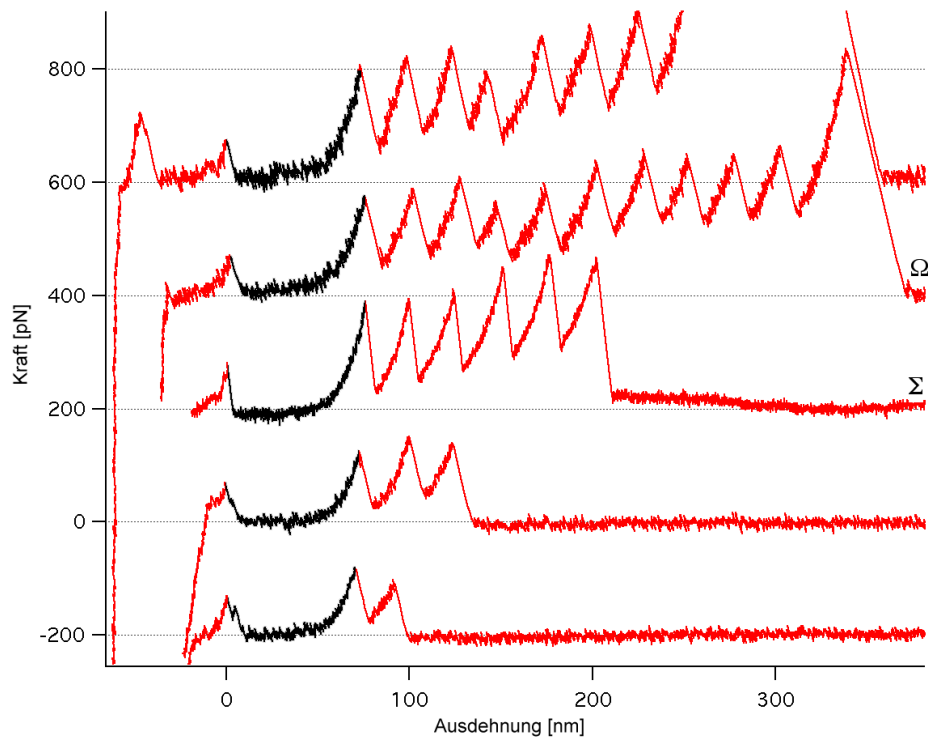


Abbildung 5.2: Auswahl verschiedener Kraftkurven, die an der Ig8-GFP-Chimäre gemessen wurden. Erläuterung im Text.

Die mit Ω gekennzeichnete Kraftkurve in Abb. 5.2 zeigt zehn Ig-Entfaltungseignisse. Dies ist mit den Eigenschaften eines einzelnen Ig8-GFP Moleküls nicht erklärbar. Ig-Domänen besitzen jedoch ins Lösungsmittel zeigende ungepaarte Cystein-Reste, über die einzelne Ig8-GFP Moleküle dimerisieren können. Mit großer Wahrscheinlichkeit zeigt Ω das Entfaltungsmuster eines Ig8-GFP-Dimers. Derartige Ereignisse sind schon beobachtet worden und wie erläutert interpretiert worden [32].

Allen in Abb. 5.2 abgebildeten Kraftkurven ist ein weiteres Muster gemeinsam (schwarz eingefärbt in Abb. 5.2). Es handelt sich hierbei um ein bisher unbekanntes Entfaltungseignis, das in Kraftkurven am reinen Ig8-Protein nicht auftritt¹ [34][32].

Passt man das WLC-Modell an die Kraftkurven an, so lässt sich die mit den Entfaltungen verbundene Zunahme der Konturlänge messen (vgl. Abb. 5.3). Die Ig-Entfaltungen

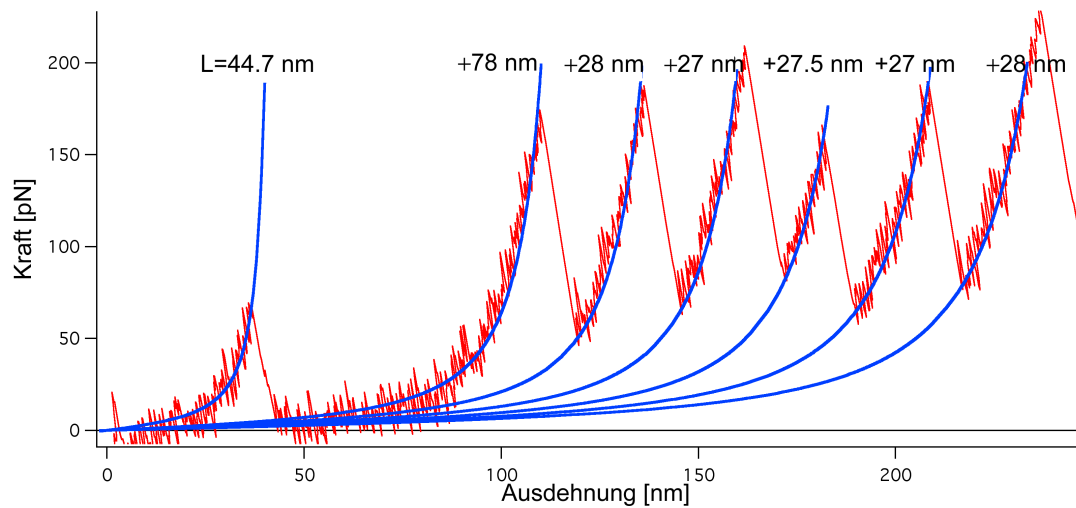


Abbildung 5.3: Anpassung von berechneten WLC-Kraftkurven (Persistenzlänge $p=0.5$ nm) an eine gemessene Ig8-GFP Kraftkurve. Erläuterung im Text.

führen zu Konturlängenzunahmen im Bereich von 28 nm. Das angesprochene Entfaltungseignis bei einer Ausdehnung von etwa 35 nm verursacht im abgebildeten Fall eine Konturlängenzunahme von 78 nm. Das unbekanntes Entfaltungseignis führt demnach zu einer um den Faktor 2.8 größeren Zunahme der Konturlänge als die Entfaltung einzelner Ig-Domänen.

Die Ig-Domänen haben eine Größe von 89 Aminosäuren, die GFP-Domäne hat eine Größe von 238 Aminosäuren. Demnach würde man nach der Entfaltung einer GFP-Domäne eine um den Faktor 2.7 größere Konturlängenzunahme als nach der Entfaltung einer Ig-Domäne erwarten. Das ist in sehr guter Übereinstimmung mit dem gemessenen Wert.

Die erste Analyse legt somit nahe, dass es sich bei dem unbekanntes Entfaltungseignis um die Entfaltung des Grün Fluoreszierenden Proteins handelt.

An der Ig8-GFP-Chimäre sind 17 verschiedene Experimente durchgeführt worden. Aus diesen Experimenten sind 21 Kraftkurven identifiziert worden, die die Entfaltung des GFP

¹Die in Anhang B vorgestellte Mustererkennung unterstützt diese Aussage auf statistischer Basis.

wie in Abb. 5.2 klar zeigen. Die GFP-Entfaltung tritt immer vor der Entfaltung von Ig-Domänen bei Kräften um 90 pN und häufig bei sehr geringen Abständen (< 30 nm) zur Oberfläche ein. Das GFP erscheint somit weniger stabil als einzelne Ig-Domänen. Dies ist unter anderem die Ursache für die geringe Zahl an Kraftkurven, die die GFP-Entfaltung zeigen. Bei geringen Abständen zur Oberfläche (< 30 nm) treten in den Kraftkurven oft noch unspezifische Wechselwirkungen auf, die ein eventuell vorhandenes GFP-Entfaltungsereignis überlagern und solche Kraftkurven für die Auswertung nicht verwendbar machen.

5.3 DdFLN(1-5)-GFP

Das GFP in ein weiteres modulares Trägerprotein einzufügen, bot die Gelegenheit, einerseits zu bestätigen, dass es sich bei dem beobachteten Entfaltungsereignis tatsächlich um die Entfaltung des GFP handelt, als auch die Eigenschaften des Trägerproteins so zu wählen, dass die Entfaltung der GFP-Domäne bei größeren Abständen (> 30 nm) zur Oberfläche geschieht. Ingo Schwaiger hat die Domänen 1 bis 5 des *Dictyostelium discoideum Filamins*, kurz DdFLN(1-5), kraftspektroskopisch untersucht. Nach seinen Erkenntnissen entfaltet Domäne 4 des DdFLN(1-5) bei Kräften im Bereich 50 bis 80 pN und zeigt einen mechanisch stabilen Zwischenzustand [16][38]. Die anderen Domänen zeigen eine leicht höhere Stabilität (Entfaltungskräfte bei 60-100 pN) [37]. In Kraftkurven an einer DdFLN(1-5)-GFP-Chimäre würde man daher öfter zunächst die Entfaltung von Domäne 4 des DdFLN(1-5) erwarten, bevor es zur Entfaltung des GFP kommt. DdFLN-Domänen haben eine Größe von 100 Aminosäuren. Deshalb wäre schon durch die Entfaltung einer einzelnen DdFLN-Domäne vor der Entfaltung des GFP ein Abstand von > 30 nm zur Probenoberfläche gesichert.

Das DdFLN(1-5) erschien somit als geeignetes Trägerprotein, um die kraftspektroskopische Untersuchung des GFP fortzuführen. Es ist daher zwischen Domäne 3 und 4 des DdFLN(1-5) mit gentechnischen Methoden eine Domäne des *Cycle3*-GFP eingefügt worden (siehe Anhang D). Die Struktur dieses Chimären-Proteins ist schematisch in Abb. 5.4 dargestellt. An der DdFLN(1-5)-GFP-Chimäre sind kraftspektroskopische Messungen durchgeführt worden. Für Einzelheiten der Versuchsdurchführung und Anbindung der Proteine an Oberflächen und Blattfeder-Spitzen sei auf Abschnitt 4.4 verwiesen.

Abb. 5.5 zeigt eine Auswahl von Kraftkurven, wie sie an einzelnen DdFLN(1-5)-GFP-Molekülen gemessen wurden. Sie zeigen Entfaltungen einzelner DdFLN-Domänen, wie sie bereits in [38] beobachtet wurden. Zusätzlich dazu tritt ein weiteres Entfaltungsereignis auf (schwarz eingefärbt), das in Messungen am reinen DdFLN(1-5)-Protein nicht zu beobachten war [37].

Das in den DdFLN(1-5)-GFP Kraftkurven enthaltene komplexe Entfaltungsmuster wird anhand von Abb. 5.6 näher erläutert. Zur Messung der mit einer Entfaltung verbundenen Zunahme der Konturlänge ist wieder das WLC-Modell an die Kraftkurven angepasst worden. Abb. 5.7 zeigt beispielhaft eine solche Anpassung.

Das neu hinzugekommene Entfaltungsereignis führt demnach zu einer um den Faktor

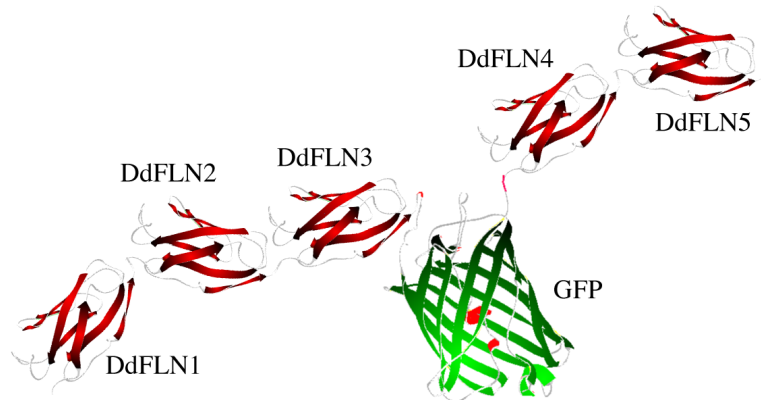


Abbildung 5.4: Schematische Darstellung der Struktur der DdFLN(1-5)-GFP-Chimäre. Rot dargestellt: DdFLN-Domänen 1-3 und DdFLN 4-5. Grün dargestellt: GFP.

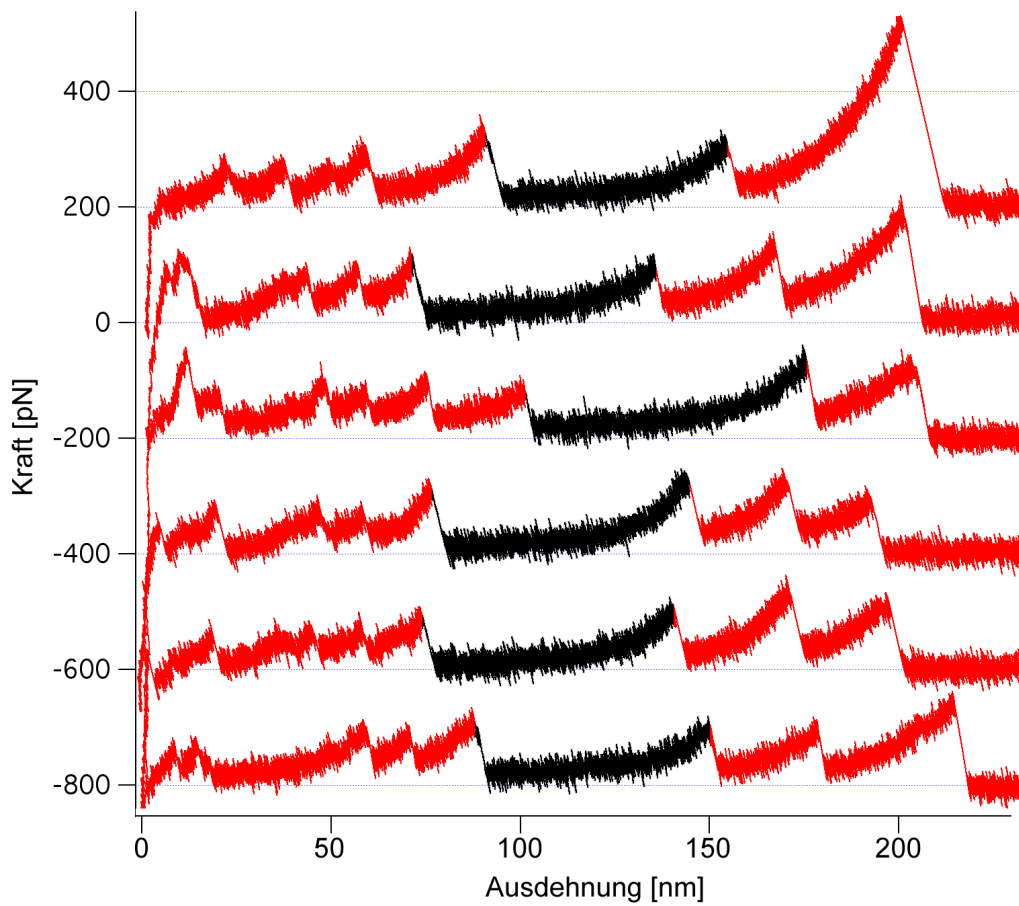


Abbildung 5.5: Auswahl an Kraftkurven, die an der DdFLN(1-5)-GFP-Chimäre gemessen wurden. Schwarz eingefärbt: Neu hinzu gekommenes Entfaltungsmuster. Die Kraftkurven zeigen die volle, für den Olympus Bio-Lever Typ A typische Rauschamplitude.

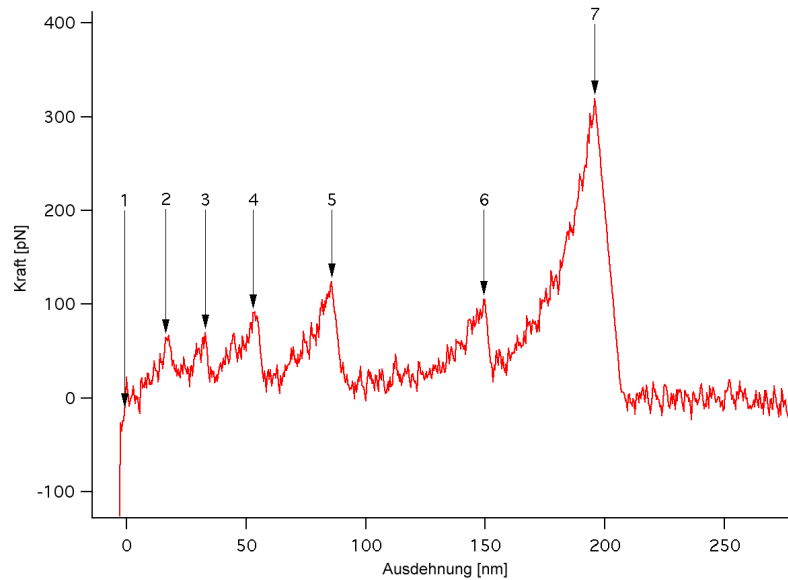


Abbildung 5.6: DdFLN(1-5)-GFP-Kraftkurve (gefiltert). (1) Ablösung der Blattfeder-Spitze von der Oberfläche. (2) Entfaltung DdFLN Domäne 4. (3) Entfaltung des stabilen Zwischenzustandes von DdFLN Domäne 4. (4) Entfaltung einer weiteren DdFLN Domäne. (5) Neu hinzugekommenes Entfaltungsmuster. (6) Entfaltung einer weiteren DdFLN-Domäne. (7) Abriss des DdFLN(1-5)-GFP-Moleküls von der Spitze oder Oberfläche.

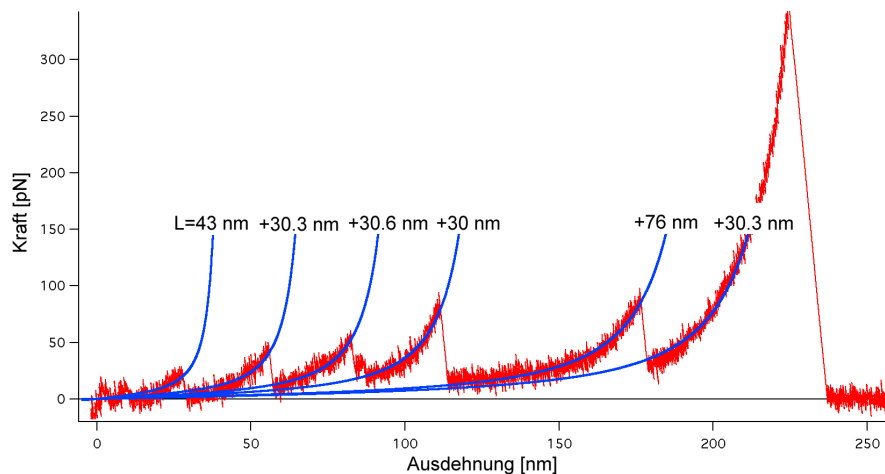


Abbildung 5.7: Anpassung von berechneten WLC-Kurven (Persistenzlänge $p=0.5$ nm) an eine DdFLN(1-5)-GFP Kraftkurve. DdFLN-Entfaltungen führen im abgebildeten Fall zu einem Konturlängenwachstum von etwa 30 nm. Das neu hinzugekommene Entfaltungsereignis verlängert das Molekül um 76 nm.

2.5 größeren Zunahme der Konturlänge als die Entfaltung der DdFLN-Domänen. DdFLN-Domänen haben eine Größe von 100 Aminosäuren, GFP umfasst 238 Aminosäuren. Man würde daher eine um den Faktor 2.4 größere Zunahme der Konturlänge nach einer GFP-Entfaltung verglichen mit der Entfaltung einer DdFLN-Domäne erwarten. Dies stimmt sehr gut mit dem gemessenen Wert überein.

Diese Analyse legt wieder nahe, dass es sich bei dem beobachteten Entfaltungsereignis um die Entfaltung der im DdFLN(1-5)-GFP-Molekül enthaltenen GFP-Domäne handelt.

Die Kraftkurven an der DdFLN(1-5)-GFP-Chimäre zeigen genau das erwartete Verhalten. Vor der Entfaltung des GFP kommt es in den meisten Fällen zur Entfaltung von zwei DdFLN-Domänen, so dass ein genügender Abstand (>30 nm) zur Oberfläche gesichert ist, um das Ereignis frei von unspezifischen Wechselwirkungen beobachten zu können. Die meisten Messungen, die für die Untersuchungen der nachfolgenden Kapitel durchgeführt worden sind, sind daher an der DdFLN(1-5)-GFP-Chimäre vorgenommen worden.

5.4 Diskussion

In den untersuchten Chimären-Proteinen Ig8-GFP und DdFLN(1-5)-GFP wurde in den Kraftkurven ein neues Entfaltungsereignis beobachtet. Bringt man DdFLN(1-5)-GFP Kraftkurven mit Ig8-GFP Kraftkurven im Bereich der GFP-Entfaltung zur Überlagerung, so wird die hohe Übereinstimmung zwischen diesen Entfaltungs-Mustern deutlich. Abb. 5.8 zeigt ein Beispiel einer solchen Überlagerung.

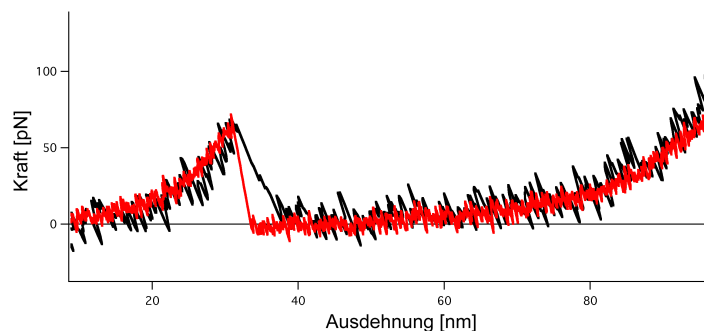


Abbildung 5.8: Schwarz: Ausschnitt aus einer Ig8-GFP-Kraftkurve. Die GFP-Entfaltung findet bei einer Ausdehnung von 30 nm statt. Rot: Ausschnitt aus einer DdFLN(1-5)-GFP-Kraftkurve. Die Kurve wurde auf der Ausdehnungsachse um ca. -90 nm verschoben, bis die GFP-Entfaltung mit der GFP-Entfaltung in der Ig8-GFP-Kurve zur Deckung kam. Die Abweichung in der Blattfeder-Relaxation rührt daher, dass die DdFLN(1-5)-GFP-Kurve mit einer härteren Blattfeder als die Ig8-GFP-Kurve gemessen wurde.

Messdaten an den reinen Ig8 und DdFLN(1-5) Proteinen zeigten dieses Muster nicht. Die einzige Gemeinsamkeit der beiden Chimären-Proteine Ig8-GFP und DdFLN(1-5)-GFP, die eine solche Übereinstimmung zwischen diesen Entfaltungsmustern verursachen könnte,

ist die jeweils enthaltene GFP-Domäne. Nach einer ersten Analyse liegt die gemessene Zunahme der Konturlänge nach diesem Entfaltungsereignis zudem in guter Übereinstimmung mit der, wie sie nach einer Entfaltung des GFP zu erwarten wäre.

Demnach wurde anhand der Messungen an der Ig8-GFP- und der DdFLN(1-5)-GFP-Chimäre gezeigt, dass das Grün Fluoreszierende Protein mechanisch entfaltet werden kann. Aus den Messungen geht hervor, dass das GFP eine geringere Stabilität als einzelne Ig-Domänen und eine vergleichbare Stabilität zu der von einzelnen DdFLN-Domänen aufweist.

Es stellt sich die Frage nach der genauen Charakteristik der mechanischen Entfaltung des GFP. Welche Kräfte sind tatsächlich für die Entfaltung nötig? Welche Bedeutung hat die mit der Entfaltung verbundene Konturlängenzunahme? Lassen sich Aussagen über natürliche Übergangsraten treffen? Treten Zwischenzustände auf dem Weg der mechanischen Entfaltung auf? Lassen sich Informationen über die Energielandschaft des GFP gewinnen?

Zur Klärung dieser Fragen ist die DdFLN(1-5)-GFP-Chimäre intensiv mit dem Kraftspektrometer untersucht worden. Die Analyse der dabei erhaltenen Ergebnisse ist Thema der folgenden Kapitel.

Kapitel 6

Natives GFP und GFP_{11-230}

6.1 Einleitung

In diesem Kapitel wird die mit der Entfaltung des GFP verbundene Zunahme der Konturlänge untersucht und interpretiert. Dabei wird sich herausstellen, dass es sich bei dem beobachteten Entfaltungsereignis nicht um die Entfaltung des nativen GFP handelt, sondern um die Entfaltung eines Zwischenzustandes. Die für die Entfaltung des Zwischenzustandes nötigen Kräfte werden analysiert. Ihre Verteilung wird mit dem in Abschnitt 3.6ff. vorgestellten linearen Zwei-Zustandsmodell erklärt. Der Übergang des nativen GFP in den Zwischenzustand konnte experimentell beobachtet werden und wird in diesem Kapitel ebenfalls untersucht und interpretiert.

6.2 Eichung von WLC-Konturlängen

Über das WLC-Modell können Konturlängen L gemessen werden. Dieser Parameter muss jedoch geeicht werden, da dessen Bedeutung, ausgedrückt in realen Größen wie einer Anzahl von Aminosäuren, in Bezug auf das verwendete Instrument unbekannt ist. Die Eichung kann über bekannte Entfaltungsereignisse geschehen. Die Konturlänge einer Proteindomäne im nativen Zustand ist durch den Abstand d_{C-N} zwischen Carboxy- und Aminoterminus der Domäne gegeben. Entfaltet die Domäne in einem Zwei-Zustandsprozess, so ist die Konturlänge L des Moleküls nach einer Entfaltung durch $L = n \cdot d_{AS}$ gegeben, wobei n die Anzahl der Aminosäuren des Moleküls ist und d_{AS} der Beitrag einer einzelnen Aminosäure zur Konturlänge ist. Die Zunahme der Konturlänge ΔL bei dieser Entfaltung ist dann durch

$$\Delta L = n \cdot d_{AS} - d_{C-N} \quad (6.1)$$

gegeben. DdFLN-Domänen entfalten (bis auf Domäne 4) in einem Zwei-Zustandsprozess [38]. Der Mittelwert der Zunahme der Konturlänge nach der Entfaltung einer DdFLN-Domäne wurde in dieser Arbeit mit dem WLC-Modell zu $\Delta L_{DdFLN} = 32.5 \pm 0.3$ nm bestimmt (Persistenzlänge $p = 0.5$ nm). DdFLN-Domänen haben eine Größe von 100

Aminosäuren. Der Abstand zwischen Amino- und Carboxy-Terminus einer nativen DdFLN-Domäne beträgt $d_{C-N} = 4 \pm 0.1$ nm [4][37]. Damit nimmt der Beitrag zur Konturlänge einer einzelnen Aminosäure, in der Eichung des verwendeten Instrumentes und bei Verwendung des WLC-Modells, den Wert

$$d_{AS} = 0.36 \pm 0.01 \text{ nm} \quad (6.2)$$

an. Mit diesem Wert werden alle im folgenden über das WLC-Modell bestimmten Konturlängen geeicht.

6.3 Längenzunahme nach GFP-Entfaltung

Zur Messung der Zunahme der Konturlänge ΔL_{GFP} nach einer GFP-Entfaltung wurde das WLC-Modell wie in Abb. 5.7 an Einzelmolekül-Kraftkurven angepasst, die an der DdFLN(1-5)-GFP-Chimäre gemessen wurden. Für die Persistenzlänge wurde ein Wert von $p = 0.5$ nm gewählt.

Die Verteilung der so gemessenen Konturlängenzunahmen ΔL_{GFP} ist in Abb. 6.1 angegeben. In die Statistik gingen 87 GFP-Entfaltungen ein. Der Mittelwert auftretender Längenzunahmen beträgt

$$\langle \Delta L_{GFP} \rangle = 76.6 \pm 0.3 \text{ nm}. \quad (6.3)$$

Passt man eine Gaußsche Glockenkurve an die Verteilung in Abb. 6.1 an, so liegt deren Zentrum bei 76.3 nm (nicht abgebildet). Die Breite σ der Glockenkurve liegt bei 2.9 nm. Die erhaltene Längenverteilung ist nahezu gleichverteilt, die Genauigkeit der Konturlängenmessung kann daher durch den Fehler des Mittelwertes (± 0.3 nm) angegeben werden.

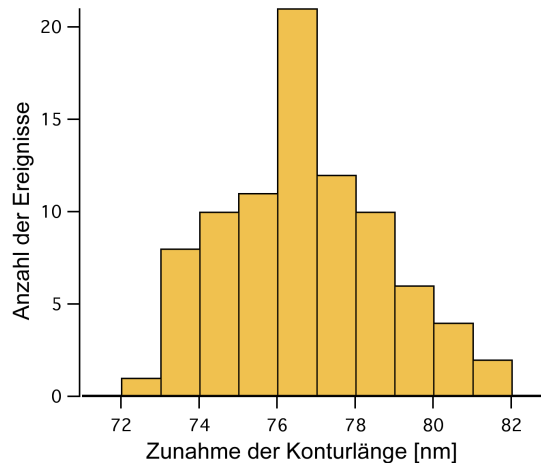


Abbildung 6.1: Verteilung der über Anpassung des WLC-Modells and Kraftkurven gewonnene Zunahme der Konturlänge ΔL_{GFP} nach der GFP-Entfaltung in DdFLN(1-5)-GFP-Kraftkurven.

Die Konturlängenzunahmen ΔL_{GFP} nach Entfaltung des GFP in Ig8-GFP-Kraftkurven sind ebenfalls über Anpassung des WLC-Modells bestimmt worden. Der Mittelwert der

Zunahmen der Konturlänge ΔL_{GFP} nach der GFP-Entfaltung in Ig8-GFP-Kraftkurven beträgt 76.9 ± 1.3 nm und ist damit in sehr guter Übereinstimmung mit dem Wert, der aus den Messungen an der DdFLN(1-5)-GFP-Chimäre bestimmt wurde.

6.4 Interpretation der Längenzunahme

Die Struktur des in dieser Arbeit untersuchten *Cycle3*-GFP ist mit Röntgen-Streuversuchen durch Battistutta *et al.* aufgelöst worden (pdb-Datei 1B9C [4]). Demnach nehmen von den insgesamt 238 Aminosäuren, aus denen das GFP besteht, nur die Aminosäuren 4 bis 230 eine feste, auflösbare Struktur an. Diese Struktur wird im folgenden als *native Struktur* S_n bzw. auch als GFP_{4-230} bezeichnet. Der entfaltete Zustand des GFP wird im folgenden als S_u bezeichnet. Die N-terminalen Aminosäuren 1 bis 3 und die C-terminalen Aminosäuren 231 bis 238 werden, da sie nicht Teil der festen Struktur des GFP sind, als *random coil* vorliegen. Die Konturlänge eines GFP-Moleküls vor der Entfaltung wäre demnach durch

$$L(1) = 11 \cdot d_{AS} + d_{4,230} \quad (6.4)$$

gegeben. Dabei ist $d_{4,230}$ der Abstand zwischen den C_α -Atomen der Aminosäuren 4 und 230 des gefaltete GFP_{4-230} . Er hat einen Wert von $d_{4,230} = 2.3 \pm 0.2$ nm (1B9C [4]). Nach der Entfaltung von GFP_{4-230} wäre die Konturlänge des GFP durch

$$L(2) = 11 \cdot d_{AS} + 227 \cdot d_{AS} \quad (6.5)$$

gegeben. Man würde daher, unter Verwendung von $d_{AS} = 0.36$ nm, eine Zunahme der Konturlänge von

$$\Delta L_{GFP} = L(2) - L(1) = 79.4 \pm 0.2 \quad (6.6)$$

erwarten. Die gemessene Konturlängenzunahme ist um 2.8 nm geringer. Diese Diskrepanz kann nicht durch die experimentelle Ungenauigkeit erklärt werden¹.

Demnach zeigen die Experimente die Entfaltung einer Teilstruktur des GFP, die kleiner als der strukturierte Teil S_n des nativen GFP ist. Die im Vergleich zu S_n fehlenden Aminosäuren können entweder vom Amino- oder vom Carboxyterminus der nativen Struktur S_n des GFP stammen. Zum Zeitpunkt der beobachteten Entfaltung müssen sie bereits als *random coil* vorgelegen haben.

Die Größe dieser Teilstruktur kann mit Hilfe der pdb-Datei 1B9C [4] bestimmt werden. Abb. 6.2 zeigt dazu schematisch, wie sich die Konturlänge des GFP im nativen Zustand S_n (a), im Zustand direkt vor der beobachteten Entfaltung (b) und im entfalteten Zustand (c) zusammensetzt.

Die Konturlänge des GFP direkt vor der beobachteten Entfaltung beträgt demnach entweder

$$L(I) = L_{1,3} + d_{C-N} + L_{j,230} + L_{231,238} \quad \text{Abb. 6.2b) (I)} \quad (6.7)$$

¹Bezieht man die Aminosäuren 1-3 und 231-238 des kompletten GFP in die Analyse mit ein, *vergrößert* sich die Diskrepanz.

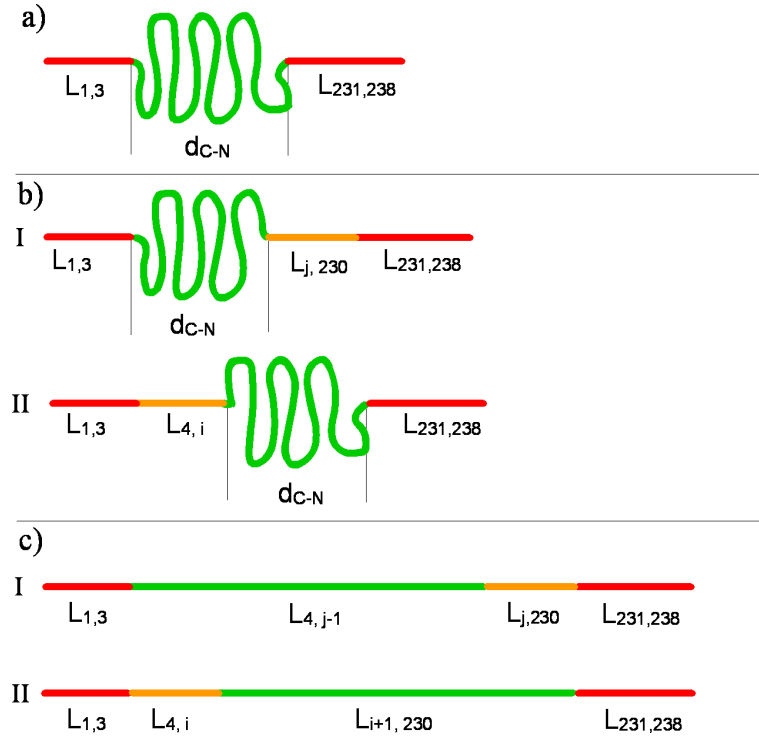


Abbildung 6.2: Schematische Darstellung der Beiträge zur Konturlänge des GFP in verschiedenen Zuständen. a) Natives GFP enthält die Struktur S_n und als *random coil* die N-terminalen Aminosäuren 1 bis 3 und die C-terminalen Aminosäuren 231 bis 238. b) Zustand des GFP direkt vor der beobachteten Entfaltung: (I) C-terminal oder (II) N-terminal liegen zusätzlich Aminosäuren als *random coil* vor. Die gefaltete Teil-Struktur ist in (I) und (II) kleiner als S_n . c) Entfalteter Zustand S_u des GFP. Die Konturlänge in (I) und (II) ist jeweils gleich. d_{C-N} bezeichnet in a) und b) den nativen Abstand zwischen N- und C-Terminus der enthaltenen gefalteten Struktur,

oder

$$L(II) = L_{1,3} + L_{4,i} + d_{C-N} + L_{231,238} \quad \text{Abb. 6.2b) (II)} \quad (6.8)$$

d_{C-N} bezeichnet dabei den Abstand zwischen Carboxy- und Aminoterminus der gefalteten Teilstruktur. Im Fall (I) wird der Aminoterminus durch die Aminosäure 4 des GFP gebildet, während der Index j der C-terminalen Aminosäure der Teilstruktur noch unbekannt ist. Im Fall (II) ist der Index i der N-terminalen Aminosäure noch unbekannt, während der Carboxyterminus durch die Aminosäure 230 des GFP gebildet wird. Die Abstände d_{C-N} sind für alle Indizes i, j aus den Abständen zwischen C_α -Atomen der jeweiligen Aminosäuren aus den Strukturdaten 1B9C [4] berechnet worden.

Die Konturlänge des GFP im komplett entfalteten Zustand beträgt $L(S_u) = 238 \cdot d_{AS}$. Der zu erwartende Konturlängenzuwachs ΔL_{GFP} nach Entfaltung der Teilstruktur lässt sich daher über $\Delta L_{GFP} = L(S_u) - L(I)$ bzw. $\Delta L_{GFP} = L(S_u) - L(II)$ berechnen und der Index j im Fall (I) bzw. der Index i im Fall (II) angeben, der zu einer Übereinstimmung mit der experimentell ermittelten Zunahme der Konturlänge führt.

Der so berechnete Längenzuwachs ΔL_{GFP} ist in Abb. 6.3 als Funktion der Anzahl von der Struktur S_n N- oder C-terminal abgetrennter Aminosäuren abgebildet.

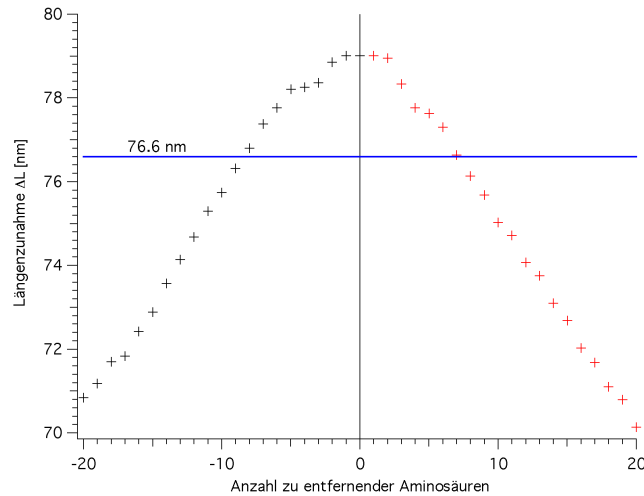


Abbildung 6.3: Erwartete Konturlängenzunahme ΔL_{GFP} nach Entfaltung des GFP als Funktion der Anzahl von der nativen Struktur S_n abgetrennter Aminosäuren. Positive Aminosäuren-Achse: Abtrennung erfolgt N-terminal. Negative Aminosäuren-Achse: Abtrennung erfolgt C-terminal. Ebenfalls eingetragen: Gemessener Konturlängenzuwachs von $\Delta L_{GFP} = 76.6 \pm 0.3$ nm

Es ergeben sich zwei mögliche Lösungen: entfernt man N-terminal sieben Aminosäuren oder C-terminal acht Aminosäuren, so würde sich die experimentell ermittelte Zunahme der Konturlänge von $\Delta L_{GFP} = 76.6 \pm 0.3$ nm im Rahmen der Messgenauigkeit ergeben. Abb. 6.4 bringt diese zwei Vorschläge für eine stabile Teilstruktur in Verbindung mit dem Strukturbild des GFP aus 1B9C [4].

Um acht C-terminale Aminosäuren (Abb. 6.4b) als bereits entfaltet anzunehmen, müsste ein β -Faltblatt zur Hälfte entfaltet werden. Dies erscheint aufgrund der vielen Wechselwirkungen mit benachbarten β -Faltblättern als unwahrscheinlich. Die N-terminale Lösung (Abb. 6.4 a)) hingegen entspricht der Entfernung einer α -helikalen Struktur, die nicht Teil des β -Fasses des GFP ist. Diesen 7 Aminosäuren (Nummer 4-10 in der Aminosäure-Sequenz des GFP, siehe Tab.E.2) wird in der pdb-Datei 1B9C kein Strukturmotiv zugeordnet, obschon sie zumindest im Strukturbild helikal angeordnet sind. Es ist anzunehmen, dass es dieser Teil der nativen Struktur S_n des GFP ist, der zum Zeitpunkt der beobachteten Entfaltung bereits fehlt.

Die bisherige Interpretation aus Kapitel 5, das beobachtete Entfaltungsereignis entspreche der Entfaltung des GFP, muss nun korrigiert werden. Es handelt sich dabei nur um die Entfaltung der reinen β -Fass-Struktur des GFP (Aminosäuren 11-230). Diese Teilstruktur des GFP wird im folgenden als GFP_{11-230} bezeichnet.

Die Aminosäuren 1-10 und 231-238 sind zum Zeitpunkt der Entfaltung von GFP_{11-230} bereits strukturlos und tragen als *random coil* zur Polypeptid-Elastizität der untersuchten Moleküle bei.

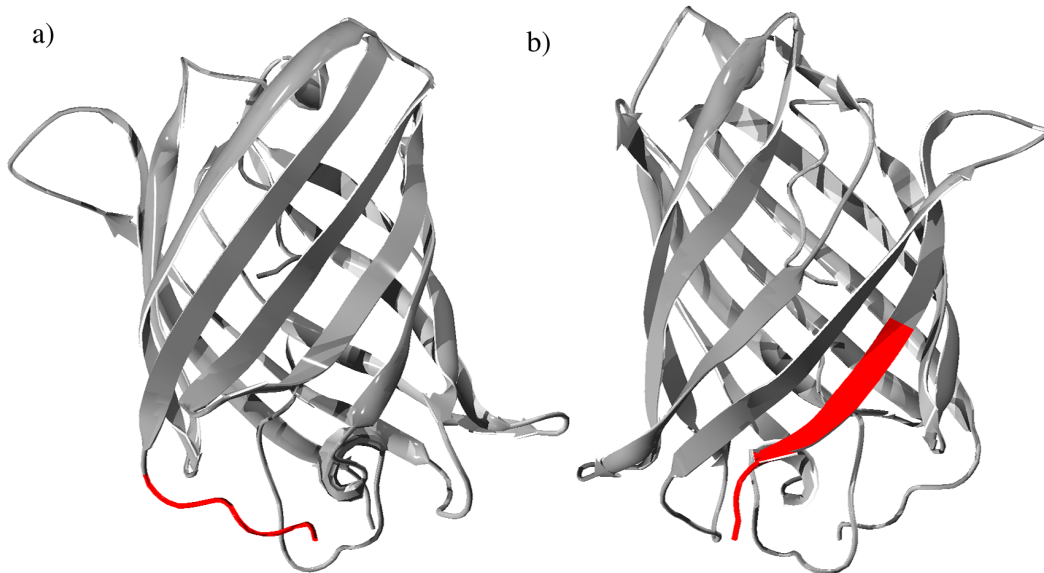


Abbildung 6.4: GFP-Struktur (pdb-Datei 1B9C) a) Rot markiert: N-terminal 7 Aminosäuren; b) Rot markiert: C-terminal 8 Aminosäuren

Es stellte sich die Frage, ob der Übergang $S_n \rightarrow GFP_{11-230}$ des GFP aus der nativen Struktur S_n in den Zustand GFP_{11-230} in den Kraftkurven beobachtbar ist.

6.5 Kraftplateaus vor der GFP_{11-230} -Entfaltung

Mit den Erkenntnissen aus dem vorherigen Abschnitt wurden die an der Ig8-GFP-Chimäre und an der DdFLN(1-5)-GFP-Chimäre gemessenen Kraftkurven auf Eigenschaften untersucht, die mit dem Übergang $S_n \rightarrow GFP_{11-230}$ in Zusammenhang stehen könnten. Es stellte sich heraus, dass es solche Eigenschaften gibt. Abb. 6.5 zeigt die GFP_{11-230} -Entfaltung in Kraftkurven, die an der Ig8-GFP-Chimäre gemessen wurden. Es sind deutlich plateauartige Abweichungen vom WLC-förmigen Verhalten vor der Entfaltung des GFP_{11-230} zu beobachten. Dieses Kraftplateau bei etwa 35 pN ist auch in Kraftkurven beobachtbar, die an der DdFLN(1-5)-GFP-Chimäre gemessen wurden, wie Abb. 6.6 zeigt.

Kraftkurven, die am reinen Ig8-Protein gemessen wurden, zeigen dieses Kraftplateau bei 35 pN nicht [32]. Es tritt ebenfalls nicht in Kraftkurven auf, die am reinen DdFLN(1-5)-Protein gemessen wurden [37]. Es kann sich daher nur um eine Eigenschaft handeln, die durch die Präsenz der GFP-Domäne in den Chimären-Proteinen verursacht wird. Die in Abb. 6.6c) gezeigte Kraftkurve zeigt beispielhaft noch eine weitere Eigenschaft dieses Kraftplateaus: es wiederholt sich vor jedem Anstieg bis zur Entfaltung von GFP_{11-230} . Danach wird es nicht mehr beobachtet.

Das hier beobachtete Kraftplateau könnte aufgrund seines Erscheinungsbildes durch nahe am thermischen Gleichgewicht stattfindende Übergänge einer in den Molekülen enthaltenen Teilstruktur verursacht worden sein [1][25][28][39][40]. Es soll daher mit einem

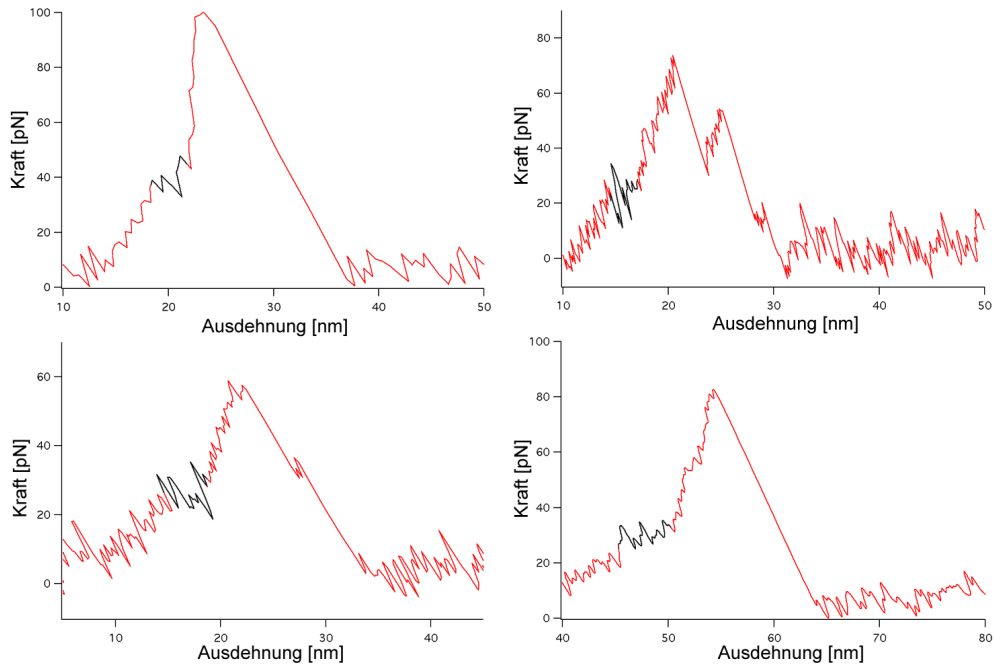


Abbildung 6.5: GFP_{11-230} -Entfaltungen in Kraftkurven, die an der Ig8-GFP-Chimäre gemessen wurden. Schwarz markiert: Plateaus in den Anstiegen vor der GFP_{11-230} -Entfaltung. Es wurden verschiedene Blattfedern und verschiedene Filterfrequenzen verwendet.

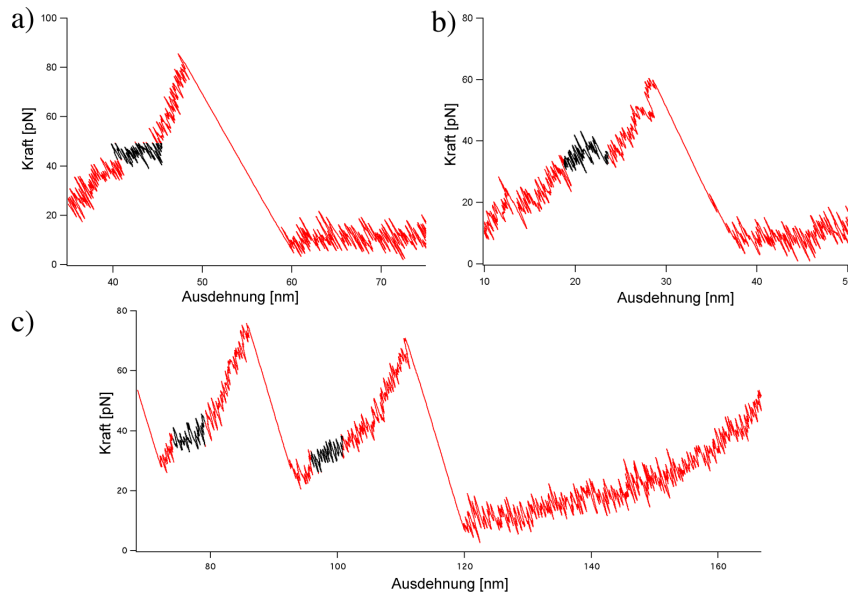


Abbildung 6.6: GFP_{11-230} -Entfaltungen in Kraftkurven, die an der DdFLN(1-5)-GFP-Chimäre gemessen wurden. a)-b) Schwarz markiert: Plateaus in den Anstiegen vor der GFP_{11-230} -Entfaltung. c) Schwarz markiert: Plateaus in den Anstiegen vor einer DdFLN-Domänen-Entfaltung und vor der GFP_{11-230} -Entfaltung.

Modell interpretiert werden, dass Übergänge zwischen zwei Zuständen im thermischen Gleichgewicht beschreibt [33][39].

6.6 Übergang native Struktur $S_n \rightarrow GFP_{11-230}$

Die Ausdehnung eines Polypeptids mit der Konturlänge L_0 , das zusätzlich auch n Segmente enthält, die sich jeweils in zwei verschiedenen Zuständen S_n und S_u befinden können, lässt sich folgendermassen angeben:

$$d(F) = d_0(F, L_0) + P_{S_n \rightarrow S_u}(F) \cdot n \cdot d_S(F, \Delta L) \quad (6.9)$$

Der Beitrag der Segmente im Zustand S_n zur Konturlänge L_0 des Moleküls wird dabei vernachlässigt. ΔL ist die Zunahme der Konturlänge des Moleküls, die durch ein Segment beim Übergang in den Zustand S_u verursacht wird. Im thermischen Gleichgewicht sind die Übergangsraten k_{off}^0 und k_{on}^0 zwischen zwei Zuständen nicht getrennt bestimmbar, sondern nur die Gleichgewichtskonstante K_{eq}^0 . Darüber ergibt sich die Wahrscheinlichkeit $P_{S_n \rightarrow S_u}(F)$ für einen Übergang $S_n \rightarrow S_u$ eines einzelnen Segmentes im thermischen Gleichgewicht und unter einer Kraft F :

$$P_{S_n \rightarrow S_u}(F) = \frac{K_{eq}(F)}{1 + K_{eq}(F)} \quad (6.10)$$

wobei die Gleichgewichtskonstante unter einer Kraft F folgendermaßen gegeben ist:

$$K_{eq}(F) = \exp\left(\frac{-\Delta G + F \cdot d_S(F, \Delta L)}{k_B T}\right) \quad (6.11)$$

Die momentane Ausdehnung d eines Polypeptids unter einer Kraft F lässt sich über die Inversion der WLC-Interpolationsformel aus Gl. 3.8 berechnen. Aus Gl. 6.9 wurden auf diese Weise Kraft-Abstandskurven berechnet und an die gemessenen Kraftplateaus angepasst. Eine solche Anpassung ist in Abb. 6.7 an zwei Ig8-GFP-Kraftkurven und in Abb. 6.8 an einer DdFLN(1-5)-GFP-Kraftkurve dargestellt, in der das Kraftplateau in zwei Anstiegen vor der Entfaltung von GFP_{11-230} aufgelöst werden konnte.

Eine gute Übereinstimmung des verwendeten Modells mit den Kraftkurven aus Messungen an der Ig8-GFP-Chimäre wie auch mit den Kraftkurven aus Messungen an der DdFLN(1-5)-GFP-Chimäre, erhält man für die Parameter:

$$n = 1 ; \Delta L = 3.2 \pm 1 \text{ nm}; \Delta G = 20 \pm 20\% k_B T \quad (6.12)$$

Da die Charakteristik des beobachteten Kraftplateaus in den Kraftkurven, die an beiden Chimären-Proteinen Ig8-GFP und DdFLN(1-5)-GFP gemessen wurden, übereinstimmt, lässt sich jetzt schliessen, dass das Kraftplateau durch einen GFP-eigenen Prozess verursacht wird.

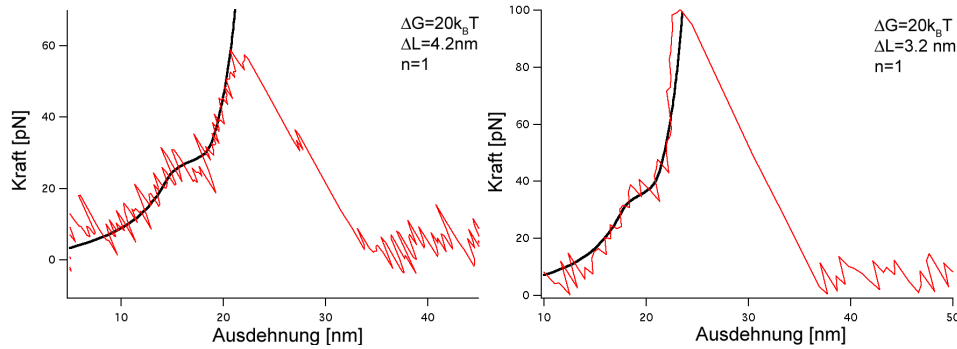


Abbildung 6.7: Anpassung des Zwei-Zustands Gleichgewichtsmodells aus Gl. 6.9 an den Anstieg vor GFP_{11-230} -Entfaltung in zwei Ig8-GFP-Kraftkurven.

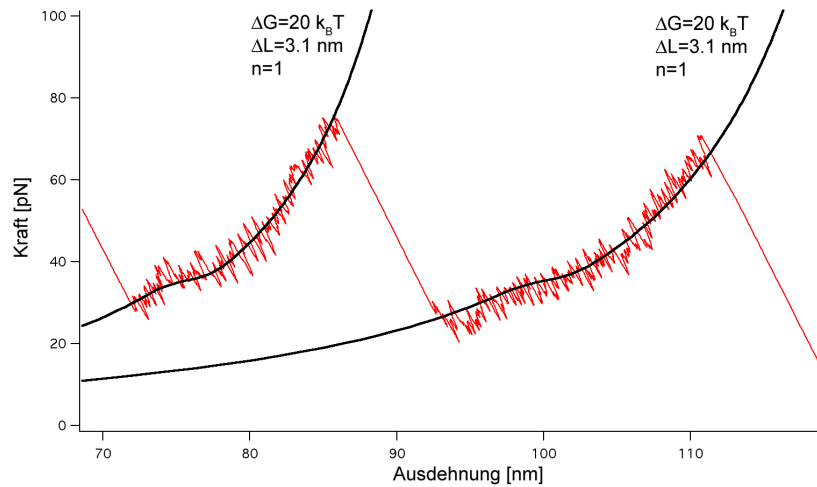


Abbildung 6.8: Anpassung des Zwei-Zustands Gleichgewichtsmodells aus Gl. 6.9 an eine DdFLN(1-5)-GFP-Kraftkurve.

Ein Längenzuwachs von $\Delta L = 3.2 \pm 1$ nm passt in ein Bild, das Gleichgewichtsübergänge der N-terminalen α -Helix (Aminosäuren 3-10) der nativen Struktur S_n des GFP als die Ursache des Kraftplateaus sieht. In einem solchen Fall würde man eine Zunahme der Konturlänge wie folgt erwarten:

$$\Delta L = L_7 + d_{11,230} - d_{4,230} \quad (6.13)$$

Dabei ist $d_{4,230}$ der Abstand zwischen Amino- und Carboxy-Terminus der nativen Struktur S_n des GFP vor dem Übergang. $d_{11,230}$ ist der Abstand zwischen Amino- und Carboxy-Terminus der nach dem Übergang verbliebenen Struktur $GFP_{11,230}$. L_7 bezeichnet die Konturlänge der bei dem Übergang entfalteten 7 Aminosäuren. Mit den Strukturdaten 1B9C [4] würde sich ein Zahlenwert von

$$\Delta L = 2.9 \pm 0.2 \text{ nm} \quad (6.14)$$

ergeben, was in sehr guter Übereinstimmung mit dem über das Gleichgewichtsmodell bestimmten Längenzuwachs ist. Die Strukturdaten von *Wildtype*-GFP (1EMB [4]) ermöglichen die Einbeziehung der N-terminalen Aminosäuren 2 und 3 des GFP in diese Analyse. Sie würden einen noch höheren Zahlenwert bei einer Entfaltung der N-terminalen α -Helix bis hin zu $\Delta L \simeq 4.5$ nm erklären können.

Diese Analyse legt somit nahe, dass das beobachtete Kraftplateau durch den Übergang $S_n \rightarrow GFP_{11-230}$ verursacht wird. Dabei wird die N-terminale α -Helix der nativen Struktur S_n des GFP entfaltet. Das verwendete Zwei-Zustandsmodell im thermischen Gleichgewicht reproduziert gut den Verlauf der gemessenen Kraftkurven im Bereich des beobachteten Kraftplateaus. Zusammen mit der Beobachtung, dass das Kraftplateau wiederholt in aufeinander folgenden Anstiegen vor der GFP_{11-230} -Entfaltung in Kraftkurven auftritt², lässt sich daher schließen, dass die Entfaltung der α -Helix nahe am thermischen Gleichgewicht stattfindet.

Aus der Anpassung des Gleichgewichts-Modells an die Kraftkurven ergibt sich, dass für den Übergang $S_n \rightarrow GFP_{11-230}$ eine Arbeit ΔG von ca. $20 k_B T$ aufgewendet werden muss. Fukuda *et al.* haben die Faltungs- und Entfaltungsraten von nativem GFP untersucht [17]. Aus der Studie geht hervor, dass bei der Bildung von nativem GFP eine Energie von ca. $16 k_B T$ frei wird. Der hier bestimmte Wert für den Übergang $S_n \rightarrow GFP_{11-230}$ liegt sogar über diesem Wert.

GFP_{11-230} ist demnach ein metastabiler Zwischenzustand der mechanischen Entfaltung des GFP, dessen Energie etwa $4 k_B T$ über der des entfalteten Zustandes S_u liegt. Zur Entfaltung von nur sieben Aminosäuren aus der nativen Struktur muss demnach mehr Energie aufgewendet werden, als bei dem gesamten Faltungsprozess des GFP frei wird.

6.7 Stabilität von GFP_{11-230}

Zur Messung der Kräfte, die zur Entfaltung der GFP_{11-230} -Struktur nötig waren, sind Einzelmolekül-Kraftkurven, die an der DdFLN(1-5)-GFP-Chimäre gemessen wurden, mit der in Abschnitt 4.5 beschriebenen Methode analysiert worden.

Abb. 6.9a) zeigt die Verteilung der GFP_{11-230} -Entfaltungskräfte. In diese Statistik sind 87 GFP_{11-230} -Entfaltungen eingegangen. Die zugehörigen Messungen sind bei konstanten externen Bedingungen (siehe Abschnitt 4.4) durchgeführt worden. Die wahrscheinlichste Kraft für eine GFP_{11-230} -Entfaltung beträgt ca. 102 pN. Die Verteilung zeigt eine Halbwertsbreite von ca. 40 pN.

Definitionsgemäß würde $F=102$ pN zugleich die Stabilität von GFP_{11-230} unter einer bestimmten Kraftladungsrate \dot{F} angeben. Diese konnte in den Messungen nur zum

²In Abb. 6.8 ist dies beispielhaft abgebildet. Die wirkenden Kräfte sind dort > 20 pN. Dies bedeutet, die Struktur hat die Rückreaktion nach der ersten Entfaltung, die zu dem Kraftplateau im ersten Anstieg führte, *gegen* eine Kraft > 20 pN durchgeführt, da ein zweites Kraftplateau (eine erneute Entfaltung) im nächsten Anstieg wieder beobachtet wird.

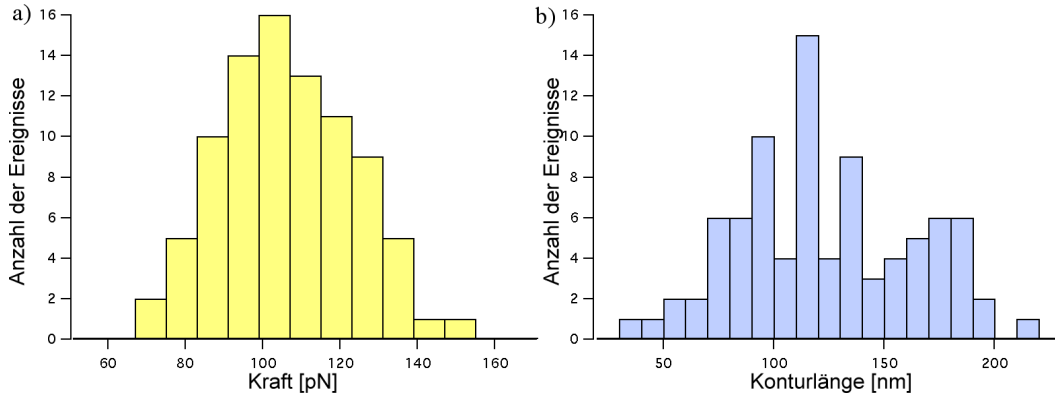


Abbildung 6.9: a) Verteilung der für die Entfaltung von GFP_{11-230} nötigen Kräfte. b) Verteilung der Konturlängen der einzelnen DdFLN(1-5)-GFP-Moleküle direkt vor der GFP_{11-230} Entfaltung.

Teil durch den Experimentator festgelegt werden (Messen mit jeweils gleicher Proben geschwindigkeit v_p und Verwendung von Blattfedern mit gleicher Federkonstante k_c). Jedoch schwankt die Kraftladungsrate \dot{F} von Entfaltungsereignis zu Entfaltungsereignis durch den Umstand, dass die Konturlänge der untersuchten Moleküle vor der Entfaltung der GFP_{11-230} -Domäne in jeder Kraftkurve unterschiedlich war (vgl. Gl. 3.32). Die Verteilung dieser Konturlängen zeigt Abb. 6.9b).

Aus der Kraftverteilung können Informationen über die natürliche Übergangsrate k_{off}^0 und die Potentialbreite $d_{GFP_{230-T_1}}$ des Zustandes GFP_{11-230} über das in Abschnitt 3.6ff. vorgestellte lineare Zwei-Zustandsmodell gewonnen werden. In eine Analyse der erhaltenen Kraftverteilung muss dabei die Verteilung der Konturlängen vor der GFP_{11-230} -Entfaltung einbezogen werden.

Mit einer Monte-Carlo-Simulation konnte das auf einfache Weise realisiert werden (vgl. Abschnitt 4.6). Abb. 6.10 zeigt eine so simulierte Kraftkurve.

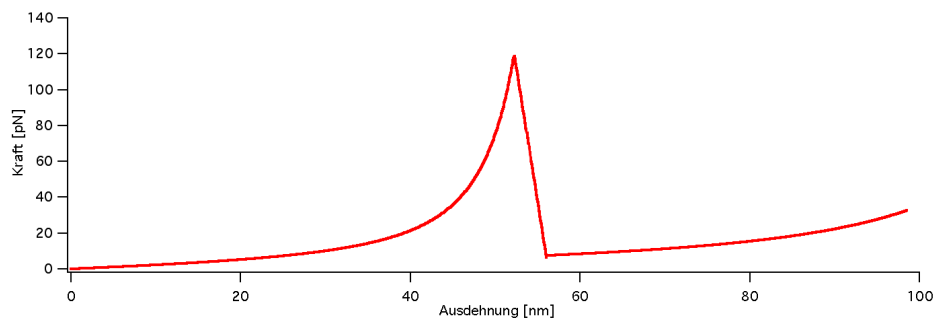


Abbildung 6.10: Simulierte Kraftkurve an einem Polypeptid der Länge $L=60$ nm, das eine virtuelle GFP_{11-230} -Domäne enthält. Bei einer Ausdehnung von 50 nm kommt es zur GFP_{11-230} -Entfaltung.

Es wurde für jede experimentell gemessene Konturlänge L und der im jeweiligen Expe-

riment verwendeten Federkonstante k_c eine Kraftkurve mit der experimentell verwendeten Geschwindigkeit v_p simuliert und die Kraft notiert, bei der die Simulation eine Entfaltung der GFP_{11-230} -Domäne voraussagte. Daraus konnte eine Kraftverteilung wie in Abb. 6.9a) angefertigt werden und mit der experimentell erhaltenen verglichen werden. Auf diese Weise konnte die in Abb. 6.9b) angegebene Längenverteilung und die für jedes tatsächlich beobachtete Entfaltungsereignis herrschenden Versuchsbedingungen einfach in eine Analyse einbezogen werden.

Durch Anpassung der Parameter $d_{GFP_{11-230-T_1}}$ und k_{off}^0 konnte die gemessene Kraftverteilung mit der Monte-Carlo-Simulation gut reproduziert werden. Die Gesamtzahl der in die experimentelle Statistik eingegangenen Entfaltungsereignisse ist 87. Es wurden keine Kräfte unterhalb von 50 pN gemessen. Die Simulation ist zur besseren Vergleichbarkeit mit dem Experiment so eingestellt worden, dass die Gesamtzahl der gezogenen Kraftkurven mit Entfaltungskräften > 50 pN genau 87 ergab. Es sind 10 solcher Simulationen durchgeführt worden und die erhaltenen Verteilungen von Entfaltungskräften gemittelt worden.

Optimale Übereinstimmung zwischen Simulation und Experiment ergab sich für die folgenden Parameter:

$$d_{GFP_{11-230-T}} = 0.28 \pm 0.03 \text{ nm}; \quad k_{off}^0 = 0.07 \pm 0.03 \text{ s}^{-1} \quad (6.15)$$

Werden die Unschärfen in Gl. 6.15 überschritten, so ergeben sich deutliche Abweichungen in der Lage des Zentrums wie auch in der Breite zwischen der simulierten Kraftverteilung und der gemessenen Kraftverteilung. Abb. 6.11a) vergleicht die durch die Simulation vorausgesagte Kraftverteilung für die angegebenen Parameter mit der experimentell erhaltenen. Halbwertsbreite und Position der beiden Verteilungen stimmen gut überein.

Das Zwei-Zustandsmodell mit einer linearen Krafframpe aus Abschnitt 3.7 kann im vorliegenden Fall näherungsweise auch direkt angewendet werden. Die Kräfte, die zur Entfaltung von GFP_{11-230} nötig waren, fallen in einen Bereich (> 50 pN), in dem das Kraft-Ausdehnungsverhalten der untersuchten DdFLN(1-5)-GFP Moleküle in erster Näherung als linear bezeichnet werden kann, also $F(d) \simeq k_p d$ (vgl. dazu Abb. 5.5). Über Gl. 3.32 ergibt sich die zeitliche Kraftabhängigkeit in dieser Näherung wie folgt:

$$F(t) = \frac{v_p}{\frac{1}{k_c} + \frac{1}{k_p}} = \eta_{eff} t \quad (6.16)$$

Daraus ergibt sich eine konstante Kraftladungsrate $\dot{F} = \eta_{eff}$ und Gl. 3.30 kann für die Berechnung einer Wahrscheinlichkeitsdichte dP/dF für das Auftreten von Entfaltungskräften angewendet werden. Setzt man $v_p=300$ nm/s, $k_c=30$ pN/nm und für $k_p = 6.3$ pN/nm, so führen die Parameter

$$d_{GFP_{11-230-T}} = 0.28 \pm 0.03 \text{ nm}; \quad k_{off}^0 = 0.07 \pm 0.03 \text{ s}^{-1} \quad (6.17)$$

zu einer guten Übereinstimmung von dP/dF sowohl mit der experimentellen Kraftverteilung als auch mit der simulierten Kraftverteilung.

Die so erhaltene dP/dF -Verteilung ist in Abb. 6.11 b) abgebildet. Zum Vergleich mit der experimentell erhaltenen Verteilung ist sie auf insgesamt 87 Entfaltungsereignisse normiert worden. Die Breite der berechneten Verteilung dP/dF ist geringer als die der experimentellen Verteilung. Das ist eine Konsequenz der breiten Mischung verschiedener Kraftladungs-Raten, die experimentell vorlag (vgl. Abb. 3.3). Die Monte-Carlo-Simulation trägt diesem Umstand Rechnung und reproduziert daher gut die Breite der experimentellen Verteilung, während es die Berechnung von dP/dF mit einer festen, effektiven Kraftladungsrate η_{eff} wie in Gl. 6.16 ignoriert.

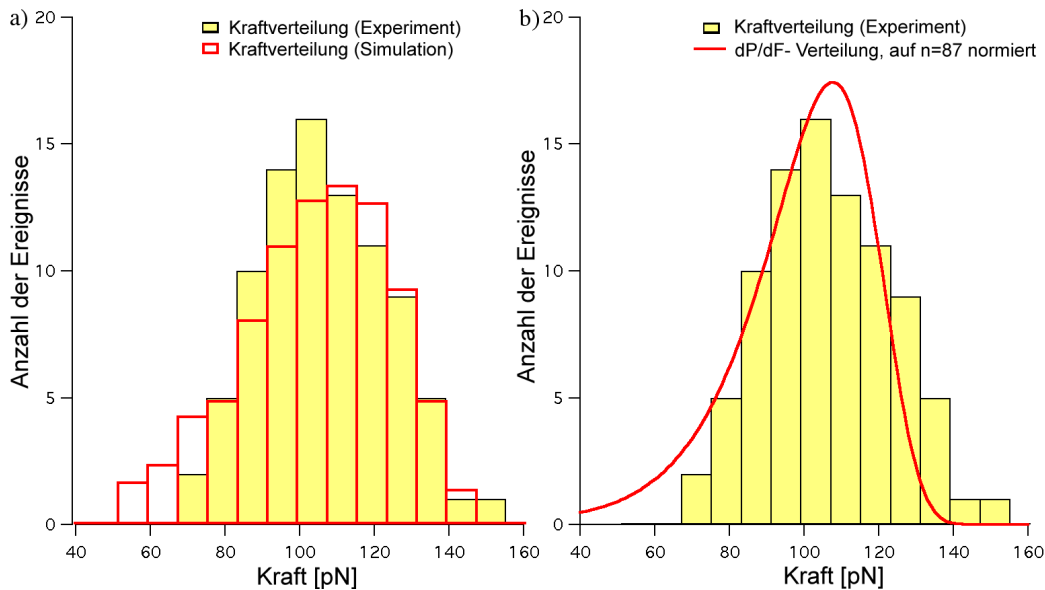


Abbildung 6.11: a) Kraftverteilung, wie sie durch Mittelung von 10 Monte-Carlo Simulationen mit jeweils 87 virtuellen Kraftkurven erhalten wurde. Parameter siehe Gl. 6.15. b) Auf 87 Entfaltungsereignisse normierte dP/dF -Verteilung. Parameter siehe Gl. 6.17; Zum Vergleich sind beide Diagramme mit der experimentell erhaltenen Kraftverteilung hinterlegt worden.

Die Analyse der für die Entfaltung von GFP_{11-230} nötigen Kräfte ergibt, dass GFP_{11-230} eine natürliche Lebensdauer von nur $\tau^0 = 14 \pm 6$ s aufweist. Aus der Literatur [17] entnimmt man Lebensdauern für natives *Cycle3*-GFP von $\tau^0 = 10^9$ s. Dies steht im deutlichen Widerspruch zu dem hier bestimmten Wert. Diese Analyse führt somit unabhängig von den Ergebnissen aus den Abschnitten 6.4 und 6.6 zu der gleichen Interpretation: Es kann sich bei GFP_{11-230} nicht um die native Struktur des GFP handeln.

6.8 Energielandschaft - Teil 1

Aus den Ergebnissen dieses Kapitels lässt sich ein Teil der Energielandschaft des GFP entlang des Pfades der mechanischen Entfaltung rekonstruieren. Die energetische Lage des Übergangszustandes T_0 zwischen S_n und GFP_{11-230} lässt sich über die Arrhenius-Gleichung

(3.17) abschätzen. Das den Übergang S_n und GFP_{11-230} kennzeichnende Kraftplateau wurde in den Kraftkurven in aufeinander folgenden Anstiegen wiederholt beobachtet. Die dazwischen liegenden Zeiträume waren geringer als 0.1 s. Das bedeutet, der Übergang $S_n \rightarrow GFP_{11-230}$ hat eine Rückrate von $k_{on}^0 \geq 10 \text{ s}^{-1}$. Nimmt man für den Vorfaktor in der Arrhenius-Gleichung 3.17 einen Wert von $k_T \simeq 10^9 \text{ s}^{-1}$ an [5][24][44], so folgt für die energetische Lage des Übergangszustandes:

$$G^0(T_0) - G^0(GFP_{11-230}) \leq 19 k_B T \quad (6.18)$$

Die energetische Lage des Übergangszustandes T_1 zwischen GFP_{11-230} und dem Zustand S_u lässt sich in der gleichen Weise über die Arrhenius-Gleichung 3.17 aus der in Abschnitt 6.7 bestimmten Übergangsrate $k_{off}^0 = 0.07 \pm 0.03 \text{ s}^{-1}$ bestimmen. Es ergibt sich:

$$G^0(T_1) - G^0(GFP_{11-230}) \simeq 23 k_B T \quad (6.19)$$

Mit dem Argument der wiederholten Beobachtung des Überganges $S_n \rightarrow GFP_{11-230}$ lässt sich auch noch unabhängig von dem in Abschnitt 6.6 verwendeten Gleichgewichtsmodell die energetische Lage des Zustandes GFP_{11-230} abschätzen. Für die Bildung der nativen Struktur $S_u \rightarrow S_n$ entnimmt man der Literatur Raten im Bereich von $\bar{k}_{S_u \rightarrow S_n}^0 \simeq 0.002 \text{ s}^{-1}$ [41]. Würde die Faltung des Proteins über den Pfad $S_u \rightarrow GFP_{11-230} \rightarrow S_n$ geschehen, so müsste die dafür gültige Rate $k_{S_u \rightarrow S_n}^0 \simeq 0.002 \text{ s}^{-1}$ sein³. Die Rate für den Übergang $GFP_{11-230} \rightarrow S_n$ ist wie eben argumentiert größer als 10 s^{-1} . Damit muss die Rate, mit der der Zustand GFP_{11-230} aus dem entfalteten Zustand S_u gebildet wird, auch bei $k_{S_u \rightarrow GFP_{11-230}}^0 \leq 0.002 \text{ s}^{-1}$ liegen. Somit liegt eine Abschätzung der Raten für die Bildung und Zerfall des Zustandes GFP_{11-230} im Vergleich zum entfalteten Zustand S_u vor. Über die Gleichgewichtskonstante $K_{eq}^0 = k_{off}^0 / k_{S_u \rightarrow GFP_{11-230}}^0$ ergibt sich daher für die energetische Lage von GFP_{11-230} :

$$G^0(GFP_{11-230}) - G^0(S_u) > 3 k_B T \quad (6.20)$$

Für den Übergang $S_n \rightarrow GFP_{11-230}$ ist laut Abschnitt 6.6 eine Arbeit von $20 k_B T$ aufzuwenden. Der Zustand GFP_{11-230} liegt nach der vorstehenden Analyse energetisch mindestens $3 k_B T$ über dem entfalteten Zustand. Damit beträgt die Energiedifferenz zwischen den Zuständen S_u und S_n mindestens $17 k_B T$. Der bereits angesprochene Literaturwert für die bei der Bildung des GFP frei werdende Energie von $16 k_B T$ stimmt im Rahmen der Messgenauigkeit damit überein [17].

Für die Konstruktion einer Energielandschaft müssen noch die Positionen der einzelnen Zustände entlang der durch den End-zu-End Abstand d_{C-N} gegebenen Reaktionskoordinate bestimmt werden. Dieser Abstand liegt parallel zur angelegten Krafrichtung. Die End-zu-End Abstände der beobachteten Zustände können aus den Messdaten direkt angegeben werden, allerdings nur bei einer endlichen Kraft F . Eine experimentell messbare Kraft ist z.B. durch 10 pN gegeben. Die Energielandschaft des GFP wird daher bei dieser Kraft konstruiert.

³Würde dieser Pfad größere Raten aufweisen, so würde er bei der Faltung auch eingeschlagen werden und der Wert für $\bar{k}_{S_u \rightarrow S_n}^0$ wäre größer gemessen worden.

Die gemessene Zunahme der Konturlänge des GFP beim Übergang $S_n \rightarrow GFP_{11-230}$ beträgt $\Delta L = 3.2$ nm. Nach dem WLC-Modell (Gl. 3.8) ergibt sich die Zunahme des End-zu-End Abstandes bei diesem Übergang unter einer Kraft von 10 pN zu $\Delta d = 1.6$ nm. Der Abstand zwischen GFP_{11-230} und dem Übergangszustand T_1 in Richtung entfalteter Zustand S_u wurde in Abschnitt 6.7 zu 0.28 nm bestimmt. Im Rahmen des verwendeten Zwei-Zustandsmodells ist dieser Wert auch bei 10 pN gültig.

Abb. 6.12 fasst die Ergebnisse dieses Kapitels in einer schematischen Darstellung der Energielandschaft des GFP entlang des Pfades der mechanischen Entfaltung des GFP bei einer Kraftbelastung von 10 pN zusammen. Die Energiedifferenz zwischen zwei Zuständen ist in dem verwendeten linearen Modell durch $\Delta G = \Delta G^0 - F\Delta d$ gegeben. Aus Gründen der Übersichtlichkeit ist die Energieachse in $\Delta G + F\Delta d$ Einheiten angegeben.

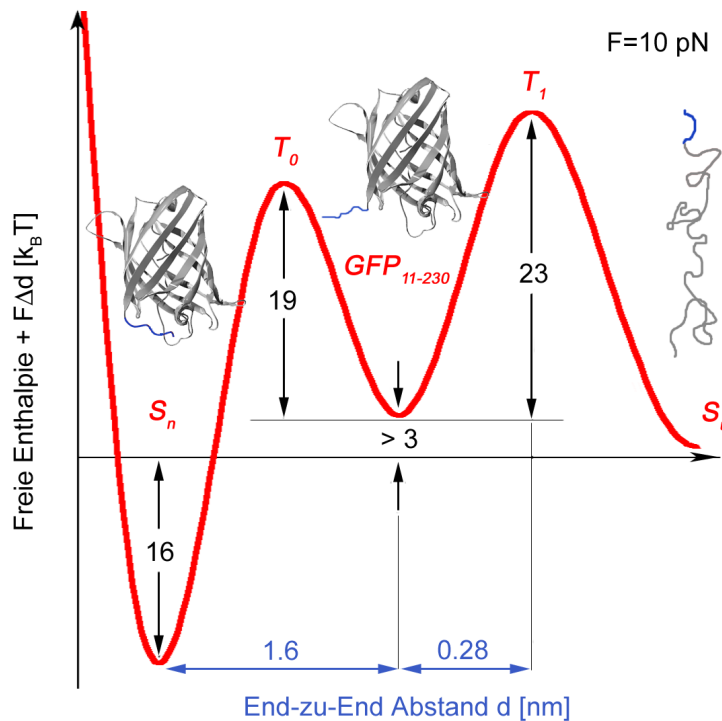


Abbildung 6.12: Schematische Darstellung der Energielandschaft des GFP entlang des Pfades der mechanischen Entfaltung mit den in diesem Kapitel beschriebenen Übergängen $S_n \rightarrow GFP_{11-230} \rightarrow S_u$ unter einer Kraft von 10 pN.

6.9 Zusammenfassung

Die mechanische Entfaltung des GFP zeigt einen stabilen Zwischenzustand: GFP_{11-230} . Dieser Zwischenzustand umfasst die 11 β -Faltblätter, die die Fass-Struktur des GFP bilden. GFP_{11-230} hat eine natürliche Lebensdauer τ^0 im Bereich von 14 s. GFP_{11-230} liegt

energetisch mindestens $3 k_B T$ über dem entfalteten Zustand S_u . Der Übergang aus der nativen GFP-Struktur S_n in den Zustand GFP_{11-230} äußert sich in den Kraftkurven in einem Kraftplateau bei ca. 35 pN vor der Entfaltung von GFP_{11-230} . Dabei wird die N-terminale α -Helix des GFP entfaltet. Dieser Übergang findet nahe am thermischen Gleichgewicht statt. Die Arbeit, die für diesen Übergang aufgewendet werden muss, liegt bei ca. $20 k_B T$.

Kapitel 7

Weitere Zwischenzustände der GFP-Entfaltung: S_1 und S_2

7.1 Einleitung

Dieses Kapitel befasst sich mit weiteren Zwischenzuständen auf dem Weg der mechanischen Entfaltung des GFP. Es wird sich zeigen, dass GFP nach der Entfaltung von GFP_{11-230} in einen weiteren Zwischenzustand mit sehr kurzer Lebensdauer übergeht. Die Lebensdauern sind gemessen worden und ihre Verteilung wird mit dem linearen Zwei-Zustandsmodell aus Abschnitt 3.5 interpretiert. Es werden zwei Vorschläge für eine mögliche Struktur des Zwischenzustandes gemacht. Weiterhin wird die Existenz eines weiteren, kurzlebigen Zwischenzustandes nachgewiesen.

7.2 Unterstrukturen der GFP-Entfaltung

Abb. 7.1 zeigt eine Zusammenstellung von Ausschnitten aus verschiedenen an der Ig8-GFP-Chimäre gemessenen Kraftkurven. Die Ausschnitte zeigen die GFP-Entfaltung.

Die Blattfeder-Relaxation ist in der Kraft-Abstandsdarstellung in guter Näherung eine Gerade mit Steigung $-k_c$, wobei k_c die Federkonstante ist. Dies ist in Abb. 7.1a) der Fall. Die in Abb. 7.1 b)-d) dargestellten Ereignisse zeigen jedoch deutliche Abweichungen davon. Abb. 7.2 zeigt Vergrößerungen der GFP_{11-230} -Entfaltung in Kraftkurven, die an der DdFLN(1-5)-GFP-Chimäre gemessen wurden. Auch hier erkennt man deutliche Abweichungen von einer normalen Relaxation der Blattfeder. Die GFP_{11-230} -Entfaltung findet in den DdFLN(1-5)-GFP Kraftkurven in genügendem Abstand (> 30 nm) zur Oberfläche statt, so dass Wechselwirkungen mit der Oberfläche als Ursache dieser Abweichungen ausgeschlossen werden können.

Es stellt sich die Frage, ob diese Abweichungen durch ein und dasselbe Objekt verursacht werden. Für Abb. 7.3 sind DdFLN(1-5)-GFP Kraftkurven ausgewählt worden, bei denen die Kräfte für die GFP_{11-230} -Entfaltung stark schwankten und die deutlich die angesprochenen Abweichungen in der Blattfeder-Relaxation zeigen. Die Kurven sind auf der

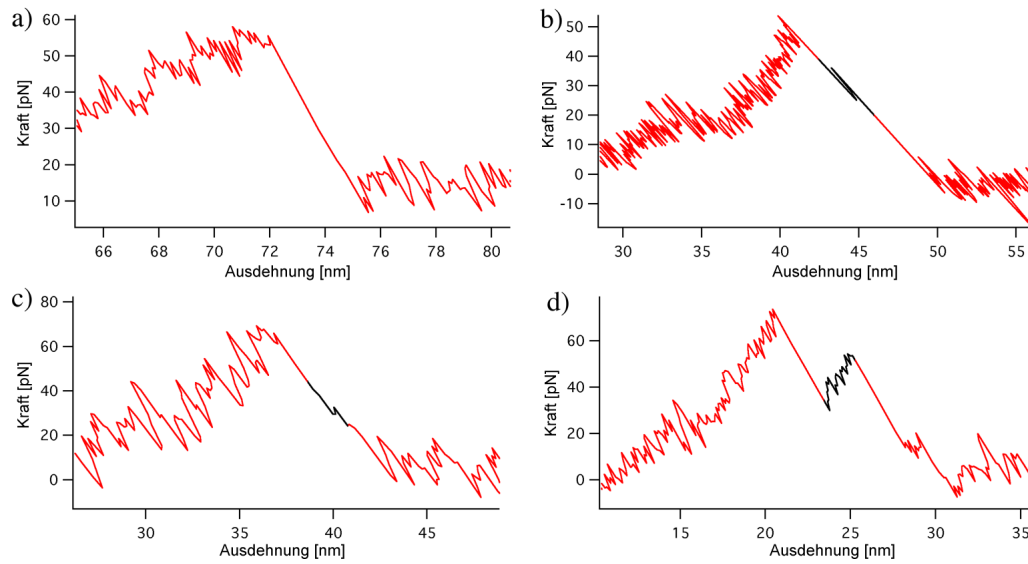


Abbildung 7.1: Vergrößerungen verschiedener GFP-Entfaltungen in Ig8-GFP Kraftkurven a) Einfache Blattfeder-Relaxation nach GFP_{11-230} -Entfaltung; b)-d) Schwarz markiert: Unregelmäßigkeiten in der Blattfeder-Relaxation

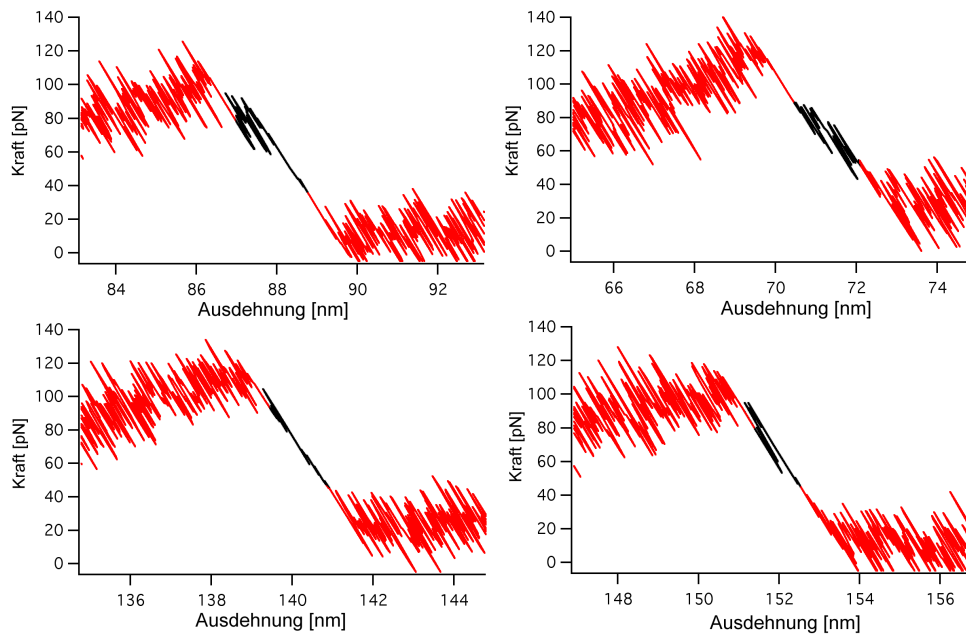


Abbildung 7.2: Vergrößerungen der GFP-Entfaltung in verschiedenen DdFLN(1-5)-GFP Kraftkurven. Schwarz eingefärbt: Unregelmäßigkeiten in der Blattfeder-Relaxationsphase.

Ausdehnungsachse so verschoben worden, dass sie sich im Bereich der GFP-Entfaltung passend überlagerten. Zusätzlich sind berechnete WLC-Kraftkurven abgebildet. Man erkennt, dass die Abweichungen den Verlauf einer Kraftkurve an einem nach der Entfaltung von GFP_{11-230} um $\Delta L = 7$ nm länger gewordenen Moleküls reproduzieren. Das spricht dafür, dass die Abweichungen tatsächlich durch ein und dasselbe Objekt verursacht worden sind.

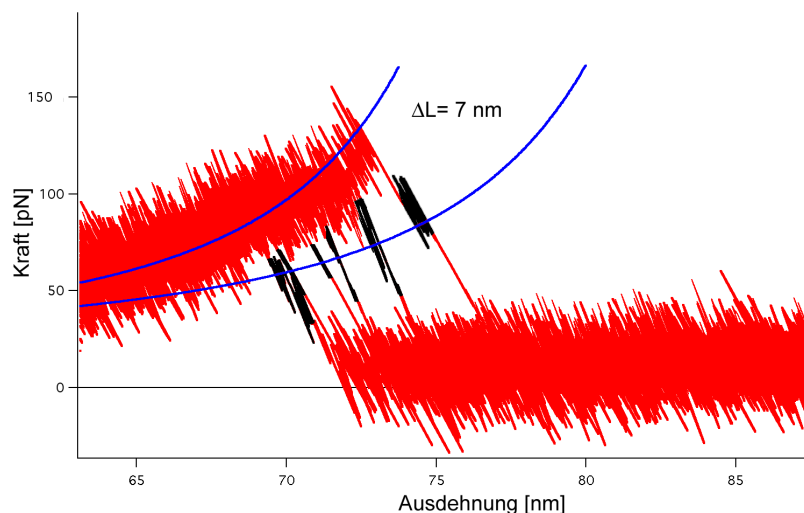


Abbildung 7.3: Sechs DdFLN(1-5)-GFP Kraftkurven sind bzgl. der GFP-Entfaltung passend überlagert worden. Blau: berechnete WLC-Kurven (Persistenzlänge 0.5 nm). Schwarz eingefärbt: Zwischenzustand.

Da die Abweichungen sowohl in Ig8-GFP und DdFLN(1-5)-GFP Kraftkurven beobachtet werden und nur in der Relaxation der Blattfeder nach der GFP_{11-230} -Entfaltung auftreten, ist anzunehmen, dass es sich bei der Ursache dieser Abweichungen um einen weiteren Zwischenzustand der mechanischen GFP-Entfaltung handelt. Aus einer Untersuchung der Lebensdauern dieses Zwischenzustandes können Informationen über Raten und Energielandschaft dieses Zwischenzustandes gewonnen werden.

7.3 Messung von Lebensdauern

Da der Zwischenzustand nicht immer zu beobachten war, und wenn, dann nur sehr kurz (einige hundert μs), weist er offenbar nur eine sehr geringe Lebensdauer auf. Für eine Untersuchung der Lebensdauern muss die Blattfederrelaxation daher möglichst schnell geschehen, damit der Zwischenzustand noch beobachtet werden kann, bevor er schon ohne Kraftbelastung zerfallen ist.

Für eine Untersuchung der Lebensdauern des Zwischenzustandes wurden hochauflösende (Samplingrate 20 kHz) Messungen mit dem Olympus Bio-Lever Typ A an der DdFLN(1-5)-GFP-Chimäre durchgeführt. Für die Durchführung der Experimente sei auf Abschnitt 4.4 verwiesen. Der Olympus Bio-Lever Typ A ist aufgrund seiner höheren Resonanzfrequenz von 8kHz für diese Untersuchungen geeignet. Die zeitliche (Kraft-)Relaxation der

Blattfeder nach der Entfaltung von GFP_{11-230} ist in den Kraftkurven nach Abweichungen von der normalen Relaxation eines überdämpften Oszillators hin analysiert worden. Bei einer normalen Zwei-Zustands-Entfaltung würde die Relaxationsphase der Blattfeder in dieser experimentellen Situation nur mit im Schnitt 2.5 Datenpunkten detektiert werden. Es sind jedoch deutliche Abweichungen von diesem Verhalten gemessen worden. Abb. 7.4 zeigt exemplarisch Ausschnitte von drei Kraft-Zeitkurven, um dies zu demonstrieren.

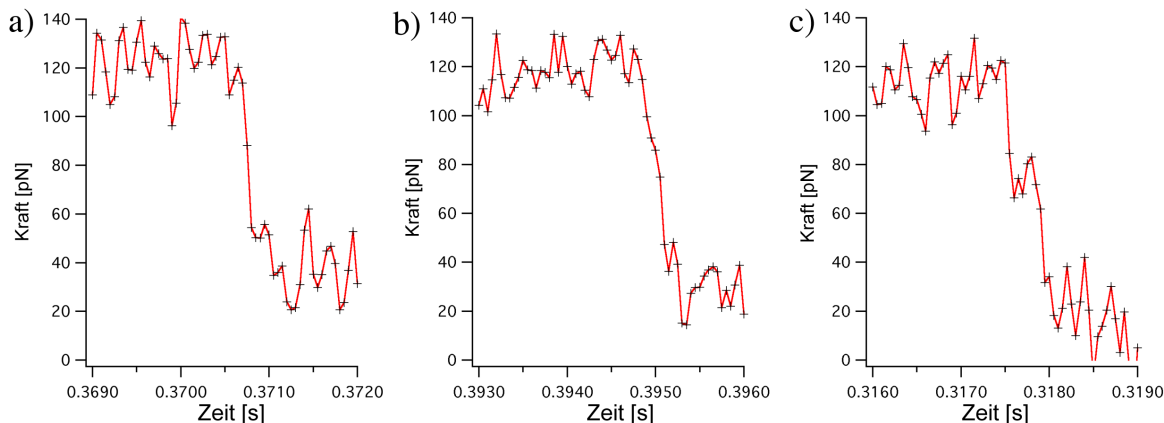


Abbildung 7.4: Die abgebildeten Kraft-Zeitsignale sind Vergrößerungen der GFP_{11-230} -Entfaltung mit anschließender Blattfeder-Relaxation. a) Normale Relaxation gemäß einem überdämpften Oszillator. Die Lebensdauer des Objektes ist hier offenbar Null. b) Die Blattfeder wird in ihrer Relaxation aufgehalten, und zwar etwa ein oder zwei Datenpunkte lang. c) Der Zwischenzustand war in diesem Fall etwa sechs Datenpunkte lang stabil und klar detektierbar. Die Rauschamplituden vor und nach der Blattfeder-Relaxation überlappen sich nicht mit der des Zwischenzustandes.

Die Lebensdauern des Zwischenzustandes in einzelnen DdFLN(1-5)-GFP-Kraftkurven sind punktweise mit einer Genauigkeit von $\pm 50\mu s$ ausgezählt worden. Abb. 7.5 zeigt die so erhaltene Verteilung der Lebensdauern.

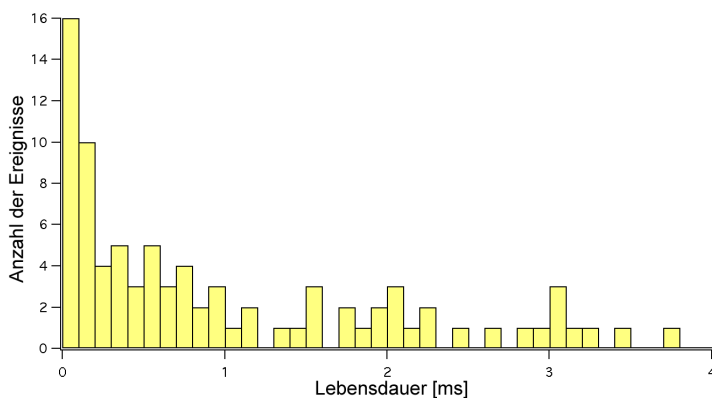


Abbildung 7.5: Verteilung gemessener Lebensdauern des Zwischenzustandes.

Der Mittelwert der gemessenen Lebensdauern liegt bei $\langle t \rangle = 1.3$ ms. Die Verteilung stellt keinen einfach-exponentiellen Verlauf dar. Abb. 7.6 verdeutlicht diesen Umstand. Dazu sind Zerfallsfunktionen der Form $P(t) = A \cdot \exp(-t/\tau)$ an verschiedene Bereiche der Lebensdauer-Verteilung angepasst worden. Um die Verteilung mit einer einzigen Zerfallskonstante τ erklären zu können, gibt es zu viele Ereignisse mit sehr kurzen Lebensdauern und zu viele Ereignisse mit vergleichsweise langen Lebensdauern.

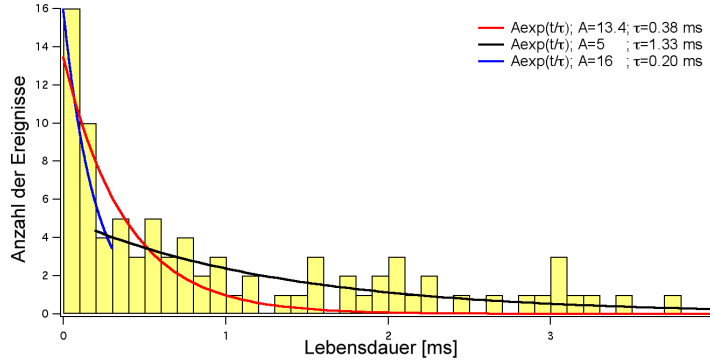


Abbildung 7.6: Anpassung einfach-exponentieller Zerfallsfunktionen an verschiedene Bereiche der gemessenen Lebensdauer-Verteilung. Die Verwendung von nur einer Zerfallskonstante τ führt zu keiner guten Übereinstimmung mit der experimentellen Verteilung.

7.4 Kraftabhängigkeit von Lebensdauern

Um die erhaltene Verteilung von Lebensdauern erklären zu können, müssen die jeweils auf den Zwischenzustand wirkenden Kräfte beachtet werden. Der Zwischenzustand wird im folgenden als S_1 bezeichnet.

In der beobachteten experimentellen Situation kann eine Näherung gemacht werden, die die Interpretation der gemessenen Lebensdauer-Verteilung vereinfacht: die an den Zwischenzustand angelegten Kräfte können als zeitlich konstant angesehen werden. Betrachtet man den Mittelwert der Lebensdauern des Zwischenzustandes, so verursacht eine über das zwischen Blattfeder und Oberfläche eingespannte Polypeptid angelegte Krafttrampe $F(t) = \eta_{eff} t$ mit den in Abschnitt 6.7 bestimmten Werten ($k_c=30$ pN/nm; $k_p=6$ pN/nm; $v_p=300$ nm/s) eine Kraftsteigerung in diesem Zeitraum von nur 2 pN. Angesichts der mindestens um eine Größenordnung größeren thermischen Rauschamplitude kann diese Steigerung vernachlässigt werden. Abschnitt 7.6 liefert ein weiteres Argument für diese Näherung.

Aus den in Abschnitt 3.6 vorgestellten Überlegungen würde man für eine zeitlich konstante Kraft eine mittlere Lebensdauer des Zwischenzustandes S_1 von

$$\tau_1(F) = \frac{1}{k_1^0} \cdot \exp\left(-\frac{F d_{S_1-T_2}}{k_B T}\right) \quad (7.1)$$

erwarten. Dabei beschreibt k_1^0 die Übergangsrate des Zustandes S_1 ohne Kraftbelastung. $d_{S_1-T_2}$ ist durch die Differenz der die Zustände S_1 und T_2 charakterisierenden End-zu-End Abstände gegeben, d.h. $d_{S_2-T_2} = d(T_2) - d(S_2)$. T_2 ist dabei der Übergangszustand für den Übergang $S_1 \rightarrow S_u$. Die Wahrscheinlichkeit $P_F(t_{exp})$ dafür, bei einer bestimmten Kraft F eine Lebensdauer t_{exp} des Zwischenzustandes zu messen, ist näherungsweise gleich zu der Wahrscheinlichkeit $P_{S_1 \rightarrow S_u}$ dafür, dass der Zwischenzustand den Übergang $S_1 \rightarrow S_u$ bis zum Zeitpunkt t unter dieser Kraftbelastung noch nicht vollzogen hat:

$$P_{S_1 \rightarrow S_u}(t, F) = \frac{1}{\tau_1(F)} \cdot \exp\left(-\frac{t}{\tau_1(F)}\right) \simeq P_F(t_{exp}) \quad (7.2)$$

Das gilt nur näherungsweise, da experimentell ein blinder Fleck vorliegt: die Zeitspanne t_{relax} , die die Blattfeder benötigt, um bis auf das durch den Zwischenzustand vorgegebene Kraftniveau zu relaxieren. D.h. die Zeit t , die der Zwischenzustand tatsächlich existierte, ist durch $t = t_{exp} + t_{relax}$ gegeben. Es können z.B für die beobachteten Ereignisse mit $t_{exp} = 0$ keine Aussagen über t (außer, dass vermutlich $0 < t < t_{relax}$ gilt) getroffen werden. Des weiteren ist die Kraftbelastung $F(t)$ im Zeitraum bis zu $t = t_{relax}$ unklar. Jedoch ist die Zeitspanne t_{relax} von der Größe der vorliegenden Ungenauigkeit der Zeitmessung ($t_{relax} \simeq 50\mu s$), so dass Gl. 7.2 gültig ist.

Liegt eine Wahrscheinlichkeitsverteilung $g(F)$ für das Auftreten, zwar zeitlich konstanter, jedoch verschieden hoher Kräfte vor, mit denen der Zustand S_1 belastet wird, so müssen die Einzelwahrscheinlichkeiten $P_F(t_{exp})$ für die Lebensdauern unter einer bestimmten Kraft entsprechend gewichtet werden. Die Wahrscheinlichkeit $P(t_{exp})$ dafür, eine bestimmte Lebensdauer des Zustandes S_1 zu messen, ist in einem solchen Fall folgendermaßen gegeben:

$$P(t_{exp}) = \frac{1}{C} \int_0^{\infty} \frac{g(F)}{\tau(F)} \exp\left(-\frac{t_{exp}}{\tau(F)}\right) dF \quad (7.3)$$

Die Konstante C sollte so gewählt werden, dass $P(t_{exp})$ z.B. auf die Gesamtzahl der in einem Experiment beobachteten Ereignisse normiert werden kann. Die Auswirkungen einer solchen Kraftverteilung $g(F)$ auf $P(t_{exp})$ werden aus Abb. 7.7 deutlich. Dort sind für drei verschieden breite Kraftverteilungen $g(F)$ mittels Gl. 7.3 die Wahrscheinlichkeiten für das Auftreten bestimmter Lebensdauern t_{exp} berechnet worden. Für eine δ -förmige Kraftverteilung erhält man erwartungsgemäß eine einfach-exponentielle Lebensdauerverteilung $P(t)$ (blaue Kurven). Mit zunehmender Breite der Kraftverteilung weicht das Verhalten von $P(t_{exp})$ jedoch deutlich davon ab. Die entstehenden Lebensdauerverteilungen könnten für doppelt-exponentiell gehalten werden. Gleiches gilt für die experimentell ermittelte Lebensdauer-Verteilung in Abb. 7.6.

Die Lage des Zentrums einer solchen Kraftverteilung $g(F)$ beeinflusst ebenfalls die Wahrscheinlichkeit, eine bestimmte Lebensdauer t_{exp} zu messen. Dies wird aus Abb. 7.8 deutlich. Je höher die Kräfte, umso wahrscheinlicher wird erwartungsgemäß das Auftreten kurzer Lebensdauern. Trotzdem verbleibt auch im Falle hoher Kräfte eine gewisse Wahrscheinlichkeit dafür übrig, unverhältnismäßig lange Lebensdauern zu beobachten.

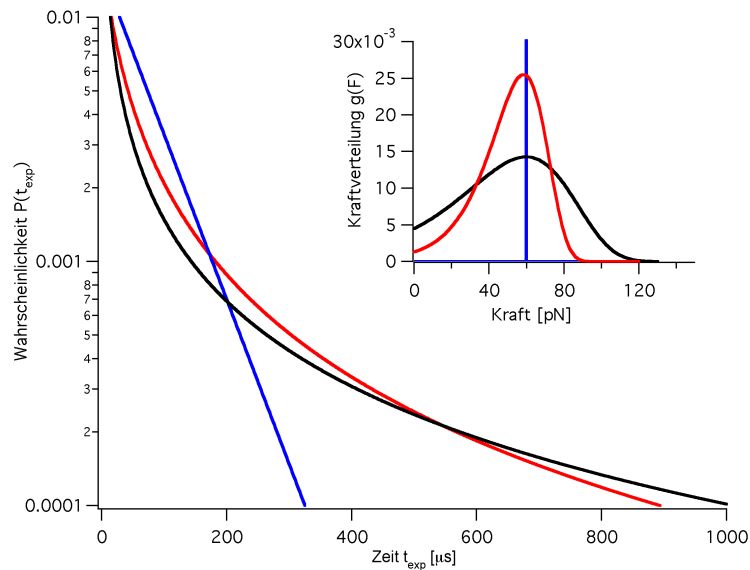


Abbildung 7.7: Die Wahrscheinlichkeiten $P(t_{exp})$ für das Auftreten bestimmter Lebensdauern t_{exp} sind mittels Gl. 7.3 berechnet worden. Dabei wurden die entsprechend farblich kodierten Kraftverteilungen $g(F)$ verwendet. Der betrachtete Zustand hat die Parameter $d=0.4$ nm und $k_{off}^0=50$ s $^{-1}$.

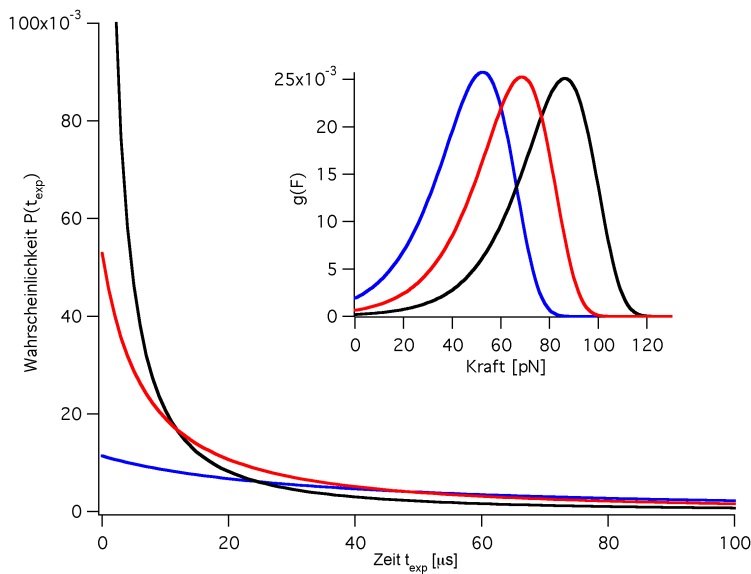


Abbildung 7.8: Wahrscheinlichkeiten $P(t_{exp})$ aus Gl. 7.3 für drei verschiedene Kraftverteilungen $g(F)$, deren Zentrum verschoben wurde.

Die in Abb. 7.5 gezeigte Verteilung von Lebensdauern zeigt ein ähnliches Verhalten wie die berechneten Verteilungen in den Abbildungen 7.7 und 7.8. Experimentell lag ja auch eine breite Kraftverteilung $g(F)$ vor: jeder Zwischenzustand wurde aufgrund der schwankenden Entfaltungskraft von GFP_{11-230} mit unterschiedlichen (konstanten) Kräften belastet. Es sollte daher möglich sein, die experimentell erhaltene Lebensdauer-Verteilung mittels Gl. 7.3 zu reproduzieren.

Dazu muß die experimentell vorliegende Kraftverteilung $g(F)$ bestimmt werden. Wie aus den bisherigen Überlegungen hervorgeht, reagieren die Lebensdauer-Verteilungen empfindlich auf Lage und Breite einer Kraftverteilung. Es ist daher wünschenswert, die Kräfte, die jeweils am Zwischenzustand wirkten, so genau wie möglich zu kennen. Die Datenpunkte, die die Existenz des Zwischenzustandes nachweisen, sind dafür nicht geeignet. Ihre geringe Anzahl würde nur eine Genauigkeit im Rahmen der thermischen Rauschamplitude (> 20 pN) erlauben.

Ausgehend von der Annahme, dass es sich bei dem beobachteten Zwischenzustand jedesmal um ein strukturell gleiches Objekt handelt, können jedoch die gut bestimmten Entfaltungskräfte von GFP_{11-230} verwendet werden: die Kraft, bei der GFP_{11-230} entfaltet, determiniert die Kraft, mit der der Zwischenzustand S_1 belastet wird. Wenn das Molekül nach der Entfaltung von GFP_{11-230} in den Zustand S_1 übergeht und sich dabei jedesmal um die gleiche Anzahl n Aminosäuren verlängert, so ist in diesem Fall die Kraft, mit der der Zustand S_1 belastet wird, durch den Schnittpunkt zwischen Blattfeder-Relaxation und einer WLC-Kurve mit einer um n Aminosäuren länger gewordenen Konturlänge bestimmt. Abb. 7.9 verdeutlicht diese Sichtweise.

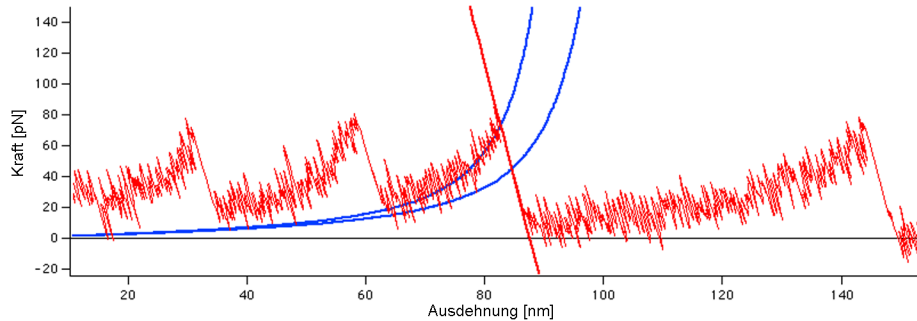


Abbildung 7.9: Der Längenzuwachs ΔL beim Übergang von GFP_{11-230} in den Zwischenzustand S_1 ist jedesmal gleich. Die auf S_1 wirkende Kraft ist daher durch den Schnittpunkt zwischen Blattfeder-Relaxation und einer um ΔL verlängerten WLC-Kurve bestimmt.

Die Blattfeder-Relaxationsphase nach einem Entfaltungsereignis bei der Kraft F_0 und der Ausdehnung d_0 ist in guter Näherung durch

$$d(F) = d_0 + \frac{1}{k_c}(F_0 - F) \quad (7.4)$$

gegeben. Das Kraft-Ausdehnungsverhalten eines bei dem Entfaltungsereignis um ΔL länger

gewordenen Polymers kann durch das WLC-Modell beschrieben werden:

$$F(d) = \frac{k_B T}{p} \left(\frac{d}{L_0 + \Delta L} + \frac{1}{4 \left(1 - \frac{d}{L_0 + \Delta L}\right)^2} - \frac{1}{4} \right) \quad (7.5)$$

Durch Einsetzen von Gl. 7.4 in Gl. 7.5 kann die Kraft bestimmt werden, bei der die Blattfeder aus der Relaxationsphase wieder in eine Phase eintritt, bei der das Molekül der Länge $L_0 + \Delta L$ wieder stetig gespannt wird. Dies ist zugleich die Kraft, mit der der Zwischenzustand S_1 belastet wird. Aus den bereits bestimmten Kräften F_0 und den Längen L_0 bei denen jeweils die Entfaltung von GFP_{11-230} beobachtet wurde, lassen sich so mit derselben Genauigkeit die Kräfte bestimmen, mit denen nachfolgend der Zwischenzustand S_1 belastet wurde. Dazu muss jedoch der Längenzuwachs ΔL , der mit dem Übergang $GFP_{11-230} \rightarrow S_1$ verbunden ist, bekannt sein.

In diesem Sinne kommunizieren die Entfaltungsereignisse miteinander: die Kraft und der Molekülzustand¹, bei dem GFP_{11-230} entfaltet, bestimmen die Lebensdauer des danach zu beobachtenden Zwischenzustandes. Je niedriger F_0 und je kleiner L_0 , umso länger kann der Zwischenzustand S_1 im Mittel leben. Diese Zusammenhänge werden im Anhang A näher diskutiert.

7.5 Längenzuwachs beim Übergang $GFP_{11-230} \rightarrow S_1$

Die Zunahme der Konturlänge ΔL beim Konformations-Übergang $GFP_{11-230} \rightarrow S_1$ ist durch Anpassen von berechneten WLC-Kurven an die Kraftkurven bestimmt worden. Die niedrige Zahl an Datenpunkten, die in jeder Kraftkurve für die Anpassung an den Zwischenzustand zur Verfügung stand, verschlechtert die Güte der Anpassung. Entsprechend breit ist die erhaltene Verteilung von Längenzunahmen ΔL für den Übergang $GFP_{11-230} \rightarrow S_1$. Sie ist in Abb. 7.10 angegeben.

Die in Abb. 7.10 gezeigte Verteilung weist ein deutliches Maximum bei $\Delta L = 7$ nm auf. Es treten jedoch auch Ereignisse auf, bei denen der gemessene Zuwachs der Konturlänge nicht mit einer Streuung um einen Mittelwert von etwa 7 nm vereinbar ist. Die Ursache dieser Asymmetrie wird in Abschnitt 7.7 besprochen. Passt man dem symmetrischen Teil der Verteilung eine Gaußsche Glockenkurve an, so liegt deren Zentrum bei einer Konturlängenzunahme von $\Delta L = 6.8 \pm 0.6$ nm.

Dieser Wert für die Zunahme der Konturlänge beim Übergang $GFP_{11-230} \rightarrow S_1$ von $\Delta L = 6.8 \pm 0.6$ nm kann mit einem Vorschlag für die Struktur des Zustandes S_1 verbunden werden. Vor dem Übergang in den Zustand S_1 befindet sich das GFP im Zustand GFP_{11-230} mit der Konturlänge:

$$L(I) = d_{11,230} \quad (7.6)$$

Bei dem Übergang in den Zustand S_1 können nur N- oder C-terminal n Aminosäuren von der GFP_{11-230} -Struktur entfaltet werden. Die Konturlänge im Zustand S_1 ist deshalb

¹Der Molekülzustand bestimmt die Konturlänge L_0 .

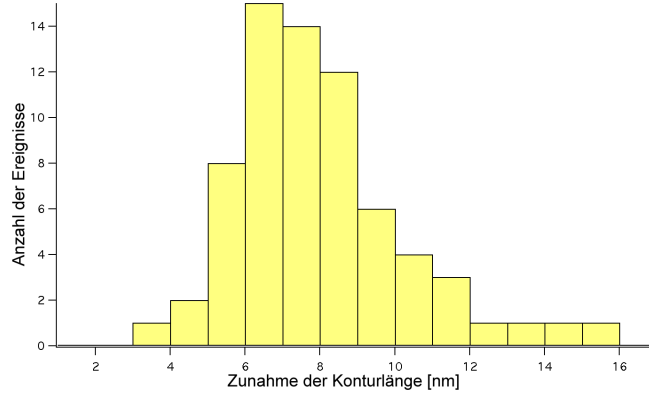


Abbildung 7.10: Verteilung der Konturlängenzunahmen beim Übergang $GFP_{11-230} \rightarrow S_1$ in DdFLN(1-5)-GFP-Kraftkurven.

folgendermaßen gegeben:

$$L(II) = n \cdot d_{AS} + d_{11+n,230} \text{ oder } L(II) = n \cdot d_{AS} + d_{11,230-n} \quad (7.7)$$

Die Abstände $d_{11+n,230}$ bzw. $d_{11,230-n}$ sind die Abstände zwischen C_α -Atomen der durch die Indizes definierten Aminosäuren des GFP. d_{AS} ist der Beitrag zur Konturlänge einer einzelnen Aminosäure (vgl. Gl. 6.2). Damit gibt $\Delta L = L(II) - L(I)$ den tatsächlich messbaren Konturlängenzuwachs wieder und kann aus den Strukturdaten 1B9C als Funktion der Anzahl entfalteter Aminosäuren n berechnet werden. Abb. 7.11 zeigt das Ergebnis dieser Berechnung.

Eine N-terminale Entfaltung von 14 ± 2 Aminosäuren oder eine C-terminale Entfaltung von 21 ± 2 Aminosäuren würde demnach den gemessenen Wert von $\Delta L = 6.8 \pm 0.6$ nm reproduzieren. Der Zustand S_1 wäre demnach entweder durch GFP_{25-230} oder durch GFP_{11-209} gegeben. Abb. 7.12 zeigt diese Strukturen.

Wie aus der Abbildung ersichtlich, ist mit beiden Vorschlägen die Entfaltung eines kompletten β -Faltblattes inklusive der darauf folgenden *loop*-Strukturen verbunden. Der Rest stellt dann die jeweils die mögliche Struktur des Zustandes S_1 dar.

7.6 Interpretation der Lebensdauern

In diesem Abschnitt soll die Verteilung der Lebensdauern des Zustandes S_1 mit dem Zwei-Zustandsmodell aus Abschnitt 3.5 interpretiert und die ihn charakterisierenden Parameter k_1^0 und $d_{S_1-T_2}$ bestimmt werden. Dabei beschreibt k_1^0 die Zerfallsrate des Zustandes S_1 in den Zustand S_u , während $d_{S_1-T_2} = d(T_2) - d(S_1)$ ist. T_2 bezeichnet den Übergangszustand zwischen S_1 und S_u .

Die erhaltene Verteilung an Lebensdauern wurde für die Interpretation bezüglich des Zustandes S_1 leicht verändert. Die Ereignisse, bei denen Längenzunahmen beim Übergang $GFP_{11-230} \rightarrow S_1$ von $\Delta L > 11$ nm verzeichnet wurden, sind aus der Statistik entfernt worden. Damit wird nur die Verteilung der Lebensdauern der Zwischenzustände interpretiert,

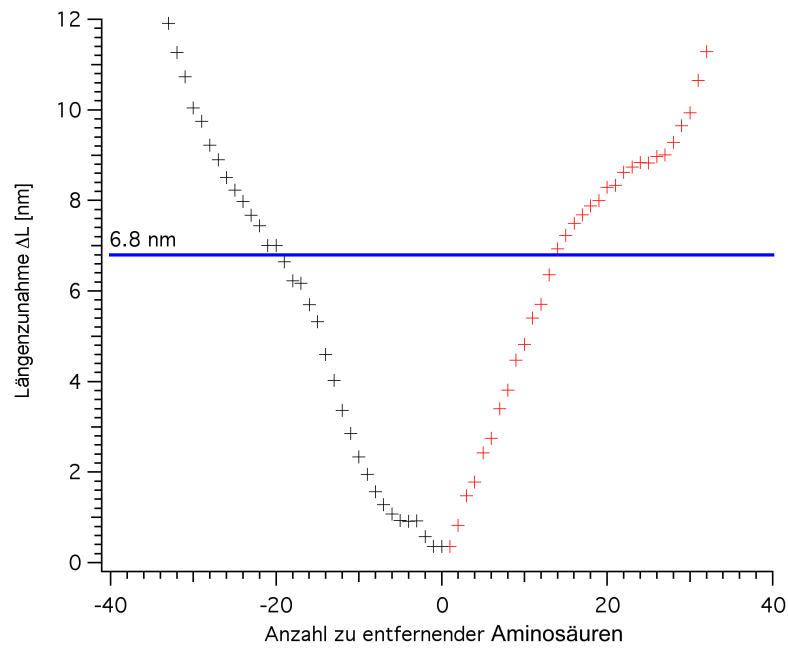


Abbildung 7.11: Erwartete Konturlängenzunahme ΔL als Funktion der Anzahl n von GFP_{11-230} abgetrennter Aminosäuren. Positive Aminosäuren-Achse: Abtrennung erfolgt N-terminal. Negative Aminosäuren-Achse: Abtrennung erfolgt C-terminal. Ebenfalls eingetragen: Gemessener Konturlängenzuwachs von $\Delta L = 6.8 \pm 0.6$ nm

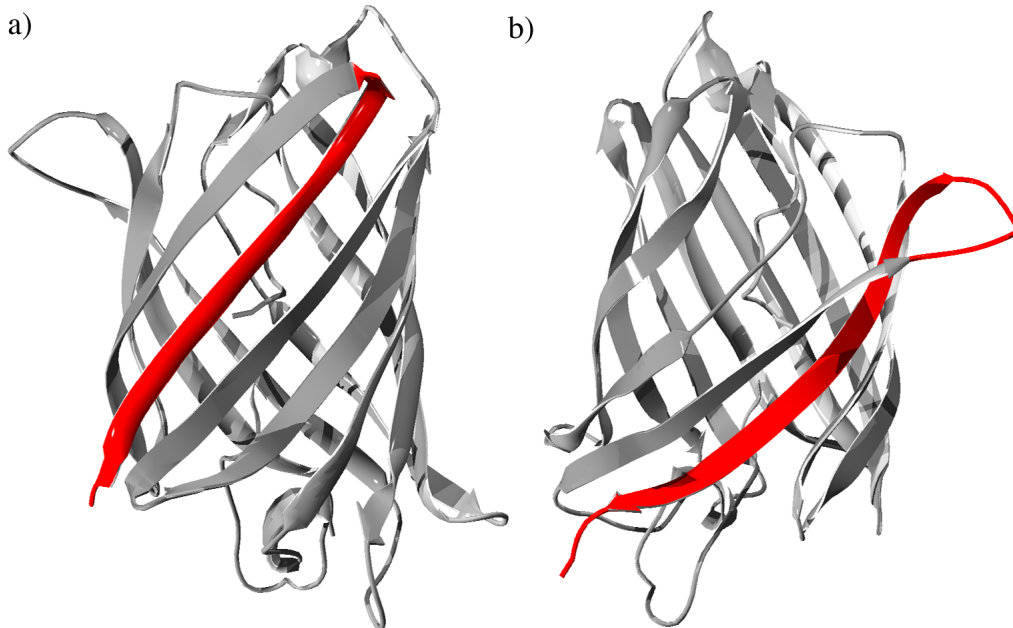


Abbildung 7.12: Strukturbild GFP_{11-230} (aus pdb-Datei 1B9C). a) Rot markiert: N-terminal 14 Aminosäuren. b) Rot markiert: C-terminal 21 Aminosäuren

bei denen die gemessene Zunahme der Konturlänge ΔL gleichverteilt um den Mittelwert 6.8 nm schwankte. Für die Anpassung einer berechneten Lebensdauer-Verteilung $P(t_{exp})$ wie in Gl. 7.2 an die Lebensdauer-Verteilung für den Zustand S_1 muss noch die für S_1 gültige Kraftverteilung $g(F)$ ermittelt werden. Mit der in Abschnitt 7.4 beschriebenen Vorgehensweise sind aus den Daten über die Kräfte F_0 und Konturlängen L_0 , bei denen eine Entfaltung des GFP_{11-230} beobachtet wurde und mit der Zunahme der Konturlänge von $\Delta L = 6.8$ nm die Kräfte berechnet worden, mit denen der Zustand S_1 jeweils belastet wurde. Ihre Verteilung $g(F)$ ist in Abb. 7.13 abgebildet. Mit der vorliegenden Kraftvertei-

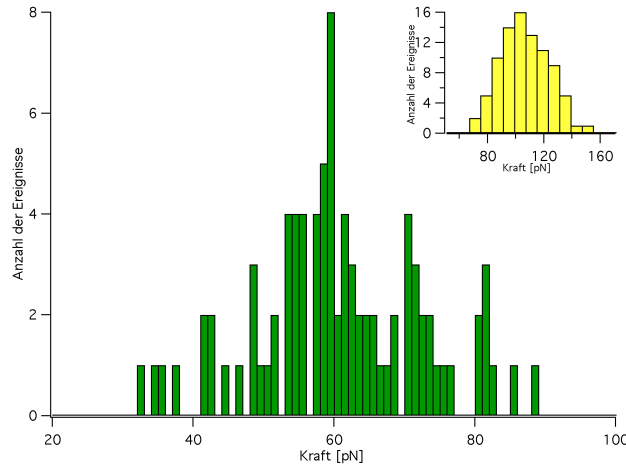


Abbildung 7.13: Verteilung $g(F)$ der Kräfte, die auf den Zustand S_1 gewirkt haben. Sie wurde rechnerisch aus den GFP_{11-230} -Entfaltungskräften gewonnen. Kleiner Graph: Verteilung GFP_{11-230} -Entfaltungskräfte.

lung wurde die Anpassung der durch Gl. 7.2 vorhergesagten Lebensdauer-Verteilung $P(t)$ an die gemessene Verteilung vorgenommen. $P(t)$ wurde dabei auf die Gesamtzahl der gemessenen Datenpunkte ($n=80$) normiert. Eine gute Übereinstimmung von $P(t)$ mit den Messdaten ergibt sich für die Parameter:

$$d_{S_1-T_2} = 0.55 \pm 0.1 \text{ nm}; \quad k_1^0 = 0.1 \text{ s}^{-1}; \quad \frac{\Delta k_1^0}{k_1^0} \simeq 10 \quad (7.8)$$

Ein Konsistenztest kann noch durchgeführt werden: wenn die Form der Verteilung der Lebensdauern tatsächlich nur auf die Kraftverteilung $g(F)$ zurückzuführen ist, so sollte sich dies in den Daten widerspiegeln, wenn sie nach der Höhe der jeweils am Zustand S_1 angreifenden Kraft sortiert würden. Für Abb. 7.15 ist dies durchgeführt worden. In Abb. 7.15a) ist die Verteilung der gemessenen Lebensdauern abgebildet, wie sie sich ergibt, wenn nur Datenpunkte benutzt werden, bei denen die wirkende Kraft kleiner als 60 pN war. In Abb. 7.15b) ist dies für die Datenpunkte getan worden, bei denen die zugehörige Kraft größer als 60 pN war. Erwartungsgemäß sind die Lebensdauern in b) sehr viel kürzer als in a). Gleichzeitig ist für diese Abbildung jeweils wieder eine Verteilung $P(t)$ gemäß Gl. 7.2 berechnet worden, wobei jeweils nur die untere (a) bzw. obere (b) Hälfte der Verteilung

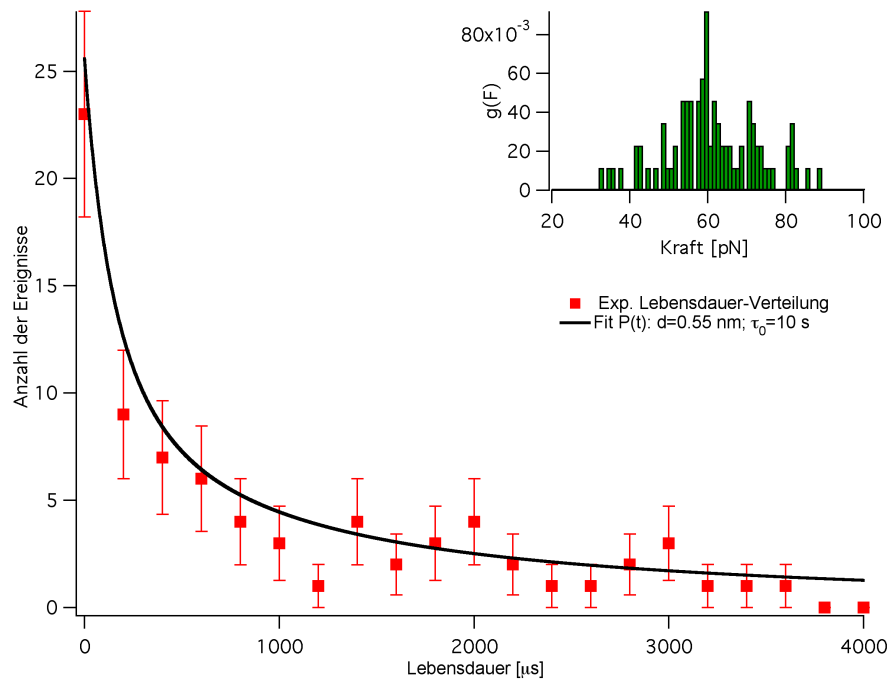


Abbildung 7.14: Anpassung einer berechneten Lebensdauer-Verteilung $P(t)$ laut Gl. 7.2 an die gemessene Verteilung. Kleiner Graph: in die Berechnung eingegangene Kraftverteilung $g(F)$.

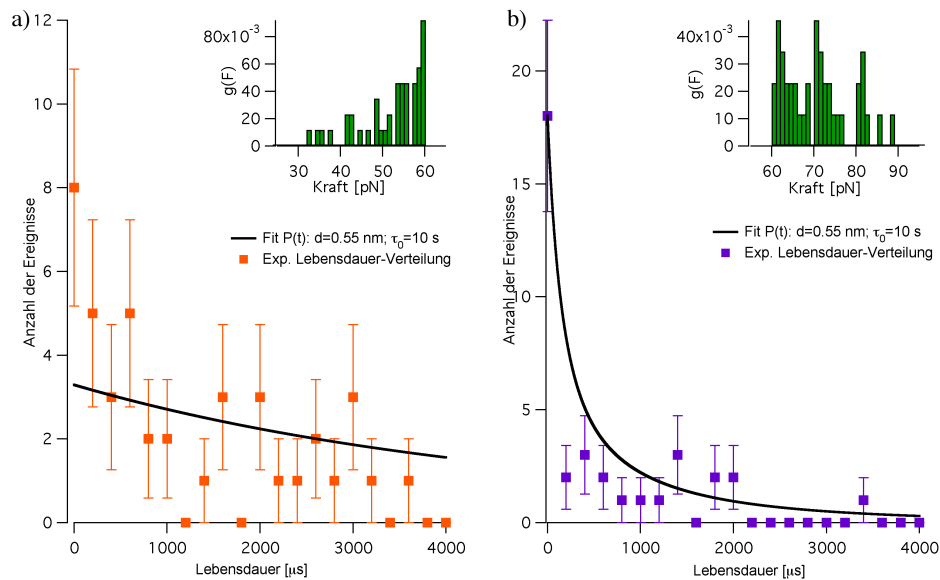


Abbildung 7.15: Anpassung einer berechneten Lebensdauer-Verteilung $P(t)$ laut Gl. 7.2 an nach Kraftbereichen sortierte experimentelle Verteilungen. Kleiner Graph: in die Berechnung eingegangene Kraftverteilung $g(F)$.

der auf S_1 wirkenden Kräfte $g(F)$ verwendet wurde. $P(t)$ wurde dabei jeweils wieder auf die Gesamtzahl der jeweiligen Datenpunkte normiert. Wählt man die gleichen Parameter wie für Abb. 7.14, so ergibt sich auch in diesen Fällen eine gute Übereinstimmung zwischen den Verteilungen $P(t)$ und den Messdaten.

Demnach konnte mit den in Gl. 7.8 angegebenen Parametern und unter Einbeziehung einer experimentell erhaltenen Kraftverteilung $g(F)$ die gemessene Verteilung der Lebensdauern des Zustandes S_1 mit einem analytischen Ausdruck reproduziert werden. Dazu wurde angenommen, der Zustand S_1 spüre aufgrund seiner kurzen Lebensdauer nur zeitlich konstante Kraftbelastungen, welche zudem über die Entfaltung von GFP_{11-230} und den dabei entstehenden Längenzuwachs ΔL determiniert werden.

Zur Überprüfung dieser Interpretation sind die Übergänge $GFP_{11-230} \rightarrow S_1 \rightarrow S_u$ mit den Methoden aus Abschnitt 4.6 simuliert worden. Die Eigenschaften der zwei Zustände GFP_{11-230} und S_1 sind dabei mit den Parametern $d_{GFP_{11-230}-T_1}$, k_{off}^0 und $d_{S_1-T_2}$, k_1^0 über Gl. 4.3 beschrieben worden. Der Konturlängenzuwachs ΔL beim Übergang $GFP_{11-230} \rightarrow S_1$ wurde zu 6.8 nm angenommen. Eine so erhaltene simulierte Kraftkurve ist in Abb. 7.16 angegeben. Die Simulation ist weitgehend auf die experimentelle Situation angepasst wor-

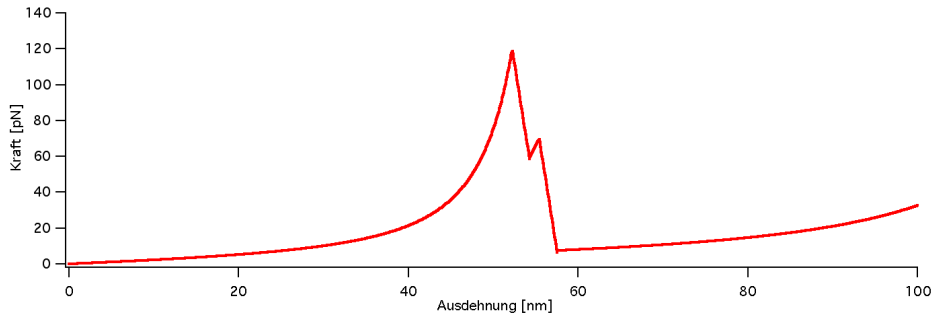


Abbildung 7.16: Simulierte Kraft-Abstandskurve an einem Polypeptid mit der Konturlänge $L_0 = 60$ nm. Es finden zwei Übergänge (Enfaltungsereignisse) statt: $GFP_{11-230} \rightarrow S_1$ und $S_1 \rightarrow S_u$.

den. Die Zeitschrittlänge Δt entsprach einer virtuellen Samplingrate von 20 kHz, wie sie auch experimentell verwendet wurde. Die Geschwindigkeit v_p wurde wie im Experiment zu 300 nm/s gewählt. Es wurden für jede experimentell beobachtete Konturlänge L der Moleküle (die Verteilung dieser Konturlängen ist in Abb. 6.9b) gezeigt) und der im jeweiligen Experiment verwendeten Federkonstante k_c virtuelle Kraftkurven erzeugt und die Kraft F_0 notiert, bei der die Simulation den Übergang $GFP_{11-230} \rightarrow S_1$ voraussagte. Ebenfalls bestimmt wurden die Kräfte $F_{intersect}$, mit denen der Zustand S_1 erstmalig belastet wurde, die Kräfte F_S , bei denen der Übergang $S_1 \rightarrow S_u$ stattfand und die Lebensdauern τ_1 des Zustandes S_1 . Unter Verwendung der bereits im vorigen Kapitel bestimmten, den Zustand GFP_{11-230} charakterisierenden Parametern $d_{GFP_{11-230}} = 0.28$ nm und $k_{off}^0 = 0.07$ s⁻¹ und mit den Parametern aus Gl. 7.8 kann die gemessene Lebensdauerverteilung des Zustandes

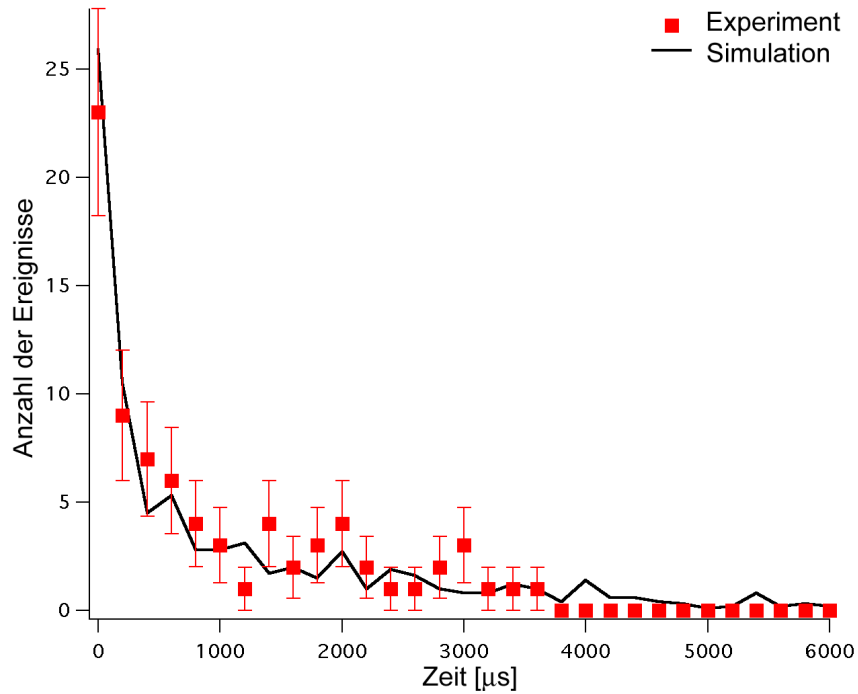


Abbildung 7.17: Aus einer Monte-Carlo Simulation erhaltene Lebensdauer-Verteilung des Zustandes S_1 im Vergleich mit der gemessenen Lebensdauer-Verteilung.

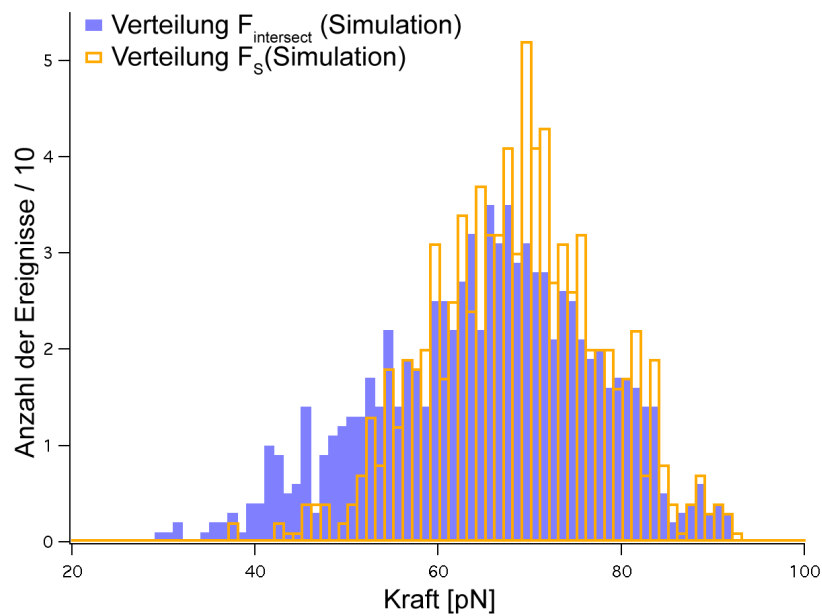


Abbildung 7.18: Aus einer Monte-Carlo Simulation erhaltene Verteilungen der Kräfte $F_{intersect}$ und der Kräfte F_S

S_1 sehr gut mit der Simulation reproduziert werden. Abb. 7.17 vergleicht das Ergebnis aus der Simulation mit den Messdaten.

Die Simulation liefert auch ein weiteres Argument für die in Abschnitt 7.4 getätigte Annahme, der Zustand S_1 spüre im wesentlichen nur zeitlich konstante Kräfte.

Dazu ist in Abb. 7.18 die Verteilung der Kräfte $F_{intersect}$, mit denen der Zustand S_1 erstmalig mit der Blattfeder auslesbar belastet wird, abgebildet. Lebt der Zustand kurz genug, so sollte sich die Verteilung der Kräfte F_S , bei denen der Zustand S_1 zerfällt, davon nur wenig unterscheiden. Dies geht auch aus Abb. 7.18 hervor. Beide Verteilungen stimmen im hohen Kraftbereich so gut wie überein, während es im niedrigen Kraftbereich in der F_S -Verteilung eine leichte Verschiebung zu höheren Kräften gibt. Das ist zu erwarten, da der Zustand S_1 bei niedrigen Kräften deutlich länger lebt und daher die Näherung einer konstanten Kraftbelastung schlechter wird. Jedoch kann im wesentlichen tatsächlich von einer zeitlich konstanten Kraftbelastung gesprochen werden.

Das Erscheinungsbild des Zustandes S_1 in den Kraftkurven vermittelt den Eindruck eines sehr instabilen Zustandes. Jedoch liegt die natürliche Lebensdauer dieses Zustandes nach den Ergebnissen dieses Abschnittes 7.6 im Bereich von 10 s und ist damit ähnlich zu der des Zustandes GFP_{11-230} . Simuliert man nur den Übergang $S_1 \rightarrow S_u$ mit den Parametern aus Gl. 7.8 in einem Polypeptid der festen Konturlänge $L_0=60$ nm, so erhält man eine Verteilung für die Entfaltungskräfte mit einem Maximum bei etwa 55 pN. Abb. 7.19 zeigt diese Verteilung.

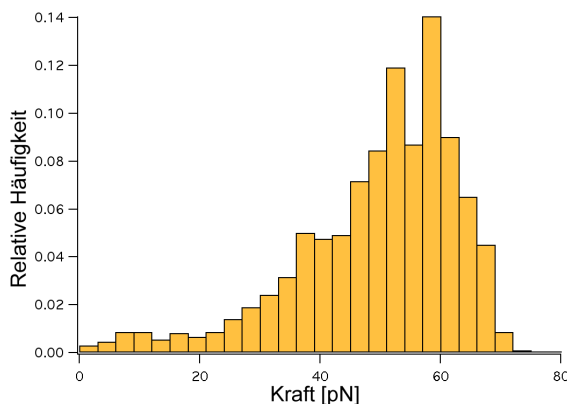


Abbildung 7.19: Aus einer Monte-Carlo Simulation erhaltene Verteilung der Entfaltungskräfte für den Zustand S_1 , wenn er einzeln auftreten könnte.

Würde GFP_{11-230} im Mittel bei niedrigeren Kräften entfalten (z.B. bei etwa 20 pN, dies könnte vielleicht durch Mutationen erreicht werden), so könnte der Zustand S_1 seine eigene, spezifische Kraftverteilung erzeugen und würde als stabiler Zustand in den Kraftkurven erscheinen. In der vorliegenden Situation wird seine Lebensdauer aber völlig durch die vorhergehende GFP_{11-230} -Entfaltung bestimmt.

7.7 Ein weiterer Zwischenzustand S_2

Die Asymmetrie der Verteilung des Längenzuwachses beim Übergang $GFP_{11-230} \rightarrow S_1$ in Abb. 7.10 wirft die Frage auf, ob es sich dabei überhaupt um nur einen Übergang handelt. Die Rauschamplitude der für die Messung der Lebensdauern verwendeten Blattfedern (Olympus Bio-Lever Typ A) verhindert die Klärung dieser Frage. Es sind daher hochauflösende Messungen (Samplingrate 20kHz) mit dem Olympus Bio-Lever Typ B an der DdFLN(1-5)-GFP-Chimäre durchgeführt worden. Die experimentellen Bedingungen sind in Abschnitt 4.4 aufgeführt. Aufgrund der um einen Faktor 5 niedrigeren Federkonstante im Vergleich zum Bio-Lever Typ A weist dieses Blattfeder-Modell ein um den Faktor $\sqrt{5}$ erniedrigtes thermisches Krafrauschen auf. Jedoch zeigt dieses Modell auch eine um den Faktor 5 niedrigere Resonanzfrequenz, so dass er für eine Analyse von Lebensdauern nicht geeignet ist - t_{relax} wäre schon im Bereich der zu messenden t_{exp} und Gl. 7.2 nicht mehr gültig. Die Frage, ob es sich nur um einen (S_1) oder mehrere Zwischenzustände auf dem Wege des Überganges $GFP_{11-230} \rightarrow S_u$ handelt, konnte jedoch durch die Verwendung dieses Blattfeder-Modells geklärt werden. Dazu sind in Abb. 7.20 Ausschnitte dreier Kraft-Zeitkurven abgebildet, die an der DdFLN(1-5)-GFP-Chimäre gemessen wurden. Abb. 7.20a) zeigt eine normale Zwei-Zustands-Entfaltung. Die in Abb. 7.20b) gezeigte

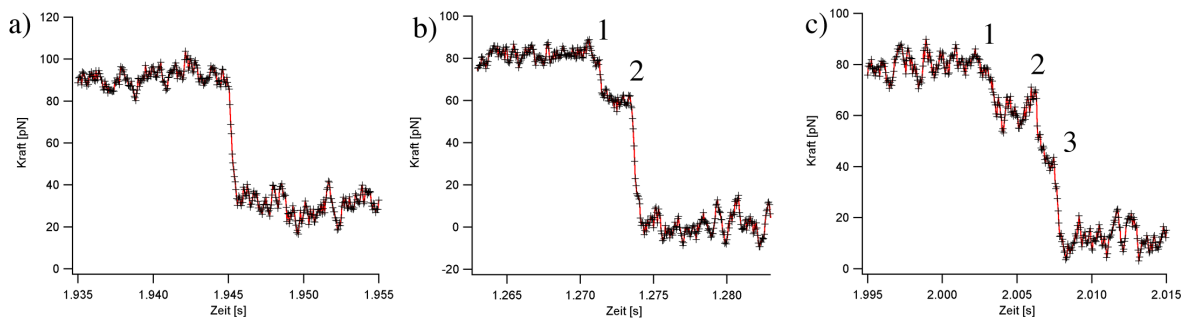


Abbildung 7.20: Vergrößerung an der DdFLN(1-5)-GFP-Chimäre mit dem Bio-Lever Typ B gemessener Kraft-Zeitsignale im Bereich von Entfaltungsereignissen - a) Die Blattfeder-Relaxation zeigt keine Unterstrukturen (Entfaltung einer DdFLN-Domäne); b) (1) Entfaltung von GFP_{11-230} - (2) Zwischenzustand; c) (1) Entfaltung von GFP_{11-230} - (2) Zwischenzustand - (3) Weiterer Zwischenzustand.

GFP_{11-230} -Entfaltung weist den Zwischenzustand S_1 nach. Etwas neues zeigt jedoch Abb. 7.20c): nach dem Zerfall des Zwischenzustandes S_1 tritt eine erneute Abweichung von einer normalen Relaxation auf. Die zugehörigen Datenpunkte können nicht durch thermisches Krafrauschen aus dem ersten Zwischenzustand erklärt werden. Es liegt auch keine Einstrreuung durch thermisches Rauschen auf dem völlig relaxierten Niveau bei 10 pN vor. Die Lebensdauer dieses zweiten Zwischenzustands beträgt im gezeigten Fall etwa 13 Datenpunkte ($650 \mu s$). Dies ist die längste Lebensdauer, die für diesen zweiten Zwischenzustand gemessen werden konnte. Dieser Zwischenzustand wird im folgenden als S_2 bezeichnet. Es wird sich hierbei um das Ereignis handeln, dessen Detektion die Asymmetrie der Längenverteilung in Abb. 7.10 produziert hat. Passt man berechnete WLC-Kurven an

Kraftkurven an, die beide Ereignisse zeigen, so kann man die jeweiligen Längenzunahmen des Moleküls messen. Dazu ist exemplarisch eine solche Anpassung in Abb. 7.21 gezeigt.

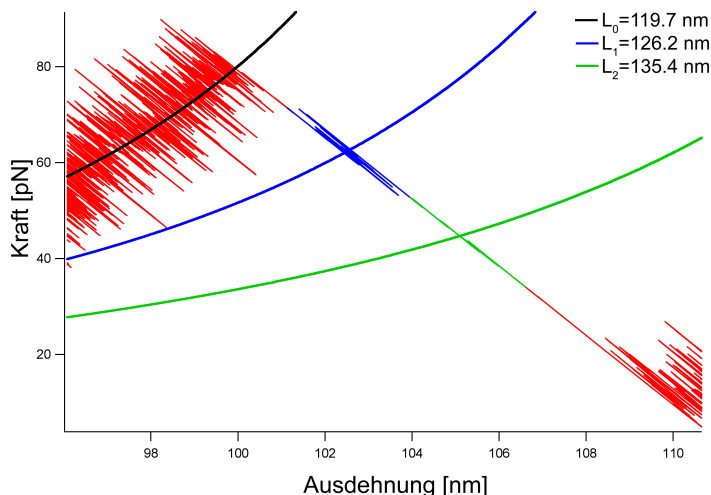


Abbildung 7.21: Anpassung des WLC-Modells an eine DdFLN(1-5)-GFP-Kraftkurve im Bereich der GFP_{11-230} -Entfaltung. Der erste Zwischenzustand folgt nach einer Längenzunahme von 6.5 nm. Der zweite Zwischenzustand folgt (von GFP_{11-230} aus berechnet) nach 15.7 nm. Passt man eine WLC-Kurve an die Datenpunkte beider Zwischenzustände an, so würde man eine Längenzunahme von etwa 10 nm messen.

Der Mittelwert der so gefitteten Längenzunahmen der Moleküle von der Entfaltung von GFP_{11-230} bis zum ersten Zwischenzustand S_1 aus den Messungen mit dem Bio-Lever Typ B beträgt 6.6 ± 0.7 nm und stimmt damit mit dem Mittelwert des symmetrischen Teils der Verteilung in Abb. 7.10 überein. Der nur sehr selten beobachtete zweite Zwischenzustand verursacht einen Konturlängenzuwachs von etwa 15 nm (gemessen vom Zustand GFP_{11-230} aus). Passt man das WLC-Modell an die Menge der Datenpunkte an, die beide Zwischenzustände umfassen, so würde man Konturlängenzunahmen im Bereich von etwa 10 nm erhalten (nicht abgebildet).

Die asymmetrische Struktur der in Abb. 7.10 gezeigten Verteilung der Längenzunahmen ΔL bis zu einem detektierten Zwischenzustand kann mit dem Auftreten zweier aufeinander folgender Zwischenzustände, S_1 und S_2 , erklärt werden. Diese sind jedoch nur mit dem Bio-Lever Typ B deutlich auflösbar. Der zweite Zwischenzustand S_2 ist zudem nur sehr selten detektierbar. Die aus den Messungen mit dem Bio-Lever Typ A erhaltene Längenverteilung in Abb. 7.10 spiegelt diesen Umstand deutlich wieder, obwohl die einzelnen Zwischenzustände in den Kraftkurven aufgrund der hohen Rauschamplitude nicht auflösbar sind. Die Verteilung zeigt zunächst einen symmetrischen Anteil. Passt man diesem symmetrischen Teil eine Gaußsche Glockenkurve an, so liegt deren Zentrum bei einer Konturlängenzunahme von $\Delta L = 6.8 \pm 0.6$ nm. Dieser Anteil wird offenbar durch die alleinige Detektion des ersten Zwischenzustandes S_1 verursacht. Die asymmetrische Schulter der Verteilung bei höheren Konturlängen beginnend bei etwa 10 nm entstand durch das

Anpassen des WLC-Modelles an Kraftkurven, bei denen die Datenpunkte in Wirklichkeit durch das Auftreten beider Zwischenzustände S_1 und S_2 verursacht worden sind. Schließlich wären da noch die sehr großen Längenzuwächse im Bereich von 12 bis 15 nm. Es stellt sich die Frage, ob bei diesen Ereignissen nicht nur der zweite Zwischenzustand S_2 aufgetreten ist, während der erste Zustand S_1 schon zerfallen war. Aus den Daten lässt sich dazu keine klare Antwort finden. Abb. 7.22 zeigt noch einmal Vergrößerungen zweier mit dem Bio-Lever Typ A gemessener GFP_{11-230} -Entfaltungen, bei denen sich die mit dem Zwischenzustand verbundene Zunahme der Konturlänge stark unterscheidet. Abb.

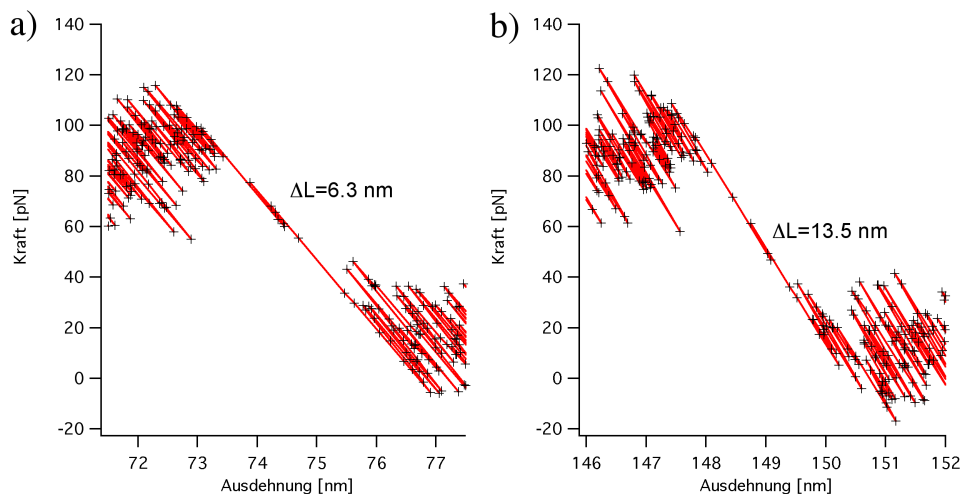


Abbildung 7.22: Zwei GFP_{11-230} -Entfaltungen, so wie sie mit dem Bio-Lever Typ A an der DdFLN(1-4)-GFP-Chimäre detektiert wurden. a) Datenpunkte des Zwischenzustandes entsprechen einem Konturlängenzuwachs von 6.3 nm; b) Datenpunkte des Zwischenzustandes entsprechen einem Konturlängenzuwachs von 13.5 nm

7.22a) zeigt wohl den ersten Zwischenzustand S_1 , was vor allem anhand des beobachteten Zuwachses der Konturlänge begründet werden kann. Ob Abb. 7.22b) nur den zweiten Zwischenzustand S_2 oder beide Zustände zeigt, ist nicht auflösbar. Die starke Zunahme der Konturlänge von 13.5 nm würde jedoch auf ersteres hindeuten.

Die gemessenen Lebensdauern t_{exp} des Zustandes S_1 sind also in Wirklichkeit Summen zweier Zeiten:

$$t_{exp} = t_1 + t_2 \quad (7.9)$$

Aus der Verteilung in Abb. 7.10 folgt, dass ungleich häufiger nur die reine Lebensdauer des Zustandes S_1 als die des Zustandes S_2 bzw. Summen beider Zeiten gemessen wurden. Das bedeutet, der Anteil von t_2 zu t_{exp} ist generell gering. Durch den Ausschluss von Datenpunkten, bei denen die zugehörige Längenzunahme ΔL größer als 11 nm war, wurde für die Analyse in Abschnitt 7.6 sichergestellt, dass die dort interpretierte Verteilung von Lebensdauern des Zustandes S_1 tatsächlich nur durch den Zustand S_1 verursacht wurde.

7.8 Energielandschaft - Teil 2

Mit den Ergebnissen aus diesem Kapitel kann ein weiterer Teil der Energielandschaft des GFP entlang des Pfades der mechanischen Entfaltung des GFP rekonstruiert werden. Es konnten die Übergänge $GFP_{11-230} \rightarrow S_1 \rightarrow S_2 \rightarrow S_u$ beobachtet werden. Unter einer Kraftbelastung von 10 pN ändert sich der End-zu-End Abstand des Moleküls beim Übergang $GFP_{11-230} \rightarrow S_1$ mit der in Abschnitt 7.5 gemessenen Längenzunahme von $\Delta L = 6.8$ nm um $\Delta d = 3.5$ nm. Die Position des Übergangszustands T_2 zwischen den Zuständen S_1 und S_2 wurde in Abschnitt 7.6 zu $d_{S_1-T_2} = 0.55$ nm bestimmt. Der Übergang $S_1 \rightarrow S_2$ führt zu einer Konturlängenzunahme von etwa 8 nm (siehe Abschnitt 7.7). Bei einer Kraft von 10 pN entspricht das einem Unterschied im End-zu-End Abstand Δd von etwa 4 nm. Damit sind die Positionen der Zustände auf der Reaktionskoordinate festgelegt.

Die energetische Lage des Übergangszustandes T_2 lässt sich wie in Abschnitt 6.8 über die Arrhenius-Gleichung (Gl. 3.17) abschätzen. Wählt man für den Vorfaktor in Gl. 3.17 wieder einen Wert von $k_T = 10^9 \text{ s}^{-1}$ [5][24][44], so ergibt sich mit der Zerfallsrate $k_1^0 = 0.1 \text{ s}^{-1}$ des Zustandes S_1 :

$$G^0(T_2) - G^0(S_1) \simeq 18 k_B T \quad (7.10)$$

Diese Ergebnisse werden zusammen mit den Ergebnissen aus Kapitel 6 im folgenden Kapitel in ein detailliertes Bild der Energielandschaft des GFP entlang des Pfades der mechanischen Entfaltung eingefügt.

7.9 Zusammenfassung

In diesem Kapitel wurden zwei weitere Zwischenzustände, S_1 und S_2 , der mechanischen GFP-Entfaltung nachgewiesen. Der Zustand S_1 wurde charakterisiert. Er tritt unter Bedingungen konstanter Kraftbelastung auf. Die Lebensdauern dieses Zustandes S_1 werden maßgeblich durch die Konturlänge L_0 des ganzen DdFLN(1-5)-GFP-Moleküls vor der Entfaltung von GFP_{11-230} und durch die Kräfte F_0 , bei denen GFP_{11-230} entfaltet, beeinflusst. Der Zustand S_1 hat eine natürliche Lebensdauer τ_1^0 von etwa 10 Sekunden. Es ergeben sich zwei mögliche Strukturen für den Zustand S_1 : bei dem Übergang $GFP_{11-230} \rightarrow S_1$ bricht entweder N- oder C-terminal ein kompletter β -Strang aus dem β -Fass des GFP heraus. Durch hochauflösende Messungen mit rauscharmen Blattfedern konnte der Zustand S_2 nachgewiesen werden.

Kapitel 8

Energielandschaft des GFP

8.1 Mechanischer Entfaltungspfad des GFP

In diesem Kapitel werden die in den vorangegangenen Kapiteln erhaltenen Ergebnisse über den Pfad der mechanischen Entfaltung des GFP in ein zusammenhängendes Bild eingefügt. Dazu ist an dieser Stelle nochmal ein Ausschnitt einer Kraftkurve (Abb. 8.1) gezeigt, bei der die Entfaltung des GFP alle in dieser Arbeit identifizierte Übergänge aufweist.

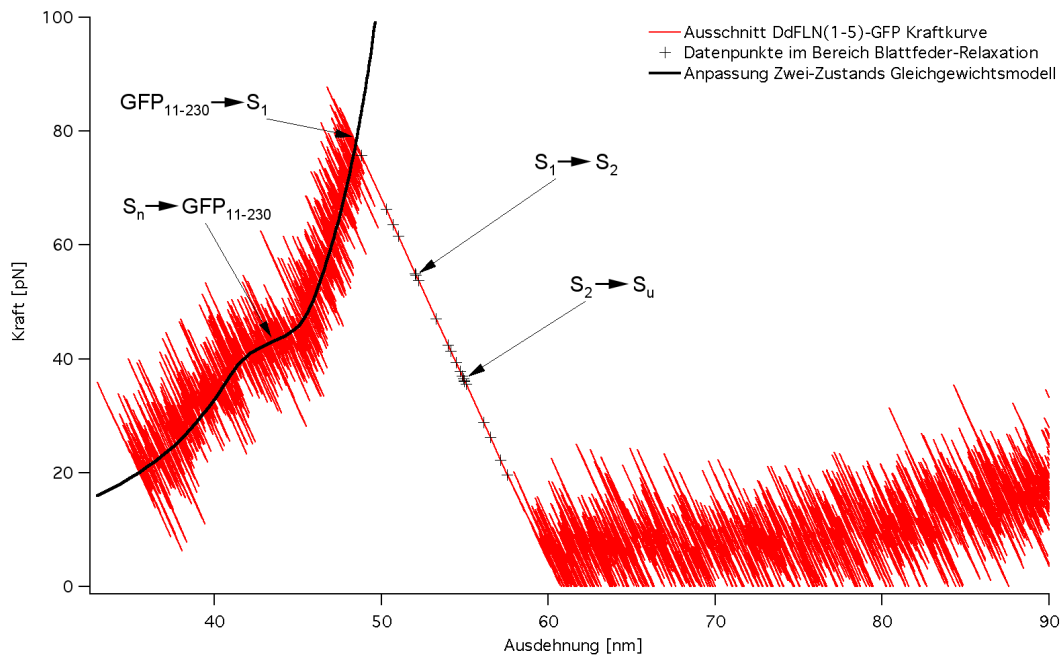


Abbildung 8.1: Ausschnitt einer DdFLN(1-5)-GFP-Kraftkurve: die Übergänge $S_n \rightarrow \text{GFP}_{11-230} \rightarrow S_1 \rightarrow S_2 \rightarrow S_u$ sind gekennzeichnet.

Demnach folgt das GFP einem komplizierten Entfaltungspfad über drei beobachtbare Zwischenzustände. Der erste Übergang entfaltet von der nativen Struktur S_n eine N-

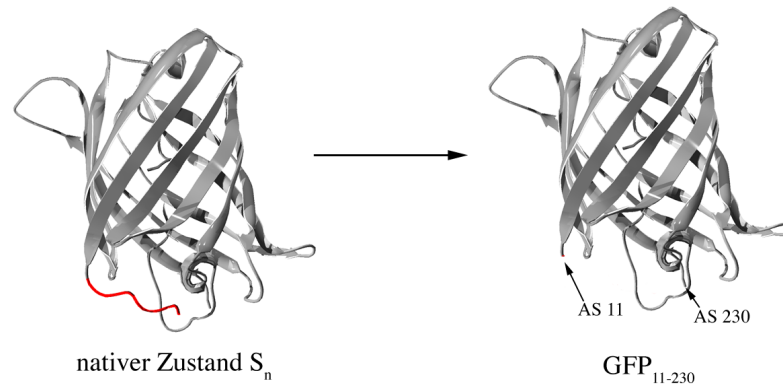


Abbildung 8.2: Schema des kraftinduzierten Überganges $S_n \rightarrow GFP_{11-230}$. *Random coil*-Elemente sind nicht abgebildet.

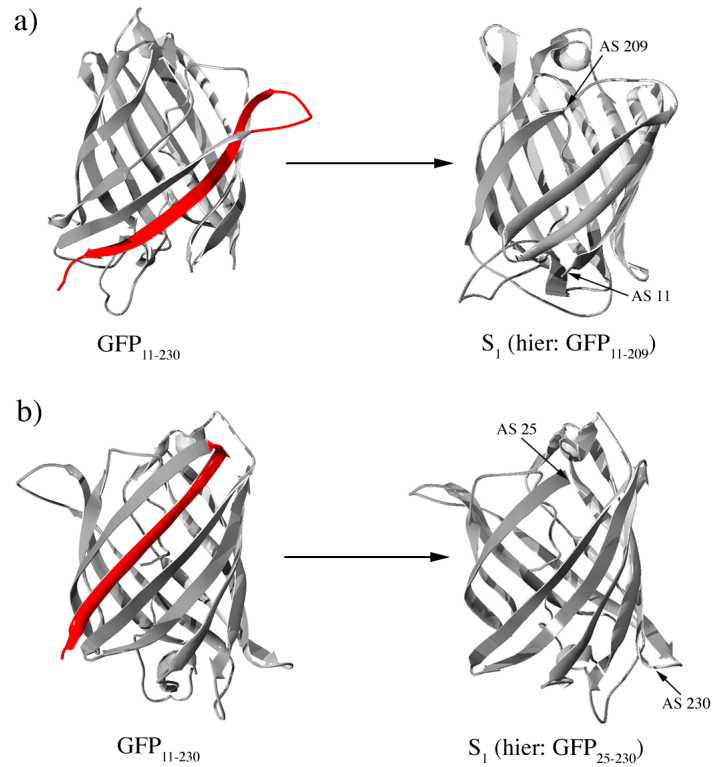


Abbildung 8.3: Schema des kraftinduzierten Überganges $GFP_{11-230} \rightarrow S_1$. Es konnten zwei mögliche Strukturen für den Zustand S_1 vorgeschlagen werden: a) GFP_{11-209} und b) GFP_{25-230} . *Random coil*-Elemente sind nicht abgebildet.

terminale α -Helix. Dieser Prozess findet nahe am thermischen Gleichgewicht statt. Die verbleibende Reststruktur wird durch GFP_{11-230} gebildet. Der Übergang ist schematisch in Abb. 8.2 gezeigt. Auf diesen Übergang folgt die Entfaltung von GFP_{11-230} in den Zustand S_1 . Für die Struktur des Zustandes S_1 konnten zwei Vorschläge gemacht werden: es handelt sich dabei entweder um GFP_{25-230} oder um GFP_{11-209} (siehe Abb. 8.3a) und b) für ein Schema dieser Übergänge). Der Übergang $GFP_{11-230} \rightarrow S_1$ findet fern vom thermischen Gleichgewicht statt und erzeugt Bedingungen konstanter Kraftbelastung für den Zustand S_1 . Schließlich konnten noch zwei weitere Übergänge beobachtet werden: $S_1 \rightarrow S_2$ und $S_2 \rightarrow S_u$. Über die Eigenschaften des Zustandes S_2 bezüglich seiner Struktur kann nur spekuliert werden. Der letzte beobachtete Übergang $S_2 \rightarrow S_u$ führt das Protein in den komplett entfalteten *random coil* Zustand.

8.2 Energielandschaft

Im Abschnitt 6.8 wurde bereits ein Teil der Energielandschaft des GFP unter einer Kraftbelastung von 10 pN rekonstruiert. Abschnitt 7.8 lieferte weitere Informationen über die Energielandschaft. Abb. 8.4 zeigt nun schematisch die komplette Energielandschaft entlang des Pfades der mechanischen Entfaltung des GFP unter einer Kraft von 10 pN, wie sie aus den Ergebnissen dieser Arbeit hervorgeht. Zur Veranschaulichung der mit den einzelnen Übergängen verbundenen strukturellen Veränderungen sind die Strukturen der jeweiligen Zustände abgebildet worden (pdb-Datei 1B9C). Für den Übergang $GFP_{11-230} \rightarrow S_1$ ist aus Gründen der Übersichtlichkeit nur eine der zwei vorgeschlagenen Strukturen von S_1 , nämlich GFP_{25-230} , abgebildet.

Es ist eine sehr raue Energielandschaft. Nach Erreichen des Zustandes S_u folgt die Energielandschaft dem durch das WLC-Modell vorgegebenen Verlauf. Die Zustände GFP_{11-230} und S_1 sind nur topologisch stabil. Die energetisch hohen Übergangszustände erklären ihre Lebensdauer unter Kraftbelastung. Bemerkenswert ist, dass die Entfaltung des stabilen Zustandes S_n aus einem Energieminimum heraus nur zu einem sehr kleinen mechanischen Effekt in den Kraftkurven führt, während die Entfaltung von GFP_{11-230} , dessen Energie über der des denaturierten Zustandes S_u liegt, zu einem großen mechanischen Effekt führt.

Der mechanische Entfaltungspfad selbst ist ebenfalls sehr rau: es ändert sich ständig die Richtung der anliegenden Kraft. Es ist davon auszugehen, dass sich das Protein bei jedem Übergang so dreht, dass die Richtung, die durch den End-zu-End Abstand der im jeweiligen Zustand noch stabilen Reststruktur des GFPs vorgegeben wird, parallel zur Richtung der anliegenden Kraft wird.

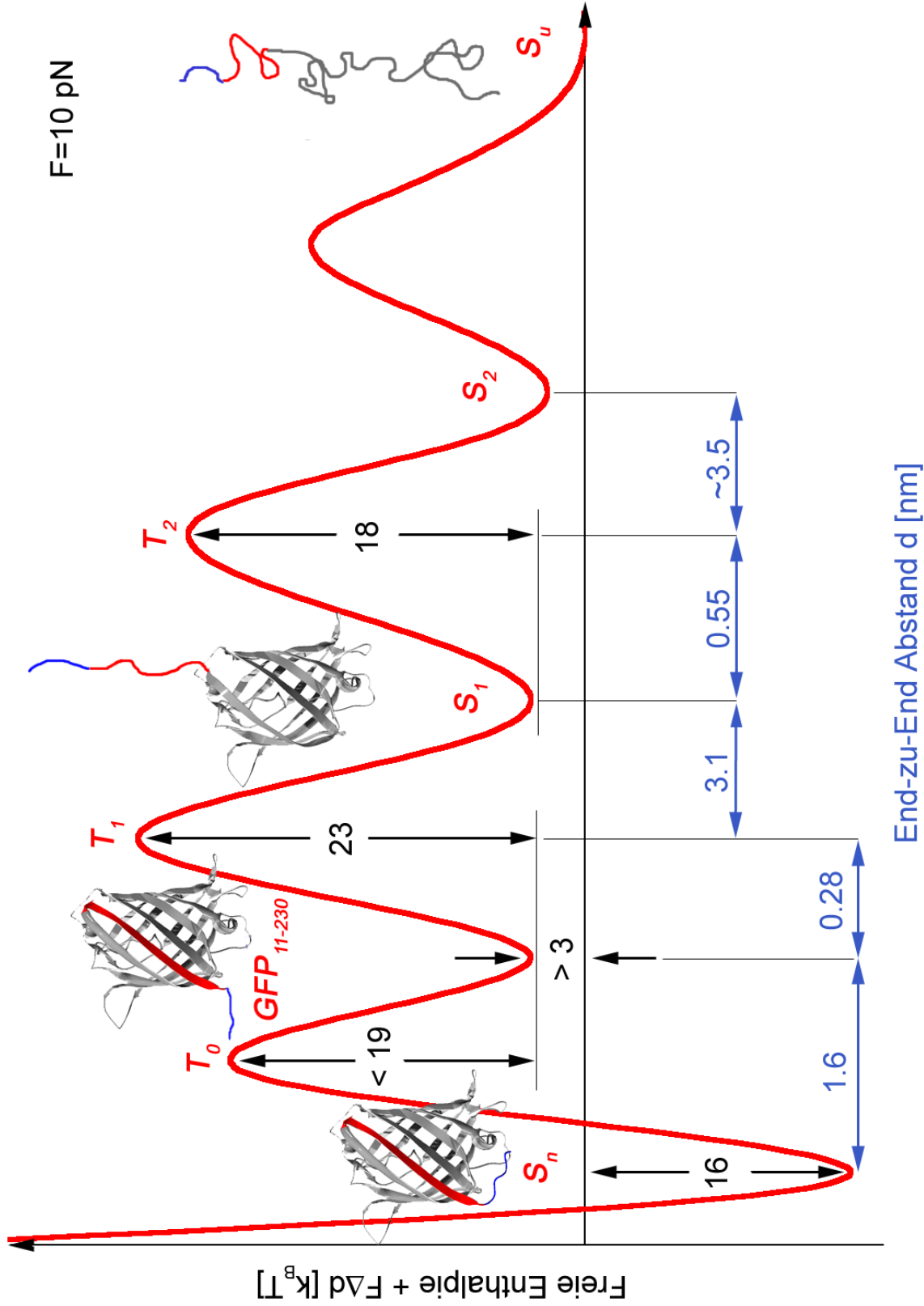


Abbildung 8.4: Schema der Energielandschaft des GFP entlang des Pfades der mechanischen Entfaltung unter einer Kraftbelastung von 10 pN. Die Energieachse ist in Einheiten $\Delta G^0 = \Delta G + F\Delta d$ angegeben. Die Reaktionskoordinate End-zu-End Abstand ist nicht maßstabsgerecht. Für den Zustand S_1 wurde aus Gründen der Übersichtlichkeit nur einer der zwei Strukturvorschlage abgebildet (GFP_{25-230}).

Kapitel 9

Ausblick

9.1 GFP als molekularer Kraftsensor

In der vorliegenden Arbeit wurde gezeigt, dass das Grün Fluoreszierende Protein mechanisch entfaltbar ist. Dies geschieht über mehrere Zwischenzustände. Mit dem Verlust der nativen Struktur geht der Verlust der Fluoreszenz des GFP einher. Dies ist auf Einflüsse des Lösungsmittels auf das Chromophor zurückzuführen. Das Lösungsmittel findet jedoch bereits Zugang zum Chromophor, wenn der erste mechanische Zwischenzustand, GFP_{11-230} , erreicht wird. Dabei wird eine α -Helix vom GFP entfaltet, die eine Art Deckel zur Fass-Struktur des GFP bildet. Führt dieser Übergang bereits zum Verlöschen der Fluoreszenz, so würde das GFP einen ausgezeichneten Kraftsensor darstellen: der Übergang ist reversibel auf einer Zeitskala von 0.1 s und findet bei Kräften im Bereich von 35 pN statt. Dies liegt im Bereich von auf molekularer Ebene physiologisch relevanten Kräften [19]. Der dabei erreichte Zwischenzustand des GFP hat eine natürliche Lebensdauer von nur 14 s: nach dieser Zeit kollabiert dieser Zustand und zerfällt in einen Zustand, der weiteren Zugang zum Chromophor für das Lösungsmittel bietet: ein β -Strang aus dem 11 β -Stränge umfassenden GFP-Fass bricht heraus. Dieser Zustand zerfällt seinerseits ohne Kraftbelastung auf einer Zeitskala von 10 s und der entfaltete Zustand des GFP wird erreicht, bei dem es definitiv zum Verlust der Fluoreszenz kommt. Es lässt sich nur spekulieren, wann es tatsächlich zum Verlöschen der Fluoreszenz kommt.

Mit den Ergebnissen aus dieser Arbeit kann das GFP als optisch auslesbarer molekularer Kraftsensor vorgeschlagen werden. Mechanische Kräfte haben auf der molekularen Ebene in der Biologie eine wichtige physiologische Bedeutung [2]. Beispielsweise sind molekulare Motoren verantwortlich für Prozesse, bei denen u.a. ganze Zellorganellen transportiert werden. Das Zusammenspiel der molekularen Komponenten des Zytoskeletts ist verantwortlich für die Fortbewegung von Zellen und spielt eine wesentliche Rolle bei der Zellteilung. Beim Ablesen von DNA durch Enzyme müssen mechanische Kräfte aufgewendet werden, um z.B. die DNA-Doppelhelix teilweise zu entwinden. Methoden, diese Prozesse *in vivo* in Bezug auf die auf molekularer Ebene wirkenden Kräfte zu untersuchen, stehen derzeit nicht zur Verfügung. Ist man jedoch in der Lage einzelne GFP-Moleküle zu beobachten, die mit

gentechnischen Methoden an geeigneter Stelle in diese Prozesse eingebunden worden sind, so könnten aus dem Verlöschen der Fluoreszenz vielleicht Rückschlüsse auf *in vivo* lokal wirkenden mechanischen Kräfte gezogen werden ziehen.

Es sind Mutanten des GFP erhältlich, die ein unterschiedliches Farbspektrum zeigen (z.B. Rot Fluoreszierendes Protein, Blau Fluoreszierendes Protein, Cyan Fluoreszierendes Protein), jedoch kaum Unterschiede in der Struktur im Vergleich zum GFP aufweisen. Ändert man z.B. durch Punktmutationen die mechanische Stabilität dieser Moleküle, so könnte auf diese Weise ein Spektrum verschieden stabiler Kraftsensoren erzeugt werden, deren Fluoreszenz als Indikator für einen bestimmten Kraftbereich fungiert.

Jedoch wäre es auch interessant, Kraftkurven an einzelnen GFP-Molekülen mit dem Fluoreszenzverhalten dieser Moleküle unter Kraft zu korrelieren. So könnte man Aufschluss über den genauen Zeitpunkt des Verlustes der Fluoreszenz einzelner GFP-Moleküle in Abhängigkeit der Kraft gewinnen und auf diesem Wege u.a. auch den Einfluss des Lösungsmittels auf das Chromophor studieren. Dies stellt jedoch eine große experimentelle Herausforderung dar.

9.2 Temperaturabhängigkeit der GFP-Stabilität

In dieser Arbeit wurden Ergebnisse nicht vorgestellt, die in einem anderen Temperaturbereich ($T \geq 30^\circ\text{C}$) erhalten wurden. Demnach zeigt das GFP in seiner Stabilität eine enorme Temperaturempfindlichkeit. Die Entfaltungskräfte für GFP_{11-230} sanken dabei im Mittel um etwa 40 pN im Vergleich zu denen bei $T = 20^\circ\text{C}$ ab. Die Breite der erhaltenen Verteilung blieb jedoch ähnlich. Die Lebensdauern des Zustandes S_1 verkürzten sich ebenfalls drastisch. Abb. 9.1 und Abb. 9.2 illustrieren dies. Der Rahmen dieser Arbeit erlaubte jedoch keine temperaturabhängige Untersuchungen. Mit solchen Untersuchungen ließen sich vielleicht noch mehr Informationen über die Energielandschaft und die grundlegenden Wechselwirkungen in der GFP-Struktur gewinnen.

Eine interessante Frage wäre z.B. das Verhalten der für die Entfaltung von GFP_{11-230} nötigen Kräfte bei physiologischen Temperaturen um 37°C . Abb. 9.1 legt nahe, dass das GFP bei solchen Temperaturen einen noch günstigeren Kraftsensor für die Detektion von Kräften auf molekularer Ebene darstellen könnte.

Ebenfalls interessant wären Untersuchungen bei niedrigen Temperaturen ($\sim 4^\circ\text{C}$). Es ist zu erwarten, dass die Lebensdauer des Zustandes S_1 stark ansteigen und der Zustand besser detektierbar wird, jedoch würden auch die Entfaltungskräfte von GFP_{11-230} ansteigen und diesen Effekt vielleicht wieder vernichten. Dies böte Gelegenheit, die in Anhang A kurz diskutierte Kommunikation zwischen aufeinander folgenden Entfaltungsereignissen näher zu untersuchen. Bei diesen niedrigen Temperaturen kann vielleicht auch der dritte hier beobachtete Entfaltungszwischenzustand S_2 näher charakterisiert werden.

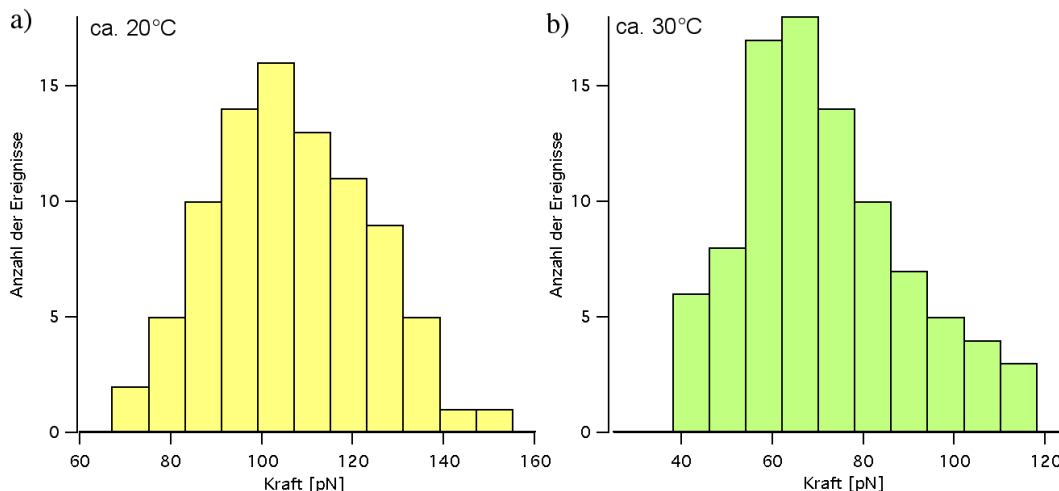


Abbildung 9.1: a) Verteilung der Entfaltungskräfte des Zustandes GFP_{11-230} bei etwa 20°C ; b) Verteilung der Entfaltungskräfte des Zustandes GFP_{11-230} bei etwa 30°C

9.3 Rückfaltung von GFP

Das GFP stellt die größte Tertiärstruktur dar, dessen Entfaltung bisher mit kraftspektroskopischen Methoden studiert wurde. Der mechanische Entfaltungspfad des GFP ist komplex und weist eine sehr raue Energielandschaft auf. Eine interessante Fragestellung wäre die Untersuchung der Rückfaltung des GFP in den nativen Zustand mit einem Kraftspektroskop. Es ist davon auszugehen, dass der dabei eingeschlagene Pfad ebenfalls Zwischenzustände aufweisen wird. Aufgrund der langsamen Faltungskinetik ($k_{on}^0 \simeq 0.002 \text{ s}^{-1}$ [41]) des Proteins ist dies aus Gründen der Anbindung an Oberfläche und AFM-Spitze experimentell allerdings sehr schwierig durchzuführen.

9.4 Weitere Stichpunkte

Der Übergang $S_n \rightarrow GFP_{11-230}$ findet nahe am thermischen Gleichgewicht statt: vielleicht ließe sich bei geeigneter Einrichtung des Messgerätes ein Hin- und Herwechseln des Proteins zwischen beiden Zuständen beobachten (wie z.B. an RNA-*hairpins* in [25]).

Die noch offene Frage, welche Struktur der Zustand S_1 tatsächlich annimmt, könnte z.B durch Mutationen in verschiedenen Regionen des GFP geklärt werden, wie es bereits in [38] gezeigt wurde. Es könnte auch untersucht werden, warum die beobachteten Zwischenzustände eine herausragende Stabilität besitzen, vielleicht könnte dies auf bestimmte Residuen in der Sequenz zurückgeführt werden.

Diese Arbeit hat ebenfalls gezeigt, dass die Untersuchung einzelner Proteindomänen als Teil eines Chimären-Proteins zwar die Untersuchung dieser Domänen mit dem Kraftspektroskop ermöglicht, jedoch mit niedriger Statistik zu kämpfen hat. Es würde sich daher empfehlen, kraftspektroskopische Messungen so weit wie möglich zu automatisieren und

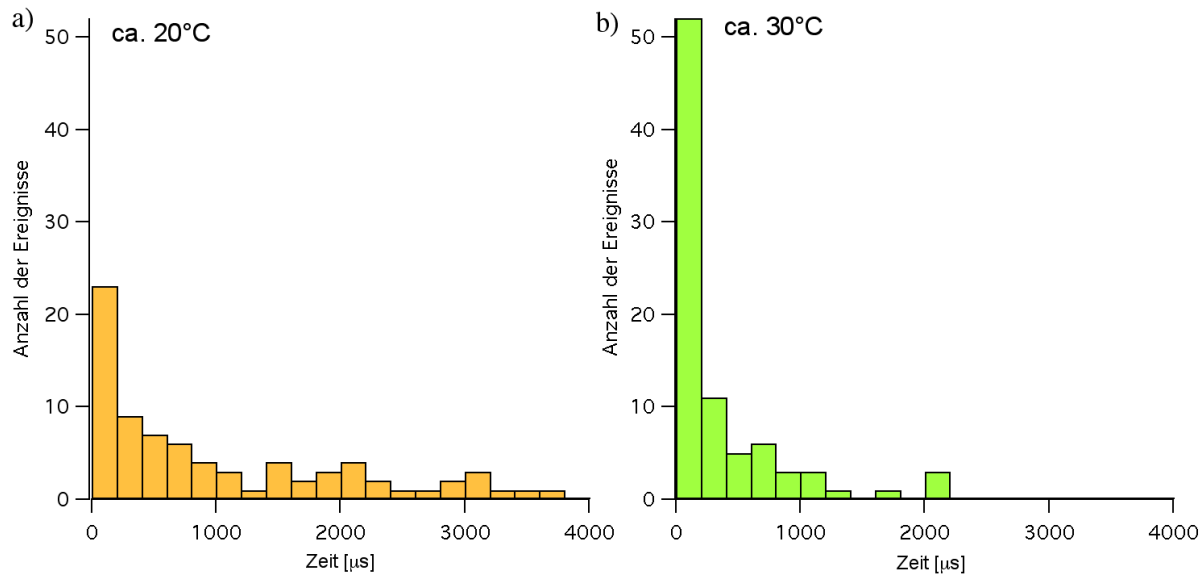


Abbildung 9.2: a) Verteilung der Lebensdauern des Zustandes S_1 bei etwa 20°C ; b) Verteilung der Lebensdauern des Zustandes S_1 bei etwa 30°C

die Messungen quasi im Dauerbetrieb durchzuführen. Für die Auswertung großer Datenmengen steht die in Anhang B vorgestellte Mustererkennung zur Verfügung.

Ein technischer Aspekt sei noch vermerkt: Die in dieser Arbeit untersuchten Effekte fanden an der Auflösungsgrenze des Instrumentes statt. Die verwendeten Blattfedern Olympus Bio-Lever Typ A und Typ B führten zu komplementären Informationen: der eine hatte die nötige Bandbreite (ω_0 gross), um kurzlebige Ereignisse zu detektieren und der andere die nötige Auflösung (k_c klein), um verschiedene Ereignisse von einander zu unterscheiden.

Für die Untersuchung der mechanischen Faltung und Entfaltung einzelner Proteine mit dem Kraftspektrometer gibt es einen klaren Bedarf an noch schnelleren und noch weicheeren Blattfedern: ohne sie ist die Grenze detektierbarer Ereignisse wie Zwischenzustände oder kurze Gleichgewichtsübergänge vermutlich mit den in dieser Arbeit aufgedeckten Übergängen so gut wie erreicht.

Anhang A

Kommunizierende Entfaltungseignisse

Aus Kapitel 7 ging hervor, dass Entfaltungseignisse nachfolgende Entfaltungen stark beeinflussen können. Im vorliegenden Fall wurden durch die Charakteristik der Entfaltung von GFP_{11-230} Bedingungen konstanter Kraft-Belastung für den Zustand S_1 erzeugt. Die Verteilung der Lebensdauern des Zustandes S_1 wurde darüber hinaus noch stark durch die Verteilung der Entfaltungskräfte von GFP_{11-230} beeinflusst. Der Zustand S_1 würde jedoch ohne die vorhergehende GFP_{11-230} -Entfaltung in Kraft-Abstandskurven wie ein normales Entfaltungseignis auftreten. Insofern hat die GFP_{11-230} -Entfaltung mit der S_1 -Entfaltung kommuniziert.

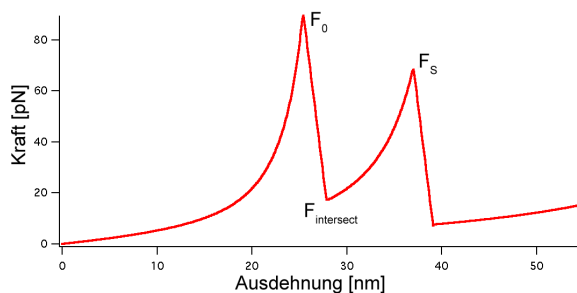


Abbildung A.1: Simulierte Kraftkurve mit zwei Entfaltungen: $S_a \rightarrow S_b$ und $S_b \rightarrow S_c$

In diesem Abschnitt soll daher qualitativ kurz diskutiert werden, inwiefern sich aufeinander folgende Konformationsübergänge, $S_a \rightarrow S_b$ und $S_b \rightarrow S_c$, in ihren Kraftverteilungen generell beeinflussen können. Als F_0 seien die Kräfte bezeichnet, bei denen der Übergang $S_a \rightarrow S_b$ stattfindet. $F_{intersect}$ seien die Kräfte, mit denen der Zustand S_b in einem Experiment erstmalig messbar belastet würde. F_S bezeichne die Kräfte, bei denen der Übergang $S_b \rightarrow S_c$ stattfindet (siehe Abb. A.1). $F_{intersect}$ stellt offensichtlich die untere Grenze der Kräfte F_S dar, bei denen überhaupt der Übergang $S_b \rightarrow S_c$ in einer Kraftkurve beobachtbar wird. Über diese Kraft kommuniziert der Zustand S_a mit dem Zustand S_b .

Der Effekt hängt stark vom relativen Konturlängenzuwachs $\Delta L/L$ des Moleküls beim

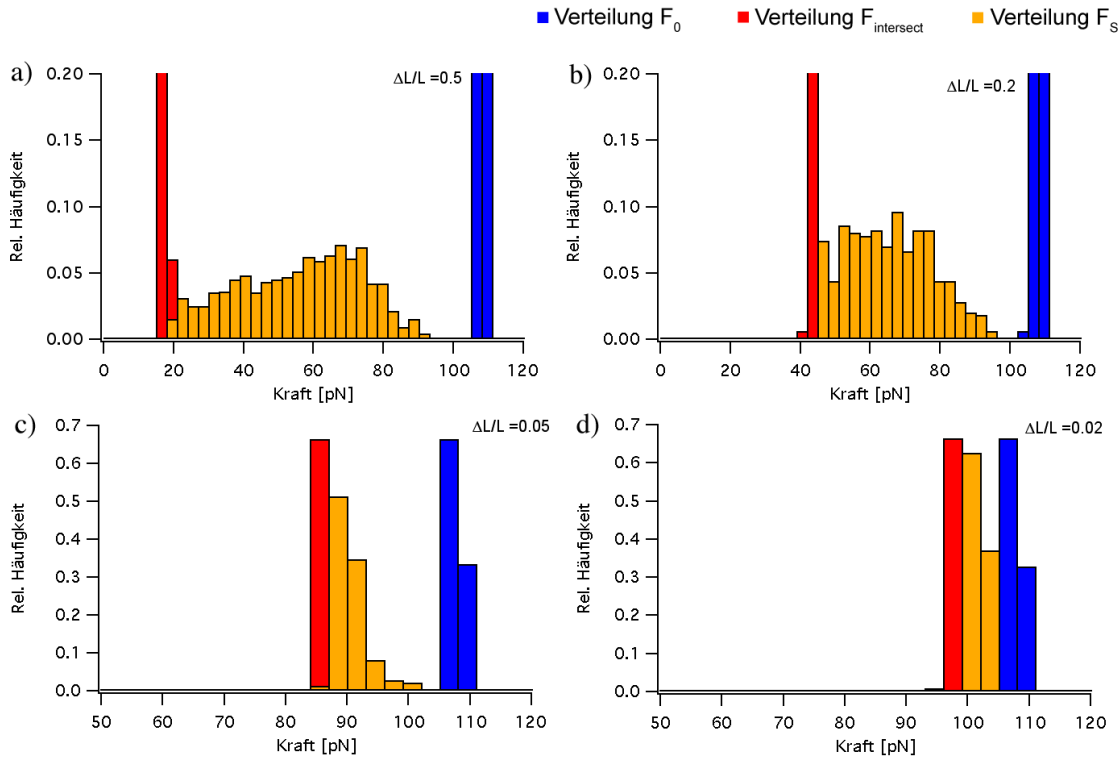


Abbildung A.2: Kraftverteilungen aus einer Simulation von zwei aufeinanderfolgenden Entfaltungen $S_a \rightarrow S_b$ und $S_b \rightarrow S_c$. Variiert wurde $\Delta L/L$ bei einer scharfen Kraftverteilung F_0 für den Übergang $S_a \rightarrow S_b$. Bild a) zeigt noch eine individuelle Kraftverteilung F_S für die Entfaltung des Zustands S_b . Die Bilder b) und c) zeigen den Übergang zu Bild d), wo die Verteilung der Kräfte F_S komplett durch die Entfaltungskräfte von S_a bestimmt wird.

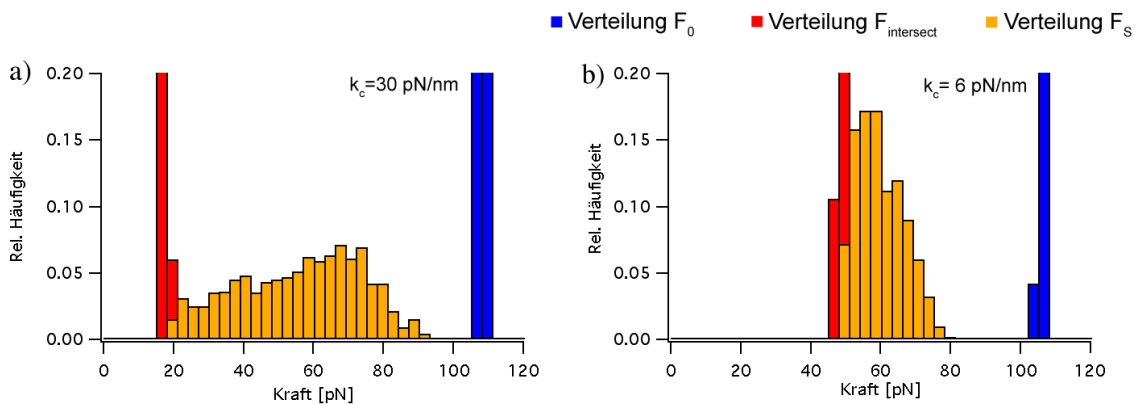


Abbildung A.3: Kraftverteilungen aus einer Simulation von zwei aufeinanderfolgenden Entfaltungen $S_a \rightarrow S_b$ und $S_b \rightarrow S_c$. Variiert wurde die relative Zunahme der Konturlänge $\Delta L/L$ nach der Entfaltung von S_a .

Übergang $S_a \rightarrow S_b$ ab. Um dies zu verdeutlichen, ist eine Simulation durchgeführt worden, bei der die Kraftverteilung F_0 für den Übergang $S_a \rightarrow S_b$ als sehr eng vorgegeben wurde (womit die Verteilung für $F_{intersect}$ ebenfalls sehr eng wird).

Verringert man $\Delta L/L$, so steigen die Kräfte $F_{intersect}$ an und führen dazu, dass die Verteilung für die Kräfte F_S der nachfolgenden Entfaltung $S_b \rightarrow S_c$ stark deformiert wird. Dies wird aus Abb. A.2a)-d) deutlich. Der Übergang von einer individuellen Kraftverteilung für die Entfaltung des Zustandes S_b , hin zu Bedingungen, wo der Zustand S_b nur noch durch den Zustand S_a bestimmte, konstante Kräfte spürt, ist deutlich zu erkennen. Aus Abb. A.2 wird weiterhin deutlich, dass es eine spürbare Beeinflussung der Verteilung der Kräfte F_S durch die Kräfte F_0 nur geben kann, wenn die Kräfte F_0 höher oder vergleichbar mit den Kräften F_S sind. Dies ist der Fall, wenn der Zustand S_a stabiler oder ähnlich stabil wie der Zustand S_b ist.

Die Federkonstante k_c der verwendeten Blattfeder hat ebenfalls einen Einfluss auf diesen Effekt. Je kleiner die Federkonstante, um so länger benötigt sie im Kraft-Abstandsbild zur Relaxation ($F_{relax} \simeq -k_c d$), womit die Kräfte $F_{intersect}$ ansteigen. Für Abb. A.3a) wurde eine Simulation mit $k_c=30$ pN/nm durchgeführt. Abb. A.3b) resultiert aus einer Simulation mit den selben Parametern wie in a), jedoch mit einer um den Faktor 5 kleineren Federkonstante, wie es z.B. auf die in dieser Arbeit verwendeten Olympus Bio-Lever Typ A und B zutrifft. Es ist deutlich zu erkennen, dass die Verteilung der Kräfte F_S für den Übergang $S_b \rightarrow S_c$ in Abb. A.3b) viel stärker durch die Kräfte F_0 dominiert wird.

Betrachtet man abschließend noch eine weniger artifizielle Situation, z.B. zwei in ihrer Stabilität identische Zustände, d.h. $d_{S_a-T_a} = d_{S_b-T_b}$ und $k_{off}^0(S_a) = k_{off}^0(S_b)$, so würde man identische Verteilungen für die Entfaltungskräfte F_0 und F_S erwarten. Erfolgen die Entfaltungen jedoch nacheinander, so führt dies zu Abweichungen in den Verteilungen. Abb. A.4 zeigt das Ergebnis einer entsprechend durchgeführten Simulation, um dies zu illustrieren. Es wurde wieder der relative Konturlängenzuwachs $\Delta L/L$ beim Übergang $S_a \rightarrow S_b$ variiert.

Die Verteilungen der Kräfte F_S in Abb. A.4b) und c) wirken völlig unbeeinflusst. Man würde in einer Analyse der Kraftverteilung dem Zustand S_b eine niedrigeren Wert für $d_{S_a-T_a}$ und eine höhere Übergangsrates $k_{off}^0(S_b)$ als für den Zustand S_a zuschreiben - obwohl die Abweichungen in den Kraftverteilungen nur durch den vorhergehenden Zustand S_a verursacht wird.

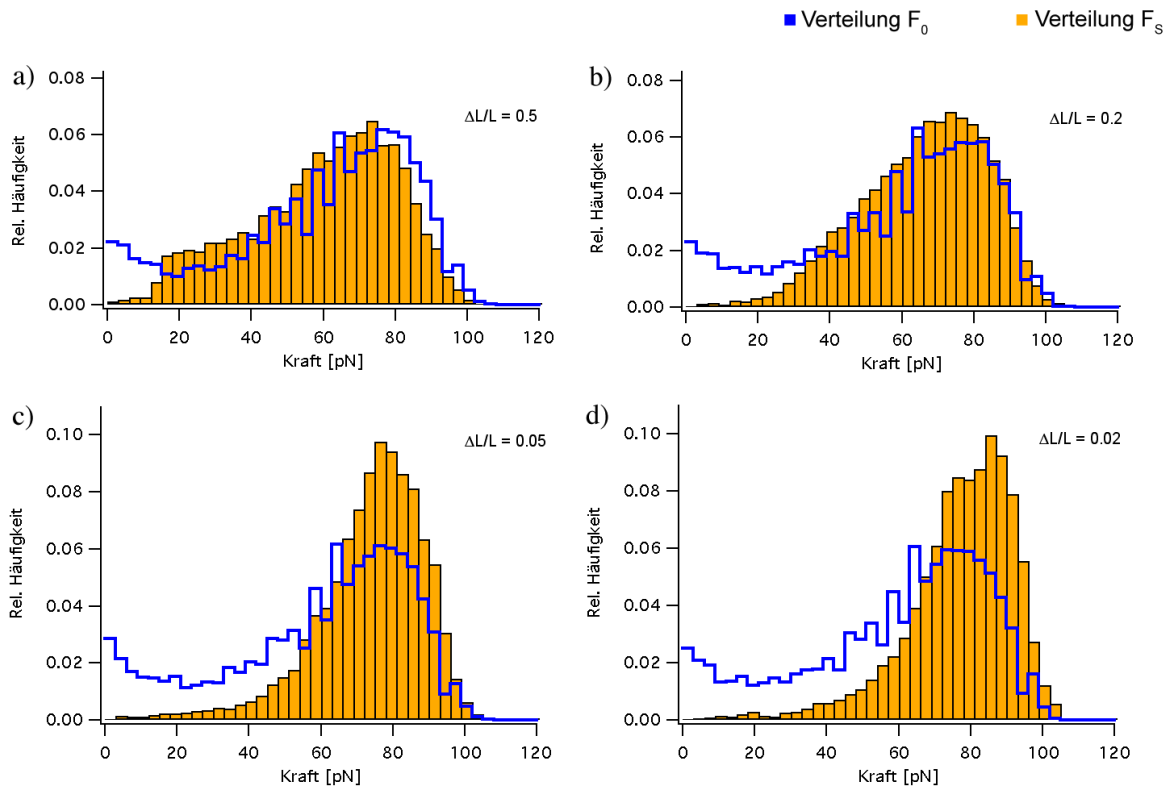


Abbildung A.4: Kraftverteilungen aus einer Simulation von zwei aufeinanderfolgenden Entfaltungen $S_a \rightarrow S_b$ und $S_b \rightarrow S_c$. Variiert wurde die Federkonstante k_c der Blattfeder. Die Verteilungen der Kräfte $F_{intersect}$ sind aus Gründen der Übersichtlichkeit nicht abgebildet.

Anhang B

Mustererkennung

B.1 Einleitung

Aufgrund an der unspezifischen Anbindung der Proteine an Substrat und Blattfederspitze ist die Ausbeute an Kraftkurven, die klare Einzelmolekülereignisse zeigen, in einem kraftspektroskopischen Experiment relativ gering. Es treten auch oft als unspezifische Wechselwirkungen bezeichnete Ereignisse auf, deren Interpretation offen bleibt. Diese Wechselwirkungen sind häufig Überlagerungen verschiedener Ereignisse wie Desorptionen, Entfaltungen und Dehnung von bereits entfaltetem Polypeptidsträngen. Im Prinzip können diese Wechselwirkungen beliebige Formen annehmen.

In früheren Kraftspektroskopie-Experimenten wurden selten sämtliche Kraftkurven, die in einem Experiment gezogen wurden, gespeichert. Dies ist u.a. darauf zurückzuführen, dass keine verlässlichen Methoden zur Verfügung standen, um aus einer großen Menge von Kraftkurven (viele tausend) nur jene herauszufiltern, die bestimmte Eigenschaften zeigen. Dieser Umstand brachte eine gewisse Subjektivität in die Messungen hinein.

In dieser Arbeit ist über das neu hinzugekommene Entfaltungsmuster des GFP in der Ig8-GFP-Chimäre gesprochen worden. Die Vermutung, dass es sich dabei um die Entfaltung des GFP handelte, konnte durch Messungen an der neu erzeugten DdFLN(1-5)-GFP-Chimäre gesichert werden. Aber einmal angenommen, man glaubt ein charakteristisches Muster für die Entfaltung eines bestimmten Proteins erkannt zu haben und es ist nicht möglich, Vergleichsmessungen an einer anderen Protein-Chimäre durchzuführen. Wie kann man dennoch zeigen, dass dieses neu gefundene Muster keine von jenen unspezifischen Wechselwirkungen ist?

Man könnte untersuchen, wie oft dieses als für das unbekannte Protein signifikant vermutete Muster in Messdaten an dieses Protein nicht enthaltenen Proben auftritt, und wie oft es in den Chimären-Daten vorkommt. Dafür ist es nötig, den Grad der Übereinstimmung eines Teils einer Kraftkurve mit einer anderen auf eine quantitative Basis zu stellen. Das kann mit einer Mustererkennung für Kraftkurven erreicht werden.

Eine solche Mustererkennung würde es zudem ermöglichen, kraftspektroskopische Messungen weitgehend zu automatisieren.

B.2 Erkennen eines Musters in Kraftkurven

Sei $g(x)$ eine Funktion, die ein bestimmtes Muster enthalte, und $f(x)$ eine Funktion, die man auf Übereinstimmung an bestimmten Orten mit der Musterfunktion $g(x)$ überprüfen möchte. Eine Kreuzkorrelationsfunktion als Funktion einer Verschiebungsvariable u ist wie in Gleichung B.1 definiert.

$$K_{g,f}(u) = \int_0^b g(x) \cdot f(x+u) dx \quad (\text{B.1})$$

Die Variable b in Gleichung B.1 steht dabei für die Breite der Musterfunktion $g(x)$. Aus der reinen Kreuzkorrelationsfunktion lernt man zunächst nichts, da sie im Prinzip durch die Werte der Funktion $f(x)$ dominiert ist. Interessant wäre aber eine Funktion $C_{g,f}(u)$, die für Verschiebungen u , die zu einer idealen Übereinstimmung zwischen $g(x)$ und $f(x)$ führen, den Wert 1 erhält und sonst immer kleinere Werte annimmt. Enthält die Funktion $f(x)$ ab einer Verschiebung $x = u$ exakt die Musterfunktion $g(x)$, so muss also die in Gleichung B.2 angegebene Bedingung gelten.

$$\lim_{f \rightarrow g} \frac{\int_0^b g(x) \cdot f(x+u) dx}{\int_0^b g(x) \cdot g(x) dx} = 1 \quad (\text{B.2})$$

Die Bedingung B.2 motiviert, eine ortsabhängige (u -abhängige) Normierung für die Kreuzkorrelationsfunktion $K_{g,f}$ einzuführen. Die Cauchy-Schwarzsche-Ungleichung hilft weiter:

$$\frac{\left(\int_0^b g(x) \cdot f(x+u) dx \right)^2}{\int_0^b g^2(x) dx \cdot \int_0^b f^2(x+u) dx} \leq 1 \quad (\text{B.3})$$

Die gesuchte ortsabhängige Normierung ist damit schon gefunden. Die gesuchte Funktion $C_{g,f}(u)$ definiert man einfach als die Quadratwurzel der linken Seite der Ungleichung B.3:

$$C_{g,f}(u) = \frac{K_{g,f}(u)}{\sqrt{\int_0^b g^2(x) dx \cdot \int_0^b f^2(x+u) dx}} \quad (\text{B.4})$$

Damit lässt sich nun der Ort u der besten Übereinstimmung einer Musterfunktion $g(x)$ mit einer Funktion $f(x)$ bestimmen, indem man das Maximum von $C_{g,f}(u)$ und das zugehörige Argument u_{max} bestimmt.

Als Resultat der obigen Berechnungen kann man nun daher den Abschnitt

$$\bar{f}(x) = f(x - u_{max}) ; x \in [0, b] \quad (\text{B.5})$$

bestimmen, der zur besten Übereinstimmung zwischen dem Muster $g(x)$ und der Testfunktion $f(x)$ geführt hat.

B.3 Bestimmung der Übereinstimmungsgüte

Es soll nun noch ein Weg gefunden werden, eine Zahl zu berechnen, die den Grad der Übereinstimmung zwischen $g(x)$ und dem identifizierten (Test-)Muster $\bar{f}(x)$ angibt. Die Zahl $C_{g,\bar{f}}(u_{max})$ aus dem vorigen Abschnitt stellt im Prinzip schon ein gutes Maß dar, es kann aber noch ein Weg gefunden werden, die Erkennung noch spezifischer zu machen.

Man betrachte dazu zwei Testfunktionen $\bar{f}(x)$ und $z(x) = \bar{f}(x) + \eta$, wobei beide die gleiche Breite b der Musterfunktion $g(x)$ haben sollen. O.B.d.A soll für das Muster $g(x)$ gelten: $\langle g(x) \rangle = 0$. Dann ergibt sich für das Verhältnis der normierten Korrelationsfunktionen aus Gl. B.4 für die Verschiebung $u = 0$:

$$\frac{C_{g,\bar{f}}(0)}{C_{g,z}(0)} = \frac{\int_0^b z^2(x) dx}{\int_0^b \bar{f}^2(x) dx} = 1 + \frac{\int_0^b 2\eta\bar{f}(x) + \eta^2 dx}{\int_0^b \bar{f}^2(x) dx} \geq 1 \quad (\text{B.6})$$

Falls $\bar{f}(x)$ eine sehr gute Übereinstimmung mit $g(x)$ aufweist, so sollte eine um η verschobene Funktion $z(x)$ zu deutlich schlechteren Werten von $C_{g,z}(0)$ führen, das Verhältnis in Ungl. B.6 sollte also immer möglichst größer 1 bleiben. Der Term $\int_0^b 2\eta\bar{f}(x) dx$ kann negativ werden und damit im schlimmstenfalls sogar ein $= 1$ in B.6 produzieren. Diesen Umstand kann man vermeiden, indem man einfach die Musterfunktion $g(x)$ (wie bereits gefordert) und die Testfunktion $\bar{f}(x)$ um ihre Mittelwerte verringert:

$$\tilde{g}(x) = g(x) - \langle g(x) \rangle \text{ und } \tilde{f}(x) = \bar{f}(x) - \langle \bar{f}(x) \rangle, \quad (\text{B.7})$$

so führt die Berechnung von $C_{\tilde{g},\tilde{f}}(0)$ zu einem Wert, der ein objektives Maß für die Übereinstimmung zwischen $g(x)$ und Abschnitt $\bar{f}(x)$ der Ausgangsfunktion $f(x)$ darstellt.

Eine kleine Korrektur ist noch durchzuführen. Falls es eine Funktion $w(x) = \eta \cdot f(x)$ gibt, so führt die bisherige Berechnungsmethode für $w(x)$ zur gleichen Übereinstimmungsgüte wie für $f(x)$. Hat z.B. $f(x)$ teilweise eine ideale Übereinstimmung mit $g(x)$ zeigt, darf dies aber für Vielfache von $f(x)$ auf keinen Fall mehr gelten. Die nötige Skalierungskorrektur definiert man am besten wie folgt:

$$s_{g,\bar{f}} = \begin{cases} \sqrt{\frac{\langle \bar{f}^2 \rangle}{\langle g^2 \rangle}} & \text{falls } \langle \bar{f}^2 \rangle \leq \langle g^2 \rangle \\ \sqrt{\frac{\langle g^2 \rangle}{\langle \bar{f}^2 \rangle}} & \text{sonst} \end{cases} \quad (\text{B.8})$$

Als praktikable Definition, jedoch nicht allgemein gültige Definition¹ eines quantitativen Übereinstimmungsmaßes γ für Abschnitte von Kraftkurven hat sich die folgende Form erwiesen:

$$\gamma := C_{g-\langle g \rangle, \bar{f}-\langle \bar{f} \rangle}^2(0) \cdot s_{g, \bar{f}} \quad (\text{B.9})$$

Aufgrund der obigen Überlegungen nimmt die Zahl γ nur positive Werte ≤ 1 an. Ist $\gamma = 1$, so liegt ideale Übereinstimmung zwischen dem Muster $g(x)$ und dem Ausschnitt $f(x)$ der untersuchten Kraftkurve $f(x)$ vor.

Die obigen Überlegungen sind in Computerprogramme umgesetzt worden (Macros für IGOR PRO 4.01, Wavemetrics). Es verbleibt, sie auf Tauglichkeit zu überprüfen.

B.4 Prinzipielle Tauglichkeit der Mustererkennung

Zunächst muss eine Musterfunktion $g(x)$ definiert werden. Es sei dies der schwarz eingefärbte Bereich in Abbildung B.1a) einer simulierten Kraftkurve (vgl. Abschnitt 4.6). Dieser Bereich entspricht der Blattfeder-Relaxation nach Entfaltung einer (simulierten) GFP-Domäne und dem anschließenden Dehnen eines um 76.6 nm verlängerten Polypeptidstranges. Als eine Menge an Testfunktionen $f(x)$ wurde der komplette Datensatz aus einem Kraftspektroskopie-Experiment an der DdFLN(1-5)-GFP-Chimäre (vgl. Kapitel 5) gewählt. Dieser Datensatz umfasst 1012 verschiedene Kraftkurven. Die Mustererkennungs-Algorithmen führen zur Ausgabe eines Wertes γ (wie in Gl. B.9 definiert) für jede einzelne Kraftkurven. Die Häufigkeit des Auftretens von γ -Werten in Intervallen von 0.02 ist in Abbildung B.1b) dargestellt.

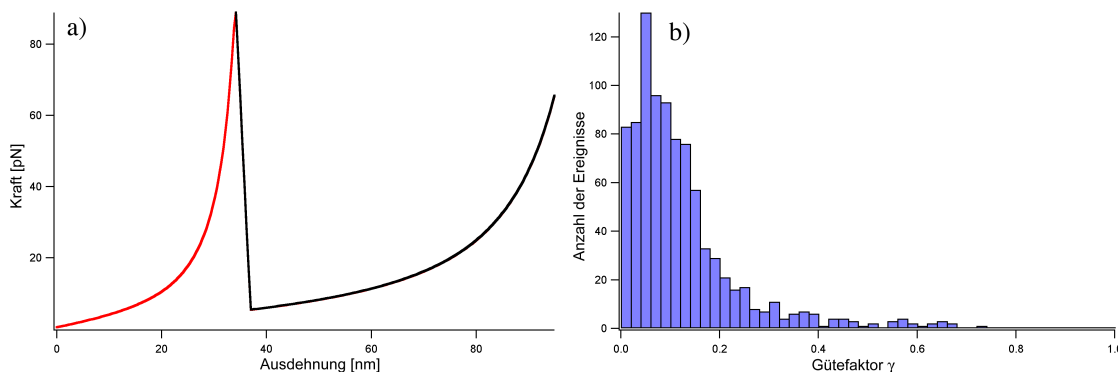


Abbildung B.1: a) - Als Musterfunktion $g(x)$ wurde der dunkle Bereich definiert. b) - Verteilung des Auftretens bestimmter γ -Werte in 1012 Kraftkurven (Probe: DdFLN(1-5)-GFP-Chimäre).

Man erkennt in Abb. B.1b), dass die Häufigkeit des Auftretens von höheren γ -Werten bereits ab $\gamma = 0.2$ sehr stark abnimmt. Es stellt sich die Frage, wie die zugehörigen Kraftkurven zu bestimmten γ -Werten aussehen. In Abb. B.2 sind willkürlich jeweils vier Kraft-

¹Diese ist durch $\gamma := C_{g, \bar{f}} \cdot s_{g, \bar{f}}$ gegeben.

kurven ausgewählt worden, die mit γ -Werten aus vier verschiedenen Intervallen bewertet wurden.

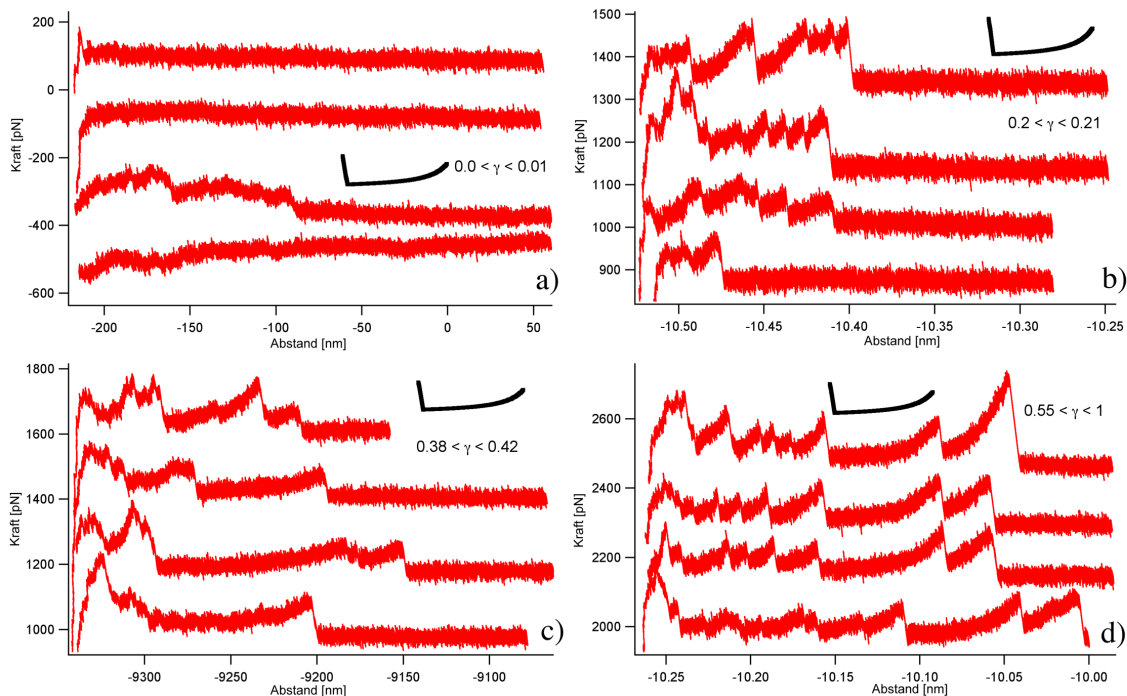


Abbildung B.2: Erläuterung im Text

Das Muster $g(x)$ ist in Abb. B.2 in a)-d) jeweils in schwarz abgebildet. In Abb. B.2a) (Bereich $0.0 \leq \gamma \leq 0.01$) erkennt man keinerlei Ähnlichkeit mit dem vorgegebenen Muster. Die zugehörigen Kraftkurven weisen keine oder kaum Adhäsionen auf. Kraftkurven, die Drifterscheinungen zeigen, fallen auch in diesen Bereich.

Abb. B.2b) (Bereich $0.2 \leq \gamma \leq 0.21$) zeigt Kraftkurven, die schon deutlich mehr Adhäsionen zeigen, es sind aber dennoch kaum Übereinstimmungen mit dem Muster zu erkennen.

Abb. B.2c) (Bereich $0.38 \leq \gamma \leq 0.42$) zeigt hingegen Kraftkurven, in denen schon zum Muster ähnliche Strukturen zu erkennen sind, diese Strukturen sind aber entweder zu lang oder weisen zu große Kräfte auf.

Abb. B.2d) (Bereich $0.55 \leq \gamma \leq 1$) hingegen zeigt Kraftkurven, die hervorragende Übereinstimmung mit dem Muster zeigen. Alle vier gezeigten Kurven besitzen einen Bereich, der sehr gut das Muster reproduziert. Eine manuelle Analyse aller 1012 Kraftkurven dieses Experimentes zeigte, dass alle Kraftkurven, die eine GFP-Entfaltung aufweisen (also das Muster $g(x)$ zeigen) von der Mustererkennungs-Routine mit $\gamma \geq 0.55$ bewertet wurden. Die Mustererkennung wurde noch mit anderen Mustern und anderen Datensätzen getestet und führte zu qualitativ ähnlichen Ergebnissen. Wählt man als Muster ein Stück einer verbrauchten, experimentell erzeugten Kraftkurve, so kann auch dieses Muster in Datensätzen (die es enthalten sollten) wiedergefunden werden.

Aus der Erfahrung weiß man, dass die Ausbeute an klaren Einzelmolekülkurven in einem Experiment meistens kleiner als 1 Prozent ist. Daher sollte man, um die Mustererkennung zum Auffinden solcher Kurven zu benutzen, ein Diagramm wie in Abb. B.1b) erzeugen und den Wert γ bestimmen, für den etwa 1 Prozent der Gesamtzahl an Kurven mit höheren γ -Werten beurteilt wurden. Dann kann man sich diese Menge an Kurven anzeigen lassen und ist sicher, alle interessanten Kurven aufgefunden zu haben. Es sollte dann noch einmal eine manuelle Auswahl erfolgen, um saubere Einzelmolekülkurven zu identifizieren.

B.5 Statistische Analyse von Datensätzen

Im vorigen Abschnitt ist die prinzipielle Tauglichkeit der Mustererkennungs-Algorithmen für das Auffinden von Kraftkurven mit bestimmten Eigenschaften vorgestellt worden. In diesem Abschnitt soll nun die Frage diskutiert werden, ob sie für einen quantitativen Vergleich von Datensätzen hinsichtlich des Auftretens bestimmter Muster geeignet wäre. Dazu sind zwei ähnlich große Datensätze aus Experimenten an zwei verschiedenen Proteinen (DdFLN(1-5)-GFP-Chimäre und Ig8-DdFLN(4)-Chimäre) mit der Mustererkennungsroutine auf das Auftreten desselben Musters $g(x)$ analysiert worden. Die experimentellen Bedingungen in beiden Experimenten waren vergleichbar, der einzige prinzipielle Unterschied liegt in der untersuchten Probe. Das vorgegebene Muster war ein Stück einer theoretisch berechneten GFP-Kraftkurve (siehe Abb. B.3) Der DdFLN(1-5)-GFP-Datensatz enthielt 1012 Kraftkurven und der Ig8-DdFLN(4)-Datensatz enthielt 934 Kraftkurven, die Datenmengen waren also vergleichbar. In Abb. B.3 sind die Ergebnisse der Analyse dargestellt.

Zunächst fällt bei Betrachtung von Abbildung B.3a) und b) auf, dass die Häufigkeit des Auftretens sehr niedriger Gütefaktoren γ in beiden Datensätzen vergleichbar hoch ist (etwa 78 %). Nach den Erkenntnissen aus dem vorigen Abschnitt bedeutet dies, dass Kraftkurven, die wie in Abb. B.2a) keine oder kaum Adhäsionen zeigen, in beiden Experimenten vergleichbar oft auftreten. Dies scheint also keine Eigenschaft der Experimentiertechnik zu sein, die von der genauen Form der Probe abhängt.

Eine Vergrößerung und Überlagerung der beiden Histogramme wie in B.3c) zeigt, dass die γ -Verteilung für den Ig8-DdFLN(4)-Datensatz extrem schnell abfällt. Ab $\gamma = 0.2$ wurden nur noch 5 (ca. 0.5% der Gesamtanzahl) Kraftkurven mit höheren Werten beurteilt. Der höchste Wert für γ liegt bei 0.46 und tritt nur einmal auf. Im Gegensatz dazu fällt die γ -Verteilung für den DdFLN(1-5)-GFP-Datensatz viel langsamer ab. Ab $\gamma = 0.2$ wurden immerhin noch 65 Kraftkurven (ca. 6.4% der Gesamtanzahl) mit höheren Gütefaktoren beurteilt. Es treten 9 Ereignisse mit $\gamma \geq 0.5$ auf. Wie im vorigen Abschnitt verdeutlicht, beginnt hier der Bereich einer hervorragenden Übereinstimmung mit dem Muster.

Im Vergleich bedeutet dies, dass das berechnete Muster $g(x)$ einer GFP-Entfaltung mit hervorragender Übereinstimmung nur im Datensatz der DdFLN(1-5)-GFP-Chimäre auftritt, während es im Datensatz der Ig8-DdFLN(4)-Chimäre völlig fehlt. Auch treten Kraftkurven mit teilweiser Übereinstimmung mit dem Muster im DdFLN(1-5)-GFP-Datensatz um eine Größenordnung häufiger auf als im Ig8-DdFLN(4)-Datensatz. Offenbar ist ein

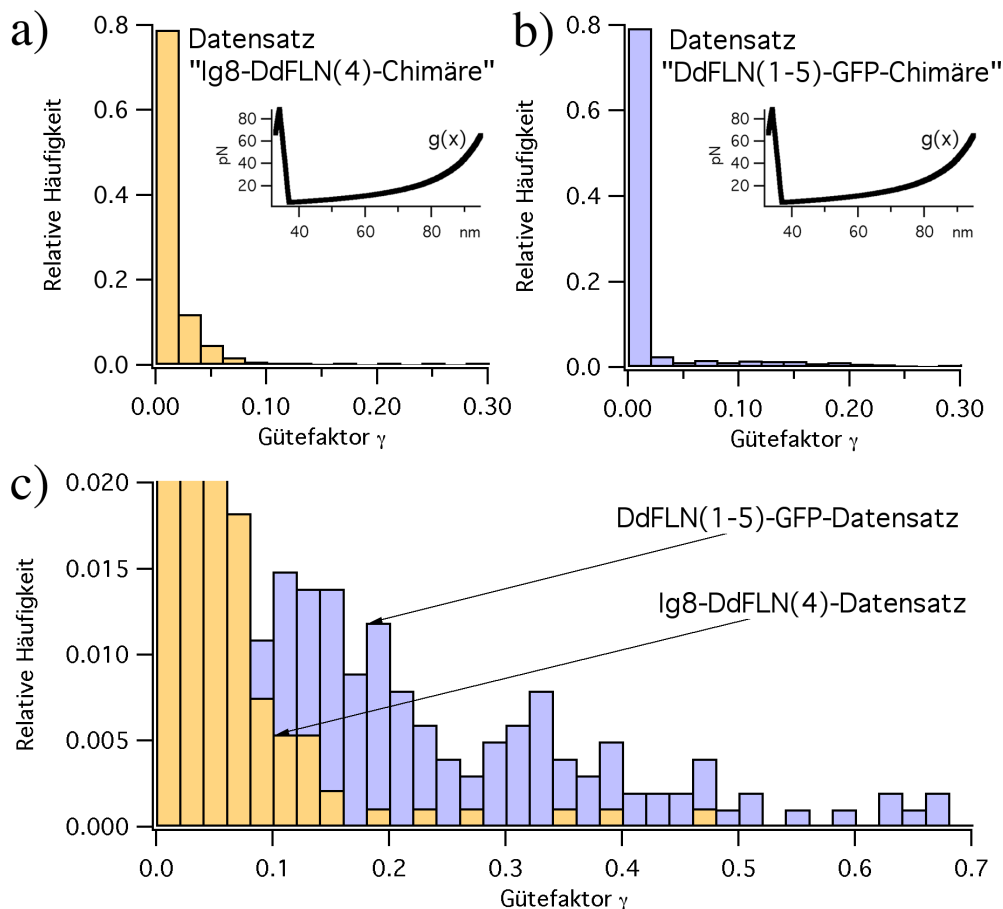


Abbildung B.3: a), b) - Histogramme der relativen Häufigkeit des Auftretens von bestimmten Gütefaktoren γ . c) Überlagerung der beiden Histogramme und Vergrößerung eines Ausschnittes. Diskussion im Text.

Entfaltungsmuster (hier GFP-Entfaltung) so hochspezifisch, dass es nur sehr selten durch zufällige, unspezifische Wechselwirkungen reproduziert wird.

Die vorgestellte Methode zur Bewertung einzelner Kraftkurven mittels eines Gütefaktors γ läßt sich auch zur quantitativen Analyse verwenden und liefert objektive Argumente, daß ein bestimmtes Muster tatsächlich eine Eigenschaft, eben ein molekularer Fingerabdruck, einer bestimmten Probe ist. Es dürfte möglich sein, auch auf statistischer Basis den Begriff *molekularer Fingerabdruck* zu bestätigen.

B.6 Zusammenfassung

Es wurde eine Methode vorgestellt, die zur Mustererkennung in Kraftkurven geeignet ist. Dazu muß ein Muster $g(x)$ definiert werden und dann für jede in einem Datensatz enthaltene Kraftkurve $f(x)$ der Abschnitt $\bar{f}(x)$ der besten Übereinstimmung mit $g(x)$ über

Identifizierung des Maximums der in Gl. B.10 definierten Funktion bestimmt werden.

$$C_{g,f}(u) = \frac{\int_0^b g(x) \cdot f(x+u) dx}{\sqrt{\int_0^b g^2(x) dx \cdot \int_0^b f^2(x+u) dx}} \quad (\text{B.10})$$

Anschließend kann der Wert eines Gütefaktors γ , der die Qualität der Übereinstimmung dieses Abschnittes $\bar{f}(x)$ mit dem Muster auf einer Skala von 0 bis 1 angibt, über Gl. B.11 bestimmt werden.

$$\gamma := C_{g-\langle g \rangle, \bar{f}-\langle \bar{f} \rangle}^2(0) \cdot s_{g, \bar{f}} \quad (\text{B.11})$$

Diese Art der Mustererkennung hat sich als tauglich erwiesen und sie kann weiterhin für statistische Analysen hinsichtlich des Auftretens bestimmter Muster in Datensätzen verwendet werden. Dabei ist γ ein objektives, quantitatives Maß für die Güte der Übereinstimmung mit dem Muster.

Mit der vorgestellten Methode wurde anhand von zwei Datensätzen verdeutlicht, dass das GFP-Entfaltungsmuster mit statistischer Relevanz nur in Kraftkurven vorkommt, die an einem GFP enthaltenden Protein gezogen wurden. Dabei ist der Eindruck entstanden, dass - abgesehen von den Kraftkurven, die hervorragende Übereinstimmung mit dem Muster zeigten - vor allem auch deutlich häufiger mittelmäßige Übereinstimmungen mit dem Muster auftreten. Es könnten dies Überlagerungen von GFP-Entfaltungen sein, die z.B. dadurch entstehen, dass mehrere Moleküle gleichzeitig gedehnt werden.

Die Methode bietet nun die Möglichkeit, kraftspektroskopische Messungen weitgehend zu automatisieren. Man kann das Messgerät selbsttätig mit einer hohen Frequenz Kraftkurven ziehen und speichern lassen und diese Kraftkurven automatisch hinsichtlich eines Musters untersuchen lassen. Auf diese Weise steht ein Werkzeug zur Verfügung, um objektiv große Datenmengen zu beurteilen.

Anhang C

Eichung der Blattfedern

Für die Eichung von Blattfedern existieren verschiedene Methoden, die z.B. in [36] zusammengefasst sind. In dieser Arbeit wurde die Eichung der Blattfedern über eine Analyse ihres thermisch angeregten Rauschens vorgenommen [14][21]. Dabei wird der Gleichverteilungssatz ausgenutzt. Bezieht man einige Subtilitäten ein - wie z.B. den Umstand, dass die Detektion der Auslenkung der Blattfeder über die Ablenkung eines Lichtzeigers geschieht [9] - so kann aus dem mittleren Ablenkungsquadrat $\langle x^2 \rangle$ der Blattfeder die Federkonstante über Gleichung C.1 bestimmt werden.

$$k_c = 0.8 \cdot \frac{k_B T}{\langle x^2 \rangle} \quad (\text{C.1})$$

Diese Gleichung gilt dann, wenn nur die erste Eigenschwingung der Blattfeder angeregt ist. Dies ist im allgemeinen nicht der Fall. Jedoch kann dieses Problem umgangen werden, in dem das Parsevalsche Theorem $\langle x^2(t) \rangle = \langle x^2(\omega) \rangle$ angewandt wird und statt $\langle x^2(t) \rangle$ die spektrale Dichte $x^2(\omega)$ des thermischen Rauschens der Blattfeder bestimmt wird und anschließend nur über die erste Eigenresonanz integriert wird. Abb. C.1 zeigt die spektrale Dichte eines Typ A Bio-Levers. Zur Bestimmung von $\langle x^2(\omega) \rangle$ würde man im abgebildeten Fall über den Frequenzbereich 0 bis 30 kHz integrieren. Um die vorgestellte Methode der Blattfeder-Eichung verwenden zu können, muss aus dem über die Ablenkung des Lichtzeigers gewonnenen Photodiodensignal eine Information über die aktuelle Auslenkung der Blattfeder gewonnen werden. Dies geschieht durch Bestimmung einer sogenannten Steigungsgerade. Drückt man die Blattfeder auf eine harte Oberfläche ein definiertes Stück ein, so erhält man bei passender Einrichtung des Instrumentes einen linearen Zusammenhang zwischen einer Länge und der zugehörigen Änderung des Photodiodensignals. Die Bestimmung dieses Parameters beschränkt hauptsächlich die Genauigkeit der Kräfteichung.

In der Praxis können mit der beschriebenen Methode Federkonstanten bis auf eine Unsicherheit von etwa $\pm 10\%$ bestimmt werden.

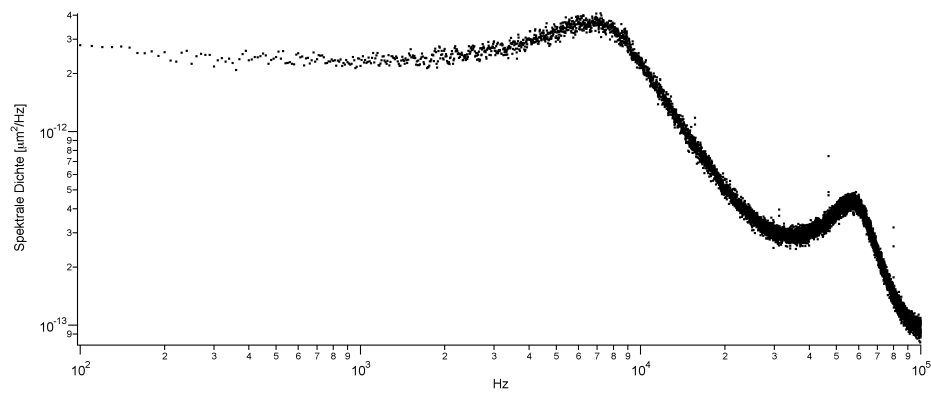


Abbildung C.1: Thermisches Rauschspektrum (*spektrale Dichte*) eines Typ A Bio-Levers in wässriger Umgebung. Deutlich zu erkennen ist die erste Resonanz der Blattfeder bei etwa 8kHz. Eine Oberschwingung der Blattfeder bei 60 kHz ist ebenfalls angeregt.

Anhang D

Molekularbiologische Methoden

Die untersuchte Protein-Chimäre DdFLN(1-5)-GFP ist in den Institutsräumen des Lehrstuhls für Angewandte Physik, erzeugt worden. Dazu mussten zunächst auf DNA-Ebene die kodierenden Sequenzen erstellt ("kloniert") werden, um die spätere Expression in Bakterien zu ermöglichen. Die folgenden zwei Abschnitte behandeln deshalb in chronologischer Form zunächst die gentechnischen Methoden und später die Protokolle, die die Isolation der Proteine ermöglichten.

D.1 Gentechnische Methoden

Ziel war es, zwischen die fünf kettenförmig aneinandergereihten Domänen des Dictostelium Filamins das Grün Fluoreszierende Protein derart einzufügen, daß die kettenartige quaternäre Struktur eines solchen Chimären-Proteins beibehalten wird.

Die DNA-Sequenzen der Domänen 1 bis 5 des DdFLN wie auch des GFP lagen zu Beginn dieser Arbeit bereits vor, und zwar als Teilstücke von in Lösung vorliegenden DNA-Plasmidringen (*Vektoren*). Die Sequenz des DdFLN wurde in Vorarbeiten durch Angelika Kardinal in das pET28a(+)-Vektorsystem eingefügt. Die Sequenz des untersuchten *Cycle3*-GFP wurde durch die Arbeitsgruppe Steipe zur Verfügung gestellt. Sie lag als Teil des pRSet5d-Vektorsystems vor. Beide Sequenzen waren mittels Restriktionsenzymen und Ligasen in die *multiple cloning site* der Vektoren eingefügt worden.

Die GFP-Sequenz trug am 5'-Ende die Sequenz des NCOI-Restriktionsenzym und am 3'-Ende die Sequenz des BSTEII-Restriktionsenzym. Eine geeignete Vorgehensweise war nun, in der Sequenz des DdFLN-Proteins zwei neue Schnittstellen für genau diese beiden Restriktionsenzyme einzufügen und jegliche sonst in der DdFLN(1-5)-Sequenz und im restlichen pET28a(+)-Vektor vorhandenen Schnittstellen für das NCOI und das BSTEII-Enzym zu deaktivieren. Mittels des kommerziell erhältlichen Multi-QuikChange Protokolls (Fa. Stratagene, Amsterdam, NL) und entsprechend synthetisierten DNA-Primern (Fa. Metabion, Ebersberg, D) wurde dies realisiert. Für die Klonierung wurden darauf speziell abgestimmte Zelllinien wie XL1-Blue, XL10-Gold und DH5 α (alle Fa. Stratagene, Amsterdam, NL) verwendet. Zur Isolation der Plasmid-DNA aus diesen Bakterien wurde das

kommerziell erhältliche QIAprep Spin Miniprep Kit Protocol (Fa. Qiagen, Konstanz, D) verwendet.

Es wurden eine BSTEII-Schnittstelle an Position 1304 und eine NCOI-Schnittstelle an Position 296 im pET28a(+)-Vektor durch Tausch einer einzelnen DNA-Base deaktiviert. Überprüft wurde dies durch Hinzufügen des jeweiligen Restriktionsenzym zur DNA-Lösung und anschließender Analyse der Verdauungsprodukte mittels Gelelektrophorese. Die zwei neuen NCOI- und BSTEII- Schnittstellen wurden zwischen den die Domänen 3 und 4 kodierenden Teilstücken der DdFLN-Sequenz eingefügt, die Erfolgsüberprüfung geschah wieder wie oben beschrieben. Nun konnte die GFP-Sequenz aus dem pRSet5d-Vektor durch Anwendung der Restriktionsenzyme NCOI und BSTEII herausgeschnitten werden. Die Verdauungsprodukte wurden über Gelelektrophorese aufgetrennt und das etwa 700bp große GFP-kodierende Sequenzstück mittels einem kommerziell erhältlichen Protokoll (Gel Extraction Kit, Fa. Qiagen, Konstanz, D) isoliert. Parallel dazu wurde der die DdFLN(1-5)-Sequenz enthaltene pET28a(+)-Vektor durch Anwendung der obigen Enzyme an den neu eingefügten Schnittstellen aufgeschnitten und die nun linearisierte DNA wieder über Gelelektrophorese analysiert und nachfolgend mit dem Qiagen Gel Extraction Kit aufgereinigt.

Im nächsten Arbeitsschritt wurde die GFP-Sequenz und der linearisierte DdFLN(1-5)-Vektor in Lösung zusammengeführt. Da beide Sequenzen an den Enden jeweils die geschnittenen, palindromischen NCOI und BSTEII Schnittstellen besaßen, besteht eine endliche Wahrscheinlichkeit für das korrekte Anlagern der GFP-Sequenz an die *sticky ends* des DdFLN(1-5)- Vektors. Durch Hinzufügen einer Ligase (T4 Ligase, Fa. Roche, Basel, CH) wurde die DNA an den Schnittstellen kovalent verbunden (*legiert*) und der DdFLN(1-5)-Vektor war um die Sequenz des GFP zwischen den DdFLN Domänen 3 und 4 erweitert. Überprüft wurde dies wieder durch Anwendung der Enzyme NCOI und BSTEII und anschließender Gelelektrophorese.

Nun konnte von der DNA-Ebene auf die Protein-Ebene übergetreten werden. Dazu wurde der neue DdFLN(1-5)-GFP-Vektor in speziell auf Proteinexpression abgestimmte Bakterien (BL21, Derivat E.coli) transformiert.

Von allen Zwischenstufen auf dem Weg zum endgültigen, für die Expression verwendeten DdFLN(1-5)-GFP-Vektor wurden Sicherheitskopien in Form von bakteriellen Glycerinkulturen bei -80°C eingelagert.

D.2 Proteinexpression

Um die DdFLN(1-5)-GFP-Chimäre zu exprimieren, wurden Standard-Protokolle benutzt, bei denen die Proteinexpression über die Zugabe eines bestimmten Zuckers (IPTG, Isopropylthiogalactoside) induziert werden kann. Dazu wurden zunächst BL21-Bakterien, die das DdFLN(1-5)-GFP-Plasmid enthielten, in einer Nährlösung bei 37°C über mehrere Stunden kultiviert. Bei einer bestimmten Populationsstärke (Optische Dichte der Lösung bei OD 0.8) wurde die Expression des Proteins durch Zugabe von 1mM IPTG induziert. Die Bakterienlösung wurde für die Expressionsphase für zwei Stunden bei 37°C inkubiert. Um

das Protein zu ernten, wurden die Bakterien in Lösung in einer *French Press* hohen Scherkräften ausgesetzt, um die Zellmembranen zu öffnen. Die Zelltrümmer wurden durch Zentrifugieren aus der Lösung entfernt. Um das DdFLN(1-5)-GFP-Protein aufzureinigen und alle anderen Verunreinigungen aus der Lösung weitestgehend zu entfernen, wurde auf Affinitätschromatographie zurückgegriffen. Das pET28a(+)-Vektorsystem enthält am Ende der einklonierten DdFLN(1-5)-GFP-Sequenz Codons, die 6 Histidine kodieren und bei Expression automatisch an den Carboxyterminus der DdFLN(1-5)-GFP-Chimäre angefügt werden. Histidine haben eine hohe Affinität zu Ni-NTA. Damit konnte das Protein durch Durchlauf durch eine Ni-NTA-Gelmatrix und anschließendem Eluieren durch Zugabe verschieden konzentriertem Imidazol (Imidazol hat eine noch höhere Affinität zu Ni-NTA als Histidine) aufgereinigt werden. Die gereinigte Proteinlösung fing nach einigen Stunden nach der Expression an zu fluoreszieren (mit dem typischen GFP-Fluoreszenzspektrum, vgl. Abb. 2.5) und wurde bei 4°C gelagert. Die korrekte Expression des Chimären-Proteins wurde mittels SDS-Gelelektrophorese hinsichtlich des Molekulargewichtes überprüft. Die Faltung in die korrekte quaternäre Kettenstruktur konnte mit den kraftspektroskopischen Messungen gezeigt werden.

Anhang E

Sequenzinformationen

Im folgenden werden die DNA- und Aminosäuresequenz der in dieser Arbeit untersuchten *Cycle3*-GFP Mutante angegeben. Es ist zu bemerken, dass das verwendete Protein im Vergleich zur Original *Cycle3*-GFP Sequenz eine weitere Punktmutation aufweist: Ser → Gly an Position 2.

E.1 GFP

CCATGGGTAA	AGGAGAAGAA	CTTTTCACTG	GAGTTGTCCC
AATTCTTGTT	GAATTAGATG	GTGATGTTAA	TGGGCACAAA
TTTTCTGTCA	GTGGAGAGGG	TGAAGGTGAT	GCAACATACG
GAAACTTAC	CCTTAAATTT	ATTTGCACTA	CTGGAAAAC
ACCTGTTCCG	TGGCCAACAC	TTGTCACTAC	TTTCTCTTAT
GGTGTTC AAT	GCTTTTCCCG	TTATCCGGAT	CATATGAAAC
GGCATGACTT	TTTCAAGAGT	GCCATGCCCC	AAGGTTATGT
ACAGGAACGC	ACTATATCTT	TCAAAGATGA	CGGGAAC TAC
AAGACGCGTG	CTGAAGTCAA	GTTTGAAGGT	GATACCCTTG
TTAATCGTAT	CGAGTTAAAA	GGTATTGATT	TTAAAGAAGA
TGGAACATT	CTCGGACACA	AACTCGAGTA	CAACTATAAC
TCACACAATG	TATACATCAC	GGCAGACAAA	CAAAAGAATG
GAATCAAAGC	TAACTTCAA	ATTCGCCACA	ACATTGAAGA
TGGATCCGTT	CAACTAGCAG	ACCATTATCA	ACAAAATACT
CCAATTGGCG	ATGGCCCTGT	CCTTTTACCA	GACAACCATT
ACCTGTCCAC	ACAATCTGCC	CTTTCGAAAAG	ATCCCAACGA
AAAGAGAGAC	CACATGGTCC	TTCTTGAGTT	TGTAACAGCT
GCTGGGATTA	CACATGGCAT	GGATGAACTA	TACAAGGGTC
ACC			

Tabelle E.1: DNA-Sequenz des verwendeten *Cycle3*-GFP. Der Leserahmen für die Expression ist 3.

MGKGEELFTG	VVPILVELDG	DVNGHKFSVS	GEGEGDATYG
KLTLKFICTT	GKLPVPWPTL	VTTFSYGVQC	FSRYPDHMKR
HDFFKSAMPE	GYVQERTISF	KDDGNYKTRA	EVKFEGDTLV
NRIELKGIDF	KEDGNILGHK	LEYNYNShNV	YITADKQKNG
IKANFKIRHN	IEDGSVQLAD	HYQQNTPIGD	GPVLLPDNHY
LSTQSALSKD	PNEKRDHML	LEFVTAAGIT	HGMDELYKGH

Tabelle E.2: Aminosäuren-Sequenz des verwendeten *Cycle3*-GFP, Die das Chromophor bildenden Aminosäuren sind Serin65 -Tyrosin66 - Glycin67.

Literaturverzeichnis

- [1] A. Ahsan, J. Rudnick und R. Bruinsma, *Biophysical Journal* **74** (1998), 132.
- [2] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts und J. Watson: *Molecular Biology of the Cell*. Garland Science, New York, 2002.
- [3] G. Bell, *Science* **200** (1978), 618.
- [4] H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov und P. Bourne, *Nucleic Acids Research* **28** (2000), 235.
- [5] O. Bieri, J. Wirz, B. Hellrung, M. Schutkowski und M. Drewello, *Proceedings of the National Academy of Science* **96** (1999), 9597.
- [6] G. Binnig, C. Quate und C. Gerber, *Physical Review Letters* **56(9)** (1986), 930.
- [7] C. Bouchiat, M. Wang, J. Alleman, T. Strick, S. Block und V. Croquette, *Biophysical Journal* **76** (1999), 409.
- [8] C. Bustamante, J. Marko, E. Siggia und S. Smith, *Science* **265** (1994), 1599.
- [9] H. Butt und M. Jaschke, *Nanotechnology* **6** (1995), 1.
- [10] M. Carrion-Vazquez, A. Oberhauser, S. Fowler, P. Marszalek, S. Broedel, J. Clarke und J. Fernandez, *Proceedings of the National Academy of Science* **96(7)** (1999), 3694.
- [11] A. Cramer, E. Whitehorn, E. Tate und W. Stemmer, *Nature Biotechnology* **14** (1996), 315.
- [12] E. Evans, *Annual Review of Biophysics and Biomolecular Structure* **30** (2001), 105.
- [13] E. Evans und K. Ritchie, *Biophysical Journal* **76** (1999), 2439.
- [14] E. Florin, M. Rief, H. Lehmann, M. Ludwig, C. Dornmair, V. Moy und H. Gaub, *Biosensors and Bioelectronics* **10** (1995), 895.
- [15] P. Flory: *Statistical Mechanics of Chain Molecules*. Hanser, München, 1988.
- [16] P. Fucini, B. Köppel, M. Schleicher, A. Lustig, T. Holak, R. Müller, M. Stewart und A. Noegel, *Journal of Molecular Biology* **291** (1999), 1017.

- [17] H. Fukuda, M. Arai und K. Kuwaijima, *Biochemistry* **39** (2000), 12025.
- [18] B. Heymann und H. Grubmüller, *Physical Review Letters* **84** (2000), 6126.
- [19] J. Howard: *Mechanics of Motor Proteins and the Cytoskeleton*. Sinauer Associates, Inc, Sunderland, Massachusetts, 2001.
- [20] T. Hugel, M. Rief, M. Seitz, H. Gaub und R. Netz, *submitted* (2003).
- [21] J. Hutter und J. Bechhoefer, *Review of Scientific Instruments* **64** (1994), 1868.
- [22] J. Kovac und C. Crabb, *Macromolecules* **15** (1982), 537.
- [23] O. Kratky und G. Porod, *Recueil Des Travaux Chimiques Des Pays-Bas* **68** (1949), 1106.
- [24] L. Lapidus, W. Eaton und J. Hofrichter, *Proceedings of the National Academy of Science* **97** (2000), 7220.
- [25] J. Liphardt, B. Onoa, S. Smith, I. Tinoco und C. Bustamante, *Science* **292** (2001), 733.
- [26] L. Livadaru, R. Netz und H. Kreuzer, *Macromolecules* **36** (2003), 3732.
- [27] J. Marko und E. Siggia, *Macromolecules* **28** (1995), 209.
- [28] P. Marszalek, H. Lu, H. Li, M. Carrion-Vazquez, A. Oberhauser, K. Schulten und J. Fernandez, *Nature* **402** (1999), 100.
- [29] V. Moy, E. Florin und H. Gaub, *Science* **266** (1994), 257.
- [30] D. Prasher, V. Eckenrode, W. Ward, F. Prendergast und M. Cormier, *Gene* **111** (1992), 229.
- [31] M. Rief: *Kraftspektroskopie an einzelnen Molekülen*. Dissertation an der Fakultät für Physik, Ludwig-Maximilians-Universität München, 1997.
- [32] M. Rief, *Persönliche Mitteilung* (2004).
- [33] M. Rief, J. Fernandez und H. Gaub, *Physical Review Letters* **81** (1998), 4764.
- [34] M. Rief, M. Gautel, F. Oesterhelt, J. Fernandez und H. Gaub, *Science* **276** (16. Mai 1997), 1109.
- [35] M. Rief, M. Gautel, A. Schemmel und H. Gaub, *Biophysical Journal* **75** (1998), 3008.
- [36] J. Sader, *Encyclopedia of Surface and Colloid Science* (2002), 846.
- [37] I. Schwaiger, *Persönliche Mitteilung* (2004).

-
- [38] I. Schwaiger, A. Kardinal, M. Schleicher, A. Noegel und M. Rief, *Nature Structural and Molecular Biology* **11** (2004), 81.
- [39] I. Schwaiger, C. Sattler, D. Hostetter und M. Rief, *Nature Materials* **1** (2002), 232.
- [40] S. Smith, Y. Cui und C. Bustamante, *Science* **271** (1996), 795.
- [41] R. Tsien, *Annual Reviews of Biochemistry* **67** (1998), 509.
- [42] M. Viani, T. Schäffer, A. Chand, M. Rief, H. Gaub und P. Hansma, *Journal of Applied Physics* **86** (1999), 2258.
- [43] T. Wilson und J. Woodland, *Annual Review of Cell and Development Biology* **14** (1998), 197.
- [44] I. Yeh und G. Hummer, *Journal of the American Chemical Society* **124** (2002), 6563.
- [45] X. Zhuang und M. Rief, *Current Opinion in Structural Biology* **13** (2003), 88.
- [46] M. Zimmer, *Chemical Reviews* **102** (2002), 759.

Danke!

An dieser Stelle möchte ich mich bei all jenen bedanken, die zum Gelingen dieser Diplomarbeit beigetragen haben.

An erster Stelle möchte ich Prof. Matthias Rief danken. Ohne die exzellente Anleitung durch ihn, seine Geduld und sein beständiges Hinterfragen meiner Ergebnisse wäre diese Arbeit nicht in dem Zustand, in dem sie ist.

Herzlich danken möchte ich auch Prof. Hermann E. Gaub für die ausgezeichnete Atmosphäre in seiner Arbeitsgruppe und die hervorragende Infrastruktur, die in seinen Laborräumen zur Verfügung steht. Unvergesslich sind das schmackhafte Essen und der gute Wein.

Ingo Schwaiger sei gedankt für die Unterstützung im Umgang mit dem von ihm konstruierten Kraftspektrometer - und für seine gute Laune in Zeiten messtechnischer Traufen.

Ein herzliches Dankeschön geht an Angelika Kardinal für die Unterstützung bei den gentechnischen Arbeiten.

Ebenfalls beteiligt waren: Tim Liedl, Thomas Bornschlögl, Stefan Kufer, Martin Benoit, Christian Kirchner, Stefan Kudera, Ferdinand Kühner, Max Kessler, Michael Schlierf, Johann Jaud und Anabel Clemen. Danke.

Schließlich möchte ich mich noch bei zwei Menschen bedanken, die mir das Studium der Physik überhaupt erst ermöglicht haben: Viola Dietz und Prof. Hans-Michael Dietz, meine Eltern.

Erklärung

Hiermit erkläre ich, die vorliegende Arbeit selbstständig und ausschließlich unter Verwendung der angegebenen Quellen und Verfasser angefertigt zu haben.

München, den

Hendrik Dietz