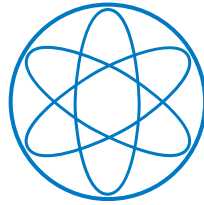


PHYSIK-DEPARTMENT



Untersuchung der Hybridisierungskinetik
kurzer DNA-Doppelstränge
mit der neuartigen Einzelmolekül-Methode
Tethered Multi-Fluorophore Motion

Dissertation von

Matthias R. J. Schickinger



TECHNISCHE UNIVERSITÄT MÜNCHEN



Technische Universität München
Fakultät für Physik
Lehrstuhl für Biomolekulare Nanotechnologie

Untersuchung der Hybridisierungskinetik
kurzer DNA-Doppelstränge
mit der neuartigen Einzelmolekül-Methode
Tethered Multi-Fluorophore Motion

Matthias R. J. Schickinger

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Martin Zacharias

Prüfer der Dissertation:

1. Prof. Dr. Hendrik Dietz

2. Prof. Dr. Matthias Rief

Die Dissertation wurde am 14.10.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 03.12.2020 angenommen.

Kurzzusammenfassung / Abstract

Diese Dissertation beschreibt die Entwicklung und Anwendung der neuartigen Einzelmolekül-Methode Tethered Multi-Fluorophore Motion (TMFM). Die Methode kombiniert Prinzipien der DNA-Nanotechnologie mit Fluoreszenzmikroskopie und Hidden-Markov-Modellierung zur Messung bimolekularer Reaktionskinetiken. Eine Besonderheit der Methode besteht in der Fähigkeit zur Auflösung heterogener Kinetik innerhalb einer Probe. TMFM wurde zur Untersuchung der Hybridisierungskinetik kurzer DNA-Doppelstränge in Abhängigkeit von verschiedenen Parametern verwendet.

This thesis reports on the development and application of the novel single-molecule method Tethered Multi-Fluorophore Motion (TMFM). The method employs principles of DNA nanotechnology along with fluorescence microscopy and Hidden-Markov modeling to measure bimolecular reaction kinetics. A special feature of the method is the ability to resolve heterogeneous kinetics within a sample. TMFM was used to study the hybridization kinetics of short DNA double strands as a function of different parameters.

Die zentralen Ergebnisse wurden im folgenden Artikel veröffentlicht:

Matthias Schickinger, Martin Zacharias, and Hendrik Dietz[°], "Tethered multifluorophore motion reveals equilibrium transition kinetics of single DNA double helices", *Proceedings of the National Academy of Sciences*, vol. 115, no. 39, pp. E7512-E7521, 2018

[°] Korrespondierender Autor

Inhaltsverzeichnis

Einleitung	1
1 DNA-Nano-Objekte für TMFM	5
1.1 Scaffolded DNA-Origami	5
1.2 Struktureller Aufbau der DNA-Nano-Objekte	6
1.3 Kontrolle der DNA-Nano-Objekte per Gel-Elektrophorese	19
1.4 Die intakte und die defekte Spezies	23
1.5 Beobachtung der DNA-Nano-Objekte im Fluoreszenzmikroskop	29
1.6 Kapitelzusammenfassung	33
2 Die Bewegung der Reporter-Einheiten	35
2.1 Wahl einer angemessenen Tracking-Prozedur	35
2.2 Positionsabfolgen und Auslenkungsabfolgen	36
2.3 Zusammenhang zwischen Helligkeit und gemessener Beweglichkeit	40
2.4 Aussortieren fehlerhafter DNA-Nano-Objekte	44
2.5 Charakterisierung der Bewegung der mobilen Reporter-Einheiten	45
2.6 Kapitelzusammenfassung	64
3 Statistische Auswertung von TMFM-Daten mit zwei Beweglichkeitszuständen	65
3.1 Hidden-Markov-Modellierung	65
3.2 Das Hidden-Markov-Modell für die TMFM-Daten	66
3.3 Weitere Schritte nach der HMM-Analyse	72
3.4 Kapitelzusammenfassung	83
4 Hybridisierungskinetik kurzer DNA-Duplexe	85
4.1 Einfluss des umgebenden Partikel-Leinen-Systems	85
4.2 Einfluss der Konzentration von Natriumionen	98
4.3 Einfluss der Basenpaarsequenz	105
4.4 Kapitelzusammenfassung	121
5 Ensemble-Kompetitions-Experimente zur Bestimmung des J-Faktors in TMFM	123
5.1 Konzept der Ensemble-Kompetitions-Experimente	123
5.2 Durchführung und Analyse der Kompetitions-Zeitreihen	125
5.3 Berechnung des J-Faktors aus dem zeitlichen Verlauf von C_{rel}	127
5.4 Resümee zu den Ensemble-Kompetitions-Experimenten	133
Fazit und Ausblick	135

A	Materialien und Methoden	141
A.1	Reagenzien	141
A.2	Laborprotokolle	142
A.3	Software	145
B	Ergänzende Informationen zu Kapitel 2	147
B.1	Kompensation des Drifts der TMFM-Probe in der xy-Ebene	147
B.2	Erprobung der HMM-Analyse an simulierten Datensätzen	149
C	Weitere ergänzende Abbildungen und Tabellen	157
D	Theoretische Ausführungen zu Kapitel 5	179
D.1	Herleitung von Modell 1 für C_{rel}	179
D.2	Theorie zur Analyse möglicher Einzelmolekül-Kompetitions-Experimente	180
	Literaturverzeichnis	183
	Danksagung	194

Einleitung

DNA kodiert das Erbgut aller bekannten Lebewesen. Eine fundamentale Eigenschaft der DNA ist die Hybridisierung zweier DNA-Einzelstränge (*single-stranded* DNA, kurz: ssDNA) revers-komplementärer Basensequenz zu einem DNA-Doppelstrang (*double-stranded* DNA, kurz: dsDNA). Die Spezifität der Basenpaarung und die reguläre Geometrie von DNA-Doppelhelices machen DNA über ihre biologische Funktion hinaus zu einem attraktiven Baumaterial für strukturierte, nanometerskalengroße Objekte. Wegweisend für die Konzeption solcher DNA-Nanoarchitekturen war die Arbeit von Nadrian Seeman aus dem Jahr 1982 [1]. Die seither andauernde Weiterentwicklung des Felds der strukturellen DNA-Nanotechnologie brachte die Realisierung einer Fülle verschiedenartiger künstlicher Nanosysteme aus DNA mit sich [2–7].

Einige dieser künstlichen Nanosysteme verwenden das Prinzip der reversiblen Hybridisierung von DNA-Strängen mit einer Länge von wenigen Basenpaaren (bp) zur Umsetzung dynamischer Funktionen. Beispiele dafür sind das Schalten zwischen verschiedenen räumlichen Konformationen [8–10], die Erzeugung gerichteter Bewegung von DNA-Walkern [11–13] oder DNA-Robotern [14], das Schließen und Öffnen von DNA-ummantelten Hohlräumen [15,16], die wiederholte ortsspezifische Arretierung beweglicher Strukturelemente [17], die Programmierung logischer Gatter [18,19] oder der ständige Austausch hoch lokalisierter Fluoreszenzmarkierungen für Super-Resolution-Mikroskopieaufnahmen [20,21].

In all diesen Mechanismen bestimmt die Kinetik der Assoziation und Dissoziation einzelner DNA-Doppelstränge die Geschwindigkeit des Gesamtprozesses. Zur Feinjustierung der Dynamik solcher Systeme oder zur Konzeption komplexerer Arrangements aus vielen, zeitlich koordinierten Komponenten wäre demnach ein detailliertes Verständnis der molekularen Kinetiken von großem Vorteil. Die Kinetik der DNA-Hybridisierung ist jedoch bis heute noch nicht komplett verstanden [22–25] und ist Gegenstand aktueller Forschung [23–44]. Für die Entwicklung verlässlicher Modelle zur Vorhersage der Hybridisierungskinetik einer beliebigen Basenpaarsequenz sind Messungen der Kinetik vieler verschiedener Sequenzvarianten notwendig [24].

Ein großer Anteil der bisherigen experimentellen Daten zur Kinetik der DNA-Hybridisierung stammt aus Ensemble-Betrachtungen von Relaxationsdynamiken aus Nichtgleichgewichtszuständen [24, 27–33, 42, 43, 45–53]. Der Prozess der Relaxation eines Ensembles in den Gleichgewichtszustand stellt eine Überlagerung vieler molekularer Assoziations- und Dissoziationsereignisse dar. Das Ensemble-Signal ist die Summe der Signale aller Moleküle der Probe. Feinheiten in den molekularen Kinetiken lassen sich daher aus Ensemble-Messungen nur schwer oder überhaupt nicht rekonstruieren.

Im Gegensatz dazu bieten Einzelmolekül-Experimente die Möglichkeit zur direkten Beobachtung einzelner molekularer Assoziations- und Dissoziationsereignisse im dyna-

mischen Gleichgewicht [20, 23, 26, 34, 35, 37, 39–41, 44, 54–56]. Die Aufzeichnung vieler einzelner Verweildauern der Interaktionspartner im assoziierten bzw. dissoziierten Zustand ermöglicht neben der Bestimmung mittlerer Übergangsraten auch die Erhebung von Wahrscheinlichkeitsverteilungen der Verweildauern in den jeweiligen Zuständen.

Desweiteren bergen Einzelmolekül-Messungen das Potential zur Identifikation von Subpopulationen mit unterschiedlicher Kinetik [54–56]. Einzelne Interaktionspartnerpaare einer kinetisch heterogenen Probe könnten anhand ihrer individuellen Verweildauerstatistik verschiedenen Subpopulationen zugeordnet werden. Allerdings korreliert die Verlässlichkeit einer solchen Zuordnung naturgemäß mit der Anzahl der detektierten Verweildauern pro Interaktionspartnerpaar. Daher ist die Auftrennung einer kinetisch heterogenen Probe mit den bisherigen Einzelmolekül-Methoden zur Messung der DNA-Hybridisierungskinetik aus zwei Gründen nur eingeschränkt realisierbar. Erstens werden in der Mehrzahl der bisherigen Methoden keine dauerhaft isolierten Paare von Interaktionspartnern betrachtet, sondern es erfolgt ein ständiger Austausch mit anderen Kopien der Interaktionspartner aus der Probenlösung. Zweitens wird die Detektion der Interaktionspartner in allen bisherigen Methoden durch Fluoreszenzmarkierung mit einzelnen Farbstoffmolekülen bewerkstelligt. Die Beobachtungsdauer jedes einzelnen Interaktionspartners ist deshalb durch die Lebensdauer eines einzelnen Farbstoffmoleküls begrenzt, welche typischerweise im Bereich von Sekunden bis wenigen Minuten liegt. Dadurch ist auch die Anzahl der detektierbaren Verweildauern pro Interaktionspartnerpaar stark beschränkt.

Etablierung einer neuartigen Einzelmolekül-Methode

Im Rahmen der hier vorgelegten Arbeit wurde eine neuartige Einzelmolekül-Methode zur Beobachtung wiederholter Assoziations- und Dissoziationsereignisse isolierter Interaktionspartnerpaare über eine ausgedehnte Zeitspanne etabliert. Diese Methode basiert auf dem Konzept der bereits existierenden Methoden *Tethered Particle Motion* (TPM) [57] und *Tethered Fluorophore Motion* (TFM) [58].

In TPM wird ein Partikel, beispielsweise ein Polystyrol-Kügelchen oder ein Gold-Nanopartikel, am Ende eines DNA-Doppelstrangs befestigt. Das andere Ende des DNA-Doppelstrangs wird an einer Oberfläche verankert. So wird der DNA-Doppelstrang zu einer molekularen Leine, welche die diffusive Bewegung des Partikels auf ein halbkugelförmiges Volumen begrenzt. Die Bewegung des Partikels wird mit einem Mikroskop aufgezeichnet. Aus der Bewegung des Partikels können Rückschlüsse auf die Gestalt und die mechanischen Eigenschaften der DNA-Leine gezogen werden. TFM funktioniert nach demselben Prinzip, wobei an Stelle des Partikels ein einzelnes Fluoreszenzfarbstoffmolekül am oberflächenfernen Ende der Leine sitzt. Zu den zahlreichen Anwendungen von TPM und TFM zählen Untersuchungen der Transkriptionskinetik [57, 59–61], der Protein-vermittelten Schlaufenbildung in dsDNA [62–77], der Aktivität von Restriktionsenzymen [78], der DNA-Mechanik [79–81], DNA-Rekombination [82–87] und DNA-Polymerisierung [58].

Eine grundlegende Neuerung der hier präsentierten Einzelmolekül-Methode gegenüber TPM und TFM ist die Verankerung zweier molekularer Interaktionspartner an spe-

zifischen Punkten des Partikel-Leinen-Systems. Die Interaktionspartner sind zueinander revers-komplementäre DNA-Einzelstränge. Durch die Verankerung im Partikel-Leinen-System werden die Interaktionspartner in einem gemeinsamen Reaktionsvolumen isoliert. Aus der Bewegung des oberflächenfernen Endes der Leine können Rückschlüsse auf den Bindezustand des Interaktionspartnerpaares gezogen werden.

Ein weiterer maßgeblicher Unterschied der hier präsentierten Methode zu TPM und TFM besteht in der Verwendung von Prinzipien der strukturellen DNA-Nanotechnologie zur Konstruktion von Reporter-Einheiten an beiden Enden der DNA-Leine. Die Reporter-Einheiten sind kompakte räumliche Gebilde aus DNA-Doppelhelices und tragen jeweils zwanzig Fluoreszenzfarbstoffmodifikationen. Dadurch kann die Bewegung des oberflächenfernen Endes der Leine über einen Zeitraum von mehr als einer Stunde im Fluoreszenzmikroskop aufgezeichnet werden.

Die Fluoreszenzfarbstoffe der Reporter-Einheit am oberflächennahen Ende der Leine haben ein anderes Emissionsspektrum als die Fluoreszenzfarbstoffe der Reporter-Einheit am oberflächenfernen Ende der Leine. Beide Enden der Leine können somit in verschiedenen spektralen Kanälen eines Fluoreszenzmikroskops abgebildet werden. Die Abbildung der Reporter-Einheit am oberflächennahen Ende der Leine ermöglicht u. a. den direkten Nachweis der korrekten Anbringung der einzelnen Partikel-Leinen-Systeme an der Probenoberfläche.

Wegen der Markierung der Partikel-Leinen-Systeme mit vielen Fluoreszenzfarbstoffmolekülen trägt die neu entwickelte DNA-Leinen-basierte Einzelmolekül-Methode den Namen *Tethered Multi-Fluorophore Motion* (TMFM).

Anwendung der neuen Methode in Einzelmolekül-Studien zur DNA-Hybridisierungskinetik

Die neu entwickelte Methode TMFM wurde für drei Studien der Hybridisierungskinetik kurzer DNA-Doppelstränge verwendet. Darin wurde die Abhängigkeit der gemessenen Hybridisierungskinetik von der Beschaffenheit des Partikel-Leinen-Systems, von der Konzentration von Natriumionen in der umgebenden Lösung und von der Basenpaarsequenz der Interaktionspartner untersucht. Die Ergebnisse der drei Studien verdeutlichen die Wichtigkeit der Berücksichtigung molekularer Heterogenität in Messungen mit chemisch synthetisierten Oligonukleotiden. Darüber hinaus leisten sie einen Beitrag zum besseren Verständnis der DNA-Hybridisierungskinetik.

Inhaltlicher Aufbau dieser Dissertation

Kapitel 1 zeigt die konzeptionellen und experimentellen Schritte der Entwicklung der Methode TMFM, vom strukturellen Design der DNA-Nano-Objekte für TMFM über die Ensemble-Charakterisierung der assemblierten Objekte bis hin zur Einzelmolekül-Beobachtung der Objekte im Fluoreszenzmikroskop.

Kapitel 2 beschreibt die Auswertung der Einzelmolekül-Mikroskopieaufnahmen. Insbesondere erfolgt hier eine detaillierte Charakterisierung der Dynamik verschiedener Partikel-Leinen-Systeme, welche die Grundlage für die weiteren Analyseschritte schafft.

Kapitel 3 beschreibt die Prozedur zum Rückschluss auf den Bindezustand eines Interaktionspartnerpaares aus der Bewegung des entsprechenden oberflächenfernen Endes der Leine. Außerdem werden verschiedene Korrekturen der experimentellen Daten vorgestellt, insbesondere die Vorgehensweise zur Berücksichtigung heterogener Hybridisierungskinetik innerhalb einer Probe von DNA-Nano-Objekten gleichen Designs.

Kapitel 4 präsentiert die Ergebnisse der drei TMFM-Studien zur Hybridisierungskinetik kurzer DNA-Doppelstränge in Abhängigkeit von der Beschaffenheit des Partikel-Leinen-Systems, von der Konzentration von Natriumionen in der umgebenden Lösung und von der Basenpaarsequenz der Interaktionspartner.

Kapitel 5 berichtet über eine ergänzende Ensemble-Kompetitions-Studie zur Bestimmung des *J-Factors*, eines im Kontext der TMFM-Messungen wichtigen Parameters.

1 DNA-Nano-Objekte für TMFM

Dieses Kapitel behandelt die Konzeption und die experimentelle Umsetzung der neuartigen Einzelmolekül-Methode Tethered Multi-Fluorophore Motion (TMFM). Zunächst erfolgt in Abschnitt 1.1 eine kurze Einführung der Technik *Scaffolded DNA-Origami* sowie damit verbundener Konzepte und Begriffe. Anschließend stellt Abschnitt 1.2 den strukturellen Aufbau der mit der DNA-Origami-Technik entwickelten Nano-Objekte vor. Abschnitte 1.3 und 1.4 zeigen die Ergebnisse der Qualitätskontrolle der assemblierten DNA-Nano-Objekte. Abschließend beschreibt Abschnitt 1.5 die Beobachtung einzelner DNA-Nano-Objekte im Fluoreszenzmikroskop.

1.1 Scaffolded DNA-Origami

Scaffolded DNA-Origami ist eine Technik zur Konzeption und Herstellung von ebenen [88] oder räumlichen [89, 90] nanometerskalengroßen DNA-Objekten definierter Struktur. Es gibt verschiedene Arten der Umsetzung dieses Prinzips. Die Ausführungen in diesem Abschnitt geben wieder, wie die Technik im Rahmen der hier vorgestellten Arbeit angewandt wurde. Für eine detailliertere Einführung in die Methodik sei beispielsweise auf die Veröffentlichung von Castro et al. aus dem Jahr 2011 [91] verwiesen.

Die mit Scaffolded DNA-Origami konzipierten DNA-Nano-Objekte bestehen aus zwei Sorten von DNA-Einzelsträngen: dem *Scaffold* und den *Staples*. Der Scaffold ist ein biotechnologisch hergestellter [92], mehrere Kilobasen (kb) langer zirkulärer Einzelstrang, dessen Basensequenz auf dem Genom des M13-Bakteriophagen basiert. Die Staples sind per chemischer Synthese hergestellte, lineare Einzelstränge mit einer Länge von ca. 20 bis 50 Nukleotiden (nt). Jedes im Rahmen dieser Arbeit konzipierte DNA-Nano-Objekt besteht aus genau einem Scaffold-Molekül und einem spezifischen Satz von etwa 200 verschiedenen Staple-Molekülen. Die Basensequenzen der Staples sind abschnittsweise komplementär zur Basensequenz des Scaffolds. Durch Hybridisierung zwischen basenkomplementären Abschnitten von Scaffold und Staples bilden sich doppelhelikale Domänen. Bindet ein Staple an zwei verschiedene Abschnitte des Scaffolds, werden dadurch zwei doppelhelikale Domänen miteinander verknüpft. Durch die Ausformung vieler solcher Verknüpfungen entsteht eine kompakte, geordnete Struktur. Ein gängiges Struktur-Element sind Helix-Bündel, d. h. Anordnungen paralleler, miteinander quervernetzter DNA-Doppelhelices in einem regelmäßigen Gitter.

In der Regel ist die Sequenz eines einzelnen Staples nicht die Komplementärsequenz eines durchgängigen Scaffold-Abschnitts. Die Mehrzahl der Staples ist derart gestaltet, dass Teile eines einzelnen Staple-Strangs an zwei oder mehrere verschiedene Abschnitte des Scaffolds binden. Ein DNA-Origami-Objekt kann als ein Geflecht aus DNA-Strängen

betrachtet werden, in welchem sich Stellen des Scaffold-Strangs in räumlicher Nähe zueinander befinden, die entlang des Scaffold-Rückgrats gemessen weit voneinander entfernt sind. Die Software caDNano [93] erleichtert die Konstruktion eines solchen Geflechts mithilfe einer grafischen Benutzeroberfläche. Ein Strang-Diagramm zeigt die Konnektivität der beteiligten DNA-Stränge. Dabei verlaufen die Achsen der Doppelhelices in horizontaler Richtung. Die Verknüpfungen zwischen verschiedenen Doppelhelices verlaufen in vertikaler Richtung (vgl. Abbn. 1.2, 1.3, 1.5 und 1.6).

Nach Eingabe der Basensequenz des Scaffolds erhält der Benutzer als Ausgabe die passenden Staple-Sequenzen. Die Synthese der Staples erfolgt üblicherweise durch einen kommerziellen Anbieter. Während oder auch nach der Synthese können die Staple-Stränge mit Modifikationen am Zucker-Phosphat-Rückgrat oder an einer der Basen versehen werden. Mehrere Staple-Stränge in den DNA-Nano-Objekten für TMFM wurden herstellerseitig mit Biotin-Gruppen oder verschiedenen Fluoreszenzfarbstoffen funktionalisiert.

Die DNA-Nano-Objekte werden nach dem Prinzip der Selbstassemblierung aus ihren einzelnen Bestandteilen zusammengefügt. Scaffold und Staples werden in salzhaltiger Lösung vermengt, wobei die Staples im Regelfall im molaren Überschuss zum Scaffold vorliegen. Die Reaktionsmischung wird nun zunächst erhitzt, um alle bereits bestehenden Basenpaarungen aufzulösen. Anschließend wird die Reaktionsmischung langsam abgekühlt. Während des Abkühlens binden die sequenzkomplementären Abschnitte des Scaffold-Strangs und der Staple-Stränge aneinander. Letztendlich formen sich vollständige DNA-Nano-Objekte aus. Details zur Herstellung der DNA-Nano-Objekte für TMFM finden sich in Anhang A.2.1.

1.2 Struktureller Aufbau der DNA-Nano-Objekte

Jedes DNA-Nano-Objekt besteht aus zwei Reporter-Einheiten und einer Leine (siehe Schemata in Abb 1.1). Die Leine ist ein DNA-Doppelstrang und verbindet die Reporter-Einheiten miteinander. Die Reporter-Einheiten sind Bündel aus DNA-Doppelhelices (vgl. Abb. 1.1 b und 1.2 bzw. 1.3). Sie sind mit Fluoreszenzfarbstoffen versehen, wobei für die beiden Einheiten zwei verschiedene Farbstoff-Sorten (Cy3 und Cy5) verwendet wurden. Aufgrund der unterschiedlichen Emissionsspektren der zwei Farbstoff-Sorten ist es möglich, die Reporter-Einheiten unabhängig voneinander in verschiedenen spektralen Kanälen eines Fluoreszenzmikroskops zu detektieren. Eine der Reporter-Einheiten trägt Biotin-Modifikationen, welche die spezifische Befestigung des Objekts an einer Oberfläche ermöglichen (Abb. 1.1 a und 1.2 c). Das Gebilde aus Reporter-Einheiten und DNA-Leine wird in dieser Arbeit auch als „Partikel-Leinen-System“ bezeichnet. Innerhalb des Partikel-Leinen-Systems können Paare von interaktionsfähigen Molekülen an definierten Positionen angebracht werden. In der hier präsentierten Arbeit wurde die Interaktion zwischen zueinander revers-komplementären DNA-Einzelsträngen untersucht (Abb. 1.1 und 1.5). Die folgenden Abschnitte erläutern Details zum Aufbau der Reporter-Einheiten und der DNA-Leine sowie zur Verankerung der Interaktionspartner.

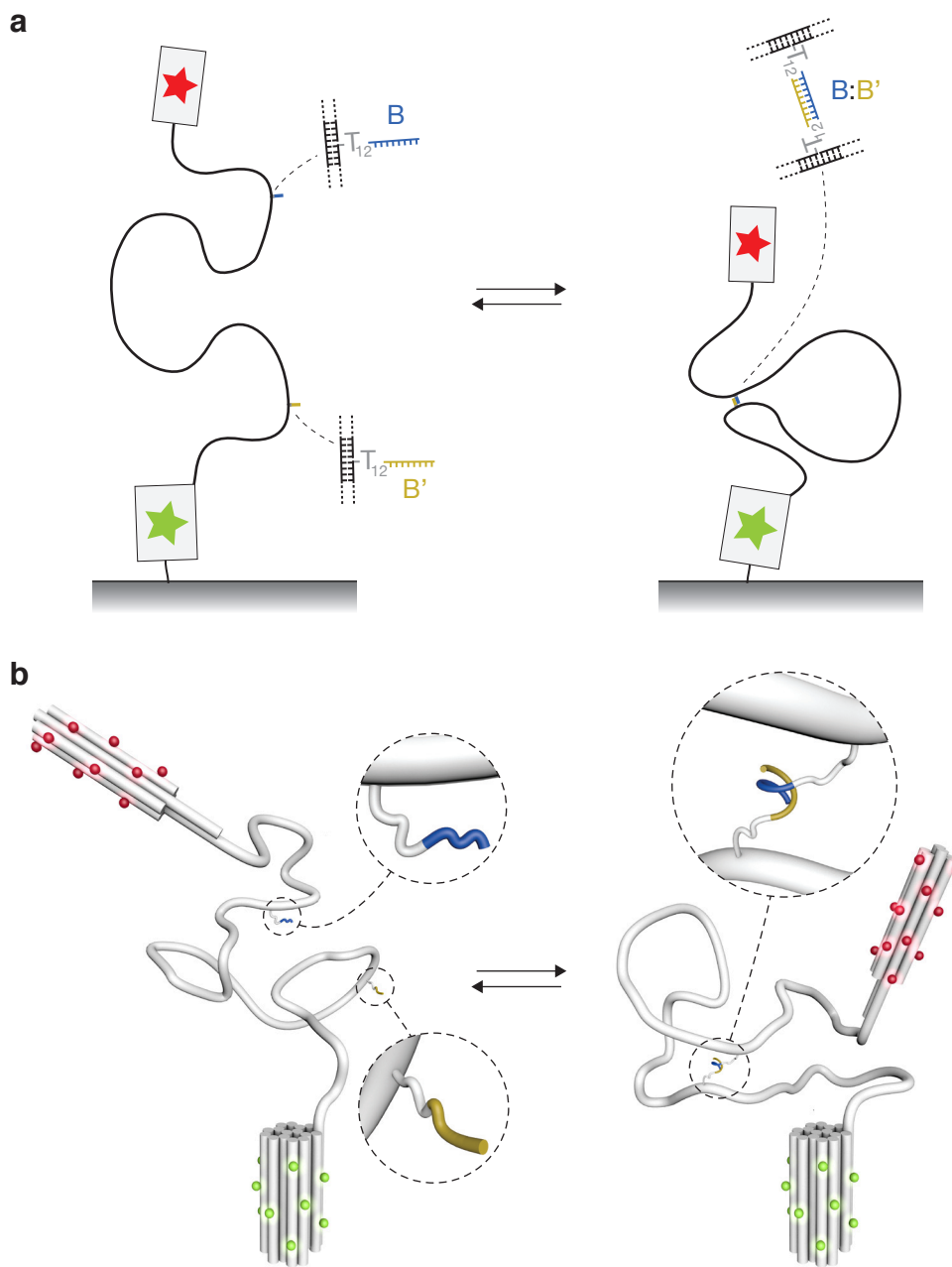


Abbildung 1.1: **Schemata eines DNA-Nano-Objekts in den Zuständen ohne Schlaufe und mit Schlaufe**
(a) Darstellung in 2D. Die grauen Rechtecke mit roten/grünen Sternen repräsentieren die Reporter-Einheiten mit verschiedenen Fluoreszenzfarbstoff-Markierungen. Die Reporter-Einheit mit dem grünen Stern ist an einer Probenoberfläche immobilisiert. Eine doppelsträngige DNA-Leine (schwarze Linie) verbindet die Reporter-Einheiten miteinander. Innerhalb der DNA-Leine zweigen Einzelstränge ab, die zueinander revers-komplementäre Sequenzen B und B' enthalten (linke Seite der reversiblen Reaktion). Mit Ausformung des Duplex B:B' entsteht eine Schlaufe in der DNA-Leine (rechte Seite der reversiblen Reaktion). **(b)** Darstellung in 3D. Dicke Zylinder repräsentieren DNA-Doppelhelices. Die Reporter-Einheiten sind Bündel aus DNA-Doppelhelices. Rote/grüne Kugeln deuten einzelne Fluoreszenzfarbstoffmoleküle an den Oberflächen der Reporter-Einheiten an. Dünne Zylinder repräsentieren die Einzelstränge mit den Interaktionssequenzen B und B' (beachte Detailansichten).

1.2.1 Die immobile Reporter-Einheit

Die Reporter-Einheit, die zur Befestigung an der Oberfläche vorgesehen ist, wird in dieser Dissertation als „immobile Reporter-Einheit“ bezeichnet. Abbildung 1.2 zeigt Details zu ihrem Aufbau. Der Korpus der immobilen Reporter-Einheit besteht aus etwa 1300 Basenpaaren, die sich auf 16 wabenförmig angeordnete Doppelhelices verteilen. Parallel zur Symmetrieachse der Helices beträgt die Länge des Korpus etwa 28 nm, wenn man von einem mittleren Beitrag pro Basenpaar von 0,335 nm ausgeht. Die größte Ausdehnung der immobilen Reporter-Einheit senkrecht zur helikalen Achse misst 15 nm, wenn man von einem mittleren Beitrag pro Doppelhelix von 2,5 nm ausgeht.

Zwanzig Staple-Moleküle der immobilen Reporter-Einheit sind in 5'-Richtung um zwei T-Nukleotide verlängert, die nicht dafür vorgesehen sind, an den Scaffold zu binden. Stattdessen sollen diese einzelsträngigen Fortsätze an oberflächenexponierten Stellen von ihren jeweiligen Doppelhelices abzweigen und in die umgebende Lösung herausragen (Abb. 1.2 c, grün eingefärbte Staples). Jeder dieser Fortsätze ist am 5'-Terminus mit einem Cy3-Farbstoff-Molekül modifiziert. Die Positionen der Cy3-modifizierten Termini sind möglichst gleichmäßig über die Oberfläche der Reporter-Einheit verteilt, um Fluoreszenzauslöschung zwischen den Farbstoff-Molekülen zu vermeiden.

An einer Stirnseite der immobilen Reporter-Einheit ragen sechs Staple-Fortsätze mit je 15 T-Basen Länge aus den Enden von Doppelhelices hervor (Abb. 1.2 c, magentafarbene Staples). Die Termini dieser Fortsätze tragen jeweils eine Biotin-TEG-Modifikation. Die Biotin-Reste dienen im TMFM-Experiment zur Verankerung der Reporter-Einheit an der Probenoberfläche. Letztere ist mit PEG-Ketten funktionalisiert, welche an einem Ende kovalent mit der Oberfläche verbunden sind. Ein Teil dieser PEG-Ketten trägt am oberflächenfernen Ende ebenfalls eine Biotin-Modifikation. Die Verankerung der immobilen Reporter-Einheiten kommt über Biotin-NeutrAvidin-Biotin-Brücken zwischen den Biotin-modifizierten Staple-Überhängen und den Biotin-modifizierten PEG-Ketten zustande. Aufgrund dieser Art der Verankerung verfügt die immobile Reporter-Einheit – entgegen ihrer Bezeichnung – über einen gewissen Bewegungsspielraum. Details zur Verankerung der DNA-Nano-Objekte finden sich in Anhang A.2.4.

An der gegenüberliegenden, im TMFM-Experiment oberflächenfernen Stirnseite der immobilen Reporter-Einheit schließt als Verlängerung von Helix Nummer 7 die DNA-Leine an die Plattform an (Abb. 1.2 c, mit Stern gekennzeichnete Stelle am rechten Rand). Die DNA-Leine mündet an ihrem anderen Ende in die mobile Reporter-Einheit.

1.2.2 Die mobile Reporter-Einheit

Diese Reporter-Einheit bewegt sich im TMFM-Experiment diffusiv in einem durch die Konturlänge der DNA-Leine gegebenen, halbkugelförmigen Volumen. Ihr Bewegungsspielraum ist deutlich größer als der Bewegungsspielraum der immobilen Reporter-Einheit. Daher wird sie in dieser Dissertation als „mobile Reporter-Einheit“ bezeichnet. Abbildung 1.3 zeigt Details zu ihrem Aufbau. Die mobile Reporter-Einheit ist ein Bündel aus sechs schachbrettartig angeordneten Doppelhelices, welches insgesamt etwa 500 Basenpaare umfasst. Die Länge des Korpus der mobilen Reporter-Einheit parallel

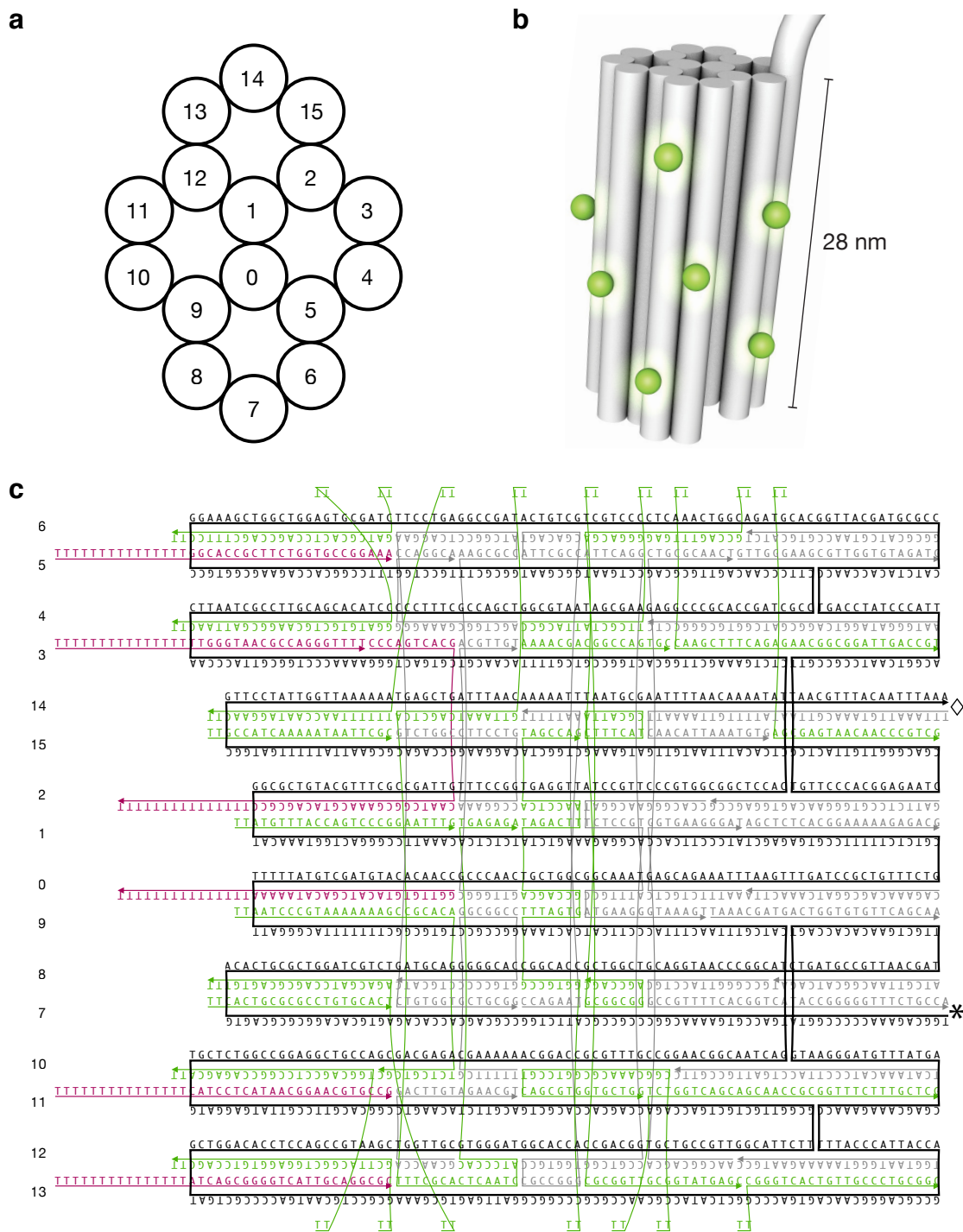


Abbildung 1.2: Details zum Aufbau der immobilen Reporter-Einheit

(a) Querschnittsfläche senkrecht zur helikalen Achse. (b) Vergrößerte Ansicht der 3D-Darstellung aus Abb. 1.1. Zylinder repräsentieren einzelne DNA-Doppelhelices. Kugeln deuten einzelne Cy3-Moleküle an. (c) Strang-Diagramm, exportiert aus caDNAo v0.2 [93]. Nummerierung der Helices wie in a. Scaffold in Schwarz. Unmodifizierte Staples in Grau, Cy3-modifizierte Staples in Grün, Biotin-TEG-modifizierte Staples in Magenta. Stern resp. Raute kennzeichnen die Anknüpfungspunkte der DNA-Leine resp. des nach dem enzymatischen Verdau verbleibenden doppelsträngigen Fortsatzes (siehe Text ab S. 11 sowie Abb. 1.4).

zur Symmetrieachse der Helices beträgt ebenfalls etwa 28 nm, wenn man von einem mittleren Beitrag pro Basenpaar von 0,335 nm ausgeht. Ihre größte Ausdehnung senkrecht zur helikalen Achse misst rund 9 nm, wenn man von einem mittleren Beitrag pro Doppelhelix von 2,5 nm ausgeht.

Die Fluoreszenzfarbstoffe sind an der mobilen Reporter-Einheit nach demselben Prinzip angebracht wie an der immobilen Reporter-Einheit: An den 5'-Enden von 20 Staple-Strängen sollen einzelsträngige Fortsätze aus zwei T-Nukleotiden von ihren jeweiligen Doppelhelices abzweigen und in die umgebende Lösung herausragen (Abb. 1.3 c, rot eingefärbte Staples). Jeder dieser Staple-Stränge ist am 5'-Terminus mit einem Cy5-Farbstoff-Molekül modifiziert. Wie bei der immobilen Reporter-Einheit sind auch bei der mobilen Reporter-Einheit die Positionen der Farbstoff-Moleküle möglichst gleichmäßig über die Oberfläche der Reporter-Einheit verteilt, um gegenseitige Fluoreszenzauslöschung zu vermeiden.

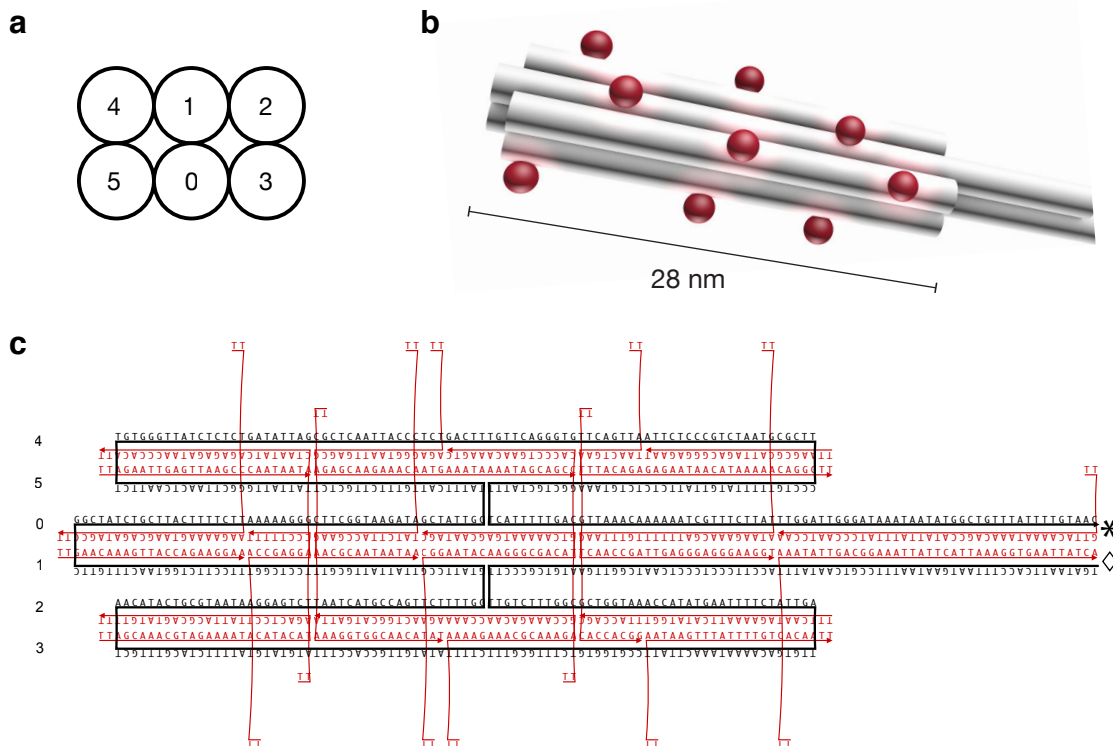


Abbildung 1.3: **Details zum Aufbau der mobilen Reporter-Einheit**

(a) Querschnittsfläche senkrecht zur helikalen Achse. (b) Vergrößerte Ansicht der 3D-Darstellung aus Abb. 1.1. Zylinder repräsentieren einzelne DNA-Doppelhelices. Kugeln deuten einzelne Cy5-Moleküle an. (c) Strang-Diagramm, exportiert aus caDNAo v0.2 [93]. Nummerierung der Helices wie in a. Scaffold in Schwarz, Staples in Rot. Jedes Staple ist Cy5-modifiziert. Stern resp. Raute kennzeichnen die Anknüpfungspunkte der DNA-Leine resp. des nach dem enzymatischen Verdau verbleibenden doppelsträngigen Fortsatzes (siehe Text ab S. 11 sowie Abb. 1.4).

1.2.3 Aufbau der DNA-Leine und Verankerung der Interaktionspartner

Die Abschnitte des Scaffolds zwischen den Reporter-Einheiten bleiben ungefaltet. Sie werden lediglich durch Zugabe von Staples mit den entsprechenden Komplementär-Sequenzen vollständig doppelsträngig gemacht. Diese Staples werden im Folgenden als „Leinen-Staples“ bezeichnet und haben fast alle eine Länge von 42 nt. Da der Scaffold zirkulär ist, existieren nach der Selbstassemblierung der DNA-Nano-Objekte zunächst zwei doppelsträngige Verbindungen zwischen den Reporter-Einheiten. Sie verbinden jeweils die zwei mit Sternen bzw. Rauten gekennzeichneten Positionen in Abb. 1.2 c und Abb. 1.3 c.

Im TMFM-Experiment sollen die Reporter-Einheiten durch eine einzelne Leine verbunden sein. Um dies zu erreichen, wird eine der zwei Verbindungen zwischen den Reporter-Einheiten durch enzymatischen Verdau herausgeschnitten. Abbildung 1.4 veranschaulicht diesen Arbeitsschritt. Das Erkennungsmotiv des Restriktionsenzym *Cla*I (Abb. 1.4 a) taucht im Scaffold genau zweimal auf. Die Sequenz des Scaffolds wurde beim Design der DNA-Nano-Objekte derart positioniert, dass diese beiden Stellen innerhalb der Verbindung zwischen Helix 14 der immobilen Reporter-Einheit und Helix 1 der mobilen Reporter-Einheit liegen (Abb. 1.4 b, links). Durch den enzymatischen Verdau mit *Cla*I wird diese Verbindung unterbrochen und ein etwa 2,9 kbp langes Fragment

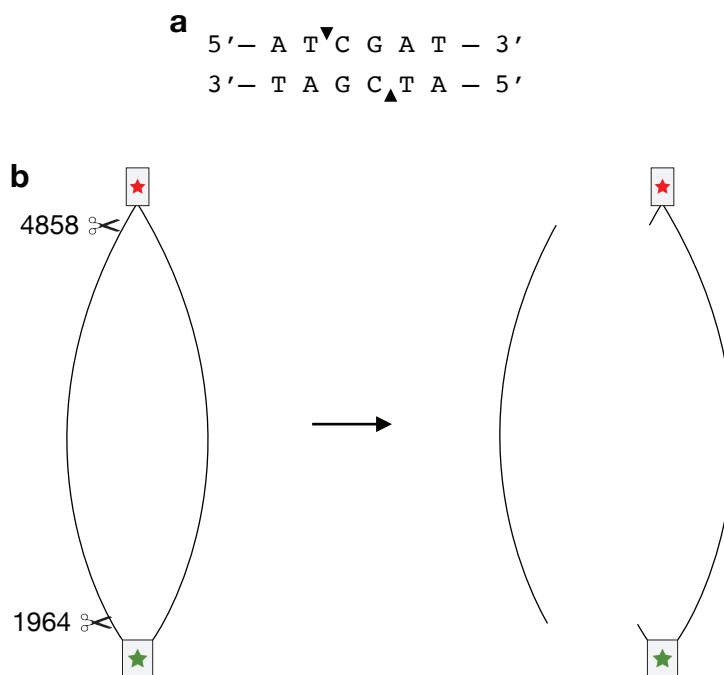


Abbildung 1.4: **Linearisierung des Scaffolds durch Verdau mit der Restriktionsendonuklease *Cla*I**

(a) Erkennungsmotiv für das Restriktionsenzym *Cla*I. Das Enzym zerschneidet das DNA-Rückgrat an den durch Pfeile gekennzeichneten Stellen. **(b)** Schema der Verdau-Reaktion. Graue Rechtecke mit roten/grünen Sternen repräsentieren die Reporter-Einheiten. Schwarze Linien repräsentieren die ungefalteten, doppelsträngigen Abschnitte des DNA-Nano-Objekts. Scheren-Symbole kennzeichnen die zwei Stellen, an denen das Erkennungsmotiv in der Scaffold-Sequenz auftritt. Die Zahlen stehen für den Index der jeweiligen C-Base im Scaffold-Strang (gezählt ab dem Startpunkt der Sequenz p8064 in caDNAo v0.2 [93]).

aus dem Scaffold-Staple-Doppelstrang herausgeschnitten (Abb. 1.4 b, rechts). Es verbleiben doppelsträngige Fortsätze an Helix 14 der immobilen Reporter-Einheit resp. Helix 1 der mobilen Reporter-Einheit mit einer Länge von 87 resp. 131 bp. An den Enden der Fortsätze befindet sich jeweils ein 5'-CG-Überhang. Es ist prinzipiell denkbar, dass diese beiden Überhänge miteinander hybridisieren und dadurch eine sehr kurze Verbindung zwischen den Reporter-Einheiten herstellen. Die Stabilität einer Duplexbindung von nur zwei GC-Basenpaaren Länge ist jedoch vergleichsweise gering. Eine solche Bindung wäre, falls sie überhaupt auftritt, sehr kurzlebig. In Kontroll-Experimenten wurden keine Anzeichen beobachtet, die auf einen potenziell störenden Einfluss dieser kurzen Überhänge auf das Mess-System hindeuten (vgl. Abschnitt 2.5).

Die Verbindung zwischen Helix 7 der immobilen Reporter-Einheit und Helix 0 der mobilen Reporter-Einheit verbleibt nach dem Verdau mit ClaI als Leine zwischen den Reporter-Einheiten. Neben der Funktion der Leine, die Reporter-Einheiten zu verbinden, bietet sie die Möglichkeit zur Verankerung molekularer Interaktionspartner. Immer dort, wo zwei benachbarte Leinen-Staples aneinander grenzen, befindet sich ein Einzelstrangbruch. Molekulare Interaktionspartner können z. B. als terminale Staple-Modifikationen

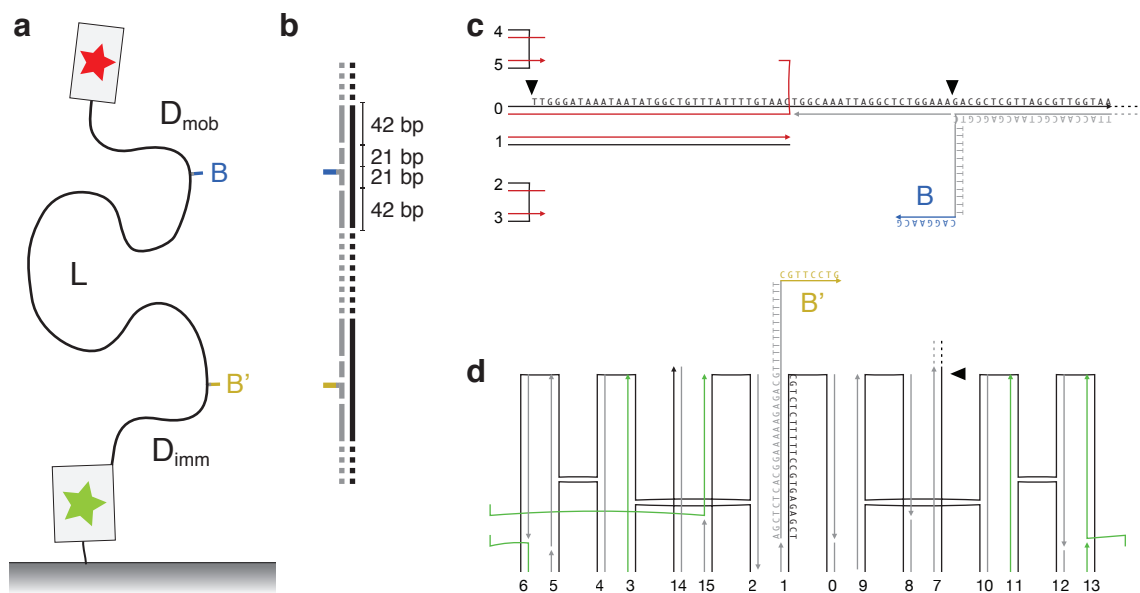


Abbildung 1.5: **Verankerung der Interaktionspartner – distale und proximale Geometrie**

(a) Verankerung der Interaktionspartner in der distalen Geometrie. Strichzeichnung eines DNA-Nano-Objekts analog zu Abb. 1.1 a. Die grauen Rechtecke mit rotem/grünem Stern repräsentieren die Reporter-Einheiten. Die schwarze Linie repräsentiert die DNA-Leine. D_{mob} : Konturlänge des Leinenstücks zwischen der mobilen Reporter-Einheit und dem Ankerpunkt von Interaktionspartner B. D_{imm} : Konturlänge des Leinenstücks zwischen der immobilen Reporter-Einheit und dem Ankerpunkt von Interaktionspartner B'. L : Konturlänge des Leinenstücks zwischen den Ankerpunkten der beiden Interaktionspartner. (b) Schema des Aufbaus der DNA-Leine mit Verankerung der Interaktionspartner in der distalen Geometrie. Scaffold in Schwarz, Staples in Grau und Farbe. Gepunktete Abschnitte: Fortsetzung einer doppelsträngigen DNA-Leine aus durchgängigem Scaffold und je 42 nt langen Staples. (c,d) Verankerung der Interaktionspartner in der proximalen Geometrie. Strang-Diagramme analog zu Abbn. 1.2 und 1.3. Scaffold in Schwarz, Staples in Grau und Farbe. Die gestrichelten Linien deuten die Fortsetzung der DNA-Leine zwischen den Reporter-Einheiten an. Die Dreiecke kennzeichnen die Bezugspunkte für die Berechnung der Konturlängen D_{mob} und D_{imm} .

an diesen Positionen verankert werden. Der Abstand zwischen zwei Basenpaaren (ca. 0,335 nm) bestimmt die räumliche Genauigkeit der Verankerung. Im Rahmen der hier präsentierten Arbeit handelt es sich bei den Interaktionspartnern um zueinander revers-komplementäre DNA-Sequenzen, welche mit B und B' bezeichnet werden (Abb. 1.5 a). Sie werden als 3'-terminale Verlängerungen von Staples am DNA-Nano-Objekt verankert (Abb. 1.5 b-d). Dabei ist jeweils zwischen dem an den Scaffold bindenden Teil und der Interaktionssequenz eine Verbinder-Sequenz aus zwölf T-Nukleotiden eingefügt (Abb. 1.5 c,d). Zur Verankerung eines Interaktionspartners in der Leine wird ein 42 nt langes Leinen-Staple im Reaktionsgemisch der Selbstassemblierung durch zwei Staples ersetzt, die jeweils 21 nt dieses 42-mers enthalten. Eines dieser zwei Staples ist zusätzlich in 3'-Richtung um den T₁₂-Verbinder und die jeweilige Interaktionssequenz verlängert. Der so entstandene einzelsträngige Fortsatz aus dem T₁₂-Verbinder und der Interaktionssequenz zweigt am 3'-Ende des an den Scaffold bindenden, 21 nt langen Abschnitts dieses Staples von der Leine ab (Abb. 1.5 b,c).

Ein Teilaspekt der Charakterisierung der Methode TMFM ist der Einfluss der Geometrie des Partikel-Leinen-Systems auf die Interaktion zwischen den Einzelsträngen B und B' (siehe Abschnitt 4.1). Zur Untersuchung dieses Einflusses wurden verschiedene Varianten von DNA-Nano-Objekten hergestellt. Konkret wurden die Konturlängen der drei Leinenstücke zwischen der mobilen Reporter-Einheit und dem Ankerpunkt von Interaktionspartner B (D_{mob}), zwischen der immobilen Reporter-Einheit und dem Ankerpunkt von Interaktionspartner B' (D_{imm}) sowie zwischen den Ankerpunkten der Interaktionspartner (L) variiert. In der „proximalen Verankerungsgeometrie“ ist D_{mob} gleich 55 bp (Abb. 1.5 c) und D_{imm} ist gleich Null, da der Einzelstrang mit der Sequenz B' aus einem Helixende der immobilen Reporter-Einheit hervorsticht (Abb. 1.5 d). In der „distalen Verankerungsgeometrie“ ist D_{mob} gleich 307 bp und D_{imm} gleich 252 bp.

Beide Verankerungsgeometrien wurden mit mehreren verschiedenen Konturlängen L kombiniert, bei jeweils gleich bleibenden Werten von D_{mob} und D_{imm} . Es ist folglich notwendig, die Gesamtlänge der DNA-Leine zwischen den Reporter-Einheiten anzupassen. Der folgende Abschnitt beschreibt den dazu verwendeten Mechanismus.

1.2.4 Mechanismus zur Anpassung der Gesamtleinenlänge

Zur besseren Vergleichbarkeit zwischen den Experimenten mit DNA-Nano-Objekten unterschiedlicher Gesamtleinenlänge ist es sinnvoll, wenn die Sequenzen der Staples der Reporter-Einheiten unverändert bleiben. Außerdem wäre eine individuelle Synthese der Farbstoff- oder Biotin-modifizierten Staples für jede Leinenlängen-Variante sehr kostspielig. Es wurde ein Mechanismus entwickelt, welcher es durch Austausch einzelner Untergruppen von Staples ermöglicht, verschiedene Gesamt-Leinenlängen einzustellen und dabei die Sequenzen aller modifizierten Staples und fast aller unmodifizierten Staples in den Reporter-Einheiten beizubehalten. Abbildungen 1.6 und 1.7 veranschaulichen den sogenannten „Aufroll-Mechanismus“, durch den die Leine ausgehend von ihrer maximalen Gesamtlänge um eines oder mehrere Segmente zu je 252 bp verkürzt werden kann. Zur Verkürzung der Leine um das erste Segment wird ein 6-Helix-Bündel mit einer totalen Scaffold-Länge von 252 Nukleotiden an Helix 7 der immobilen

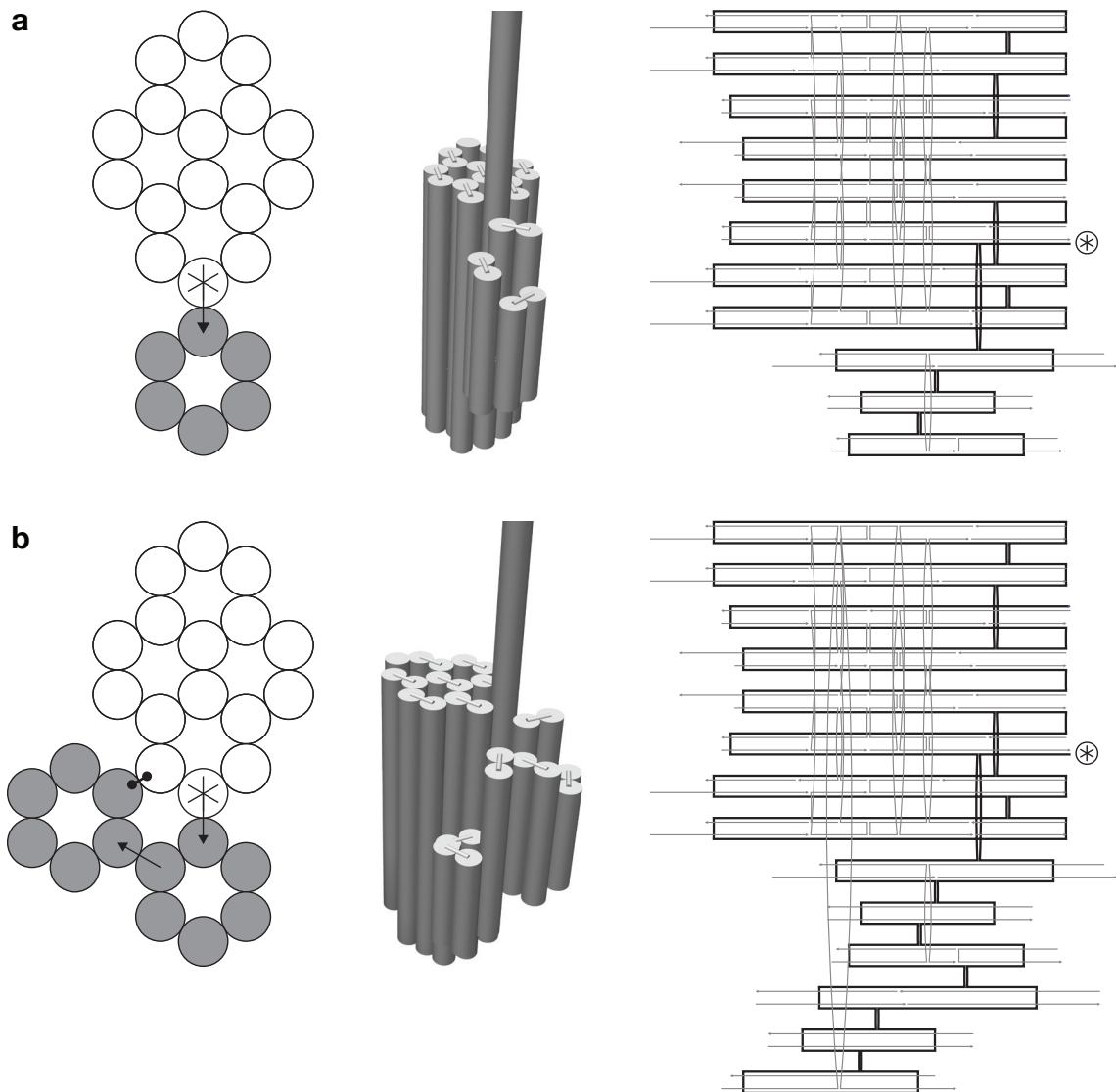


Abbildung 1.6: **Mechanismus zur Anpassung der Leinenlänge: Verkürzung um ein bis zwei Segmente**
 Funktionsprinzip der Verkürzung der Leine um **(a)** ein und **(b)** zwei 252-bp-Segmente. Links: Querschnittsflächen der immobilen Reporter-Einheit (weiß) mit angrenzenden aufgerollten Leinensegmenten (grau). Der Stern kennzeichnet jeweils Helix 7. Pfeile deuten Scaffold-Verbindungen zwischen den Bündeln und der Reporter-Einheit an. Die hantelförmige Verbindungslinie in b deutet die Verknüpfung zweier Helices durch Staples an (oberste und unterste Helix im Strang-Diagramm). Mitte: 3D-Darstellung der immobilen Reporter-Einheit mit je einem angrenzenden kurzen 6-Helix-Bündel pro aufgerolltem Leinen-Segment. Zylinder repräsentieren einzelne DNA-Doppelhelices. Grafiken wurden mit CanDo [94] erstellt. Rechts: Strang-Diagramm, exportiert aus caDNAo v0.2 [93]. Scaffold in Schwarz, Staples in Grau. Sterne kennzeichnen den Anknüpfungspunkt der Leine an Helix 7.

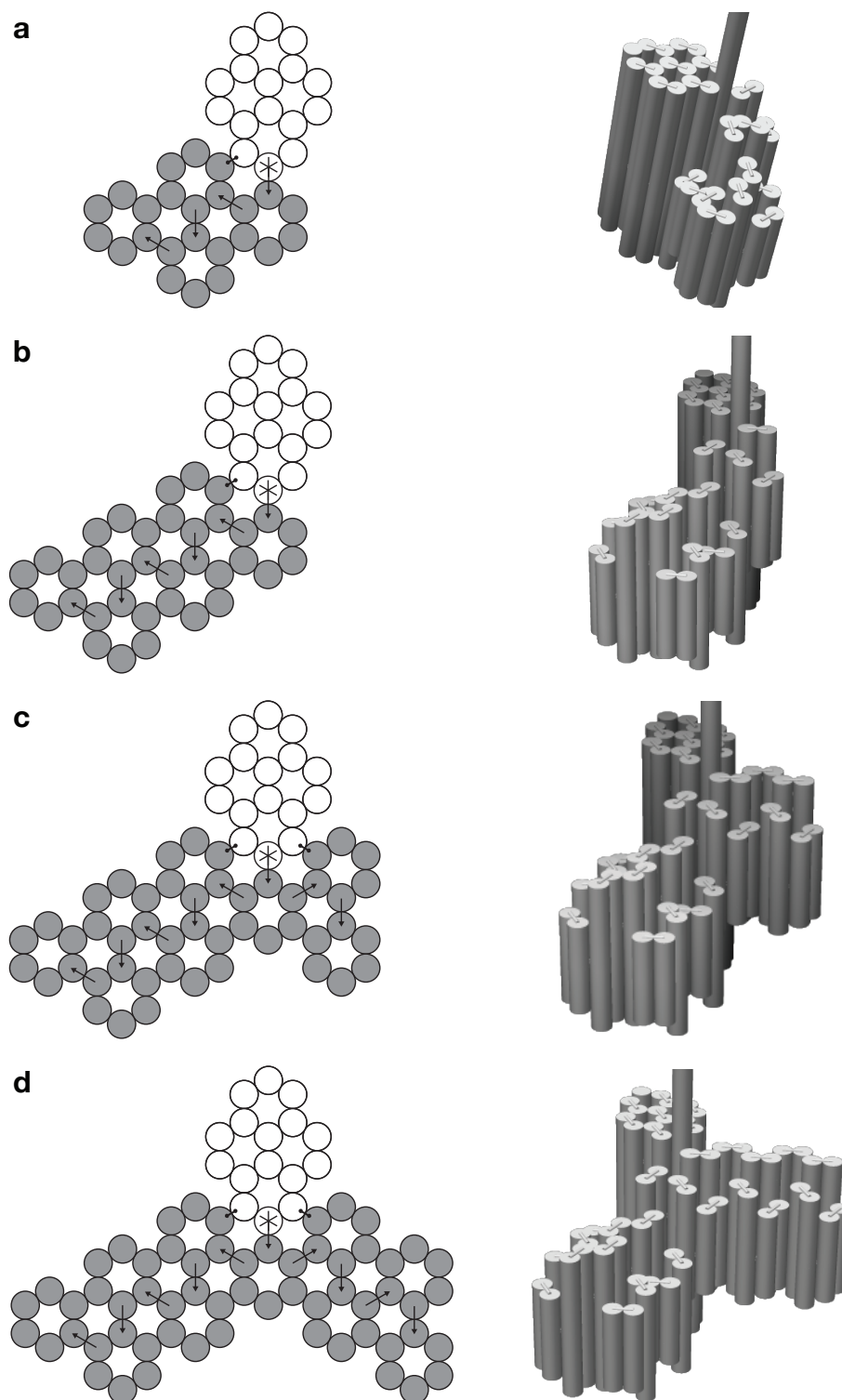


Abbildung 1.7: **Mechanismus zur Anpassung der Leinenlänge: Verkürzung um bis zu zehn Segmente**
 Verkürzung der Leine um (a) vier, (b) sechs, (c) acht oder (d) zehn Segmente. Jeweils links: Querschnittsflächen der immobilen Reporter-Einheit (weiß) mit angrenzenden aufgerollten Leinensegmenten (grau). Der Stern kennzeichnet jeweils Helix 7. Pfeile deuten Scaffold-Verbindungen an. Hantelförmige Verbindungslinien deuten Verknüpfungen zweier Helices durch Staples an. Jeweils rechts: Entsprechende 3D-Darstellungen, erstellt mit CanDo [94]. Zylinder repräsentieren einzelne DNA-Doppelhelices.

Reporter-Einheit abgezweigt (Abb. 1.6 a). Der freie Teil der Leine verkürzt sich dementsprechend um 252 bp. Zur Verkürzung um eine weitere Stufe wird ein weiteres 6-Helix-Bündel an Helix 7 der immobilen Reporter-Einheit abgezweigt (Abb. 1.6 b). Das erste Bündel zweigt nun nicht mehr an Helix 7 der immobilen Reporter-Einheit ab, sondern an einer Helix des Bündels der zweiten Stufe. Auf diese Art und Weise entsteht pro 252-bp-Segment, um das die Leine verkürzt wird, ein zusätzliches kurzes 6-Helix-Bündel am Rand der immobilen Reporter-Einheit (vgl. Abb. 1.7). Die kürzeste Gesamt-Leinenlänge aller im Rahmen dieser Arbeit für TMFM-Experimente verwendeten Objekte hatte das Objekt mit $L = 546$ bp und der proximalen Verankerungsgeometrie. Dazu wurden zehn Segmente, d. h. insgesamt 2520 bp, in zehn kurzen 6-Helix-Bündeln am Rand der immobilen Reporter-Einheit aufgerollt (Abb. 1.7 d).

1.2.5 Referenzobjekte für die Zustände ohne und mit Schlaufe in der Leine

Die DNA-Nano-Objekte mit Interaktionspartnern können entweder ohne oder mit einer Schlaufe in der DNA-Leine vorliegen (vgl. Abb. 1.1). Sind die Interaktionspartner B und B' voneinander gelöst, so hat das Objekt keine Schlaufe in der Leine. Mit Ausformung des Duplex B:B' entsteht eine Schlaufe in der Leine. Für beide Zustände wurden Referenz-Objekte entworfen, die bei korrekter Selbstassemblierung dauerhaft in einem der beiden Zustände verweilen sollen.

Als Referenz für den Zustand ohne Schlaufe in der Leine wurden Objekte entworfen, bei denen eine Interaktionssequenz oder beide Interaktionssequenzen fehlen (Abb. 1.8). An den Ankerpunkten der Interaktionspartner in der distalen Verankerungsgeometrie befinden sich in keinem dieser Objekte einzelsträngige Staple-Überhänge (Abb. 1.8 a). Die i_0 -Referenz hat an den Ankerpunkten der Interaktionspartner in der proximalen Verankerungsgeometrie jeweils einen 3'-terminalen, zwölf T-Nukleotide langen einzelsträngigen Staple-Überhang. Diese zwei Staple-Überhänge tragen jedoch keine Interaktionssequenzen (Abb. 1.8 b, c, d). Die B_0 -Referenz hat am Ankerpunkt nahe der mobilen Reporter-Einheit in der proximalen Verankerungsgeometrie einen einzelsträngigen 3'-terminalen Staple-Überhang mit T_{12} -Verbinder und Interaktionssequenz B (Abb. 1.8 e, f). Am Ankerpunkt an Helix 1 der immobilen Reporter-Einheit hat dieses Referenz-Objekt jedoch keinen einzelsträngigen Staple-Überhang (Abb. 1.8 e, i). Die 0-Referenz hat gar keine einzelsträngigen Überhänge an möglichen Ankerpunkten der Interaktionspartner (Abb. 1.8 a, g, h, i).

Die Objekte der i_0 -Referenz, der 0-Referenz und der B_0 -Referenz sollten keine Schlaufe in der DNA-Leine ausformen können. Träte in diesen Objekten doch eine Schlaufenbildung auf, würde das auf einen weiteren Mechanismus der Schlaufenbildung neben der Bildung des Duplex B:B' hindeuten.

Als Referenzen für den Zustand mit Schlaufe wurden Objekte entworfen, bei denen die Scaffold-Abschnitte nahe der Ankerpunkte der Interaktionspartner durch Staples fest miteinander verbunden werden (Abb. 1.9). Die F_p -Referenz ahmt den Zustand mit Schlaufe bei Verankerung der Interaktionspartner in der proximalen Verankerungsgeometrie nach (Abb. 1.9 a, b). Die F_d -Referenz ahmt den Zustand mit Schlaufe bei Verankerung der Interaktionspartner in der distalen Verankerungsgeometrie nach (Abb. 1.9 c-e).

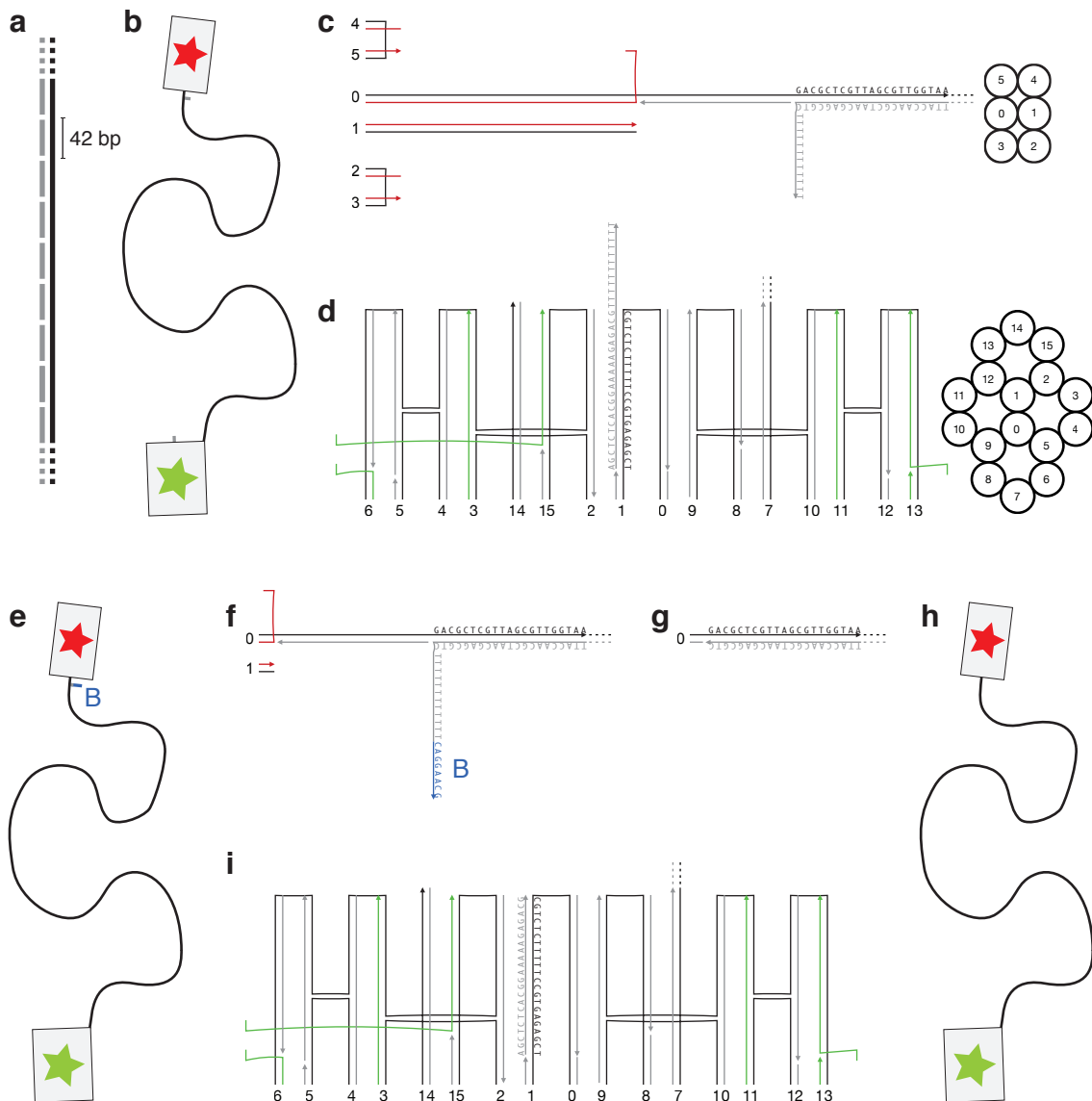


Abbildung 1.8: **Referenzobjekte ohne Schlaufe in der Leine: i0-Referenz, 0-Referenz und B0-Referenz**
(a) Schema des Aufbaus der DNA-Leine ohne Staple-Überhänge an den Ankerpunkten der Interaktionspartner in der distalen Verankerungsgeometrie. Scaffold in Schwarz, Staples in Grau. Gepunktete Abschnitte: Fortsetzung einer doppelsträngigen DNA-Leine aus durchgängigem Scaffold und je 42 nt langen Staples.
(b) Schema der i0-Referenz. Die grauen Rechtecke mit roten / grünen Sternen repräsentieren die Reporter-Einheiten. Die schwarze Linie repräsentiert die DNA-Leine. Die kurzen grauen Linien deuten einzelsträngige T_{12} -Überhänge an. **(c,d)** Strang-Diagramme für die Verankerung der T_{12} -Überhänge in der i0-Referenz: **(c)** Leinenende nahe der mobilen Reporter-Einheit, **(d)** Seite der immobilen Reporter-Einheit mit dem Ankerpunkt des Interaktionspartners in der proximalen Verankerungsgeometrie. Scaffold in Schwarz, Staples in Grau und Farbe. Die gestrichelten Linien deuten die Fortsetzung der DNA-Leine zwischen den Reporter-Einheiten an. Jeweils ganz rechts: Querschnitte der jeweiligen Reporter-Einheit senkrecht zur helikalen Achse, mit derselben Nummerierung der Helices wie im entsprechenden Strang-Diagramm. **(e)** Schema der B0-Referenz, analog zu a. Die kurze blaue Linie deutet die Interaktionssequenz B an. **(f,g)** Strang-Diagramme des Leinenendes nahe der mobilen Reporter-Einheit, jeweils analog zu c, für **(f)** die B0-Referenz (Interaktionssequenz B in Blau) und **(g)** die 0-Referenz. **(h)** Schema der 0-Referenz, analog zu a. **(i)** Strang-Diagramm eines Teils der immobilen Reporter-Einheit für die 0- und B0-Referenzen, analog zu d.

In jeder der Referenzen für den Zustand mit Schlaufe sollen die Ankerpunkte der Interaktionspartner durch zwei bestimmte Staples fest miteinander verknüpft werden. Jedes dieser Staples kann mit einem Teil an einen Scaffold-Abschnitt nahe des Ankerpunktes eines Interaktionspartners binden und mit dem restlichen Teil an einen Scaffold-Abschnitt nahe des Ankerpunktes des anderen Interaktionspartners binden (Abb. 1.9b,d,e). Dabei besteht allerdings auch eine gewisse Wahrscheinlichkeit, dass

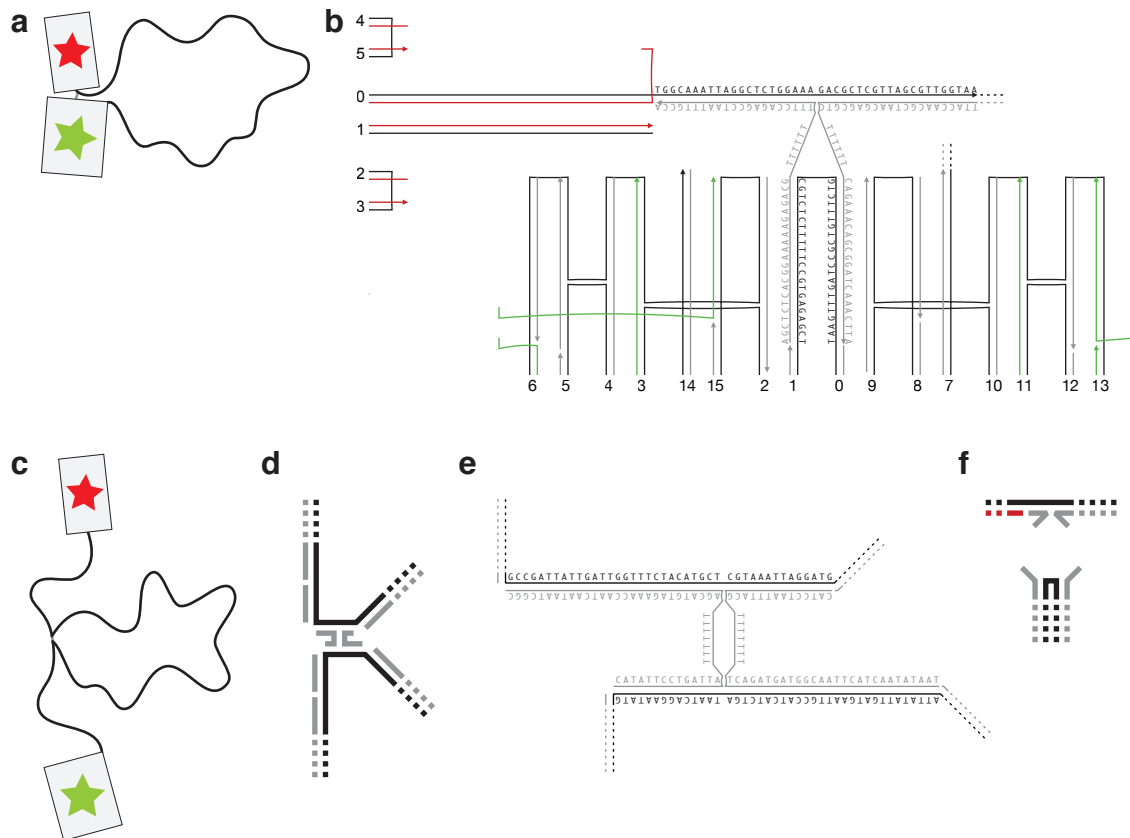


Abbildung 1.9: **Referenzobjekte mit fester Schlaufe in der Leine: F_p -Referenz und F_d -Referenz**

(a) Schema der F_p -Referenz. Die grauen Rechtecke mit roten / grünen Sternen repräsentieren die Reporter-Einheiten. Die schwarze Linie repräsentiert die DNA-Leine. Die kurze graue Linie deutet die Verknüpfung zwischen den Ankerpunkten der Interaktionspartner in der proximalen Verankerungsgeometrie an. (b) Strang-Diagramm der Verknüpfung für die F_p -Referenz. Gezeigt sind das Leinenende nahe der mobilen Reporter-Einheit und die Seite der immobilen Reporter-Einheit mit dem Ankerpunkt des Interaktionspartners in der proximalen Verankerungsgeometrie. Scaffold in Schwarz, Staples in Grau und Farbe. Die gestrichelten Linien deuten die Fortsetzung der DNA-Leine zwischen den Reporter-Einheiten an. (c) Schema der F_d -Referenz, analog zu a. Die kurze graue Linie deutet die Verknüpfung zwischen den Ankerpunkten der Interaktionspartner in der distalen Verankerungsgeometrie innerhalb der DNA-Leine an. (d) Schematischer Aufbau der Verknüpfung innerhalb der DNA-Leine für die F_d -Referenz. Scaffold in Schwarz, Staples in Grau. Gepunktete Abschnitte: Fortsetzung einer doppelsträngigen DNA-Leine aus durchgängigem Scaffold und je 42 nt langen Staples. (e) Strang-Diagramm der Verknüpfung innerhalb der DNA-Leine für die F_d -Referenz. Scaffold in Schwarz, Staples in Grau. Die gestrichelten Linien deuten die Fortsetzung der DNA-Leine in Richtung Reporter-Einheiten bzw. in der Schlaufe an. (f) Mögliche Fehlassemblierung der F_p -Referenz durch doppelte Besetzung des Scaffolds durch die Verknüpfungs-Staples. Scaffold in Schwarz, Staples in Grau und Rot. Gepunktete Abschnitte: Fortsetzung von DNA-Doppelhelices aus Scaffold und Staples.

jeweils zwei Kopien der zur Verknüpfung der Ankerpunkte gedachten Staples an den Scaffold binden (vgl. Veranschaulichung für den Fall der F_p -Referenz in Abb. 1.9 f). Diese Wahrscheinlichkeit steigt mit dem molaren Überschuss der Staples zum Scaffold in der Reaktionsmischung der Selbstassemblierung. Im Falle einer doppelten Besetzung durch beide Verknüpfungstaples entstehen fehlgefaltete Referenz-Objekte ohne die gewollte feste Schlaufe.

1.3 Kontrolle der DNA-Nano-Objekte per Gel-Elektrophorese

Bei der Gel-Elektrophorese werden Proben in eine Gel-Matrix eingebracht und durch Anlegen eines elektrischen Feldes über das Gel aufgetrennt. Aufgrund des äußeren elektrischen Feldes bewegen sich elektrisch geladene Bestandteile der Probe durch die Gel-Matrix. Die relative Laufweite verschiedener Objekte im Gel kann Aufschluss über ihre relative Größe und Kompaktheit geben. In der Regel migrieren DNA-Nano-Objekte gleichen Molekulargewichts umso schneller im Gel, je kompakter ihre Form ist. DNA-Stränge gleicher Topologie (linear/zyklisch, einzel-/doppelsträngig) migrieren umso schneller im Gel, je geringer ihr Molekulargewicht ist. Die DNA-Nano-Objekte und die überschüssigen Staple-Stränge sind in den Gelen durch Zugabe des interkalierenden Farbstoffs Ethidiumbromid (EtBr) oder durch Fluoreszenzfarbstoffmodifikationen an den Staple-Strängen mit einem Gel-Scanner abbildbar.

DNA-Nano-Objekte mit verschiedenen Leinenlängen und Interaktionslängen (siehe Tabelle 1.1) wurden nach Protokoll A.2.1 assembliert und anschließend dem Verdau mit ClaI unterzogen. Die erste Analyse der so hergestellten DNA-Nano-Objekte fand stets per Elektrophorese in Agarose-Gelen statt. Abbildungen 1.10 und 1.11 zeigen Scans von Gelen mit Proben assemblierter DNA-Nano-Objekte sowohl vor als auch nach dem enzymatischen Verdau. Die überschüssigen Staples liegen nach der Selbstassemblierungs-Reaktion weiterhin als kurze Einzelstränge vor und migrieren mit Abstand am schnell-

Tabelle 1.1: **Interaktionssequenzen von 7 bp bis 14 bp**

Die Sequenzen sind in Richtung 5'-Ende nach 3'-Ende angegeben. ΔG_{NN} : Nach dem Nearest-Neighbor-Modell mithilfe des DINAmelt Servers [95] vorhergesagte Gibbs'sche freie Energie der Hybridisierung für eine Reaktionstemperatur von 23 °C, $[Na^+] = 0,5 M$ und $[Mg^{++}] = 0 M$. Dabei wurden die Sequenzen der Interaktionspartner inkl. T_{12} -Verbinder eingegeben.

Länge (bp)	Interaktionssequenz B	Interaktionssequenz B'	ΔG_{NN} (kcal/mol)
7	CAGAACG	CGTTCTG	-9,6
8	CAGGAACG	CGTTCCTG	-11,7
9	CCAGAGACG	CGTCTCTGG	-13,5
10	CCAGAGACAG	CTGTCTCTGG	-14,2
11	CCAAGAGACAG	CTGTCTCTTGG	-15,4
12	CCAAGAGACACG	CGTGTCTCTTGG	-18,1
13	CCACAGAGACACG	CGTGTCTCTGTGG	-20,2
14	CCACAGAGAGCACG	CGTGCTCTCTGTGG	-22,5

ten (diffuse Bande am unteren Rand der Gel-Scans, Kennzeichnung: St). Ihnen folgen mit deutlichem Abstand scharfe, helle Banden, deren Laufweite für unverdaute Proben je nach Leinenlänge im Bereich zwischen der 8-kbp-Bande und der 4-kbp-Bande des aufgetragenen Markers liegt (jeweils linke Hälfte der Gele in Abb. 1.10). Diese Banden der Probe werden im Folgenden als „Hauptbanden“ bezeichnet. Sie enthalten Monomere von assemblierten DNA-Nano-Objekten. Weitere Banden mit geringerer Laufweite können durch die Formation von Objekten höherer Ordnung wie z. B. Dimeren von DNA-Nano-Objekten entstehen.

Zusätzlich zu den Proben der Objekte mit Interaktionspartnern wurden in jedem der in Abb. 1.10 gezeigten Gele je eine Probe der entsprechenden Referenzobjekte ohne Schlaufe und mit fester Schlaufe (i0-Referenz resp. F_p-Referenz, vgl. Abschnitt 1.2.5) aufgetragen. Vor dem enzymatischen Verdau (linke Hälfte der Gele in Abb. 1.10) zeigt die i0-Referenz im Agarose-Gel eine einzelne helle Hauptbande. Die F_p-Referenz hat auf gleicher Höhe eine schwache Bande sowie eine dazu vorgelagerte, deutlich hellere Bande. Die Objekte mit 7 oder 8 bp langer Interaktionssequenz haben eine einzelne helle Hauptbande auf gleicher Höhe mit der Hauptbande der i0-Referenz. Die Objekte mit mindestens 9 bp langer Interaktionssequenz zeigen zwei Hauptbanden mit vergleichbarer Helligkeit: Eine davon läuft auf Höhe der Hauptbande der i0-Referenz, die andere auf Höhe der vorgelagerten, helleren Bande der F_p-Referenz.

Dieses Bandenmuster lässt sich wie folgt interpretieren: Objekte mit Schlaufe in der DNA-Leine weisen eine kompaktere Form auf und migrieren somit schneller im Gel als Objekte ohne Schlaufe. Die F-Referenz enthält hauptsächlich Objekte mit Schlaufe, aber auch einen geringen Anteil von fehlerhaft assemblierten Objekten ohne Schlaufe. Ist die Duplexbindung zwischen den Interaktionspartnern schwach (7 oder 8 bp), so verbringt das System die meiste Zeit im Zustand dissoziierter Einzelstränge B und B' und migriert im Gel ähnlich wie die 0-Referenz. Dahingegen verbringen Systeme mit ausreichend stabiler Duplexbindung zwischen den Interaktionspartnern (ab 9 bp) den größten Teil der Zeit im Zustand des hybridisierten Duplex B:B' und ausgeformter Schlaufe. Sie migrieren im Gel ähnlich schnell wie die vordere, helle Hauptbande der F-Referenz. Ein gewisser Prozentsatz jeder Probe von Objekten mit Interaktionspartnern ist defekt in dem Sinn, dass keine korrekte Duplexbindung zwischen den Interaktionspartnern zustande kommen kann. In den Proben von Objekten mit stabiler Interaktion migrieren derart defekte Objekte mit ähnlicher Geschwindigkeit durch das Gel wie die 0-Referenz und die langsamere, schwache Bande der F-Referenz. Abschnitt 1.4 zeigt experimentelle Befunde, welche die Hypothese der Existenz einer defekten Spezies untermauern, und diskutiert mögliche Ursachen.

Nach dem Verdau mit ClaI (rechte Hälfte der Gele in Abb. 1.10) bleibt das eben erläuterte Bandenmuster für dieselbe Abfolge von Objekt-Varianten erhalten, ist jedoch insgesamt zu einer größeren Laufweite hin verschoben. In jeder verdauten Probe erscheint eine zusätzliche helle Bande (Kennzeichnung: C), die im Gel leicht schneller läuft als die 3-kbp-Bande des Markers. Es ist evident, dass die C-Bande das etwa 2,9 kbp lange DNA-Fragment enthält, welches vom Restriktionsenzym aus dem DNA-Nano-Objekt herausgeschnitten wird. Das Fehlen dieses Fragments in den DNA-Nano-Objekten führt dazu, dass sie im Gel schneller migrieren als ihre unverdauten Pendants.

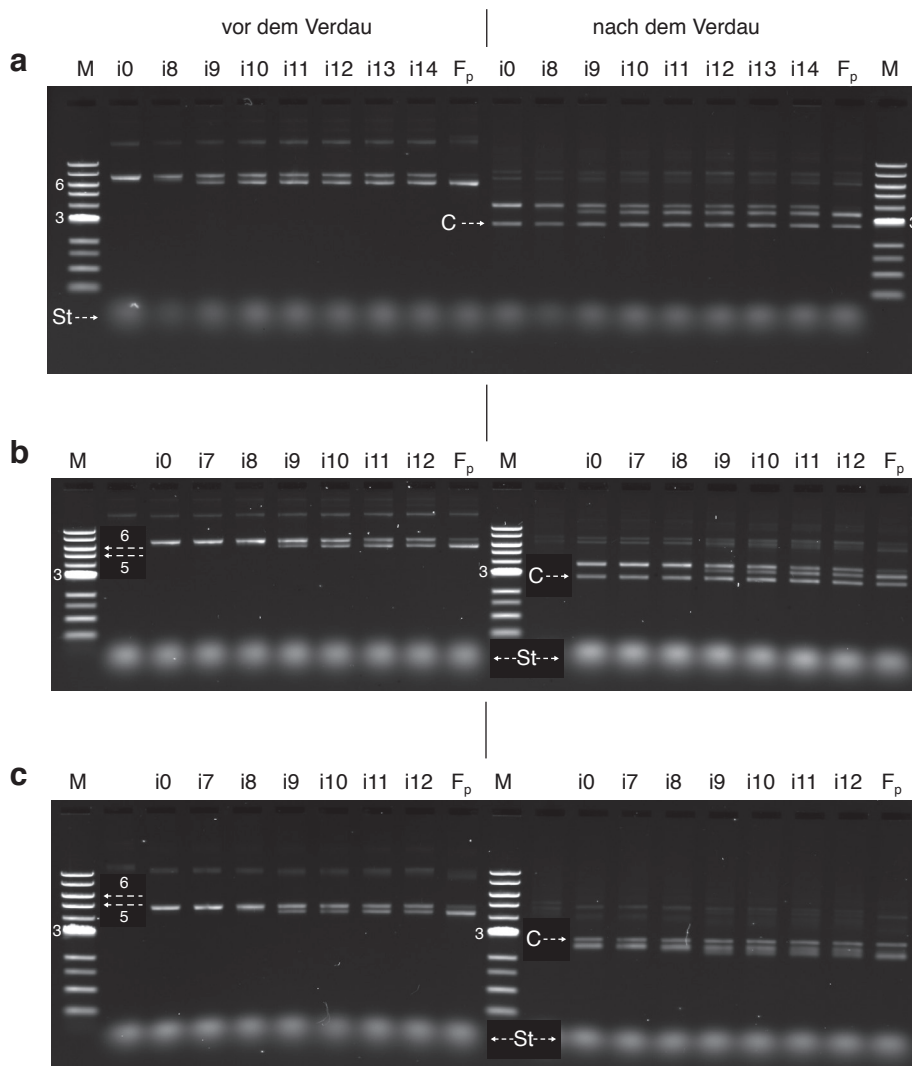


Abbildung 1.10: **Gelelektrophoretische Untersuchung der Produkte der Selbstassemblierungs-Reaktion vor und nach dem Verdau mit dem Restriktionsenzym Clal**

Scans von Gelen mit Proben von DNA-Nano-Objekten mit **(a)** voller Leinenlänge, **(b)** um ein 252-bp-Segment verkürzter Leine und **(c)** um zehn 252-bp-Segmente verkürzter Leine. Die Gele enthielten EtBr und wurden im EtBr-Kanal des Scanners abgebildet. Jeweils linke Hälfte der Gele: Proben von Objekten nach der Selbstassemblierungs-Reaktion, aber vor dem Verdau mit Clal. Jeweils rechte Hälfte der Gele: Proben der gleichen Objekte nach dem Verdau mit Clal. **M**: Marker mit Fragmenten doppelsträngiger DNA mit Länge 10/8/6/5/4/3/2/1,5/1/0,5kbp (von oben nach unten). **3, 5, 6**: Marker-Banden mit Fragmenten der Länge 3, 5 und 6 kbp. **i0**: Referenz ohne Interaktionspartner (vgl. Abb. 1.8 a-d). **Fp**: Referenz mit fester Schlaufe in der proximalen Verankerungsgeometrie (vgl. Abb. 1.9 a, b). **i7, i8, ... , i13, i14**: Objekte mit Interaktionslänge 7, 8, ... , 13, 14 bp. Verankerung der Interaktionspartner in der proximalen Verankerungsgeometrie. **C**: Banden mit dem von Clal herausgeschnittenen DNA-Fragment. **St**: Banden mit den überschüssigen Staple-Molekülen.

Sowohl vor als auch nach dem Verdau sollten die DNA-Nano-Objekte umso kompakter sein, je mehr Segmente der DNA-Leine am Rand der immobilen Reporter-Einheit in kurzen 6-Helix-Bündeln aufgerollt sind und je kürzer der freie, ungefaltete Teil der DNA-Leine ist (siehe Abschnitt 1.2.4 mit Abbn. 1.6 und 1.7). Mit abnehmender freier Leinenlänge der Objekte sollte folglich die Migrationsgeschwindigkeit der entsprechenden Banden im Gel zunehmen. Die Veränderung der Migrationsgeschwindigkeit im Gel mit der freien Leinenlänge lässt sich in den Gel-Scans in Abb. 1.10 mithilfe des Markers und der C-Bande als Referenzen nachvollziehen. Insbesondere laufen nach dem enzymatischen Verdau Objekte mit voller oder nur leicht verringerter Leinenlänge langsamer als das herausgeschnittene DNA-Fragment (Abb. 1.10 a,b), wohingegen Objekte mit stark verkürzter Leine im Gel schneller laufen als das herausgeschnittene DNA-Fragment (Abb. 1.10 c). Abbildung 1.11 zeigt direkte Vergleiche der Migrationsgeschwindigkeit von Objekten verschiedener Leinenlänge im selben Gel. Mit zunehmender Verkürzung der Leine steigt die Laufweite der Objekte monoton an. Die Variante mit der 8-bp-Interaktionssequenz hat für alle Leinenlängen eine einzelne Hauptbande, welche allerdings bei manchen Leinenlängen etwas breiter erscheint als bei anderen Leinenlängen. Die Variante mit der 9-bp-Interaktionssequenz hat für fast alle Leinenlängen zwei unterscheidbare Hauptbanden. Bei einer Verkürzung der Leine um elf Segmente (das entspricht in der proximalen Verankerungsgeometrie einer Länge des Leinenstücks zwischen den Interaktionspartnern von $L = 294$ bp bei einer Gesamt-Leinenlänge von rund 400 bp) sind die zwei Banden nicht mehr einzeln aufzulösen.

Die Ergebnisse der Gel-Elektrophorese-basierten Analyse lassen die Folgerung zu, dass die Selbstassemblierung und der anschließende enzymatische Verdau der DNA-Nano-Objekte ihrer Konzeption entsprechend funktionieren. Es wurden keine Anzeichen beobachtet, die auf Fehler im Design hindeuten – insbesondere, was den Aufroll-Mechanismus zur Einstellung der Leinenlänge angeht.

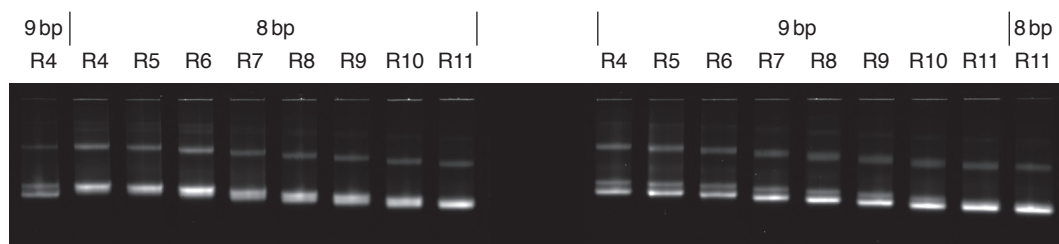


Abbildung 1.11: **Gelelektrophoretische Untersuchung von Objekten mit verschiedener Leinenlänge**
 Scan eines Gels mit Proben von Varianten assemblierter DNA-Nano-Objekte mit der proximalen Verankerungsgeometrie vor dem enzymatischen Verdau. Die Varianten unterscheiden sich durch die Interaktionslänge (8 oder 9 bp) und durch die Konturlänge L des Leinenstücks zwischen den Ankerpunkten der Interaktionspartner (vgl. Abb. 1.5): 2058 (R4), 1806 (R5), 1554 (R6), 1302 (R7), 1050 (R8), 798 (R9), 546 (R10) oder 294 bp (R11).

1.4 Die intakte und die defekte Spezies

Anm.: Der Großteil der in diesem Abschnitt gezeigten Ergebnisse wurde im Rahmen eines Praktikums von Sebastian Grün am Lehrstuhl von Prof. Dietz erarbeitet. Herrn Grün sei an dieser Stelle erneut für die gute Zusammenarbeit gedankt, sowie für die gewissenhafte Protokollierung seiner experimentellen Ergebnisse.

Im vorangegangenen Abschnitt wurde die Hypothese formuliert, dass die Proben von Objekten mit Interaktionspartnern zwei Spezies von Objekt-Monomeren enthalten: eine intakte Spezies und eine defekte Spezies. Bei der intakten Spezies ist die vollständige, korrekte Ausformung des Objekt-internen Duplex B:B' und der damit verbundenen Schlaufe in der Leine möglich, bei der defekten Spezies jedoch nicht.

Eine Ursache für das Auftreten der defekten Spezies kann sein, dass bei der Staple-Synthese der Anteil von Strängen mit vollständiger, korrekter Sequenz nicht 100 % beträgt. Aufgrund von Fehlstellen oder Verkürzungen in den Sequenzen der Interaktionspartner bilden sich im Vergleich zum vollständigen Duplex B:B' weniger korrekte Basenpaare aus, was zu einer deutlich weniger stabilen Bindung führt. Grundsätzlich können derartige Synthesefehler für alle Staple-Sorten des DNA-Nano-Objekts auftreten. In der Regel liegen in der Selbstassemblierungs-Reaktion die Staples im Überschuss zum Scaffold vor. In der Konkurrenz um die Bindung an den Scaffold sollten sich die korrekt synthetisierten Kopien gegen die fehlerhaften Kopien durchsetzen, sofern das betreffende Staple mit seiner vollen Länge an den Scaffold bindet. Diese Bedingung ist zwar für die meisten Staples der hier vorgestellten Objekte erfüllt¹, nicht aber für die interaktionspartnertragenden Staples. Insbesondere sollten Synthesefehler in den Interaktionssequenzen B oder B' keinen Einfluss auf die Bindung des Staples an den Scaffold bzw. auf die Konkurrenz mit Kopien vollständiger, korrekter Basensequenz haben.

Ein weiterer Hinderungsgrund für die Bildung von Objekt-internen Duplexen können überschüssige interaktionspartnertragende Staples in der Probenlösung sein. Deren Interaktionssequenzen B' resp. B können Duplexbindungen mit den im Objekt verankerten Sequenzen B resp. B' eingehen und dadurch die Bildung Objekt-interner Duplexe hemmen. Derart blockierte Objekte unterscheiden sich von der dauerhaft defekten Spezies dadurch, dass eine Objekt-interne Duplexbindung wieder uneingeschränkt möglich wird, sobald sich die überschüssigen Staples von den im Objekt verankerten Interaktionspartnern lösen.

Um die Zwei-Spezies-Hypothese zu überprüfen, wurden schnelle und langsame Banden separat aus Agarose-Gelen extrahiert und die Extrakte durch erneute Gel-Elektrophorese untersucht. Proben von DNA-Nano-Objekten mit Interaktionspartnern wurden über eine Laufzeit von sechs Stunden elektrophoretisch aufgetrennt, um eine möglichst große Separation der Banden zu erreichen (Abb. 1.12 a, d, g). Anschließend wurden die Banden manuell ausgeschnitten (Abb. 1.12 b, e, h), die DNA-Nano-Objekte aus den ausgeschnittenen Gel-Stücken extrahiert und die Extrakte in neuen Agarose-Gelen aufgetrennt (Abb. 1.12 c, f, i). Die Extrakte aus den Aufreinigungen der schnellen Banden zeigen im neuen Gel jeweils eine einzelne Bande auf derselben Höhe wie die schnelle

¹Abgesehen von den T₂- resp. T₁₂-Fortsätzen der Farbstoff- resp. Biotin-modifizierten Staples

Bande der entsprechenden unaufgereinigten Probe, mit einem schwachen nachgelagerten Schweif. Die Extrakte aus den Aufreinigungen der langsamen Banden zeigen im neuen Gel jeweils zwei Banden auf derselben Höhe wie die zwei Banden der entsprechenden unaufgereinigten Proben. Die schnelleren Banden dieser Extrakte sind deutlich schwächer ausgeprägt als die dazugehörigen langsameren Banden, wohingegen in den

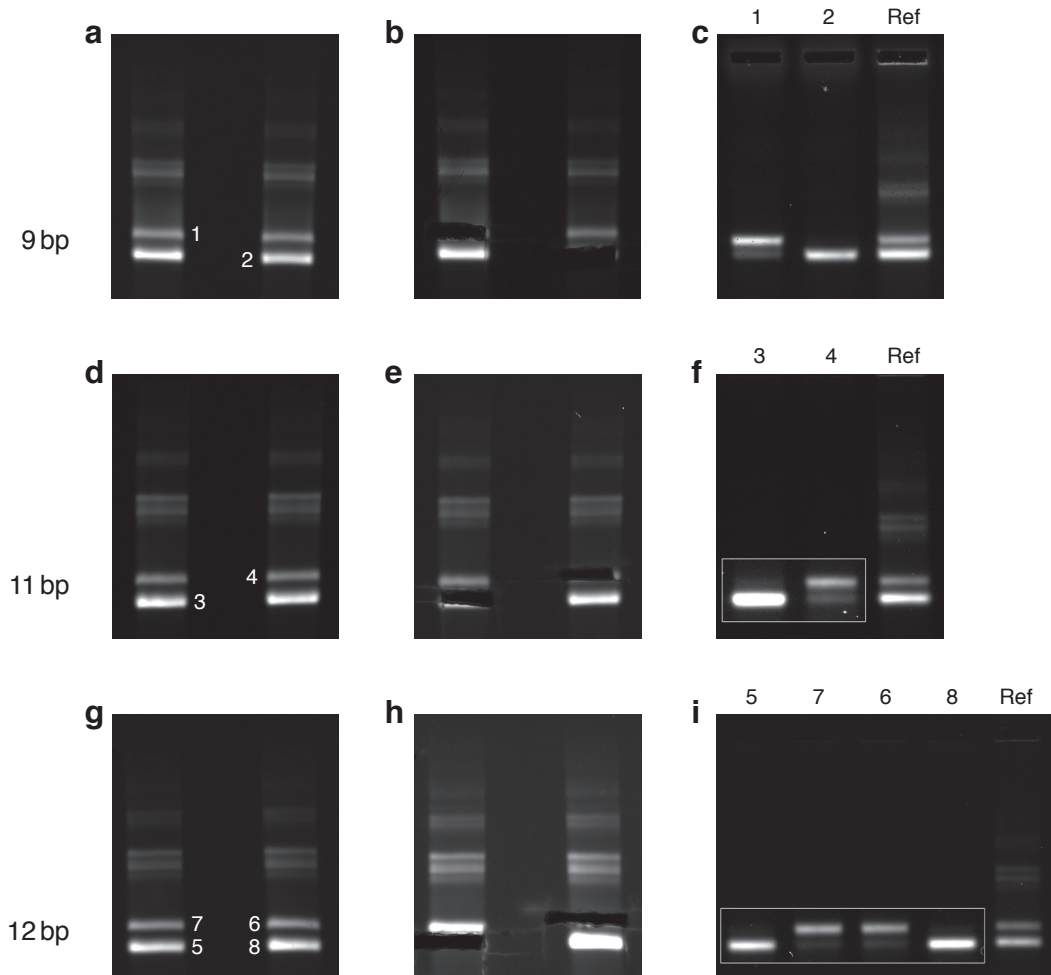


Abbildung 1.12: **Separate Aufreinigung und anschließende Elektrophorese der Extrakte**

(a) Scan eines Agarose-Gels mit zwei Proben von DNA-Nano-Objekten mit 9 bp Interaktionslänge und einer Konturlänge des Leinenstücks zwischen den Ankerpunkten der Interaktionspartner von $L = 2058$ bp in der proximalen Verankerungsgeometrie. Die Proben stammen aus derselben Selbstassemblierungs- und Verdau-Reaktion. Der Scan des Gels erfolgte nach sechs Stunden Laufzeit bei 75 V angelegter Spannung. (b) Scan desselben Gels nach dem Ausschneiden einzelner Banden. (c) Scan eines Agarose-Gels mit Extrakten aus den ausgeschnittenen Gel-Stücken. **1, 2**: Zuordnung ausgeschnittener Banden aus dem ersten Gel zu aufgetragenen Extrakten im neuen Gel. **Ref**: Referenz-Probe der entsprechenden unaufgereinigten DNA-Nano-Objekte. Der Scan des Gels erfolgte nach drei Stunden Laufzeit bei 75 V angelegter Spannung. (d-f) Wie a-c, für Proben von DNA-Nano-Objekten mit 11 bp Interaktionslänge. (g-i) Wie a-c, für Proben von DNA-Nano-Objekten mit 12 bp Interaktionslänge. Der in h gezeigte Scan wurde nach dem Ausschneiden der Banden 5 und 6, aber vor dem Ausschneiden der Banden 7 und 8 aufgenommen. Zur deutlicheren Sichtbarkeit der Banden wurde innerhalb der weißen Rahmen in f und i der Kontrast zusätzlich erhöht.

entsprechenden unaufgereinigten Proben die schnelleren Banden deutlich stärker ausgeprägt sind als die langsameren Banden.

Das Ergebnis der separaten Gel-Aufreinigung deutet darauf hin, dass das Auftreten zweier Hauptbanden tatsächlich eine Unterteilung der Objekt-Monomere in eine intakte und eine defekte Spezies widerspiegelt. Das Auftreten schwacher, schneller Banden auch in den Extrakten aus der Aufreinigung der langsamen Banden kann verschiedene Ursachen haben:

1. Intakte, aber durch überschüssige interaktionspartnertragende Staples blockierte Objekte migrieren im Gel etwa mit der gleichen Geschwindigkeit wie die Objekte der defekten Spezies und befinden sich zum Zeitpunkt des Ausschneidens in der langsamen Bande. Die blockierenden Staples lösen sich nach der Extraktion aus dem Gel von den Objekten. Verglichen zu den Konzentrationen während der Assemblierung und des Verdauens sind die Konzentrationen der Objekte und der überschüssigen interaktionspartnertragenden Staples im Extrakt aus dem Gel deutlich geringer. Eine erneute Blockade ist deshalb deutlich unwahrscheinlicher als die ursprüngliche Blockade bei höheren Konzentrationen von Objekten und überschüssigen Staples. Im neuen Gel migrieren die intakten, nun nicht mehr blockierten Objekte mit der mittleren Geschwindigkeit der schnellen Hauptbande.
2. Lösen sich blockierende Staples bereits im Aufreinigungsgel von eigentlich intakten Objekten, so migrieren Staples und Objekte mit deutlich verschiedener Geschwindigkeit, weshalb eine erneute Blockade auch in diesem Fall unwahrscheinlich ist. Die ehemals blockierten Objekte migrieren nun zwar schneller durch das Gel als noch im blockierten Zustand, doch womöglich reicht die restliche Laufzeit des Gels nicht aus, um eine deutliche Separation von den Objekten der defekten Spezies zu erlangen. Es werden gemeinsam mit den defekten Objekten auch intakte, nicht mehr blockierte Objekte extrahiert, die dann im neuen Gel bereits von Anfang an mit der mittleren Geschwindigkeit der schnellen Hauptbande migrieren.
3. Auch ohne den Einfluss von überschüssigen interaktionspartnertragenden Staples kann es passieren, dass ein Teil der intakten Objekte gemeinsam mit der langsamen Bande aus dem Gel extrahiert wird. Der Anteil der Zeit, den die Objekte der intakten Spezies im Zustand mit Schlaufe verbringen, ist statistisch verteilt. Einzelne Objekte verbringen nur einen verhältnismäßig geringen Anteil der Laufzeit in diesem Zustand, sodass ihre mittlere Migrationsgeschwindigkeit nicht viel größer ist als die der Objekte der defekten Spezies. Selbst Objekte, die denselben Anteil der Laufzeit im Zustand mit Schlaufe verbringen, migrieren nicht zwingend mit derselben mittleren Geschwindigkeit durch das Agarose-Gel. Es kann vorkommen, dass sich Objekte vorübergehend in der Gel-Matrix verfangen und dadurch stärker zurückgehalten werden als andere, gleichartige Objekte [96,97].

Mit dem letzten Punkt können auch die nachgelagerten Schweife der Banden der Extrakte aus den Aufreinigungen der schnellen Banden begründet werden. Die Argumentation der ersten beiden Punkte wird durch die in Abb. 1.13 gezeigten Ergebnisse gestützt. DNA-Nano-Objekte mit unterschiedlicher Interaktionslänge wurden mit hohem Überschuss interaktionspartnertragender Staples (rechnerisch 20-fache Scaffold-

Konzentration je Staple) assembliert und verdaut. Die verschiedenen Reaktionsgemische wurden nach dem Verdau in Agarose-Gelen elektrophoretisch aufgetrennt. Jedes der Gele wurde zu vier verschiedenen Zeitpunkten gescannt: nach 1:30 h, 3:00 h, 4:30 h und 6:00 h Laufzeit. Zu jeder Haupt-Doppelbande wurde durch Summierung der Pixelhelligkeiten über die ganze Breite der Gel-Bahn ein Helligkeitsprofil erstellt (Abb. 1.13).

Die langsameren Banden sind in allen Gel-Scans deutlich heller als die dazugehörigen schnelleren Banden. Mit steigender Interaktionslänge verschiebt sich das Helligkeitsverhältnis weiter und weiter zur langsameren Bande. Für die Objekte mit der längsten Interaktionssequenz (12 bp) ist die schnelle Bande nur mehr sehr schwach ausgeprägt. Das Verhältnis der Peak-Höhen bleibt für diese Probe über die gesamte Laufzeit des Gels nahezu gleich und die zwei Banden sind selbst nach mehreren Stunden Laufzeit scharf, definiert und deutlich voneinander getrennt. Anders bei den Proben der Objekte mit kürzeren Interaktionssequenzen: Hier nähern sich die Peak-Höhen der zwei Banden über die Laufzeit des Gels einander nach und nach an und das Signal zwischen den Peaks erscheint mit der Zeit zunehmend verschmiert. Besonders deutlich ist diese Verschiebung der Helligkeit für die Proben der Objekte mit 9 oder 10 bp langer Interaktionssequenz zu erkennen. Auch bei genauerer Betrachtung der Helligkeitsverteilung in der Probe der Objekte mit 11 bp langer Interaktionssequenz nach 6:00 h Gel-Laufzeit

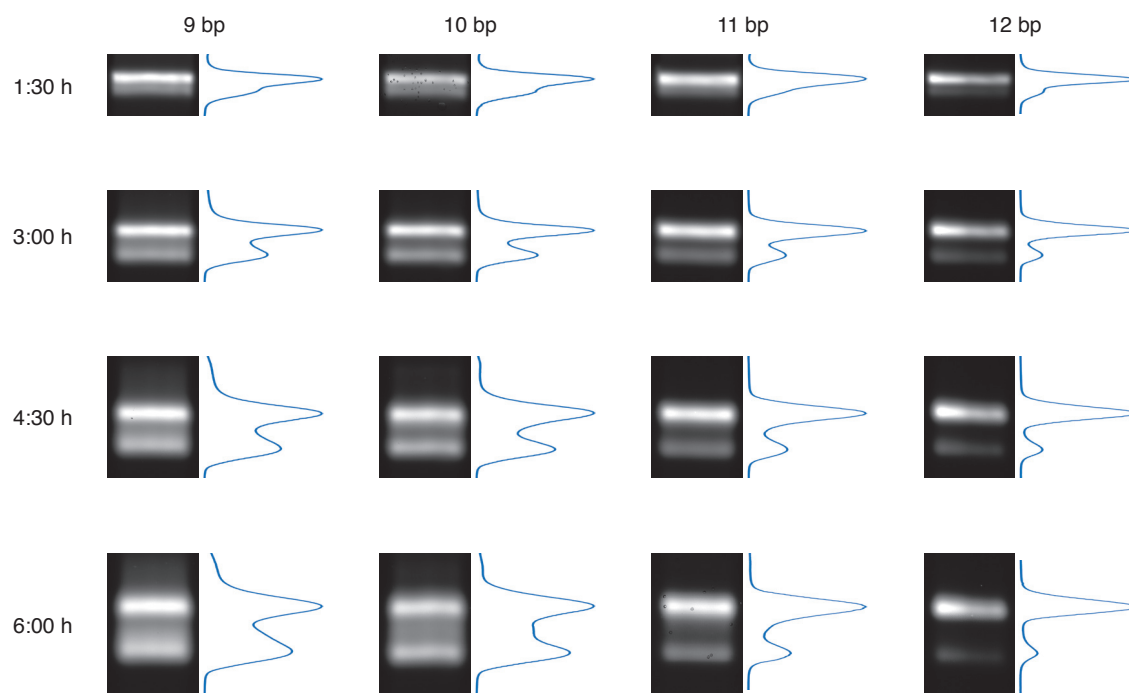


Abbildung 1.13: **Assemblierung mit hohem Überschuss interaktionspartnertragender Staples**

Spalten: Proben von DNA-Nano-Objekten mit einer Konturlänge des Leinenstücks zwischen den Ankerpunkten der Interaktionspartner von $L = 2058$ bp in der proximalen Verankerungsgeometrie und verschiedenen Interaktionslängen, wie oberhalb angegeben. Reihen: Laufzeit der Gele bis zu den jeweiligen Scans, wie linkerhand angegeben. Gezeigt sind jeweils ein Bildausschnitt aus dem Scan um die Doppel-Hauptbande und das dazugehörige Helligkeitsprofil. Die Profile entstehen durch Summierung der Pixelhelligkeiten senkrecht zur Laufrichtung des Gels über die Breite der entsprechenden Gel-Bahn.

fällt auf: Die Front der langsameren Bande ist nicht scharf, sondern zur schnelleren Bande hin verschmiert. Zudem liegt der Minimalwert der Helligkeit im Bereich zwischen den Banden deutlich über dem Hintergrund-Niveau des Scans, im Gegensatz zur direkt rechts daneben gezeigten Probe der Objekte mit 12 bp langer Interaktionssequenz.

Diese Beobachtungen untermauern erstens die These, dass überschüssige interaktionspartnertragende Staples die Bildung Objekt-interner Duplexe in eigentlich intakten Objekten hemmen, und zweitens die These, dass derart blockierte Objekte durch Abfallen der blockierenden Staples im Gel die Fähigkeit zur Objekt-internen Duplex-Bildung wiedererlangen können. Je stärker die jeweilige Duplexbindung ist, umso langsamer löst sich bei sonst gleichen Bedingungen die Blockade eigentlich intakter Objekte auf. Für die Objekte mit 12 bp Interaktionslänge ist die Bindung so stabil, dass selbst nach mehreren Stunden im Gel keine sichtbare Auflösung der Blockade stattfindet.

In einer weiteren Reihe von Reaktionsansätzen wurde das rechnerische Konzentrationsverhältnis zwischen interaktionspartnertragenden Staples und Scaffold in der Nähe des Werts 1 variiert. Bei kleinen rechnerischen Konzentrationsverhältnissen, d. h. relativem Scaffold-Überschuss, war die schnellere Hauptbande nur schwach ausgeprägt, die langsamere Hauptbande dafür umso stärker. Diese Beobachtung passt ins Gesamtbild, da der Anteil der Spezies, welche nicht zur Schlaufenbildung ausgestattet ist, durch einen Mangel an interaktionspartnertragenden Staples steigen sollte. Bei rechnerischen Konzentrationsverhältnissen $\gg 1$ nahm ebenfalls die Helligkeit der langsameren Hauptbande zu und die Helligkeit der schnelleren Hauptbande ab – in Übereinstimmung mit der weiter oben diskutierten Beobachtung bei sehr hohem Überschuss von interaktionspartnertragenden Staples. Der relative Anteil der Objekte in der schnellen Hauptbande an der Gesamtheit der Monomere wurde durch Vergleich der summierten Pixelhelligkeiten der zwei Hauptbanden abgeschätzt. Der größte Anteil intakter Monomere in den vier verschiedenen Reihen (9, 10, 11 und 12 bp Interaktionslänge) lag bei etwa 70 %.

Als Konsequenz der in den vorangegangenen Absätzen dargelegten Ergebnisse wurden bei der Assemblierung der Objekte für die TMFM-Experimente die interaktionspartnertragenden Staples stets mit rechnerisch 1,25-fachem Überschuss je Staple zum Scaffold eingesetzt. Einerseits sollte somit auch im Rahmen des Pipettierfehlers für jedes Scaffold-Molekül mindestens ein Paar interaktionspartnertragender Staples im Reaktionsgemisch vorhanden sein, andererseits sollte eine übermäßige Blockade seitens überschüssiger interaktionspartnertragender Staples verhindert werden.

Während einer bereits fortgeschrittenen Phase des TMFM-Projekts wurde erprobt, die DNA-Nano-Objekte mittels Größenausschlusschromatographie (engl. *Size Exclusion Chromatography*, kurz: SEC) aufzureinigen. Per SEC war eine reproduzierbar reinere Auftrennung der Probenbestandteile möglich als durch manuelles Ausschneiden der Banden aus Agarose-Gelen. In der SEC-Säule werden die Bestandteile der eingespülten Probe ihrer Größe entsprechend unterschiedlich lange zurückgehalten. Größere Objekte durchlaufen die Säule schneller, kleinere Objekte langsamer. Chromatogramme wurden durch Aufzeichnung des Fluoreszenzsignals der aufgetrennten Probe im Bereich der Emission von Cy5 mit einem der Säule nachgelagerten Fluoreszenzdetektor angefertigt. Ein Beispiel für ein solches Chromatogramm ist in Abb. 1.14 a abgebildet. Die Extraktion der Probenbestandteile in der SEC-Apparatur erfolgt in Fraktionen, welche jeweils in

einem eigenen Reservoir aufgefangen werden. Proben der Fraktionen wurden mitsamt einer Referenzprobe von unaufgereinigten Objekten auf ein Agarose-Gel aufgetragen und elektrophoretisch aufgetrennt (Abb. 1.14b). Anhand der Analyse der Fraktionen im Agarose-Gel lässt sich folgern, dass der erste und niedrigste der drei Peaks im Chromatogramm Objekte höherer Ordnung repräsentiert. Es folgen die Objekte der langsamen Gel-Bande (die defekte Spezies) und zuletzt die im Mittel kompakteren Objekte der schnellen Gel-Bande (die intakte Spezies). Die Fraktionen zwischen den zwei hohen Peaks weisen nach Auftrennung im Gel sowohl eine schnelle als auch eine langsame Bande auf.

Die Beobachtungen bei der Erprobung der SEC untermauern weiter die Hypothese der Unterteilung in eine intakte und eine defekte Spezies von DNA-Nano-Objekten.

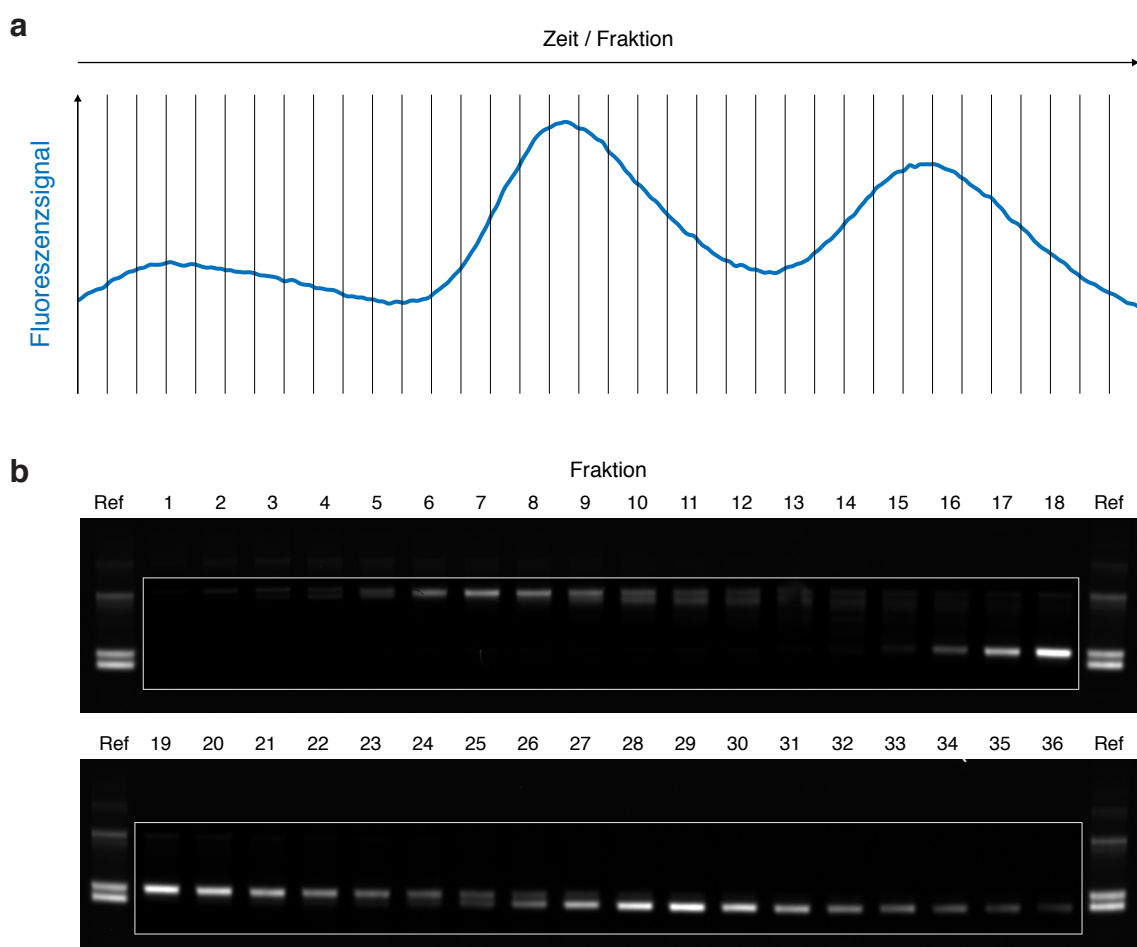


Abbildung 1.14: **Auftrennung der DNA-Nano-Objekte per Größenausschlusschromatographie**

(a) Blau: Ausschnitt des Chromatogramms einer Probe von Objekten mit 12 bp Interaktionslänge und einer Konturlänge des Leinenstücks zwischen den Ankerpunkten der Interaktionspartner von $L = 1050$ bp in der proximalen Verankerungsgeometrie. Schwarz: angedeutete Fraktionsgrenzen. **(b)** Scan eines Agarose-Gels mit Proben der extrahierten Fraktionen. Zur deutlicheren Sichtbarkeit der Banden wurde innerhalb der weißen Rahmen jeweils der Kontrast erhöht. **Ref:** Proben der entsprechenden unaufgereinigten Objekte. **1-36:** Serie von Fraktionen im Bereich des in a gezeigten Ausschnitts.

Folgen des Auftretens einer defekten Spezies für die Aufreinigung der DNA-Nano-Objekte

Die Aufreinigung der DNA-Nano-Objekte mit 8 oder 9 bp Interaktionslänge für die TMFM-Experimente erfolgte in Agarose-Gelen (für Details siehe Anhang A.2.2). Im Fall der Objekte mit 9 bp Interaktionslänge wurde versucht, die vorgelagerte Hauptbande der intakten Spezies einzeln aus den Gelen zu extrahieren. Allerdings ist eine reproduzierbar sortenreine Extraktion der intakten Spezies aus dem Gel selbst mit hoher Sorgfalt und erheblichem Aufwand nahezu unmöglich. Im Fall der Objekte mit 8 bp Interaktionslänge konnten die intakte Spezies und die defekte Spezies im Gel ohnehin nicht aufgetrennt werden. Die unausweichliche molekulare Heterogenität der aufgereinigten Proben birgt letztlich die Gefahr, dass die Ergebnisse der Einzelmolekül-Fluoreszenzmikroskopie-Experimente durch Mitglieder der defekten Spezies verfälscht werden. Im Rahmen der Analyse der TMFM-Daten wurde dieser Umstand berücksichtigt und eine auf statistischen Überlegungen basierende Prozedur zur Kompensation molekularer Heterogenität entwickelt (siehe Abschnitt 3.3.3).

Die Aufreinigung der Objekte für die Ensemble-Kompetitions-Experimente (siehe Kapitel 5) erfolgte per SEC, da für diese Experimente größere Volumina von Proben mit möglichst ausschließlich intakten DNA-Nano-Objekten benötigt wurden. Details zu dieser Form der Aufreinigung der Objekte finden sich in Anhang A.2.3.

1.5 Einzelmolekülbeobachtung der DNA-Nano-Objekte im Zwei-Kanal-Fluoreszenzmikroskop

Zur Beobachtung im Fluoreszenzmikroskop wurden die aufgereinigten DNA-Nano-Objekte in speziell angefertigten Probenkammern verankert. Das Boden-Deckglas jeder Probenkammer ist mit Polyethylenglykol-Ketten funktionalisiert, von denen ein bestimmter Anteil (in der Regel 0,1 % oder 1 %) Biotin-Modifikationen trägt. So wird einerseits der Großteil der Glasoberfläche gegen unspezifische Interaktion mit den DNA-Nano-Objekten passiviert, andererseits wird an bestimmten Stellen die Verankerung der immobilen Reporter-Einheit per Biotin-NeutrAvidin-Biotin-Brücken ermöglicht. Details zur Probenpräparation finden sich in Anhang A.2.4.

Es wurde ein *Total Internal Reflection*-Fluoreszenzmikroskop (TIRFM) mit invertiertem Objektiv verwendet. Während der Aufnahme ist die Probenkammer derart in einem Probenhalter positioniert, dass ihr Boden-Deckglas auf der Vorderseite des Objektivs aufliegt. Die Blickrichtung auf die DNA-Nano-Objekte im Mikroskop zeigt durch den Boden der Probenkammer, d. h. senkrecht zu der Ebene, in der die immobilen Reporter-Einheiten verankert sind. Im Detektions-Strahlengang auf der Rückseite des Objektivs spaltet ein dichroitischer Spiegel die Signale der Fluoreszenzfarbstoffe in zwei Kanäle auf: den Cy3-Kanal und den Cy5-Kanal. Jeder Kanal verfügt über eine eigene Kamera vom Typ EMCCD (*Electron-Multiplying Charge-Coupled Device*), wobei die zwei Kameras durch eine zentrale Ansteuerungs-Elektronik zeitlich synchronisiert werden. Die Kameras nehmen abwechselnd Bilder mit derselben Verschlusszeit auf. Während der Aufnahmen der Kamera des Cy3-Kanals wird die Probe ausschließlich mit einem 532-nm-Laser

beleuchtet (Anregung der Cy3-Farbstoffe), während der Aufnahmen der Kamera des Cy5-Kanals ausschließlich mit einem 640-nm-Laser (Anregung der Cy5-Farbstoffe). Somit können die beiden Reporter-Einheiten in den zwei Kanälen getrennt voneinander beobachtet werden. Die Verschlusszeit jedes Kamerabilds betrug in der Regel 50 ms, woraus eine effektive Bildrate von 10 Bildern pro Sekunde in jedem Kanal resultiert.

Abbildung 1.15 a und b zeigen einen Ausschnitt eines Einzelbilds im Cy3-Kanal und den entsprechenden Ausschnitt des darauffolgenden Einzelbilds im Cy5-Kanal. In beiden Bildern erscheinen helle Flecke vor einem dunklen Hintergrund. Die Positionen der Flecke in den zwei Kanälen stimmen paarweise überein. Jedes Paar von Flecken repräsentiert das auf die Sensoren der Kameras projizierte Signal der Fluorophore einer immobilen Reporter-Einheit im Cy3-Kanal und der dazugehörigen mobilen Reporter-Einheit im Cy5-Kanal. Bei Betrachtung der Aufnahmen im Verlauf der Zeit beobachtet man eine unterschiedliche Dynamik der Reporter-Einheiten in den zwei Kanälen: Während die große Mehrzahl der immobilen Reporter-Einheiten nahezu stationär erscheint, vollführt die Mehrzahl der mobilen Reporter-Einheiten eine deutlich sichtbare Zitter-

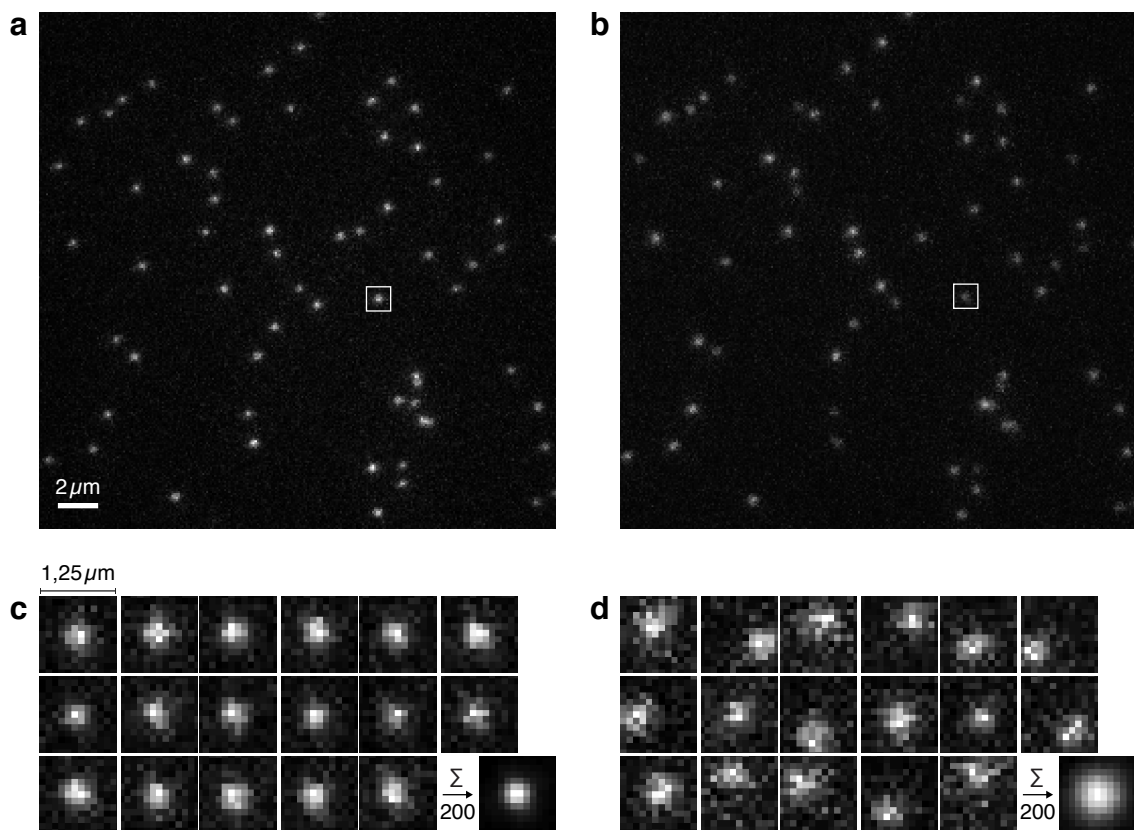


Abbildung 1.15: **Zwei-Kanal-Fluoreszenzmikroskopie und Bewegung der Reporter-Einheiten**

Bildausschnitte einzelner Kamera-Aufnahmen **(a)** im Cy3-Kanal bei Beleuchtung mit einem 532-nm-Laser und **(b)** im Cy5-Kanal bei Beleuchtung mit einem 640-nm-Laser. Weiße Rahmen kennzeichnen die Regionen der Nahaufnahmen in c und d. **(c)** Nahaufnahmen des Fluoreszenzsignals einer immobilen Reporter-Einheit im Cy3-Kanal. Gezeigt ist eine Sequenz von 17 aufeinanderfolgenden Einzelbildern und eine Überlagerung aus 200 Einzelbildern (unten rechts). **(d)** Wie c, für die dazugehörige mobile Reporter-Einheit im Cy5-Kanal.

bewegung um jeweils individuelle Mittelpunkte. In Zeitreihen von Nahaufnahmen der Flecke lässt sich die Fluktuation der Position der jeweiligen Reporter-Einheiten verfolgen (Abb. 1.15 c und d).

Die Bewegung der Reporter-Einheiten über den Verlauf von Aufnahmen mit vielen Einzelbildern stellt die primäre Messgröße einer TMFM-Messung dar. An diesem Punkt setzt Kapitel 2 an. Die hier folgenden Absätze erläutern Besonderheiten in der Aufnahmetechnik, welche es ermöglichten, TMFM-Aufnahmen mit mehreren Zehntausend Einzelbildern und über einer Stunde Beobachtungsdauer anzufertigen.

Vermeidung von Bleichen und Blinken der Fluoreszenzfarbstoffe

Bei Experimenten mit organischen Fluoreszenzfarbstoffen wie Cy3 und Cy5 sind die Effekte des „Bleichens“ und des „Blinkens“ zu berücksichtigen. Die Farbstoffe haben eine begrenzte Lebensdauer. Nach einer gewissen Anzahl von Anregungs- und Relaxationszyklen bleichen sie irreversibel aus, d. h. sie verlieren dauerhaft ihre Eigenschaft, zu fluoreszieren [98]. Oxidation der Farbstoffmoleküle durch Sauerstoff in der Pufferlösung trägt maßgeblich zum Bleichprozess bei [99]. Der Effekt des Blinkens entsteht dadurch, dass die Elektronen-Systeme der Farbstoffmoleküle bei Anregung nicht nur in einen kurzlebigen Singulett-Zustand, sondern auch in einen langlebigen Triplett-Zustand übergehen können. Die mittlere Lebensdauer eines Singulett-Zustands liegt auf der Größenordnung von Nanosekunden, die eines Triplett-Zustands auf der Größenordnung von Millisekunden bis zu mehreren Sekunden [98]. Nach dem Übergang in einen Triplett-Zustand kann ein Farbstoffmolekül somit über einen Zeitraum von einigen Kamerabildern in einem dunklen Zustand verweilen und währenddessen keine Photonen emittieren.

Um Bleichen und Blinken entgegenzuwirken, wurde ein *Oxygen Scavenging System* aus den Enzymen Glukose-Oxidase und Catalase (kurz: GOC) und der *Triplet State Quencher* 6-Hydroxy-2,5,7,8-Tetramethylchroman-2-Carboxylsäure (auch als „Trolox“ bezeichnet) verwendet [99]. Es ist anzumerken, dass sich der hier verwendete Aufnahmebuffer in zwei wesentlichen Punkten von dem in [99] verwendeten Aufnahmebuffer unterscheidet:

1. Es wurde eine höhere Konzentration von Glukose-Oxidase eingesetzt (442 U/ml statt 165 U/ml).
2. Da die enzymatischen Reaktionen von GOC nach einer gewissen Zeit zu einer Absenkung des pH-Werts führen [100], wurde eine höhere Konzentration von Tris-HCl mit pH 8 eingesetzt (100 mM statt 10 mM), um den pH-Wert im Aufnahmebuffer über die gesamte Dauer einer TMFM-Messung zu stabilisieren.

Weitere Details zur Zusammensetzung des Fluoreszenzmikroskopie-Aufnahmebuffers finden sich in Anhang A.1.2.

Stroboskopische Beleuchtung

Eine Besonderheit der hier angewandten Aufnahmetechnik bestand darin, dass in jedem aufgenommenen Kamerabild die Beleuchtungsdauer der Probe mit dem jeweiligen Laser deutlich kürzer war als die Belichtungszeit der Kamera. Die Belichtungszeit jedes Kamerasensors pro Einzelbild betrug in der Regel knapp 50 Millisekunden. Wird die Probe während der gesamten Belichtungszeit des Sensors beleuchtet, so detektiert man das summierte Fluoreszenzsignal der Farbstoffe an den Reporter-Einheiten in diesem gesamten Zeitfenster. Bei einer derartigen, kontinuierlichen Beleuchtung erscheint das von beweglichen Objekten herrührende Signal verschmiert. Die Positionen der Objekte und damit auch ihre Bewegung von Einzelbild zu Einzelbild sind dann sowohl mit bloßem Auge als auch mit dem Tracking-Algorithmus weniger genau auszumachen. Dieser Effekt wird im Englischen auch als *motion blur* bezeichnet und wird u. a. in den Referenzen [58,80,101] näher thematisiert.

Um die Reporter-Einheiten möglichst genau lokalisieren zu können, wurde die Probe innerhalb jedes Kamerabilds nur eine Millisekunde lang mit hoher Intensität des jeweiligen Lasers beleuchtet. Das detektierte Fluoreszenzsignal der Farbstoffe an den Reporter-Einheiten entstammt dann auch nur diesem kurzen Zeitfenster. Jedes Kamerabild beinhaltet sozusagen „Momentaufnahmen“ der diffusiven Bewegung der Reporter-Einheiten innerhalb eines 1 ms kurzen Intervalls.

Ein weiterer positiver Effekt der stroboskopischen Beleuchtung lag darin, dass die Farbstoffe selbst bei (rechnerisch) gleicher zeitlich gemittelter Beleuchtungsenergie im Durchschnitt erst nach deutlich längerer Beobachtungsdauer blichen als bei kontinuierlicher Beleuchtung.

Stabilisation der Probe in der Fokusebene des Objektivs

Die Probenkammer ist während des Experiments in einem verfahrbaren, elektronisch ansteuerbaren Halter befestigt. Zu Anfang jeder TMFM-Messung wurde die Position der Probenkammer über dem Objektiv so eingestellt, dass das Signal der immobilen Reporter-Einheiten maximal scharf im Cy3-Kanal abgebildet wurde. Während der langen Messdauern trat mitunter starker Drift der Probenkammer sowohl senkrecht zur Blickrichtung, als auch parallel dazu auf. Ohne manuelles Nachjustieren bewegte sich die Probenoberfläche dementsprechend nach und nach aus der Fokusebene des Objektivs heraus. Da der Mikroskopie-Aufbau über keinen Hardware-basierten Autofokus-Mechanismus verfügte, wurde eine Feedback-Schleife zur Fokus-Stabilisierung in der zentralen Ansteuerungs-Software implementiert. Als Maß für die Schärfe der Abbildung wurde gemäß [102] die Standardabweichung σ_1 der Helligkeitswerte einer Region von 256×256 Pixeln in der Bildmitte des Cy3-Kanals verwendet. Der Wert von σ_1 wurde während der Messung in jedem Kamerabild dieses Kanals ausgelesen und der Probenhalter automatisch verfahren, um stets einen möglichst hohen Wert von σ_1 und somit ein möglichst scharfes Bild herzustellen.

1.6 Kapitelzusammenfassung

Die DNA-Nano-Objekte wurden mit der Methode Scaffolded DNA-Origami konzipiert. Dementsprechend besteht jedes DNA-Nano-Objekt aus einem mehrere Kilobasen langen Scaffold-Molekül und ca. 200 jeweils 20 bis 50 Basen langen Staple-Molekülen.

Jedes DNA-Nano-Objekt besitzt zwei Reporter-Einheiten, die durch eine doppelsträngige DNA-Leine miteinander verbunden sind. Jedes einzelne Objekt wird daher auch als „Partikel-Leinen-System“ bezeichnet. Die Reporter-Einheiten sind kompakte Helix-Bündel mit einem jeweiligen Gyrationradius von ca. 14 nm. Die Reporter-Einheiten tragen zur Beobachtung im Fluoreszenzmikroskop Farbstoff-Modifikationen. Eine der Reporter-Einheiten trägt zusätzlich Biotin-Modifikationen zur Befestigung an einer Probenoberfläche. Durch Austauschen bestimmter Gruppen von Staples im Reaktionsgemisch der Selbstassemblierung können DNA-Nano-Objekte mit unterschiedlicher Leinenlänge hergestellt werden.

An der DNA-Leine oder an den Reporter-Einheiten können Interaktionspartner verankert werden. Im Fall der hier vorgelegten Arbeit handelt es sich bei den Interaktionspartnern um zwei kurze, zueinander revers-komplementäre DNA-Einzelstränge mit den Interaktionssequenzen B und B'. Mit der Hybridisierung der Interaktionspartner zum Duplex B:B' geht die Ausformung einer Schlaufe in der DNA-Leine einher. Es wurden zwei verschiedene Verankerungsgeometrien untersucht. In der proximalen Verankerungsgeometrie befinden sich die Ankerpunkte der Interaktionspartner nahe an den Reporter-Einheiten. In der distalen Verankerungsgeometrie haben die Ankerpunkte der Interaktionspartner einen größeren Abstand zu den Reporter-Einheiten. Folglich ist die freie Leinenlänge zwischen den Reporter-Einheiten im Zustand des hybridisierten Duplex B:B' und ausgeformter Schlaufe in der Leine in der distalen Verankerungsgeometrie größer als in der proximalen Verankerungsgeometrie.

Für beide Verankerungsgeometrien wurden Referenzobjekte hergestellt, die dauerhaft entweder den Zustand des Partikel-Leinen-Systems ohne Schlaufe in der Leine oder den Zustand des Partikel-Leinen-Systems mit Schlaufe in der Leine nachahmen.

Die korrekte Assemblierung der DNA-Nano-Objekte wurde per Gel-Elektrophorese nachgewiesen. In den Agarose-Gelen zeigten sich zwei verschiedene Spezies von Objekten. Die Objekte der intakten Spezies können durch Hybridisierung der Interaktionspartner B und B' zum Duplex B:B' eine Schlaufe ausbilden, die Objekte der defekten Spezies dagegen nicht. Fehler in den Interaktionssequenzen sind eine plausible mögliche Ursache für dieses heterogene Verhalten der Objekte.

Die TMFM-Experimente wurden mit Gel-aufgereinigten DNA-Nano-Objekten in einem TIRF-Mikroskop durchgeführt. Dabei wurden die immobile Reporter-Einheit und die mobile Reporter-Einheit in verschiedenen Kanälen getrennt voneinander beobachtet. Eine Routine zur Fokus-Stabilisierung, stroboskopische Beleuchtung sowie die Verwendung eines Oxygen Scavenging Systems ermöglichten TMFM-Aufnahmen von über einer Stunde Beobachtungsdauer, bei einer Bildrate von 10 Hz je Kanal.

2 Die Bewegung der Reporter-Einheiten

Dieses Kapitel widmet sich der primären Messgröße der TMFM-Experimente: der Bewegung der Reporter-Einheiten. Abschnitte 2.1 bis 2.4 legen die wesentlichen Schritte der rechnergestützten Bewegungsanalyse dar, diskutieren grundsätzliche Beobachtungen und thematisieren verschiedene in den TMFM-Messungen auftretende Artefakte. Abschnitt 2.5 zeigt die Ergebnisse der detaillierten Charakterisierung der Bewegung der mobilen Reporter-Einheiten, auf welchen die weitere Analyse (siehe Kapitel 3) aufbaut.

2.1 Wahl einer angemessenen Tracking-Prozedur

Das Signal eines fluoreszenzmarkierten Partikels wie einer Reporter-Einheit erzeugt auf dem Sensor einer Kamera des Fluoreszenzmikroskops eine pixelierte Helligkeitsverteilung (Abb. 2.1 a, b). Aus dieser Helligkeitsverteilung kann mit einer Genauigkeit von Bruchteilen der Pixelgröße die auf den Sensor der Kamera projizierte Position des Partikelmittelpunkts abgeschätzt werden (Abb. 2.1 c, d). In der hier vorgelegten Arbeit wurde dazu der *Virtual Window Center of Mass-Algorithmus* (kurz: VWCM-Algorithmus) nach [103] adaptiert. Dieser Algorithmus ermittelt den „Schwerpunkt“ der Helligkeits-

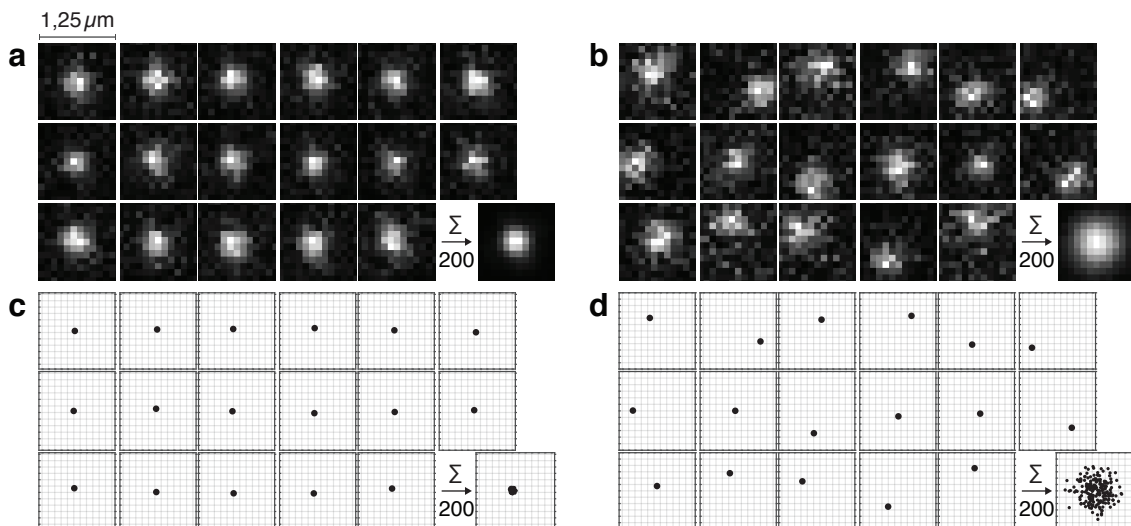


Abbildung 2.1: **Positionsabschätzung am Beispiel eines Paares von Reporter-Einheiten**

(a) Nahaufnahme des Fluoreszenzsignals einer immobilen Reporter-Einheit im Cy3-Kanal und (b) der dazugehörigen mobilen Reporter-Einheit im Cy5-Kanal. Gezeigt ist jeweils eine Sequenz von 17 aufeinanderfolgenden Einzelbildern und eine Überlagerung aus 200 Einzelbildern (jeweils unten rechts). (c) Geschätzte xy-Positionen des Mittelpunkts der immobilen Reporter-Einheit in den 17 bzw. 200 (unten rechts) Einzelbildern. Die Gitterlinien entsprechen den Begrenzungen der Kamerapixel. (d) wie c, für die mobile Reporter-Einheit.

verteilung eines Partikels unter Berücksichtigung des Hintergrundsignals. Diese Prozedur ist weniger rechenintensiv und deshalb deutlich schneller als die gängige Methode, eine zweidimensionale Gauß-Funktion an die pixelierte Helligkeitsverteilung zu fitten.

Der VWCM-Algorithmus unterscheidet sich von der 2D-Gauß-Fit-Methode in einer weiteren Eigenschaft, aufgrund derer er sich besser für das Tracking der Reporter-Einheiten in TMFM eignet: Es wird keinerlei Annahme über die Form der Partikel getroffen. Die bis zu 20 Fluoreszenzfarbstoffe jeder Reporter-Einheit sind nicht kugelsymmetrisch angeordnet, weshalb die Querschnittsfläche einer Reporter-Einheit senkrecht zur optischen Achse von ihrer räumlichen Orientierung abhängt. Desweiteren sind nicht zwingend alle 20 Farbstoffmoleküle an jeder Reporter-Einheit vorhanden und selbst die vorhandenen Farbstoffmoleküle können trotz der entsprechenden Gegenmaßnahmen zeitweise in einem dunklen Zustand verweilen oder dauerhaft bleichen (vgl. Abschnitt 1.5). Folglich kann sich die Form des Abbilds einer Reporter-Einheit in verschiedenen Kamerabildern stark unterscheiden. Insbesondere wäre die Annahme nicht gerechtfertigt, dass die Form des Abbilds einer Reporter-Einheit grundsätzlich einer zweidimensionalen Normalverteilung ähnelt.

2.2 Positionsabfolgen und Auslenkungsabfolgen

Die Tracking-Prozedur liefert letztendlich für jedes Kamerabild die geschätzten Projektionen der Positionen der Reporter-Einheiten auf die Ebene senkrecht zur optischen Achse des Mikroskops. Diese Ebene wird hier als xy -Ebene definiert. Die gemessene Bewegung einer Reporter-Einheit über den Verlauf einer TMFM-Aufnahme mit einer Gesamtzahl von T Kamerabildern wird somit durch Abfolgen von x - und y -Koordinaten zu insgesamt T äquidistanten Zeitpunkten dargestellt (vgl. Abb. 2.2):

$$(x(t), y(t)), \quad t = 1, 2, 3, \dots, T$$

Der zeitliche Abstand zwischen zwei aufeinanderfolgenden Elementen der Abfolgen ist durch die Bildperiode der TMFM-Aufnahme gegeben. Diese betrug in jedem Kanal standardmäßig 100 ms (vgl. Abschnitt 1.5).

Die gemessene Bewegung der Reporter-Einheiten ist eine Überlagerung der Bewegung der Reporter-Einheiten im Bezug auf die jeweiligen Ankerpunkte der DNA-Nano-Objekte an der Probenoberfläche mit der Relativbewegung zwischen Probenkammer und dem Sensor der Kamera [79]. In vielen TMFM-Aufnahmen trat eine langsame, unperiodische Bewegung der Probenkammer auf dem Objektiv des Mikroskops auf, die üblicherweise als „Drift“ bezeichnet wird. Deutlicher Drift ist in den Positionsabfolgen mit bloßem Auge sichtbar (Abb. 2.2). Zur Driftkorrektur wurde von jeder Positionsabfolge die Abfolge der über ein Zeitfenster von ± 50 Bildperioden gemittelten Position der jeweiligen Reporter-Einheit subtrahiert (für Details siehe Anhang B.1).

Nach der Driftkorrektur liegen für jede Reporter-Einheit Abfolgen von Auslenkungen aus ihrer gemittelten Position in x - und y -Richtung vor (vgl. Abb. 2.3 a, b, d, e):

$$(\Delta x(t), \Delta y(t)), \quad t = 1, 2, 3, \dots, T$$

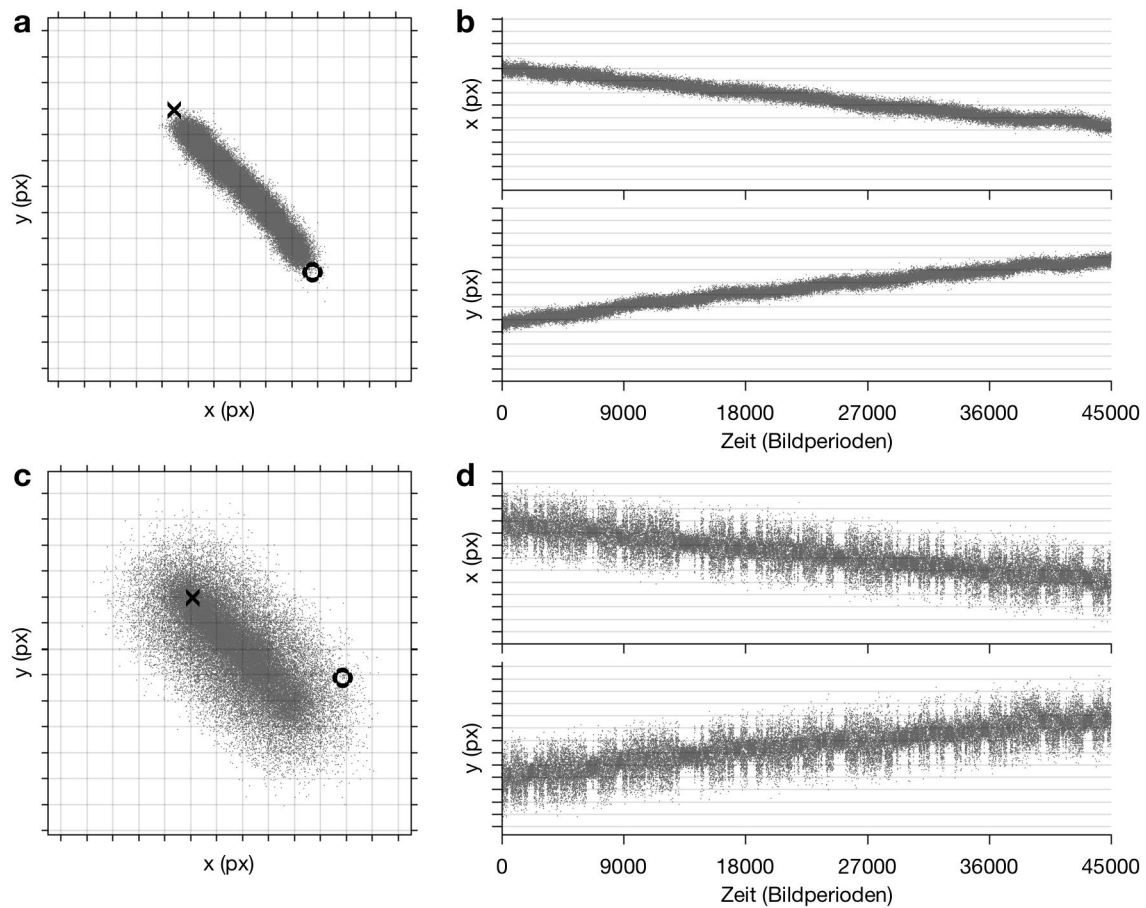


Abbildung 2.2: **Positionsabfolgen der Reporter-Einheiten eines DNA-Nano-Objekts**

Objekt mit Interaktionssequenz CAGGAACG:CGTTCCTG, Konturlänge des Leinenstücks zwischen den Interaktionspartnern $L = 1554$ bp und distaler Verankerungsgeometrie. Anzahl der Kamerabilder: $T = 45000$ je Kamera (Aufnahmedauer: 75 Minuten). Die Skaleneinteilungen und Gitterlinien der x- und y-Achsen entsprechen den Begrenzungen der Kamerapixel. **(a)** Streudiagramm der geschätzten Positionen der immobilen Reporter-Einheit des Objekts. Kreis: erste Position der Abfolge. Kreuz: letzte Position der Abfolge. **(b)** Abfolgen der geschätzten x- und y-Koordinaten der immobilen Reporter-Einheit über der Zeit. **(c, d)** Wie a und b, für die mobile Reporter-Einheit des Objekts.

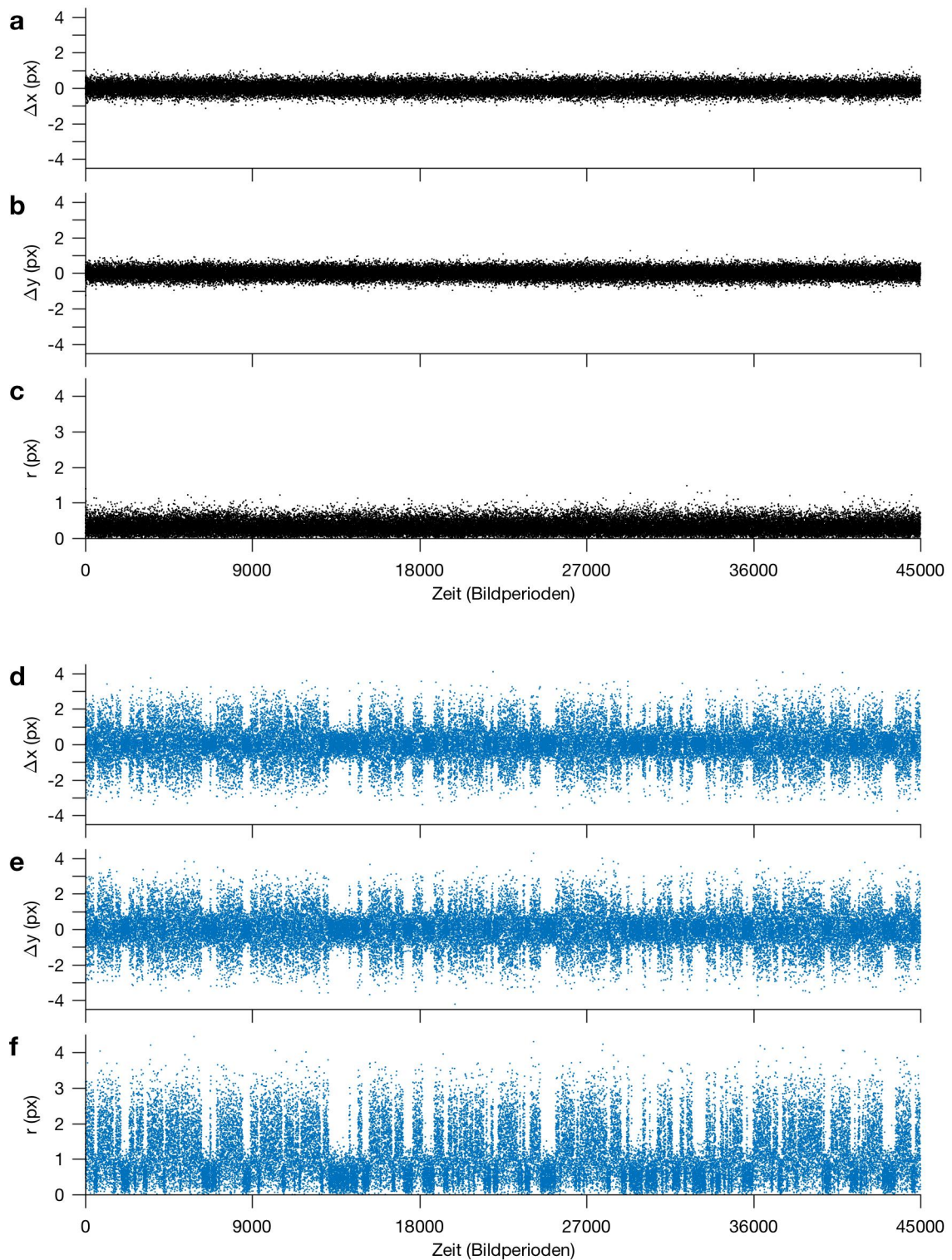


Abbildung 2.3: **Auslenkungsabfolgen der Reporter-Einheiten eines DNA-Nano-Objekts**

Es handelt sich um dasselbe Objekt, dessen Positionsabfolgen in Abb. 2.2 gezeigt sind. Abfolgen **(a)** der Auslenkung in x-Richtung Δx , **(b)** der Auslenkung in y-Richtung Δy sowie **(c)** der radialen Auslenkung r nach Gleichung (2.1) für die immobile Reporter-Einheit des Objekts. **(d-f)** Wie a-c, für die mobile Reporter-Einheit des Objekts.

Zur einfacheren grafischen Darstellung bietet sich die eindimensionale, radiale Auslenkung an (vgl. Abb. 2.3 c, f):

$$r(t) = \sqrt{(\Delta x(t))^2 + (\Delta y(t))^2} \quad (2.1)$$

Die immobile Reporter-Einheit ist nicht vollkommen starr auf der Probenoberfläche verankert (vgl. Abschnitt 1.2.1). Sie sollte bei korrekter Verankerung des Objekts daher über einen gewissen Bewegungsspielraum verfügen. Im TMFM-Experiment zeigten fast alle immobilen Reporter-Einheiten eine verglichen zur mobilen Reporter-Einheit kleine Beweglichkeit. Die Auslenkungen Δx und Δy des hier gezeigten Beispielobjekts (Abb. 2.3 a und b) liegen fast ausschließlich im Bereich zwischen -1 und 1 Pixel. Der Maximalwert der radialen Auslenkung (Abb. 2.3 c) beträgt rund 1,3 Pixel. Aus dem Ausmaß und der Form der Bewegung der immobilen Reporter-Einheit lässt sich folgern, dass sie über den Verlauf der Messung stabil an der PEG-Beschichtung der Probenoberfläche verankert ist.

Für die mobile Reporter-Einheit ist bei korrekter Verankerung des Objekts ein deutlich größerer Bewegungsradius zu erwarten als für die immobile Reporter-Einheit. Darüber hinaus sollte die mobile Reporter-Einheit je nach Bindezustand der Interaktionspartner eine unterschiedliche Beweglichkeit besitzen: Im dissoziierten Zustand ist der mögliche Bewegungsradius der mobilen Reporter-Einheit durch die Gesamtlänge der DNA-Leine gegeben (vgl. linke Seite in Abb. 1.1). Im hybridisierten Zustand ist eine Schlaufe in der DNA-Leine ausgeformt (vgl. rechte Seite in Abb. 1.1). Dadurch verkürzt sich im Vergleich zum dissoziierten Zustand die freie Leinenlänge und mit ihr der mögliche Bewegungsradius der mobilen Reporter-Einheit. Im TMFM-Experiment erreichen die Auslenkungen der mobilen Reporter-Einheit (Abb. 2.3 d-f) deutlich höhere Absolutwerte als die der immobilen Reporter-Einheit. Der Maximalwert der radialen Auslenkung der mobilen Reporter-Einheit (Abb. 2.3 f) liegt bei rund 4,5 Pixeln. In den Auslenkungsabfolgen der mobilen Reporter-Einheit wechseln sich Intervalle höherer und geringerer Breite ab.

Der quadratische Mittelwert (engl. *root mean square*, kurz: RMS) der radialen Auslenkung über ein Zeitfenster der Breite $\pm \vartheta$ Bildperioden

$$r_{\text{RMS},\vartheta}(t) = \sqrt{\frac{1}{2\vartheta + 1} \cdot \sum_{t'=t-\vartheta}^{t+\vartheta} (r(t'))^2} \quad (2.2)$$

gibt weiteren Aufschluss über die Beweglichkeit der Reporter-Einheiten. Der Wert von $r_{\text{RMS},\vartheta}$ ist ein Maß für die Breite der Verteilung von r im Intervall $[t - \vartheta, t + \vartheta]$: Je breiter die Verteilung von r in diesem Intervall, umso höher ist der Wert von $r_{\text{RMS},\vartheta}$.

Abbildung 2.4 illustriert diesen Zusammenhang für dasselbe Beispielobjekt, dessen Positions- und Auslenkungsabfolgen in Abb. 2.2 und 2.3 gezeigt sind. Die Abfolge und Häufigkeitsverteilung von $r_{\text{RMS},20}$ der immobilen Reporter-Einheit weisen ein definiertes Niveau auf. Die Abfolge und Häufigkeitsverteilung von $r_{\text{RMS},20}$ der mobilen Reporter-Einheit weisen zwei definierte Niveaus auf.

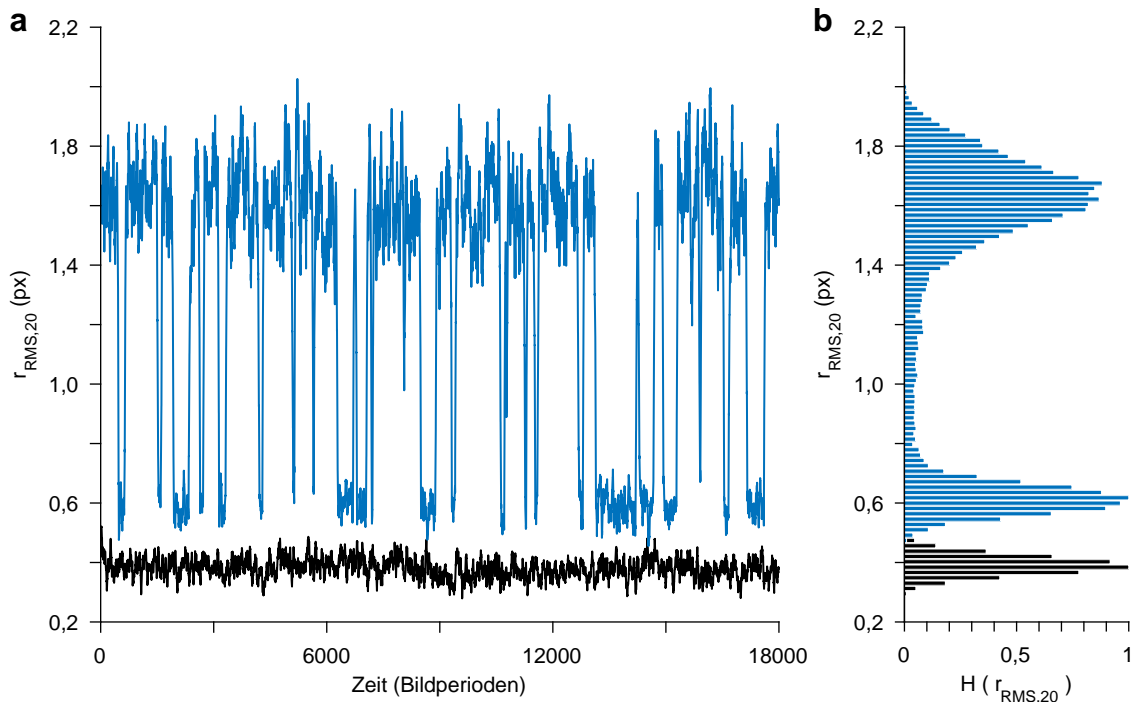


Abbildung 2.4: **Beweglichkeitsniveaus der Reporter-Einheiten eines DNA-Nano-Objekts**

Es handelt sich um dasselbe DNA-Nano-Objekt, dessen Positions- und Auslenkungsabfolgen in Abb. 2.2 und Abb. 2.3 gezeigt sind. **(a)** Quadratischer Mittelwert der radialen Auslenkung $r_{\text{RMS},20}$ nach Gl. (2.2) der immobilen Reporter-Einheit (schwarz) sowie der mobilen Reporter-Einheit (blau) über die ersten 18000 Bildperioden der Messung. **(b)** $H(r_{\text{RMS},20})$: Häufigkeiten von $r_{\text{RMS},20}$, normiert auf den jeweiligen Maximalwert, für die immobile Reporter-Einheit (schwarze Balken) sowie die mobile Reporter-Einheit (blaue Balken) über den gesamten Zeitraum der Messung (45000 Bildperioden).

Die einzelnen Beweglichkeitsniveaus der Reporter-Einheiten des hier gezeigten Objekts bleiben über den Verlauf der Messung gleich. Die mobile Reporter-Einheit wechselt zwar zwischen zwei Zuständen, die Breite der Abfolgen Δx , Δy und r in jedem der Zustände ändert sich jedoch nicht. In dieser Eigenschaft ist dieses Objekt allerdings *nicht* repräsentativ für die Allgemeinheit aller Objekte. Bei vielen Objekten trat über den Verlauf einer Messung eine Verschiebung der Beweglichkeitsniveaus beider Reporter-Einheiten und insbesondere der mobilen Reporter-Einheit auf. Mit diesem Charakteristikum der TMFM-Daten beschäftigt sich der folgende Abschnitt.

2.3 Zusammenhang zwischen der Helligkeit und der gemessenen Beweglichkeit der Reporter-Einheiten

Aus den TMFM-Aufnahmen wurde neben der Position ebenfalls die Helligkeit der Reporter-Einheiten erfasst. Dazu wurden für jede Reporter-Einheit in jedem Kamerabild die Intensitäten aller Pixel in einer Pixelregion fester Größe um die Position der Reporter-Einheit summiert. Somit erhält man für jede Reporter-Einheit analog zu den

Positions- und Auslenkungsabfolgen eine Helligkeitsabfolge

$$I(t), \quad t = 1, 2, 3, \dots, T.$$

Die in einem Kamerabild detektierte Helligkeit einer Reporter-Einheit hängt ab

1. von der Gesamtintensität I_{FL} des Fluoreszenzlichts, das die Farbstoffmoleküle dieser Reporter-Einheit während der Bildperiode emittieren, und
2. vom Anteil q_{proj} des emittierten Fluoreszenzlichts, der scharf auf den Sensor der Kamera abgebildet wird.

Für I_{FL} spielt neben dem Bleichzustand der Farbstoffmoleküle auch der Abstand der Reporter-Einheit von der Probenoberfläche während der Bildperiode eine Rolle. Das evaneszente Feld der TIRF-Illumination klingt exponentiell mit dem Abstand zur Probenoberfläche ab. Deshalb sinkt die Wahrscheinlichkeit der Anregung der noch nicht gebliebenen Farbstoffmoleküle mit steigendem Abstand der Reporter-Einheit zur Probenoberfläche. Auch für q_{proj} ist der Abstand zwischen Reporter-Einheit und Probenoberfläche zum Zeitpunkt der Emission ein entscheidender Faktor: Je weiter die Reporter-Einheit von der Probenoberfläche und somit von der Fokusebene des Objektivs entfernt ist, umso kleiner ist q_{proj} .

Aufgrund dieser Zusammenhänge sind die Helligkeiten der Reporter-Einheiten, insbesondere der mobilen Reporter-Einheiten, im Verlauf einer Messung nicht konstant (Abb. 2.5, 2.6, C.1 und C.2, jeweils graue Kurven in a). Durch sukzessives Bleichen der Farbstoffmoleküle [100] sinkt über die Dauer der Messung nach und nach die mittlere Helligkeit der Reporter-Einheit ab. Außerdem haben die mobilen Reporter-Einheiten im Zustand geringerer Beweglichkeit im Mittel einen kleineren Abstand zur Probenoberfläche und im Zustand höherer Beweglichkeit im Mittel einen größeren Abstand zur Probenoberfläche. Bei gleicher Bleichstufe führt folglich ein Wechsel vom Zustand höherer Beweglichkeit in den Zustand geringerer Beweglichkeit im Mittel zu einer Erhöhung der Helligkeit der Reporter-Einheit. Analog führt ein Wechsel vom Zustand geringerer Beweglichkeit in den Zustand höherer Beweglichkeit im Mittel zu einer Verringerung der Helligkeit der Reporter-Einheit. Bedingt durch die schnelle Diffusion der mobilen Reporter-Einheiten kann sich auch ohne einen Zustandswechsel der Abstand einer mobilen Reporter-Einheit zur Probenoberfläche in zwei aufeinanderfolgenden Kamerabildern stark unterscheiden. Infolgedessen unterliegt die Helligkeit einer Reporter-Einheit auch auf derselben Bleichstufe und im selben Beweglichkeitsniveau starken Schwankungen zwischen einzelnen Kamerabildern. Die Helligkeitsabfolgen $I(t)$ erscheinen stark verrauscht, weshalb es zweckmäßig ist, für jede Reporter-Einheit auch die über ein Zeitfenster von ± 10 Bildperioden Median-gefilterte Helligkeitsabfolge zu betrachten (Abb. 2.5, 2.6, C.1 und C.2, jeweils schwarze Kurven in a):

$$I_{med,10}(t) = \text{med}(I(t-10), \dots, I(t+10)) \quad (2.3)$$

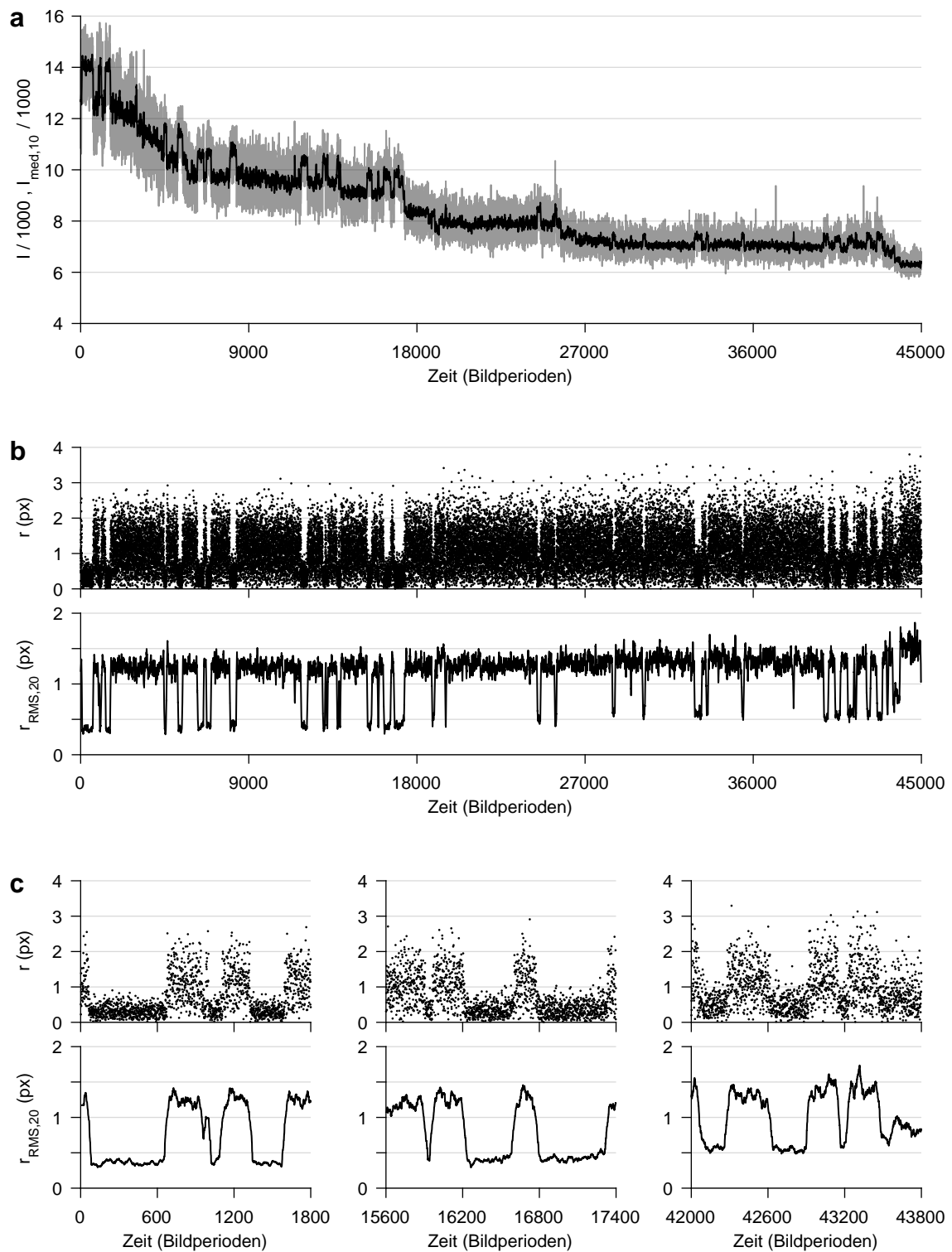


Abbildung 2.5: **Helligkeit und gemessene Beweglichkeit einer mobilen Reporter-Einheit, Bsp. 1**
 Objekt mit Interaktionspartnern in proximaler Verankerungsgeometrie, Konturlänge $L = 1050$ bp. **(a)** Grau: Helligkeitsabfolge $I(t)$. Schwarz: Median-gefilterte Helligkeitsabfolge $I_{med,10}$ nach Gl. (2.3). **(b)** Abfolgen der radialen Auslenkung r nach Gleichung (2.1) und von $r_{RMS,20}$ nach Gleichung (2.2) über die gesamte Dauer der Messung. **(c)** Drei verschiedene jeweils 3 Minuten lange Ausschnitte der Abfolgen von r und $r_{RMS,20}$.

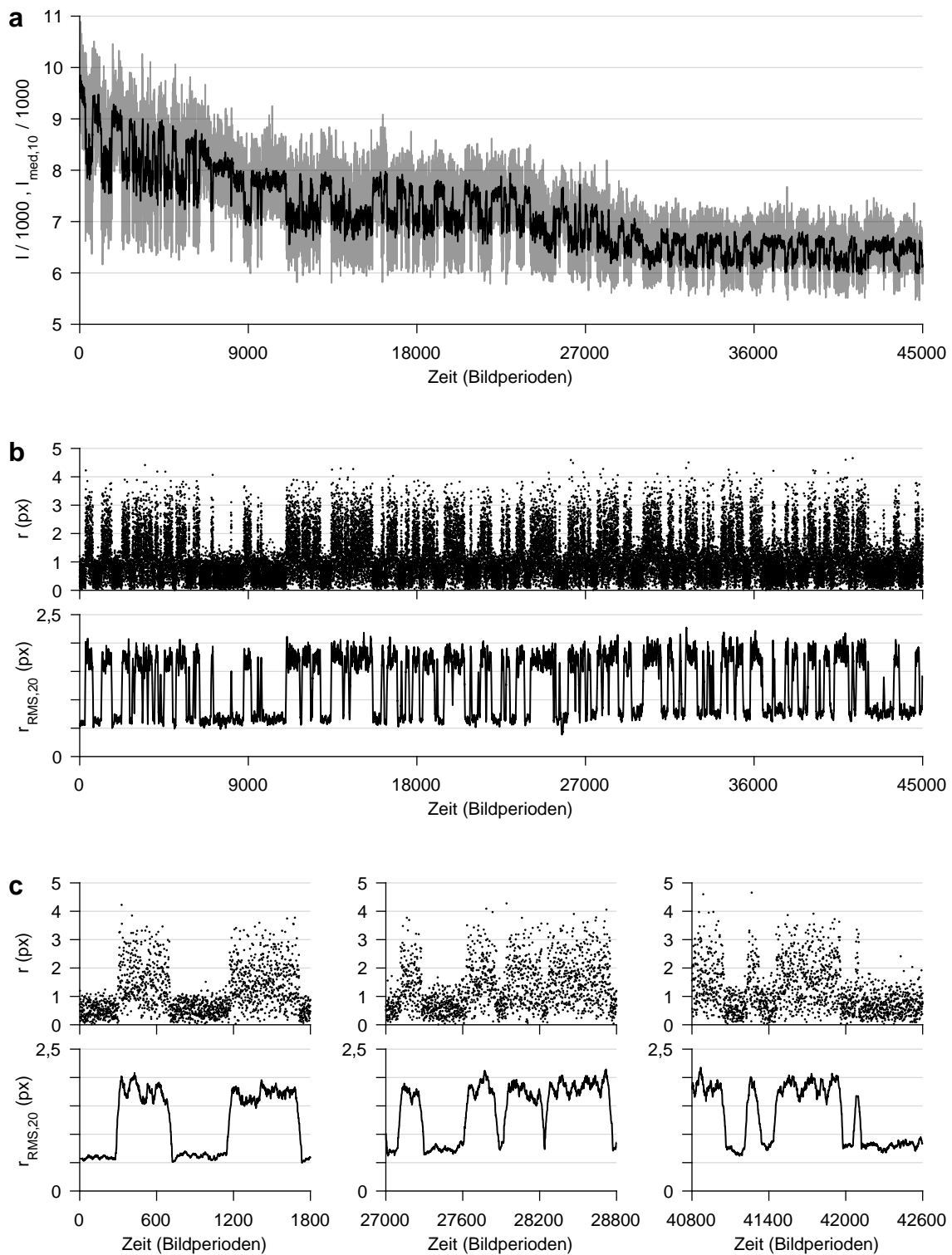


Abbildung 2.6: **Helligkeit und gemessene Beweglichkeit einer mobilen Reporter-Einheit, Bsp. 2**

Objekt mit Interaktionspartnern in distaler Verankerungsgeometrie, Konturlänge $L = 1554$ bp. **(a)** Grau: Helligkeitsabfolge $I(t)$. Schwarz: Median-gefilterte Helligkeitsabfolge $I_{med,10}$ nach Gl. (2.3). **(b)** Abfolgen der radialen Auslenkung r nach Gleichung (2.1) und von $r_{RMS,20}$ nach Gleichung (2.2) über die gesamte Dauer der Messung. **(c)** Drei verschiedene jeweils 3 Minuten lange Ausschnitte der Abfolgen von r und $r_{RMS,20}$.

Die Helligkeit der Reporter-Einheit beeinflusst die Breite der Auslenkungsabfolgen

In der Gegenüberstellung der Helligkeitsabfolgen (Abb. 2.5, 2.6, C.1, C.2, jeweils a) und der Abfolgen von $r(t)$ sowie $r_{\text{RMS},20}(t)$ (Abb. 2.5, 2.6, C.1, C.2, jeweils b und c) fällt auf, dass mit der Abnahme der Helligkeit ein gradueller Anstieg der gemessenen Beweglichkeit der Reporter-Einheiten einhergeht. Die Abfolgen von r werden breiter, $r_{\text{RMS},20}(t)$ verschiebt sich hin zu höheren Werten.

Diese Beobachtung kann damit erklärt werden, dass die statistische Ungenauigkeit der Positionsabschätzung vom Signal-zu-Rausch-Verhältnis (*Signal-to-Noise Ratio*, kurz: SNR) abhängt [103]. Je geringer die Helligkeit einer Reporter-Einheit, desto kleiner das SNR und desto größer die Ungenauigkeit der Positionsabschätzung. Dadurch wirkt die Beweglichkeit erhöht. Im Rahmen der Charakterisierung der Bewegung der mobilen Reporter-Einheiten (siehe Abschnitt 2.5) wurde dieser Zusammenhang weiter untersucht.

2.4 Aussortieren fehlerhafter DNA-Nano-Objekte

In der Regel eigneten sich nicht alle DNA-Nano-Objekte aus einer TMFM-Messung zur weiteren Analyse. Die Einzelmolekül-Daten aller Objekte wurden betrachtet und manuell aussortiert, falls sie als fehlerbehaftet eingestuft wurden. Daten von DNA-Nano-Objekten wurden hauptsächlich dann als fehlerbehaftet eingestuft, wenn

- der Abstand zwischen benachbarten DNA-Nano-Objekten weniger als etwa 12 px betrug, was Fehler in der Positionsabschätzung durch den VWCM-Algorithmus zur Folge hatte.
- dauerhaftes oder wiederholtes transientes Kleben der mobilen Reporter-Einheit an der Probenoberfläche beobachtet wurde.
- das Fluoreszenzsignal mindestens einer der Reporter-Einheiten nach weniger als 10000 Bildperioden ab Anfang der Messung durch Bleichen der Farbstoffe oder Ablösung des DNA-Nano-Objekts von der Oberfläche verschwunden war.
- die immobile Reporter-Einheit mehr als ein definiertes Beweglichkeitsniveau aufwies
- bei DNA-Nano-Objekten mit Interaktionspartnern die mobile Reporter-Einheit nur ein definiertes Beweglichkeitsniveau oder mehr als zwei definierte Beweglichkeitsniveaus aufwies.
- bei DNA-Nano-Objekten der Referenzen für einen der beiden Zustände (vgl. Abschnitt 2.5) die mobile Reporter-Einheit mehr als ein definiertes Beweglichkeitsniveau aufwies.

In etwa 80 % der Datensätze aus den TMFM-Messungen wurden jeweils mindestens 20 % der DNA-Nano-Objekte als fehlerfrei eingestuft und einer weiteren Analyse unterzogen. In etwa 10 % der Datensätze lag der Anteil der als fehlerfrei eingestuften DNA-Nano-Objekte sogar über 50 %.

2.5 Charakterisierung der Bewegung der mobilen Reporter-Einheiten

Zur Charakterisierung der Bewegung der mobilen Reporter-Einheiten für die distale Verankerungsgeometrie der Interaktionspartner wurden Referenzobjekte hergestellt, die bei korrekter Selbstassemblierung jeweils dauerhaft einen der zwei Zustände nachahmen. Abschnitt 2.5.1 zeigt die Ergebnisse der Charakterisierung der Bewegung der mobilen Reporter-Einheiten dieser Referenzobjekte für den Zustand mit fester Schlaufe in der Leine und für den Zustand ohne Schlaufe in der Leine.

Zur Charakterisierung der Bewegung der mobilen Reporter-Einheiten im Zustand mit Schlaufe in der Leine für die proximale Verankerungsgeometrie der Interaktionspartner wurden Abschnitte von Auslenkungsabfolgen von Objekten mit 9 nt langen Interaktionspartnern verwendet (siehe Abschnitt 2.5.2).

2.5.1 Charakterisierung der Bewegung der mobilen Reporter-Einheiten der Referenzobjekte mit fester Schlaufe und ohne Schlaufe in der Leine

Zur Charakterisierung der Bewegung der mobilen Reporter-Einheit im Zustand mit Schlaufe in der Leine in der distalen Verankerungsgeometrie diente die F_d -Referenz (Abb. 2.7 a). Bei diesen Referenzobjekten beträgt die Konturlänge L in der Schlaufe

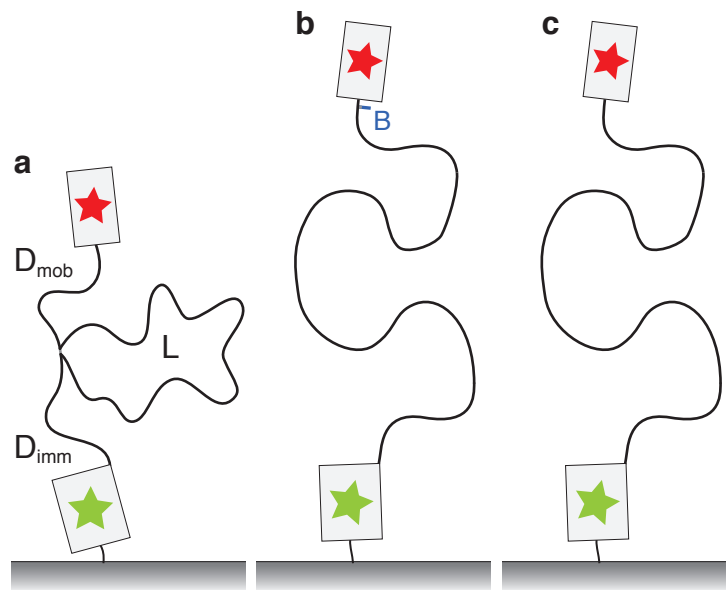


Abbildung 2.7: **Schemata der F_d -Referenz, der B_0 -Referenz und der 0-Referenz**

(Detailliertere Darstellungen dieser Referenzobjekte finden sich in Abb. 1.8 e-i und Abb. 1.9 c-e.)

Die grauen Rechtecke mit rotem/grünem Stern repräsentieren die Reporter-Einheiten. Die schwarze Linie repräsentiert die DNA-Leine. Detailliertere Darstellungen finden sich in Abbildungen 1.8 und 1.9. **(a)** Schema der F_d -Referenz mit fester Schlaufe in der Leine. D_{imm} resp. D_{mob} bezeichnen die Konturlängen der Leinstücke zwischen der immobilen resp. mobilen Reporter-Einheit und dem Punkt der festen Verknüpfung in der Leine. L bezeichnet die Konturlänge der Schlaufe, die durch die Verknüpfung in der Leine entsteht. **(b)** Schema der B_0 -Referenz mit nur dem Interaktionspartner B (blau) in proximale Verankerungsgeometrie nahe der mobilen Reporter-Einheit. **(c)** Schema der 0-Referenz ohne Interaktionspartner.

1008 bp. Die freie Leinenlänge $D_{\text{imm}} + D_{\text{mob}}$ beträgt 601 bp und ist damit um 42 bp länger als im Zustand mit Schlaufe in der Leine bei den Objekten mit Interaktionspartnern in distaler Verankerungsgeometrie.

Zur Charakterisierung der Bewegung der mobilen Reporter-Einheit im Zustand ohne Schlaufe in der Leine wurden zwei Sorten der Referenzobjekte verwendet, bei denen eine Interaktionssequenz oder beide Interaktionssequenzen fehlen: Die B0-Referenzen (Abb. 2.7b) und die 0-Referenzen (Abb. 2.7c). Im Rahmen dieser Arbeit wurden Varianten der B0- und 0-Referenzen mit insgesamt sechs verschiedenen Leinenlängen untersucht. Die verschiedenen Varianten werden hier mit den Kürzeln B0-Ln oder 0-Ln bezeichnet. Der Parameter $n \in \{2, 4, 6, 8, 10, 12\}$ gibt die Gesamtlänge $D_{\text{imm}} + L + D_{\text{mob}}$ der DNA-Leine zwischen den Reporter-Einheiten in der jeweiligen Variante gemäß der folgenden Regel an:

$$D_{\text{imm}} + L + D_{\text{mob}} = 97 \text{ bp} + n \cdot 252 \text{ bp} \quad (2.4)$$

Objekte der Referenz B0-Ln oder 0-Ln haben dieselbe Gesamtleinenlänge wie Objekte mit Interaktionspartnern in proximaler Verankerungsgeometrie und

$$L = 42 \text{ bp} + n \cdot 252 \text{ bp}$$

und wie Objekte mit Interaktionspartnern in distaler Verankerungsgeometrie und

$$L = 42 \text{ bp} + (n - 2) \cdot 252 \text{ bp}, \text{ für } n > 2.$$

Anm.: Die Experimente mit den Referenzobjekten F_d , 0-L8, 0-L12 sowie erste Analyseschritte wurden im Rahmen der Bachelor-Arbeit und HiWi-Tätigkeit von Dominik Voglmaier am Lehrstuhl von Prof. Dietz durchgeführt. Herrn Voglmaier sei an dieser Stelle erneut für die gute Zusammenarbeit gedankt.

Unimodalität der Bewegung der mobilen Reporter-Einheiten

Korrekt assemblierte Referenzobjekte sollten im TMFM-Experiment nur ein Beweglichkeitsniveau aufweisen. Für die Objekte der F_d -Referenz sollte die Schlaufe fest geschlossen sein und sich nicht öffnen können. Umgekehrt verhält es sich für die Objekte der B0-Referenz und der 0-Referenz: Da diese Objekte nicht über ein vollständiges Paar von Interaktionspartnern verfügen, sollte für sie laut Design keine Schlaufenbildung in der DNA-Leine möglich sein.

Die Auslenkungsabfolgen der großen Mehrzahl der mobilen Reporter-Einheiten der F_d -, B0- und 0-Referenzen zeigten im TMFM-Experiment nur ein Beweglichkeitsniveau. Nur vereinzelte Referenzobjekte wiesen mehr als ein Beweglichkeitsniveau der mobilen Reporter-Einheit auf. Der geringe Anteil dieser Objekte an der Gesamtheit lässt darauf schließen, dass die Ursache für das Auftreten von zwei Beweglichkeitsniveaus in der fehlerhaften Assemblierung einzelner Objekte liegt. Aufgrund von Assemblierungsfehlern können beispielsweise einzelsträngige Motive in den DNA-Nano-Objekten offen liegen, die laut Design doppelsträngig sein sollten [104]. Liegen in einem Objekt der B0- oder 0-Referenz mehrere solche einzelsträngige Motive mit einer gewissen Sequenzkomple-

mentarität zueinander vor, können diese aneinander binden und so kann eine Schlaufe in der Leine entstehen.

Solche alternative Wege der Schleifenbildung könnten dann auch in vereinzelt Objekten mit Interaktionspartnern auftreten. Falls die Konturlänge einer defektbedingten Schlaufe in einem ähnlichen Bereich liegt wie die Konturlänge der durch Bildung des Duplex B:B' gebildeten Schlaufe, hätte die mobile Reporter-Einheit in den beiden Zuständen mit Schlaufe ein ähnliches Beweglichkeitsniveau. Das DNA-Nano-Objekt würde dann zunächst nicht als fehlerhaft identifiziert. Falls sich allerdings durch die alternative Schleifenbildung die Statistik der Verweildauern dieses Objekts in den Zuständen mit und ohne Schlaufe maßgeblich verändert, würde das im Rahmen der Prozedur zur Detektion molekularer Heterogenität erkannt (siehe Abschnitt 3.3.3). Das Objekt würde dann von der weiteren Analyse ausgeschlossen.

Zusammenhang zwischen Helligkeit und gemessener Beweglichkeit der Reporter-Einheiten

Wie bei den DNA-Nano-Objekten mit Interaktionspartnern (vgl. Abschnitt 2.3) wurde auch bei den Referenzobjekten mit fester Schlaufe in der Leine und ohne Schlaufe in der Leine beobachtet, dass die gemessene Beweglichkeit der mobilen Reporter-Einheiten mit abfallender Helligkeit ihres Fluoreszenzsignals ansteigt. Abbildungen 2.8 und 2.9 sowie Tabelle 2.1 verdeutlichen diesen Zusammenhang am Beispiel eines Objekts der F_d -Referenz¹.

Die Median-gefilterte Abfolge der Helligkeit um die Position der Reporter-Einheit $I_{\text{med},10}$ (Abb. 2.8 a) fällt von einem anfänglichen Niveau bei etwa 13000 im Verlauf der Messung nach und nach ab und liegt zuletzt bei einem Wert von knapp unter 6000. Die Helligkeitsabnahme vollzieht sich dabei in mehreren Stufen, welche auf das Bleichen einzelner Farbstoffmoleküle zurückzuführen sind.

Die Breite der Auslenkungsabfolgen (Abb. 2.8 b) nimmt über den Verlauf der Messung zu, insbesondere in der zweiten Hälfte. Für eine genauere Betrachtung wurden fünf Zeitintervalle auf vier verschiedenen Intensitätsniveaus ausgewählt (Abb. 2.8 a). Jedes dieser Intervalle hat eine Länge von 1000 Bildperioden. Somit ergeben sich fünf Stichproben von je 1000 Werten von $I_{\text{med},10}$, je 1000 Werten von Δx und je 1000 Werten von Δy . Tabelle 2.1 gibt für jedes Intervall das Maximum, den Median und das Minimum der Stichprobe von $I_{\text{med},10}$ sowie das arithmetische Mittel und die Standardabweichung der Stichproben von Δx und Δy an. Mit steigendem Index der Intervalle und der sukzessiven Helligkeitsabnahme geht eine Zunahme der Standardabweichungen der Auslenkungsstichproben einher. Die Werte der Standardabweichungen der Stichproben aus Intervallen 2 und 3 liegen um etwa 10 bis 15 Prozent über den Werten aus Intervall 1. Die Werte aus Intervall 4 liegen um mehr als 20 Prozent über den Werten aus Intervall 1. Die Werte aus Intervall 5 liegen um etwa 40 Prozent über den Werten aus Intervall 1.

Für jedes der fünf Intervalle wurden die Verteilungen der Stichproben von Δx und Δy betrachtet (Abb. 2.9). Die jeweilige Wahrscheinlichkeitsdichteverteilung (PDF) wurde einerseits durch Bildung eines Histogramms der relativen Häufigkeit geschätzt, an-

¹Abb. C.3 und C.4 sowie Tabelle C.2 zeigen eine analoge Betrachtung für ein Objekt der Referenz B0-L6.

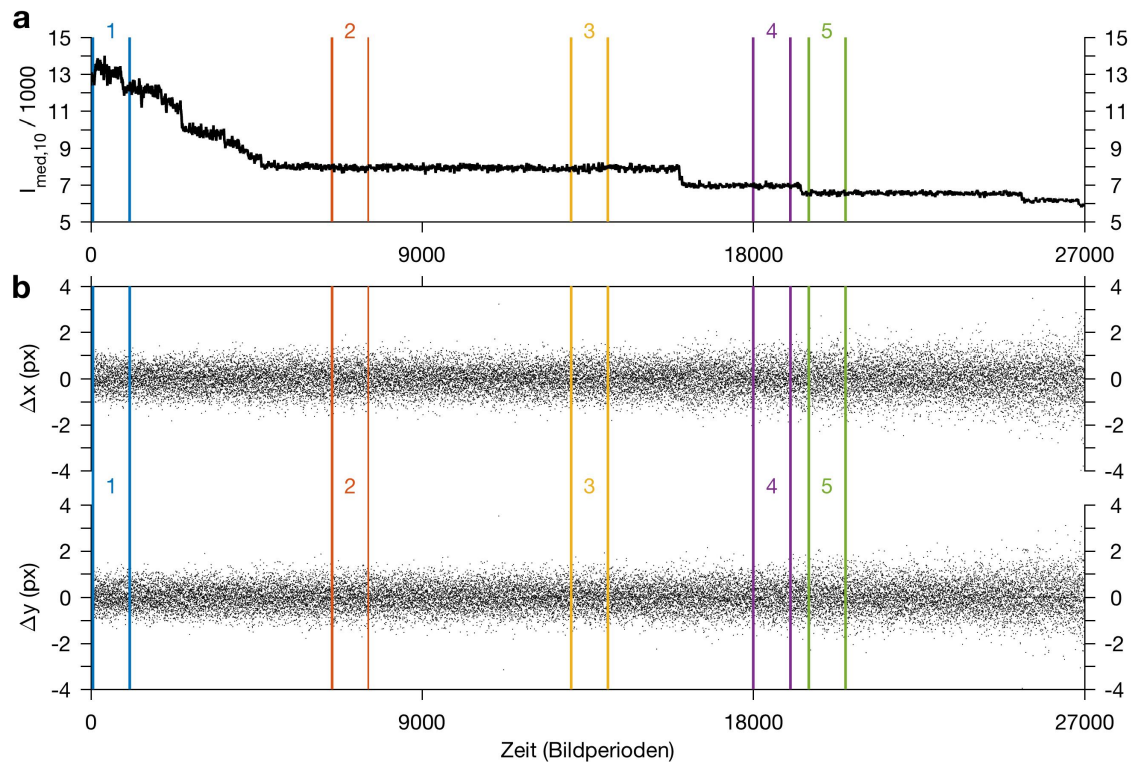


Abbildung 2.8: **Helligkeits- und Auslenkungsabfolgen einer mobilen Reporter-Einheit, F_d -Referenz**

(a) Schwarze Linie: $I_{med,10}$, Median-gefilterte Abfolge der summierten Pixelhelligkeit um die Position der Reporter-Einheit nach Gl. (2.3). (b) Punkte: Δx und Δy , Abfolgen der Auslenkungen in x- und y-Richtung nach Gl. (2.2). In Farbe: Grenzen und Nummerierung der fünf Zeitintervalle, denen die Kennzahlen und Verteilungen in Tab. 2.1 und Abb. 2.9 entstammen.

Tabelle 2.1: **Kennzahlen der Helligkeit und der Auslenkungen für die in Abb. 2.8 markierten Intervalle**
 k: Nummer des Intervalls. $I_{med,10}$: Median-gefilterte summierte Pixelhelligkeit nach Gl. (2.3). Jeweils gerundete Werte des Maximums, Medians und Minimums im k-ten Intervall. $m_{x,k} / m_{y,k}$: Arithmetisches Mittel der Stichprobe von $\Delta x / \Delta y$ im k-ten Intervall. $s_{x,k} / s_{y,k}$: Standardabweichung der Stichprobe von $\Delta x / \Delta y$ im k-ten Intervall.

k	$I_{med,10}$			$m_{x,k}$	$m_{y,k}$	$s_{x,k}$	$s_{y,k}$
	Maximum	Median	Minimum				
1	14006	12993	11822	0,021	0,017	0,503	0,482
2	8173	7917	7619	0,011	0,016	0,563	0,551
3	8237	7892	7562	0,008	-0,013	0,554	0,554
4	7246	6940	6706	0,016	0,004	0,619	0,598
5	6771	6569	6330	-0,007	0,005	0,688	0,672

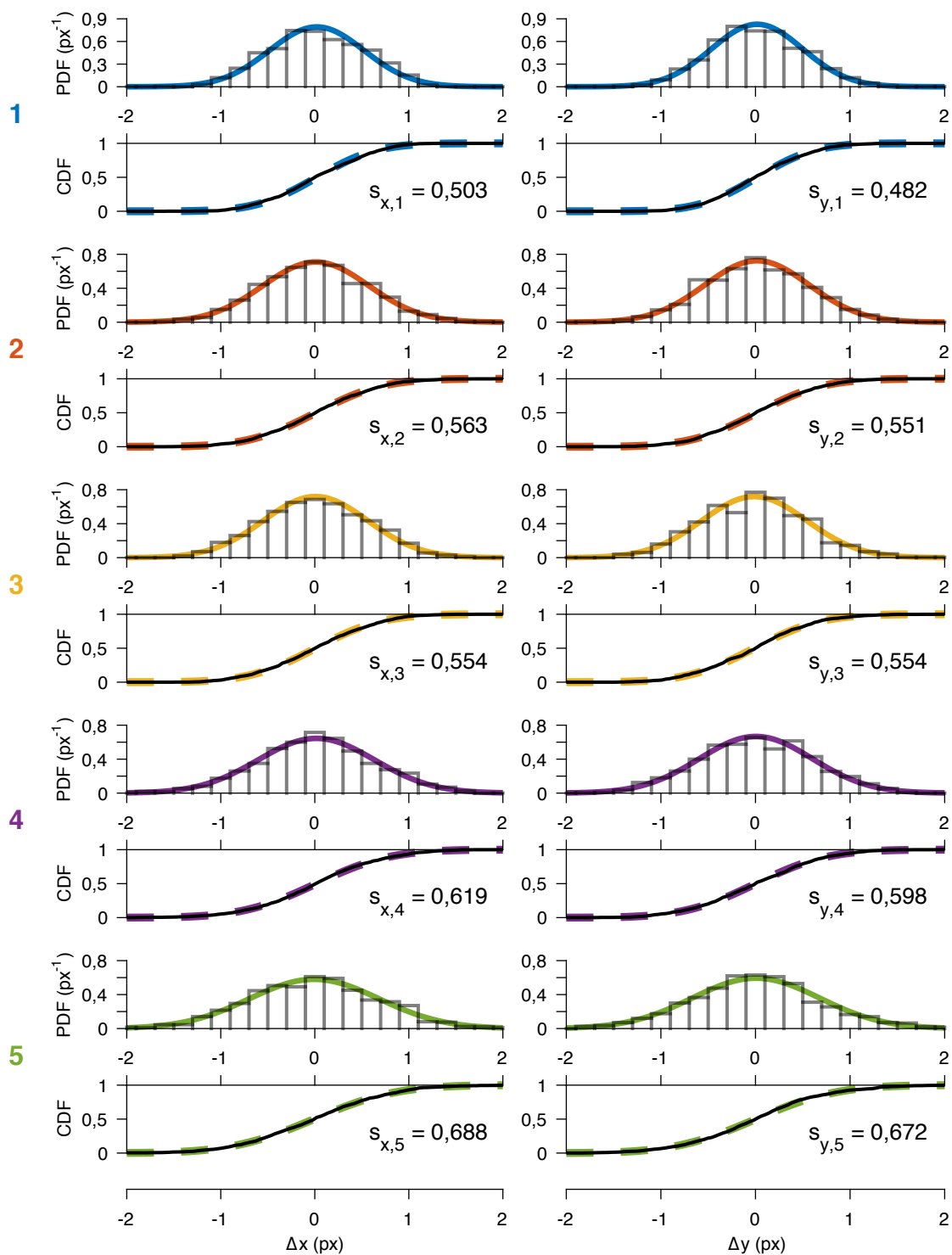


Abbildung 2.9: Verteilungen der Auslenkungsstichproben aus den fünf Zeitintervallen

Verteilungen von Δx (linke Spalte) und Δy (rechte Spalte). Histogramme: Aus der relativen Häufigkeit ermittelte Wahrscheinlichkeitsdichte. Schwarze Linien: Kumulierte relative Häufigkeit. Farbige durchgezogene resp. gestrichelte Linien: Normalverteilte Wahrscheinlichkeitsdichte resp. Verteilungsfunktion mit Erwartungswert $m_{x,k} / m_{y,k}$ (arithmetisches Mittel der Stichprobe) und Standardabweichung $s_{x,k} / s_{y,k}$ (Standardabweichung der Stichprobe), $k \in \{1, \dots, 5\}$.

dererseits durch die kontinuierliche normalverteilte Dichtefunktion der Form

$$p(\delta) = \frac{1}{s\sqrt{2\pi}} \cdot \exp\left(-\frac{(\delta - m)^2}{2s^2}\right). \quad (2.5)$$

Dabei bezeichnet m das Stichprobenmittel, s die Standardabweichung der Stichprobe.

Zwischen den Histogrammen und den Verläufen der normalverteilten Dichtefunktionen p wurde für alle Stichproben eine gute Übereinstimmung festgestellt. Ebenso ergab der Vergleich der Verteilung der kumulierten relativen Häufigkeit mit der zur PDF gehörigen Verteilungsfunktion (CDF) für jede Stichprobe eine gute Übereinstimmung zwischen den Verläufen der beiden Kurven. Diese große Ähnlichkeit der Auslenkungsverteilungen zu Normalverteilungen wird an den jeweiligen Enden der Abschnitte 2.5.1 und 2.5.2 erneut aufgegriffen.

Autokorrelation der Auslenkungsabfolgen der mobilen Reporter-Einheiten

Ein weiterer Teil der Charakterisierung der Bewegung der mobilen Reporter-Einheiten bestand in der Betrachtung der Autokorrelation der Auslenkungsabfolgen. In dieser Arbeit wird die normierte Autokorrelation $R(v, \Delta t)$ einer beliebigen Abfolge

$$v(t), t \in \{1, 2, \dots, \ell\}$$

bei der zeitlichen Verschiebung Δt wie folgt definiert:

$$R(v, \Delta t) = \frac{\frac{1}{\ell - \Delta t} \sum_{t=1}^{\ell - \Delta t} v(t) \cdot v(t + \Delta t)}{\frac{1}{\ell} \sum_{t=1}^{\ell} v(t)^2} \quad (2.6)$$

Bei der Betrachtung der normierten Autokorrelation der Auslenkungsabfolgen Δx und Δy zeigte sich eine periodische Abhängigkeit von der zeitlichen Verschiebung Δt . Abbildung 2.10 illustriert dies anhand der normierten Autokorrelationen von Ausschnitten der Auslenkungsabfolgen einer Auswahl mobiler Reporter-Einheiten der F_d -Referenz innerhalb eines 500 Bildperioden langen Zeitintervalls. Die Periodizität in den Verläufen von $R(\Delta x, \Delta t)$ über der zeitlichen Verschiebung Δt ist klar erkennbar (Abb. 2.10 a). Die Periodizität in $R(\Delta y, \Delta t)$ ist weniger stark ausgeprägt, tritt jedoch in der Mittelwertskurve (schwarze Linie) deutlich zutage (Abb. 2.10 b).

Um der Frage nachzugehen, ob eine Periodizität von $R(\Delta x, \Delta t)$ und $R(\Delta y, \Delta t)$ nur vereinzelt oder global in den TMFM-Messungen auftritt, wurden die normierten Autokorrelationen vieler jeweils 500 Bildperioden langer Ausschnitte aus den Auslenkungsabfolgen der Reporter-Einheiten aller Referenzobjekte ermittelt. Dieses Vorgehen liefert für jeden Wert von Δt Verteilungen der Werte von $R(\Delta x, \Delta t)$ und $R(\Delta y, \Delta t)$. Aus diesen Verteilungen wurden jeweils das arithmetische Mittel sowie der Wertebereich zwischen dem 0,005-Quantil und dem 0,995-Quantil betrachtet. Eine derartige globale Betrachtung

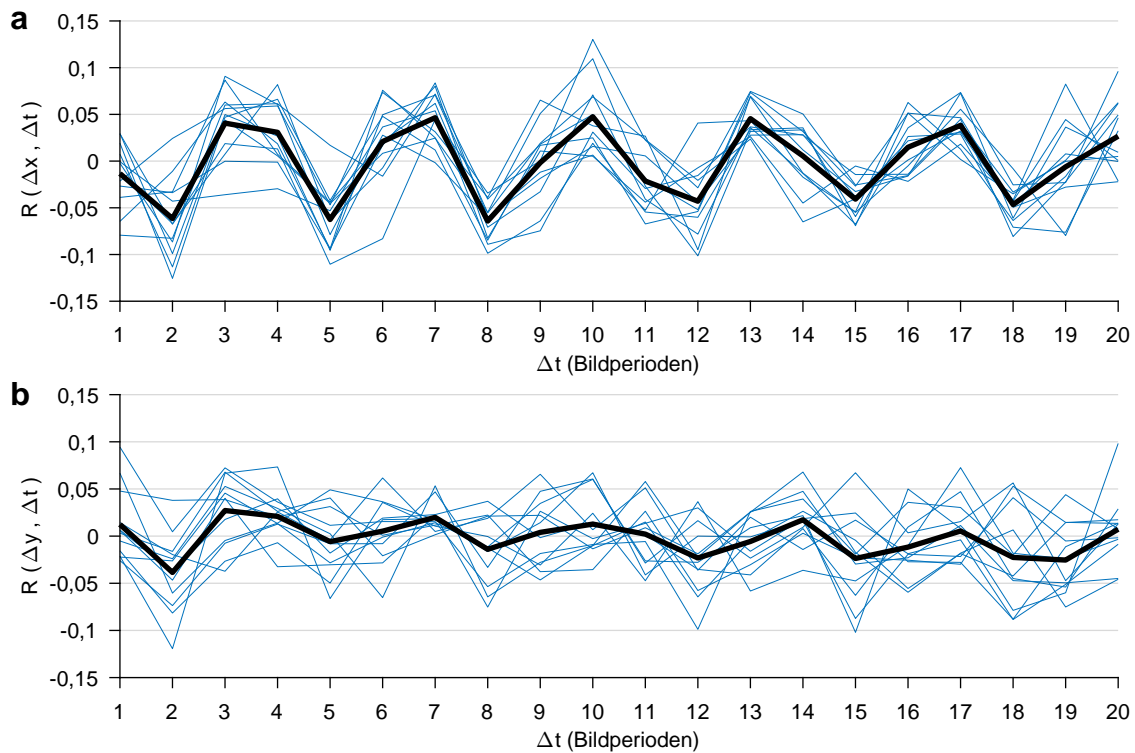


Abbildung 2.10: **Normierte Autokorrelationen von Ausschnitten der Abfolgen Δx und Δy mobiler Reporter-Einheiten von zwölf Beispielobjekten der F_d -Referenz**

Die Ausschnitte der Auslenkungsabfolgen entstammen demselben 500 Bildperioden langen Zeitintervall der TMFM-Messung. **(a)** $R(\Delta x, \Delta t)$ und **(b)** $R(\Delta y, \Delta t)$ nach Gl. (2.6) für zeitliche Verschiebungen Δt um 1 bis 20 Bildperioden. Blaue Kurven: jeweils normierte Autokorrelationen der Auslenkung einzelner Reporter-Einheiten. Schwarze Kurven: jeweils arithmetisches Mittel aller blauen Kurven.

für die F_d -Referenz (Abb. 2.11) liefert ein qualitativ gleiches Ergebnis wie die Betrachtung der zwölf einzelnen Ausschnitte (Abb. 2.10): eine deutlich erkennbare Periodizität im Verlauf von $R(\Delta x, \Delta t)$ (Abb. 2.11 a) und eine schwächere, aber dennoch erkennbare Periodizität im Verlauf von $R(\Delta y, \Delta t)$ (Abb. 2.11 c).

In den Auslenkungsabfolgen tritt folglich systematisch eine periodische Bewegungskomponente mit einer Frequenz von etwa 3 Hz auf. Die Ursache dafür liegt aller Wahrscheinlichkeit nach in einer Vibration der Probe relativ zur Aufnahmeoptik. Die zur Korrektur instrumentellen Drifts verwendete Prozedur (siehe Anhang B.1) entspricht letztlich einem Tiefpass-Filter und entfernt nur Vibrationen mit Frequenzen $\leq 0,1$ Hz.

Zur nachträglichen Vibrationskorrektur der Auslenkungsabfolgen wurde eine Prozedur ähnlich zu der in [79] beschriebenen Vibrationskorrektur implementiert. Für jeden Zeitpunkt der Messung wurde durch Mittelung über die Auslenkungen der Reporter-Einheiten mehrerer DNA-Nano-Objekte ein Schätzwert für die Auslenkung des Schwerpunkts der gesamten Probe $(\Delta x_{CM}, \Delta y_{CM})$ ermittelt. Subtraktion der Abfolge der ge-

schätzten Schwerpunktauslenkung von den Abfolgen $(\Delta x, \Delta y)$ liefert eine Schätzung für die Abfolge der tatsächlichen xy -Auslenkungen der Reporter-Einheiten $(\Delta x_{RE}, \Delta y_{RE})$ im Bezug auf den Ankerpunkt des DNA-Nano-Objekts an der Probenoberfläche:

$$(\Delta x_{RE}, \Delta y_{RE}) = (\Delta x, \Delta y) - (\Delta x_{CM}, \Delta y_{CM}) \quad (2.7)$$

Die normierten Autokorrelationen der vibrationskorrigierten Abfolgen $R(\Delta x_{RE}, \Delta t)$ und $R(\Delta y_{RE}, \Delta t)$ wurden ebenfalls global betrachtet. Sie wurden aus denselben Sätzen von jeweils 500 Bildperioden langen Zeitintervallen ermittelt wie $R(\Delta x, \Delta t)$ und $R(\Delta y, \Delta t)$. Die Verläufe des arithmetischen Mittels und der 0,005- bzw. 0,995-Quantile von $R(\Delta x_{RE}, \Delta t)$ und $R(\Delta y_{RE}, \Delta t)$ über der zeitlichen Verschiebung für die F_d -Referenz zeigen keine ausgeprägte Periodizität (Abb. 2.11 b und d).

Eine gleichartige Betrachtung der normierten Autokorrelationen von $(\Delta x, \Delta y)$ und $(\Delta x_{RE}, \Delta y_{RE})$ erfolgte für die Referenzobjekte für den Zustand ohne Schlaufe in der Lei-

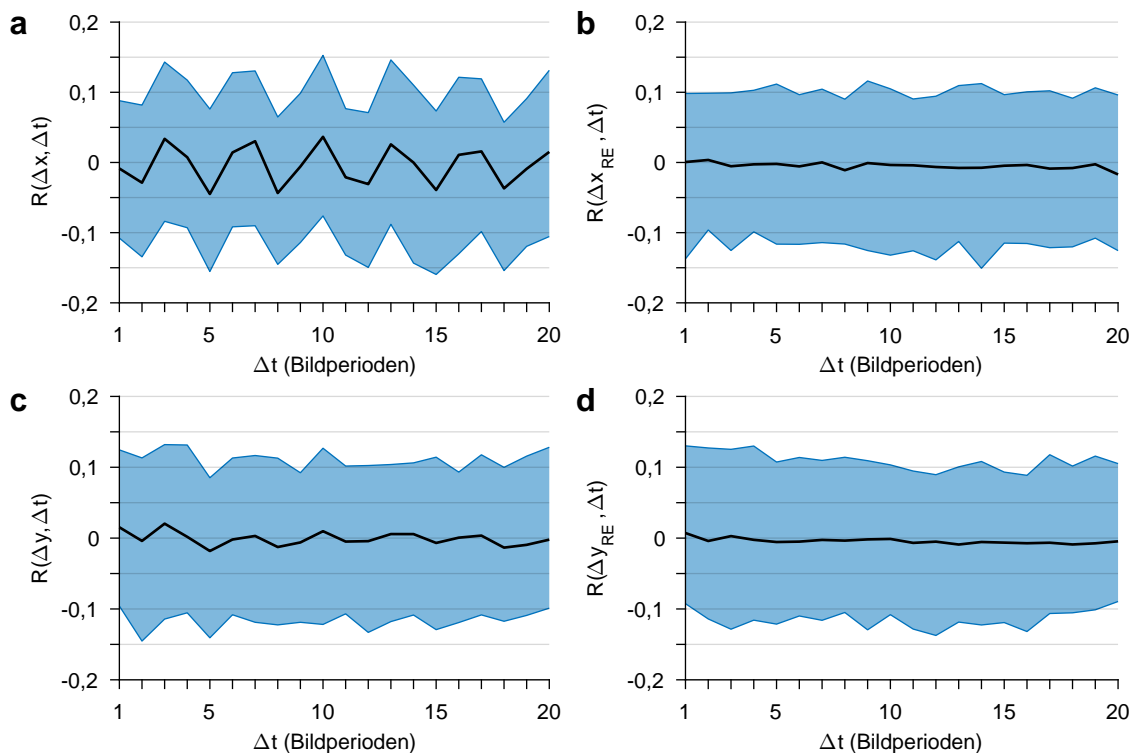


Abbildung 2.11: **Normierte Autokorrelationen der Abfolgen Δx , Δx_{RE} , Δy und Δy_{RE} mobiler Reporter-Einheiten der F_d -Referenz in vielen ($N = 274$) jeweils 500 Bildperioden langen Ausschnitten** (a) $R(\Delta x, \Delta t)$, (b) $R(\Delta x_{RE}, \Delta t)$, (c) $R(\Delta y, \Delta t)$ und (d) $R(\Delta y_{RE}, \Delta t)$ nach Gl. (2.6) für zeitliche Verschiebungen Δt um 1 bis 20 Bildperioden. Die blauen Flächen werden begrenzt durch die 0,005- und 0,995-Quantile der Verteilungen aus allen 274 Intervallen. Schwarze Kurven: jeweils arithmetisches Mittel der Werte aus allen 274 Intervallen.

ne B0-L2, B0-L4 und B0-L6 (Abb. 2.12) sowie 0-L8, B0-L10 und 0-L12 (Abb. 2.13). Die Periodizität von $R(\Delta x, \Delta t)$ ist für die mobilen Reporter-Einheiten der Referenzobjekte mit den zwei kürzesten Gesamtleinenlängen (Abb. 2.12 a, e) ungefähr gleich stark ausgeprägt wie im Fall der F_d -Referenz. Mit weiter steigender Gesamtleinenlänge ist die Periodizität von $R(\Delta x, \Delta t)$ immer weniger stark ausgeprägt (Abb. 2.12 i, Abb. 2.13 a, e, i). Die Periodizität von $R(\Delta y, \Delta t)$ ist nur bei der kürzesten Leinenlänge deutlich auszumachen (Abb. 2.12 c), bei allen höheren Gesamtleinenlängen jedoch kaum bis gar nicht (Abb. 2.12 g, k, Abb. 2.13 c, g, k). Die Verläufe von $R(\Delta x_{RE}, \Delta t)$ und $R(\Delta y_{RE}, \Delta t)$ zeigen für alle Gesamtleinenlängen keine ausgeprägte Periodizität (jeweils rechte Spalten in Abb. 2.12 und Abb. 2.13).

Das Ausmaß der zeitlich unkorrelierten Brown'schen Bewegung der mobilen Reporter-Einheiten nimmt mit steigender freier Leinenlänge zu. Die oben dargelegten Beobachtungen sind demnach konsistent mit der Annahme, dass die periodische Komponente der Auslenkungsabfolgen $(\Delta x, \Delta y)$ von einer systematischen Vibration in der Messapparatur herrührt, deren Amplitude in allen Messungen ungefähr gleich groß ist. Der relative Anteil der periodischen Komponente an der Gesamtbewegung schrumpft dann mit wachsender Leinenlänge, was sich in einer weniger ausgeprägten Periodizität niederschlägt. Unabhängig von der Leinenlänge konnte der periodische Bewegungsanteil mithilfe der Vibrationskorrektur erfolgreich entfernt oder zumindest auf ein sehr geringes Maß verringert werden.

Zur quantitativen Einordnung der Mittelwerte und Wertebereiche der normierten Autokorrelation der vibrationskorrigierten Auslenkungsabfolgen wurde ein Vergleich zu simulierten Daten gezogen. Der simulierte Datensatz bestand aus 10000 Abfolgen Δ_{SIM} von normalverteilten Pseudozufallszahlen mit einer jeweiligen Länge der Abfolgen von $\ell = 500$. Für jede dieser Abfolgen wurde die normierte Autokorrelation $R(\Delta_{SIM}, 1)$ nach Gl. (2.6) berechnet. Aus der Verteilung der 10000 Einzelwerte für $R(\Delta_{SIM}, 1)$ wurden die Vergleichswerte für die Mittelwerte und Quantile der experimentell erhaltenen Verteilungen bestimmt.

Tabelle 2.2 zeigt eine Aufstellung der Statistik aus den experimentellen Daten aller Referenzobjekte im Vergleich mit der Statistik aus dem simulierten Datensatz. Die Mittelwerte und Quantile der experimentellen Daten liegen etwa im selben Bereich wie die aus dem simulierten Datensatz gewonnenen Vergleichswerte. Dementsprechend können die Abfolgen der Auslenkungen im Zustand ohne Schlaufe in der Leine und im Zustand mit Schlaufe in der Leine bei distaler Verankerungsgeometrie als zeitlich unkorreliert angesehen werden.

Die zur statistischen Auswertung der Zwei-Zustands-Abfolgen verwendete Methode (siehe Kapitel 3) fordert, dass die Auslenkung zu jedem Zeitpunkt eine unabhängige Zufallsvariable sein muss. Demnach darf der Wert der Auslenkung zu einem bestimmten Zeitpunkt der Abfolge nicht vom Wert zu früheren Zeitpunkten abhängen. Aus diesem Grund wurden für die Analyse die vibrationskorrigierten Abfolgen $(\Delta x_{RE}, \Delta y_{RE})$ verwendet.

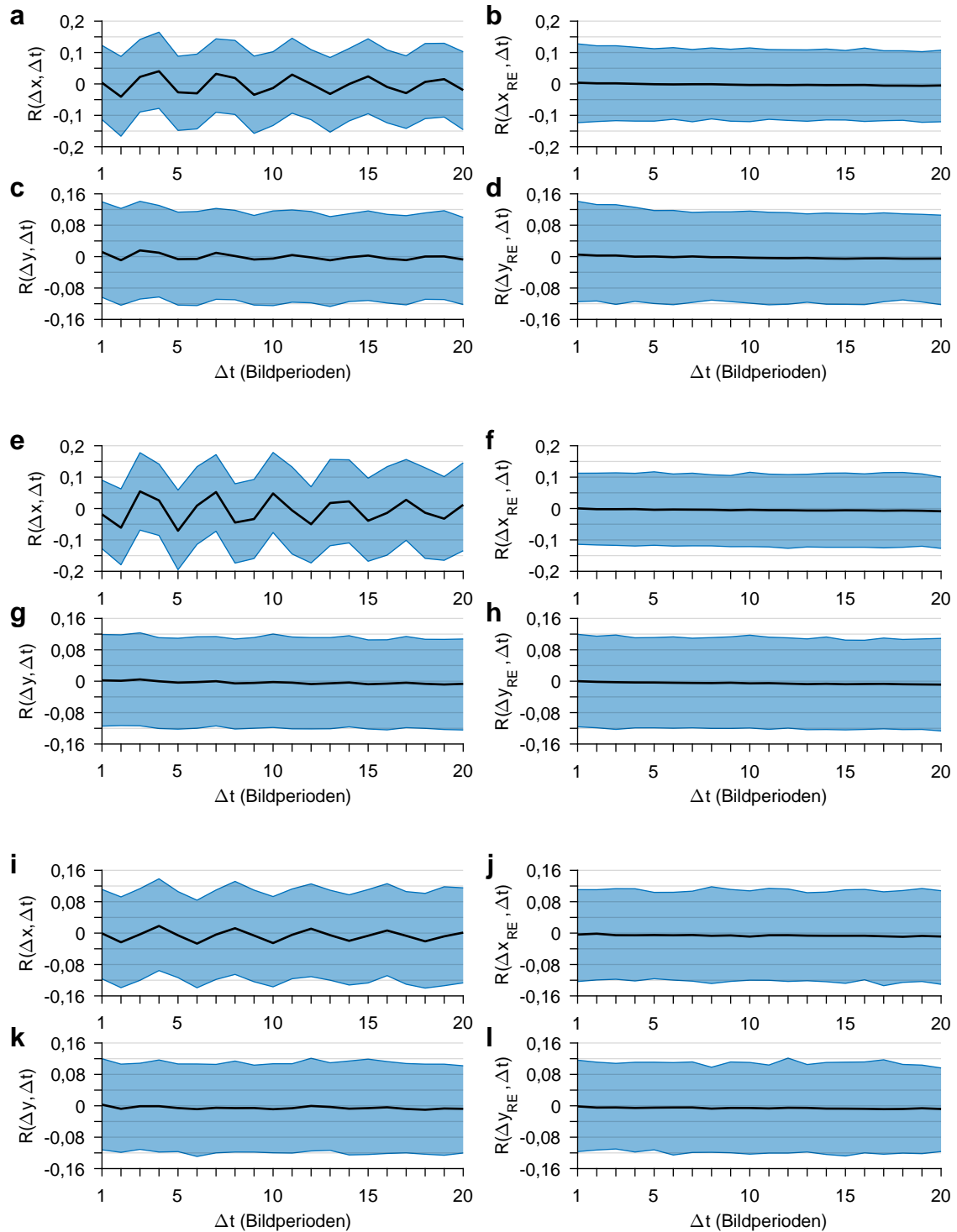


Abbildung 2.12: **Normierte Autokorrelationen der Abfolgen Δx , Δx_{RE} , Δy und Δy_{RE} mobiler Reporter-Einheiten der Referenzen B0-L2, B0-L4 und B0-L6 in vielen je 500 Bildperioden langen Ausschnitten (a-d) Für Referenz B0-L2: (a) $R(\Delta x, \Delta t)$, (b) $R(\Delta x_{RE}, \Delta t)$, (c) $R(\Delta y, \Delta t)$ und (d) $R(\Delta y_{RE}, \Delta t)$ nach Gl. (2.6) für zeitliche Verschiebungen Δt um 1 bis 20 Bildperioden. Die blauen Flächen werden begrenzt durch die 0,005- und 0,995-Quantile der Verteilungen aus allen $N=5772$ Intervallen. Schwarze Kurven: jeweils arithmetisches Mittel der Werte aus allen $N=5772$ Intervallen. (e-h) Wie a-d, für Referenz B0-L4 ($N=6593$). (i-l) Wie a-d, für Referenz B0-L6 ($N=2306$).**

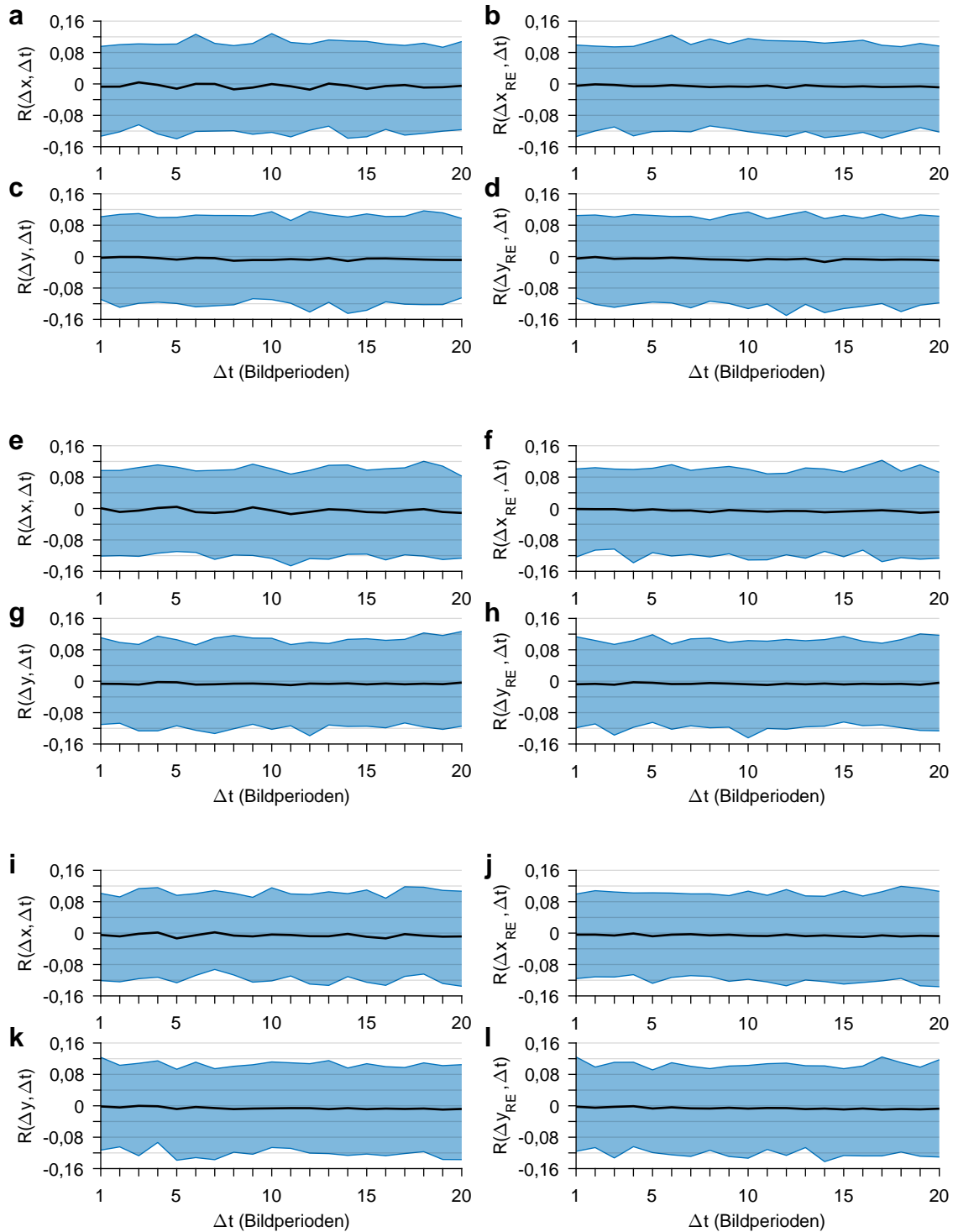


Abbildung 2.13: **Normierte Autokorrelationen der Abfolgen Δx , Δx_{RE} , Δy und Δy_{RE} mobiler Reporter-Einheiten der Referenzen 0-L8, B0-L10 und 0-L12 in vielen je 500 Bildperioden langen Ausschnitten (a-d) Für Referenz 0-L8: (a) $R(\Delta x, \Delta t)$, (b) $R(\Delta x_{RE}, \Delta t)$, (c) $R(\Delta y, \Delta t)$ und (d) $R(\Delta y_{RE}, \Delta t)$ nach Gl. (2.6) für zeitliche Verschiebungen Δt um 1 bis 20 Bildperioden. Die blauen Flächen werden begrenzt durch die 0,005- und 0,995-Quantile der Menge der Werte aus allen $N=503$ Intervallen. Schwarze Kurven: jeweils arithmetisches Mittel der Verteilungen aus allen $N=503$ Intervallen. (e-h) Wie a-d, für Referenz B0-L10 ($N=494$). (i-l) Wie a-d, für Referenz. 0-L12 ($N=506$).**

Tabelle 2.2: **Statistik über die global betrachteten normierten Autokorrelationen $R(\Delta x_{RE}, 1)$ und $R(\Delta y_{RE}, 1)$ der mobilen Reporter-Einheiten aller Referenzobjekte im Vergleich zu simulierten Daten**
 $R(\Delta x_{RE}/\Delta y_{RE}, 1)$: Normierte Autokorrelation in vielen je 500 Bildperioden langen Ausschnitten der vibrationskorrigierten Auslenkungsabfolgen Δx_{RE} bzw. Δy_{RE} nach Gl. (2.6). \bar{R} : Arithmetisches Mittel der Einzelwerte im Datensatz. N: Anzahl der Einzelwerte im Datensatz. Δ_{SIM} : Abfolgen normalverteilter Pseudo-Zufallszahlen mit $\ell=500$ (siehe Text für nähere Erläuterung).

Datensatz		Quantile von $R(\Delta x_{RE}/\Delta y_{RE}, 1)$				\bar{R}	N
		0,005-Q.	0,25-Q.	0,75-Q.	0,995-Q.		
F _d	Δx_{RE}	-0,137	-0,030	0,033	0,098	0,0006	274
F _d	Δy_{RE}	-0,093	-0,027	0,035	0,130	0,0072	274
B0-L2	Δx_{RE}	-0,123	-0,027	0,035	0,128	0,0039	5772
B0-L2	Δy_{RE}	-0,115	-0,027	0,036	0,141	0,0051	5772
B0-L4	Δx_{RE}	-0,114	-0,030	0,032	0,112	0,0005	6593
B0-L4	Δy_{RE}	-0,116	-0,031	0,031	0,119	0,0001	6593
B0-L6	Δx_{RE}	-0,123	-0,034	0,026	0,111	-0,0036	2306
B0-L6	Δy_{RE}	-0,117	-0,032	0,029	0,116	-0,0015	2306
0-L8	Δx_{RE}	-0,134	-0,037	0,028	0,099	-0,0047	503
0-L8	Δy_{RE}	-0,105	-0,033	0,025	0,105	-0,0048	503
B0-L10	Δx_{RE}	-0,124	-0,032	0,031	0,101	-0,0013	494
B0-L10	Δy_{RE}	-0,119	-0,035	0,020	0,113	-0,0078	494
0-L12	Δx_{RE}	-0,116	-0,035	0,023	0,100	-0,0039	506
0-L12	Δy_{RE}	-0,116	-0,032	0,026	0,124	-0,0023	506
SIM	Δ_{SIM}	-0,114	-0,031	0,029	0,115	-0,0006	10000

Form der vibrationskorrigierten Auslenkungsverteilungen der mobilen Reporter-Einheiten

Im Rahmen der statistischen Auswertung der Zwei-Zustands-Abfolgen (siehe Kapitel 3) ist es nützlich, eine Annahme über die Wahrscheinlichkeitsverteilung der Auslenkungen treffen zu können. Für die nicht vibrationskorrigierten Auslenkungsabfolgen Δx und Δy wurde bereits festgestellt, dass ihre Wahrscheinlichkeitsverteilungen eine große Ähnlichkeit zu Normalverteilungen aufweisen. Ausgehend von dieser Beobachtung wurden auch die Verteilungen der vibrationskorrigierten Auslenkungen Δx_{RE} und Δy_{RE} mobiler Reporter-Einheiten der F_d-, B0- und 0-Referenzen stichprobenweise untersucht und mit Normalverteilungen verglichen. Abbildungen 2.14, 2.15 und C.5-C.9 zeigen für die verschiedenen Referenzobjekte Wahrscheinlichkeitsdichteverteilungen in jeweils 1000 Bildperioden langen, zufällig ausgewählten Stichproben aus den vibrationskorrigierten Auslenkungsabfolgen. Dabei sind nur die Verteilungen von Δx_{RE} gezeigt. Ein qualitativ gleiches Bild ergab sich für die entsprechenden Stichproben von Δy_{RE} .

Die Untersuchung lieferte für alle Referenzobjekte dasselbe Ergebnis:

1. Die Form der Verteilungen in jeder Stichprobe der vibrationskorrigierten Auslenkungen entspricht annähernd der Form einer Normalverteilung.
2. Es bestehen teils deutliche Unterschiede in den Breiten der Verteilungen in verschiedenen Stichproben ein- und derselben Referenz, was auf unterschiedliche Signal-zu-Rausch-Verhältnisse in den Zeitintervallen zurückzuführen ist, denen die Auslenkungsstichproben entstammen.

2.5.2 Charakterisierung der Bewegung der mobilen Reporter-Einheit im Zustand mit Schlaufe in der Leine bei proximaler Verankerungsgeometrie

Zur Charakterisierung der Bewegung der mobilen Reporter-Einheiten im Zustand mit Schlaufe in der Leine bei proximaler Verankerungsgeometrie der Interaktionspartner wurden keine Referenzobjekte mit nur einem Beweglichkeitsniveau verwendet. Stattdessen wurden Zeitintervalle in Auslenkungsabfolgen von Objekten mit 9 nt langen Interaktionspartnern (CCAGAGACG:CGTCTCTGG) ausgewählt, während derer die mobilen Reporter-Einheiten eine geringe Beweglichkeit zeigten. Dazu wurde ein Schwellwert ρ für die radiale Auslenkung r definiert. Alle Zeitpunkte einer Auslenkungsabfolge mit $r \geq \rho$ oder einem Abstand von höchstens 50 Bildperioden zum nächsten Zeitpunkt mit $r \geq \rho$ wurden aussortiert. Alle verbleibenden Intervalle der Abfolge mit einer Länge von mindestens 100 Bildperioden gingen in die Charakterisierung der Bewegung der mobilen Reporter-Einheit ein. Abbildung 2.16 veranschaulicht die Vorgehensweise und zeigt ihr Ergebnis für Beispielabfolgen der Referenz i9p-L6 (Konturlänge der Schlaufe $L = 1554$ bp) bzw. i9p-L4 ($L = 1050$ bp).

Bei dieser Vorgehensweise zur Auswahl der Intervalle kann es vorkommen, dass kurze Aufenthalte im Zustand ohne Schlaufe nicht aussortiert werden, sofern das Ausmaß der Bewegung darin nicht viel größer ist als im Zustand mit Schlaufe. Eine dadurch mögliche Verfälschung der Charakterisierung der Bewegung im Zustand mit Schlaufe wird als vernachlässigbar eingestuft.

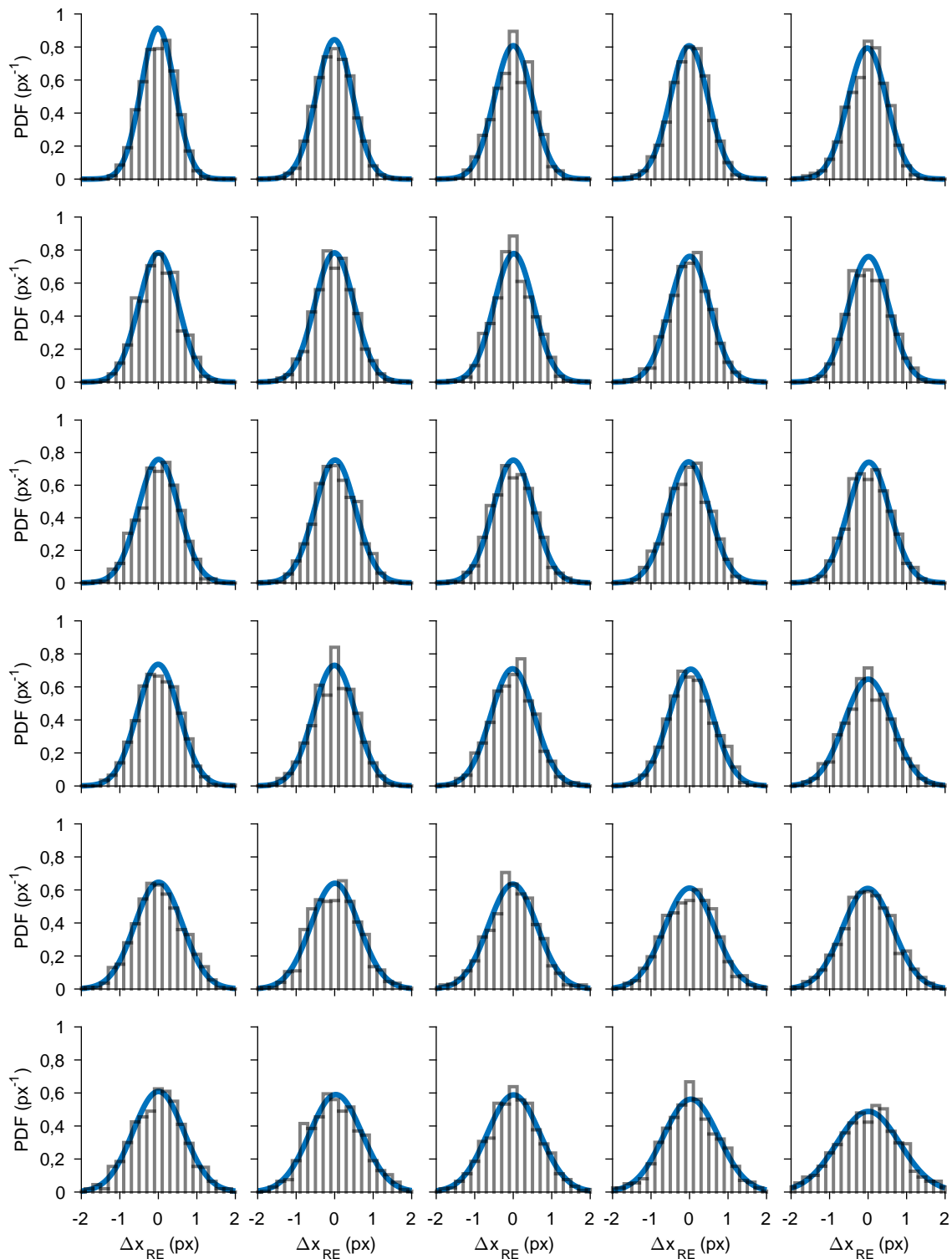


Abbildung 2.14: **Wahrscheinlichkeitsdichteverteilungen von Stichproben aus Δx_{RE} von 30 mobilen Reporter-Einheiten der F_d -Referenz** Histogramme: Aus der relativen Häufigkeit geschätzte Wahrscheinlichkeitsdichte. Blaue Linien: Normalverteilte Wahrscheinlichkeitsdichte nach Gl. (2.5). Die Stichproben sind nach dem Wert der Standardabweichung sortiert. Stichprobenumfang: je 1000 Bildperioden.

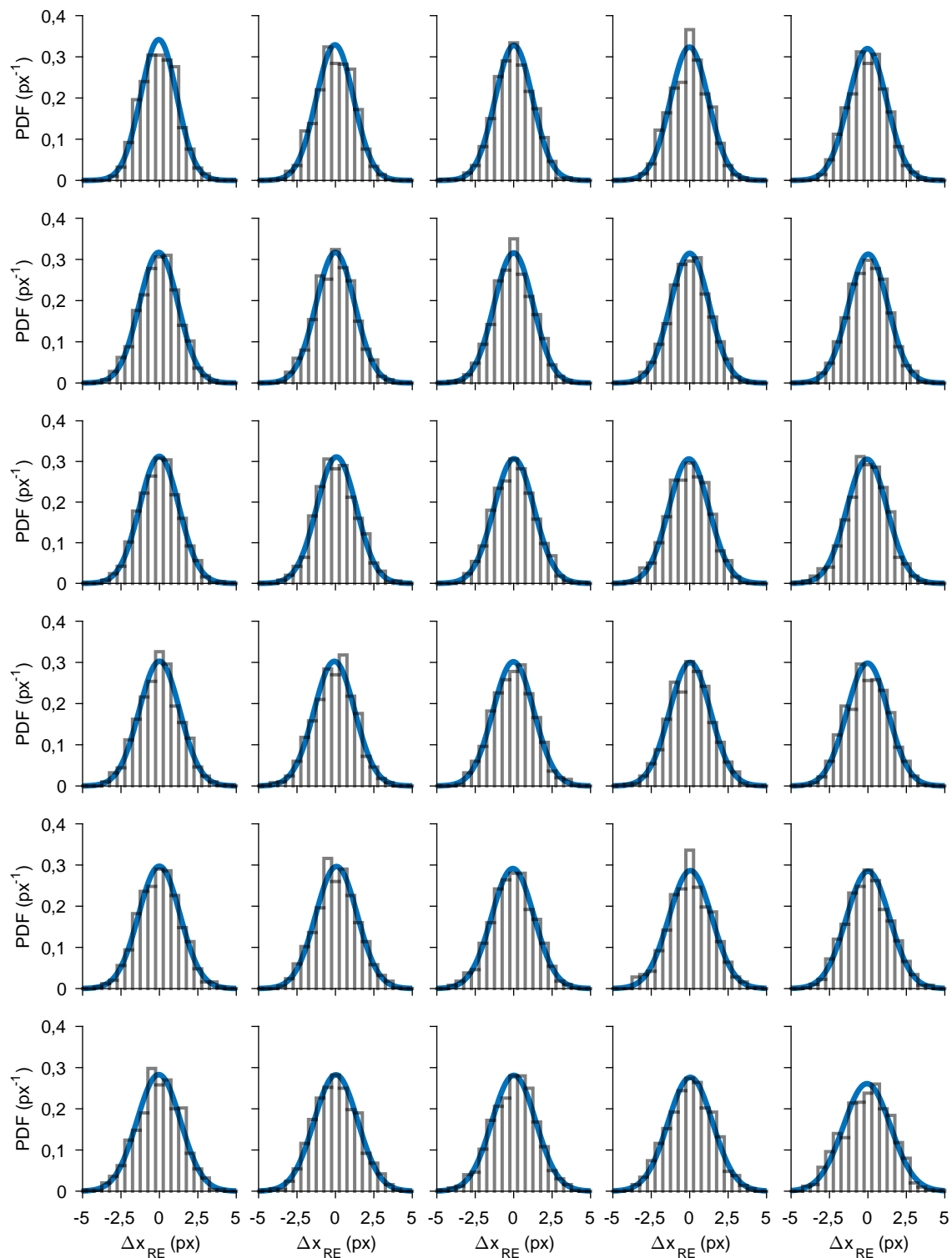


Abbildung 2.15: **Wahrscheinlichkeitsdichteverteilungen von Stichproben aus Δx_{RE} von 30 mobilen Reporter-Einheiten der Referenz 0-L8** Histogramme: Aus der relativen Häufigkeit geschätzte Wahrscheinlichkeitsdichte. Blaue Linien: Normalverteilte Wahrscheinlichkeitsdichte nach Gl. (2.5). Die Stichproben sind nach dem Wert der Standardabweichung sortiert. Stichprobenumfang: je 1000 Bildperioden.

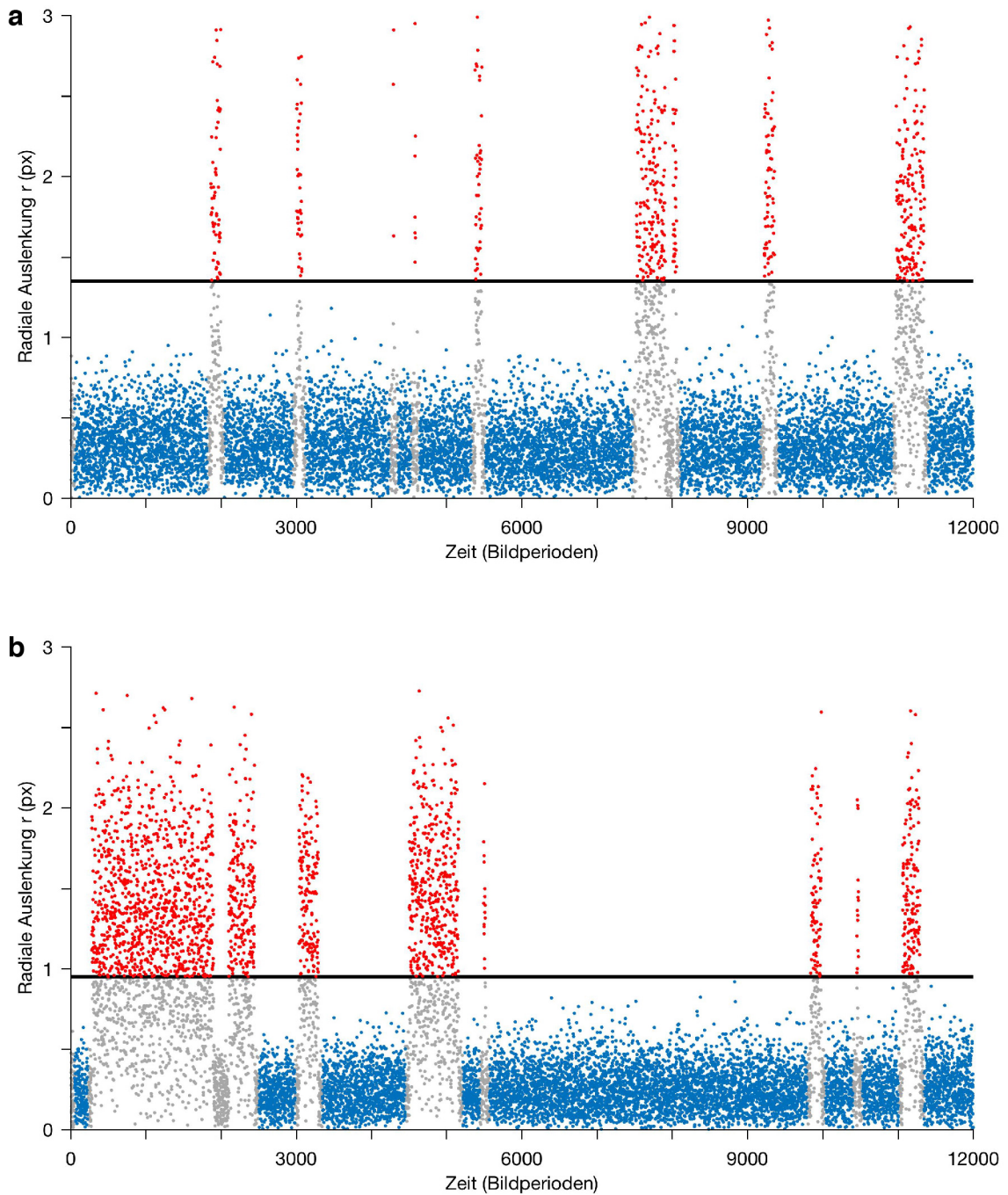


Abbildung 2.16: **Ergebnis der Auswahl der Referenzintervalle für den Zustand mit Schlaufe in der Leine bei proximaler Verankerungsgeometrie der Interaktionspartner** am Beispiel je einer Auslenkungsabfolge von (a) Referenz i9p-L6 und (b) Referenz i9p-L4. Schwarze Linie: jeweiliger Schwellwert p . Punkte: Radiale Auslenkung r über der Zeit in Bildperioden. Rot: Zeitpunkte mit $r \geq p$. Grau: Zeitpunkte mit $r < p$, die wegen ihrer Nähe zu Zeitpunkten mit $r \geq p$ aussortiert wurden. Blau: Intervalle der jeweiligen Abfolge, die in die Charakterisierung der Bewegung mit eingingen.

Autokorrelation der Auslenkungsabfolgen der mobilen Reporter-Einheiten

Wie bei den Referenzobjekten mit fester Schlaufe und ohne Schlaufe in der Leine wurde auch für die Referenzen i9-L6 und i9-L4 eine Periodizität in den Verläufen der normierten Autokorrelation der Auslenkungsabfolgen Δx und Δy der mobilen Reporter-Einheiten festgestellt (Abb. 2.17 und Abb. C.11, jeweils a und c).

Auch hier wurde die Methode zur Vibrationskorrektur durch Subtraktion der geschätzten Schwerpunktauslenkung angewandt. In den Verläufen der Autokorrelation der vibrationskorrigierten Auslenkungsabfolgen $R(\Delta x_{RE}, \Delta t)$ und $R(\Delta y_{RE}, \Delta t)$ ist keine ausgeprägte Periodizität erkennbar (Abb. 2.17 und Abb. C.11, jeweils b und d). Allerdings ist die normierte Autokorrelation der Abfolgen Δx_{RE} und Δy_{RE} bei $\Delta t = 1$ (Tab. 2.3) leicht höher als bei den F_d -, B0- und 0-Referenzen (vgl. Tab. 2.2). Dies gilt sowohl für die Mittelwerte als auch für die 0,005- und 0,995-Quantile der Verteilung der Werte von $R(\Delta x_{RE}, \Delta t)$ und $R(\Delta y_{RE}, \Delta t)$.

Im Zustand mit Schlaufe in der Leine bei proximaler Verankerungsgeometrie der Interaktionspartner befindet sich die mobile Reporter-Einheit nahe an der immobilen Reporter-Einheit und der Probenoberfläche. Eine mögliche Erklärung für die leicht erhöhte normierte Autokorrelation bei kleinen zeitlichen Verschiebungen könnten Wechselwirkungen zwischen der mobilen Reporter-Einheit und der immobilen Reporter-

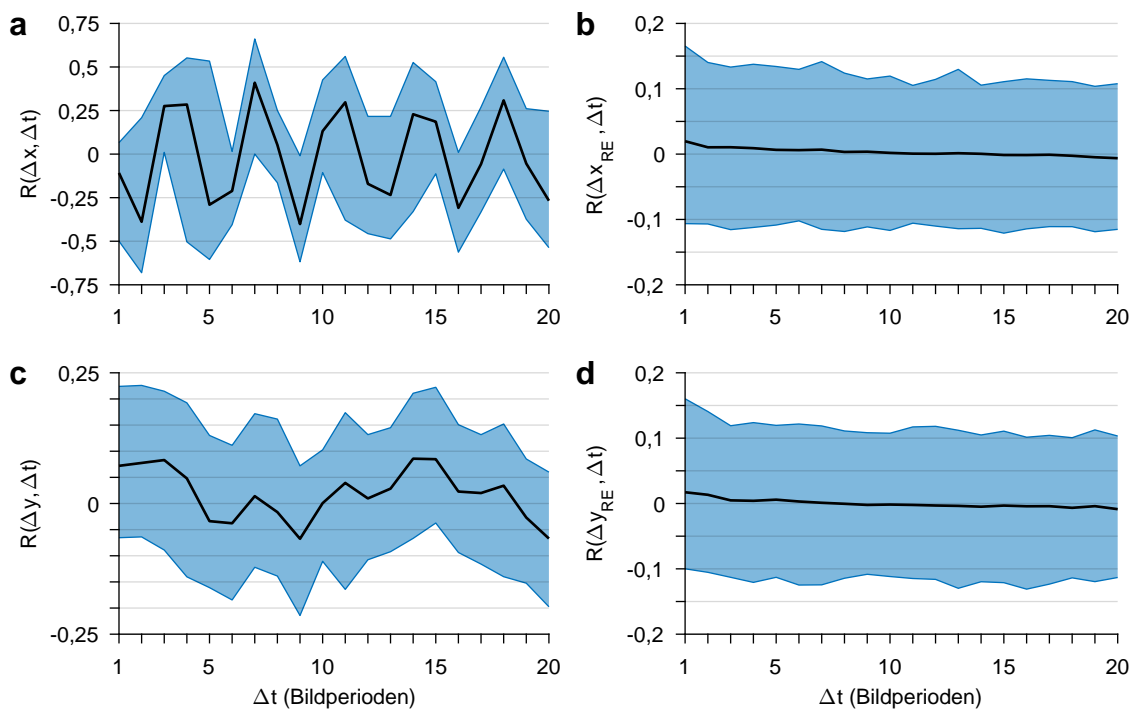


Abbildung 2.17: **Normierte Autokorrelationen der Abfolgen Δx , Δx_{RE} , Δy und Δy_{RE} mobiler Reporter-Einheiten der Referenz i9p-L6 in vielen ($N = 1396$) jeweils 500 Bildperioden langen Ausschnitten** (a) $R(\Delta x, \Delta t)$, (b) $R(\Delta x_{RE}, \Delta t)$, (c) $R(\Delta y, \Delta t)$ und (d) $R(\Delta y_{RE}, \Delta t)$ nach Gl. (2.6) für zeitliche Verschiebungen Δt um 1 bis 20 Bildperioden. Die blauen Flächen werden begrenzt durch die 0,005- und 0,995-Quantile der Verteilungen aus allen 1396 Intervallen. Schwarze Kurven: jeweils arithmetisches Mittel der Werte aus allen 1396 Intervallen.

Einheit bzw. der Probenoberfläche sein. Gelegentliches Festkleben oder Verheddern der mobilen Reporter-Einheit an einer Stelle für eine Dauer von wenigen Bildperioden könnte im Mittel zu einer höheren registrierten Korrelation der Bewegung der Reporter-Einheit auf kurzen Zeitskalen führen.

Nichtsdestotrotz sind $R(\Delta x_{RE}, 1)$ und $R(\Delta y_{RE}, 1)$ auch im Fall der Referenzen i9p-L6 und i9p-L4 gering. Die Breite der Wertebereiche unterscheidet sich nicht drastisch vom Wertebereich der Vergleichswerte aus dem Datensatz von 10000 simulierten normalverteilten Abfolgen mit $\ell = 500$ (letzte Zeile in Tab. 2.3, zur Erklärung siehe Abschnitt 2.5.1). Für den Zweck der statistischen Auswertung der Zwei-Zustands-Abfolgen (siehe Kapitel 3) wird die Forderung nach zeitlicher Unkorreliertheit der Auslenkungsabfolgen als erfüllt angesehen.

Tabelle 2.3: **Statistik über die global betrachteten normierten Autokorrelationen $R(\Delta x_{RE}, 1)$ und $R(\Delta y_{RE}, 1)$ der mobilen Reporter-Einheiten von Ref. i9p-L6 und i9p-L4 im Vergleich zu $R(\Delta x_{SIM}, 1)$**
 $R(\Delta x_{RE}/\Delta y_{RE}, 1)$: Normierte Autokorrelation in vielen je 500 Bildperioden langen Ausschnitten der vibrationskorrigierten Auslenkungsabfolgen Δx_{RE} bzw. Δy_{RE} nach Gl. (2.6). \bar{R} : Arithmetisches Mittel der Einzelwerte im Datensatz. N: Anzahl der Einzelwerte im Datensatz. Δ_{SIM} : Abfolgen normalverteilter Pseudo-Zufallszahlen mit $\ell=500$ (siehe Text in Abschnitt 2.5.1 für nähere Erläuterung).

Datensatz		Quantile von $R(\Delta x_{RE}/\Delta y_{RE}, 1)$				\bar{R}	N
		0,005-Q.	0,25-Q.	0,75-Q.	0,995-Q.		
i9p-L6	Δx_{RE}	-0,106	-0,013	0,051	0,165	0,0197	1396
i9p-L6	Δy_{RE}	-0,100	-0,014	0,048	0,160	0,0172	1396
i9p-L4	Δx_{RE}	-0,109	-0,013	0,052	0,171	0,0219	1154
i9p-L4	Δy_{RE}	-0,105	-0,012	0,056	0,208	0,0239	1154
SIM	Δ_{SIM}	-0,114	-0,031	0,029	0,115	-0,0006	10000

Breite und Form der Auslenkungsverteilungen der mobilen Reporter-Einheiten

Auch im Zustand mit Schlaufe in der Leine bei proximaler Verankerungsgeometrie der Interaktionspartner wurde beobachtet, dass mit abnehmender Helligkeit der einzelnen Reporter-Einheiten die jeweilige Breite ihrer Auslenkungsabfolgen zunahm. Die Form der Verteilungen in mehreren Stichproben aus den vibrationskorrigierten Auslenkungsabfolgen Δx_{RE} entspricht annähernd der Form einer Normalverteilung (Abb. 2.18 und Abb. C.10). Dasselbe gilt für die entsprechenden Stichproben von Δy_{RE} .

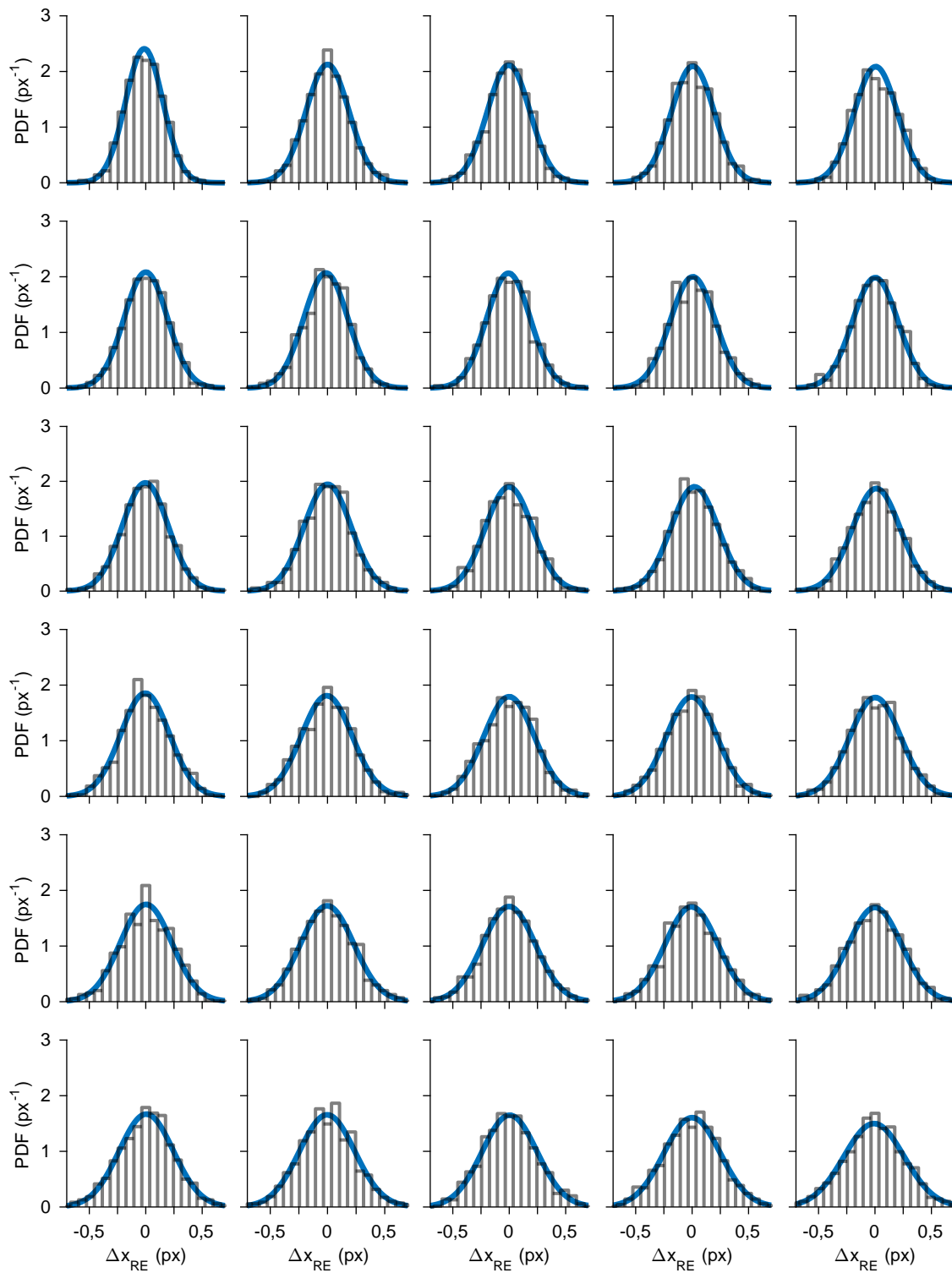


Abbildung 2.18: **Wahrscheinlichkeitsdichteverteilungen von Stichproben aus Δx_{RE} von 30 mobilen Reporter-Einheiten der Referenz i9p-L6** Histogramme: Aus der relativen Häufigkeit geschätzte Wahrscheinlichkeitsdichte. Blaue Linien: Normalverteilte Wahrscheinlichkeitsdichte nach Gl. (2.5). Die Stichproben sind nach dem Wert der Standardabweichung sortiert. Stichprobenumfang: je 1000 Bildperioden.

2.6 Kapitelzusammenfassung

Aus den TMFM-Aufnahmen wurden per Partikel-Tracking Abfolgen der Positionen der Reporter-Einheiten im Verlauf der einzelnen TMFM-Messungen extrahiert. Die gemessene Bewegung der Reporter-Einheiten ist eine Überlagerung der Bewegung der Reporter-Einheiten im Bezugssystem der Probenkammer mit der Relativbewegung zwischen der Probenkammer und dem Sensor der Kamera. Mit Korrekturen der Positionsabfolgen um Drift und Vibration der Messapparatur erhält man Abfolgen der geschätzten Auslenkungen der Reporter-Einheiten um die Ankerpunkte der DNA-Nano-Objekte an der Probenoberfläche.

Der Bewegungsradius einer mobilen Reporter-Einheit hängt davon ab, ob sich das Partikel-Leinen-System im Zustand ohne Schlaufe in der Leine oder im Zustand mit Schlaufe in der Leine befindet. In den Auslenkungsabfolgen der mobilen Reporter-Einheiten von Objekten mit Interaktionspartnern wechseln sich Intervalle höherer und niedrigerer Beweglichkeit ab. Die momentane Beweglichkeit einer mobilen Reporter-Einheit ist ein Indikator für den momentanen Bindezustand der Interaktionspartner.

Neben den Positions- bzw. Auslenkungsabfolgen wurde für jede Reporter-Einheit auch die zeitliche Abfolge der summierten Helligkeit einer Pixelregion fester Größe um die Position der Reporter-Einheit erfasst. Es besteht ein Zusammenhang zwischen der Helligkeit und der gemessenen Beweglichkeit der Reporter-Einheiten: Mit abnehmender Helligkeit einer Reporter-Einheit verringert sich das Signal-zu-Rausch-Verhältnis und damit auch die Genauigkeit der Positionsabschätzung. Die Breiten der Auslenkungsverteilungen der Reporter-Einheiten nehmen deshalb allgemein über den Verlauf einer Messung zu. Dies gilt sowohl für den Zustand niedrigerer Beweglichkeit, als auch für den Zustand höherer Beweglichkeit.

Die Bewegung der mobilen Reporter-Einheiten in den Zuständen mit Schlaufe in der Leine und ohne Schlaufe in der Leine wurde für verschiedene Leinenlängen und für beide Verankerungsgeometrien charakterisiert. Insbesondere wurde dabei die zeitliche Korrelation der Auslenkungsabfolgen und die Form der Auslenkungsverteilungen betrachtet. Die vibrationskorrigierten Auslenkungsabfolgen $(\Delta x_{RE}, \Delta y_{RE})$ der mobilen Reporter-Einheiten können als zeitlich unkorreliert angesehen werden. Die Verteilungen der vibrationskorrigierten Auslenkungen in Stichproben aus den Auslenkungsabfolgen der mobilen Reporter-Einheiten weisen eine große Ähnlichkeit zu Normalverteilungen auf.

3 Statistische Auswertung von TMFM-Daten mit zwei Beweglichkeitszuständen

Es ist nicht möglich, aus den TMFM-Messdaten den Zustand eines Interaktionspartnerpaares zu einem beliebigen Zeitpunkt direkt auszulesen. Aus dem Ausmaß der Bewegung der mobilen Reporter-Einheit kann allerdings indirekt darauf geschlossen werden, in welchem Zustand das Interaktionspartnerpaar vorliegt, da je nach Zustand des Interaktionspartnerpaares eine Schlaufe in der DNA-Leine ausgeformt ist oder nicht. Die Problemstellung lautet folglich, aus der zeitlichen Abfolge von Intervallen unterschiedlich hoher Beweglichkeit der mobilen Reporter-Einheit auf die zeitliche Abfolge von Aufenthalten des Interaktionspartnerpaares in den Zuständen „B+B’“ oder „B:B’“ rückzuschließen. Zur Lösung dieser Aufgabe wurde die Methode der *Hidden-Markov-Modellierung* [105–108] angewandt.

Abschnitt 3.1 erläutert das Prinzip der Hidden-Markov-Modellierung. Anschließend beschreibt Abschnitt 3.2 das zur Analyse der TMFM-Daten verwendete Hidden-Markov-Modell (HMM). Abschnitt 3.3 lenkt schließlich den Blick auf die Verweildaueranalyse und die darin vorgenommenen Korrekturen der experimentellen Daten, insbesondere den Umgang mit dem Auftreten heterogener Hybridisierungskinetik in der Probe.

3.1 Hidden-Markov-Modellierung

Ein Hidden-Markov-Modell [105–108] beschreibt ein System, welches im Verlauf der Zeit stochastisch zwischen einer bestimmten Anzahl von definierten, für den Beobachter jedoch verborgenen Zuständen hin- und her wechselt. Bezeichne $\zeta(t)$ den verborgenen Zustand des Systems zum Zeitpunkt t . Eine fundamentale Annahme des Modells besteht darin, dass die zeitliche Abfolge von Aufenthalten in den verschiedenen Zuständen eine Markov-Kette erster Ordnung ist: Die Wahrscheinlichkeit, dass das System zum Zeitpunkt $t + 1$ in einem bestimmten Zustand verweilen wird, hängt einzig und allein vom Zustand des Systems zum Zeitpunkt t ab und nicht vom Zustand des Systems zu früheren Zeitpunkten $0, \dots, t - 1$. Im mathematischen Formalismus des Modells äußert sich diese Annahme in einer zeitlich konstanten Übergangsmatrix \mathcal{T} . Für ein System mit M möglichen Zuständen ist \mathcal{T} eine $(M \times M)$ -Matrix. Dementsprechend ist \mathcal{T} im Fall eines Zwei-Zustands-Systems eine (2×2) -Matrix:

$$\mathcal{T} = \begin{pmatrix} \mathcal{T}_{11} & \mathcal{T}_{12} \\ \mathcal{T}_{21} & \mathcal{T}_{22} \end{pmatrix} \quad (3.1)$$

Der Matrix-Eintrag \mathcal{T}_{ij} gibt die Wahrscheinlichkeit für einen Übergang des Systems vom Zustand i zum Zeitpunkt t in den Zustand j zum Zeitpunkt $t + 1$ an:

$$\mathcal{T}_{ij} = p(\zeta(t+1) = j \mid \zeta(t) = i) \quad (3.2)$$

Der für einen äußeren Beobachter sichtbare Teil des Systems wird als Observable bezeichnet. Die Observable ist eine Zufallsvariable, deren Werte in jedem Zustand nach einer für diesen Zustand charakteristischen Wahrscheinlichkeitsverteilung verteilt sind. Man spricht in diesem Kontext auch von „Emissionswahrscheinlichkeiten“ der Observablen. Im Zustand i des Systems könnte die Wahrscheinlichkeitsverteilung für die Observable beispielsweise eine Normalverteilung mit Erwartungswert μ_i und Standardabweichung σ_i sein. Verweilt das System zum Zeitpunkt t im Zustand i , so gilt für die Emissionswahrscheinlichkeit eines bestimmten Werts $o(t)$ der Observablen:

$$\varphi(o(t) \mid \zeta(t) = i) = \frac{1}{\sigma_i \sqrt{2\pi}} \cdot \exp \left\{ -\frac{(o(t) - \mu_i)^2}{2\sigma_i^2} \right\} \quad (3.3)$$

Mithilfe eines HMM ist es möglich, zu einer gegebenen Abfolge der Observablen $o(1), \dots, o(T)$ jene Werte für die Modellparameter zu ermitteln, welche die Wahrscheinlichkeit des Auftretens genau dieser Observablenabfolge maximieren. In diesem Fall spricht man auch von einem *Maximum Likelihood Hidden Markov Model* (MLHMM) [107]. Am Anfang der MLHMM-Analyse werden die Modellparameter, d. h. die Übergangs- und Emissionswahrscheinlichkeiten sowie die Startbedingungen des Systems, grob geschätzt. Danach folgt eine iterative Prozedur, die in jeder Iteration vereinfacht dargestellt zwei Schritte beinhaltet:

1. Berechnung der Aufenthaltswahrscheinlichkeiten des Systems in den verschiedenen Zuständen für jeden Zeitpunkt, basierend auf der aktuellen Schätzung der Modellparameter.
2. Erneute Schätzung der Modellparameter, basierend auf den im ersten Schritt berechneten Wahrscheinlichkeiten.

Im Idealfall konvergiert diese Prozedur zu einer stabilen Lösung für die Modellparameter. Abschließend kann mit dem sog. Viterbi-Algorithmus basierend auf den finalen Modellparametern die wahrscheinlichste Zustandsabfolge $z(1), \dots, z(T)$ zur Observablenabfolge $o(1), \dots, o(T)$ bestimmt werden.

Für eine genauere Behandlung von Grundlagen zur Hidden-Markov-Modellierung und zu den bei der Umsetzung verwendeten Algorithmen sei auf [105–108] verwiesen.

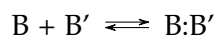
3.2 Das Hidden-Markov-Modell für die TMFM-Daten

Die Implementierung der HMM-Analyse erfolgte in MATLAB. Dazu wurden Teile des Funktionspakets von Chodera et al. [107] übernommen und für die Auswertung der TMFM-Daten angepasst. Die maßgeblichen selbst eingebrachten Veränderungen und Neuerungen im Programmcode betreffen die Berechnung der Emissionswahrschein-

lichkeiten (vgl. Abschnitt 3.2.2). Der Programmcode ist über das GitHub-Repositorium <https://github.com/mschickinger/TMFM> frei zugänglich.

3.2.1 Definition der Zustände und dabei getroffene Annahmen

Die Anschauung der Hybridisierung zweier DNA-Einzelstränge zu einem Duplex



als ein Zwei-Zustands-System ist lediglich eine Vereinfachung der wahren Gegebenheiten [36,42,109]. Sowohl der Zustand der dissoziierten Einzelstränge B und B', als auch der Zustand des Duplex B:B' spalten sich in unterschiedliche Mikrozustände auf. Einerseits stellen sich in den dissoziierten Einzelsträngen energetisch distinkte Zustände durch die Bildung von intra-Strang-Sekundärstrukturen ein [42]. Andererseits können sich in einem ausgeformten DNA-Duplex einzelne Basenpaare öffnen und schließen, ohne dass es dadurch sofort zur vollständigen Dissoziation der Einzelstränge kommt.

Der Übergang zwischen den zwei Extremen „vollständige Dissoziation“ und „vollständige Basenpaarung im Duplex“ vollzieht sich nicht auf einmal, sondern vielmehr über eine Reihe von Zwischenzuständen. Das *Nucleation-Zippering*-Modell der DNA-Hybridisierung [109] unterteilt die Ausformung eines stabilen Duplex in zwei Phasen: Am Anfang formiert sich ein Hybridisierungsnukleus aus wenigen Basenpaaren. Ist dieser Nukleus ausreichend stabil, kann eine Vervollständigung des Duplex durch sukzessive Bildung weiterer Basenpaare (Zippering) stattfinden. Als ratenlimitierender Schritt wird dabei die Nukleation angesehen, da das Zippering sehr viel schneller als die Nukleation vonstattengeht. Es kann jedoch auch der Fall eintreten, dass die Lebensdauer eines Nukleus zu kurz ist und sich die Einzelstränge wieder voneinander lösen, bevor sich das Zippering vollzieht [36]. Gemäß dem Prinzip des detaillierten Gleichgewichts erfolgt die Dissoziation eines vollständigen Duplex, indem sich zunächst einzelne Basenpaare öffnen, bis zu wenige intakte Basenpaarungen vorhanden sind, um eine stabile Bindung zwischen den Einzelsträngen aufrechtzuerhalten [36].

Die Observable in den TMFM-Daten ist die geschätzte xy-Auslenkung $(\Delta x_{RE}, \Delta y_{RE})$ der mobilen Reporter-Einheit um den Ankerpunkt des DNA-Nano-Objekts an der Probenoberfläche. Die Wahrscheinlichkeitsverteilung von $(\Delta x_{RE}, \Delta y_{RE})$ richtet sich danach, ob eine Schlaufe in der Leine ausgeformt ist oder nicht. Es wird die Annahme getroffen, dass im Zustand mit Schlaufe in der Leine die Wahrscheinlichkeitsverteilung von $(\Delta x_{RE}, \Delta y_{RE})$ unabhängig ist vom Bindungszustand einzelner Basenpaare im Duplex B:B', solange es nicht zur vollständigen Dissoziation des Duplex kommt. Ebenso wird die Annahme getroffen, dass im Zustand ohne Schlaufe in der Leine die Wahrscheinlichkeitsverteilung von $(\Delta x_{RE}, \Delta y_{RE})$ unabhängig ist von der Bildung und Lösung von intra-Strang-Sekundärstrukturen in den Einzelsträngen B und B'.

Die zwei verborgenen Zustände des beobachteten Systems werden somit definiert als der Zustand mit Schlaufe in der DNA-Leine (Zustand 1) und der Zustand ohne Schlaufe in der DNA-Leine (Zustand 2). Übertragen auf das Interaktionspartnerpaar repräsentiert Zustand 1 die Situation „teilweise oder vollständig **hybridisiert**“ und Zustand 2

die Situation „vollständig **dissoziiert**“. Deshalb werden die zwei Zustände in dieser Arbeit auch als der „hybridisierte Zustand“ und der „dissoziierte Zustand“ bezeichnet. Um die Voraussetzung einer zeitlich konstanten Übergangsmatrix \mathcal{T} zu erfüllen, wird dabei die Annahme getroffen, dass in jedem dieser zwei Zustände die jeweiligen intramolekularen Umordnungsprozesse auf einer sehr viel schnelleren Zeitskala als die Bildrate der TMFM-Aufnahmen (10 Bilder pro Sekunde) stattfinden. Außerdem muss die Wahrscheinlichkeit der Kolo-kalisation der beiden im Partikel-Leinen-System verankerten Interaktionspartner B und B' während Aufenthalt im dissoziierten Zustand auf der Zeitskala der Beobachtung konstant sein. Basierend auf den in Abschnitt 2.5 gezeigten Ergebnissen kann die mittlere Geschwindigkeit der Bewegung der Interaktionspartner als deutlich schneller als die Bildrate der TMFM-Aufnahme angenommen werden. Folglich kann auch ein konstanter Wert der Kolo-kalisationswahrscheinlichkeit von B und B' in jeder Bildperiode angenommen werden.

3.2.2 Modellierung der Emissionswahrscheinlichkeiten

Eine Erkenntnis der Charakterisierung der Bewegung der mobilen Reporter-Einheiten (Abschnitt 2.5) lag darin, dass die Formen der Verteilungen von Δx_{RE} und Δy_{RE} jeweils gut mit Normalverteilungen übereinstimmen. Darauf basierend wurde die Emissionswahrscheinlichkeit eines bestimmten Werts von $(\Delta x_{RE}, \Delta y_{RE})$ zum Zeitpunkt t im Zustand i durch eine bivariate Normalverteilung mit Erwartungswerten $\mu_{x,i}, \mu_{y,i}$ und Standardabweichungen $\sigma_{x,i}, \sigma_{y,i}$ modelliert:

$$\begin{aligned} \varphi((\Delta x_{RE}(t), \Delta y_{RE}(t)) | \zeta(t) = i) &= \\ &= \frac{1}{2\pi\sigma_{x,i}\sigma_{y,i}} \cdot \exp \left\{ -\frac{1}{2} \left(\frac{(\Delta x_{RE}(t) - \mu_{x,i})^2}{\sigma_{x,i}^2} + \frac{(\Delta y_{RE}(t) - \mu_{y,i})^2}{\sigma_{y,i}^2} \right) \right\} \end{aligned} \quad (3.4)$$

Aufgrund der Symmetrie des Systems gilt für die Erwartungswerte in beiden Zuständen

$$\mu_{x,i} = \mu_{y,i} = 0 \quad , \quad i \in \{1, 2\} \quad (3.5)$$

Demnach besteht kein systematischer Unterschied zwischen den Erwartungswerten in Zustand 1 und Zustand 2. Das alleinige Unterscheidungskriterium zwischen den Verteilungen der Koordinaten in den zwei verschiedenen Zuständen ist folglich die Standardabweichung der Verteilung, die in Zustand 2 größer ist als in Zustand 1. Hidden-Markov-Modellierung wird in der Biophysik normalerweise zur Analyse von Systemen verwendet, in dessen Zuständen die Observable unterschiedliche Erwartungswerte hat. Die Zustände stellen beispielsweise verschiedene Konformationen eines Proteinmoleküls dar. Die dazugehörige Observable ist dann für gewöhnlich der mit einer optischen Falle gemessene End-zu-End-Abstand des Moleküls oder die FRET-Effizienz zwischen zwei Fluorophoren an verschiedenen Punkten der Proteinkette [108]. Dass die einzigen zwei Zustände eines Systems in einem HMM mit demselben Erwartungswert modelliert werden, ist im Kontext biophysikalischer Einzelmolekül-Studien unüblich. Daher wurde die hier verwendete HMM-Analyse zunächst an simulierten Daten erprobt und die Verlässlichkeit der Analyse charakterisiert (siehe Anhang B.2).

In einem klassischen HMM sind die Verteilungen der Emissionswahrscheinlichkeiten der Observablen zeitlich konstant. Die Forderung der zeitlichen Konstanz impliziert insbesondere die Unabhängigkeit der Wahrscheinlichkeitsverteilung der Observablen zum Zeitpunkt t vom Wert der Observablen zu früheren Zeitpunkten. Die Observablenabfolgen müssen demnach zeitlich unkorreliert sein, damit eine klassische HMM-Analyse anwendbar ist. Die Charakterisierung der Bewegung der mobilen Reporter-Einheiten (Abschnitt 2.5) ergab, dass die Abfolgen von $(\Delta x_{RE}, \Delta y_{RE})$ als zeitlich unkorreliert betrachtet werden können. Dieser Aspekt der Forderung nach zeitlicher Konstanz der Emissionswahrscheinlichkeiten ist somit erfüllt. Allerdings zeigte dieselbe Charakterisierung, dass mit der Abnahme der Helligkeit einer Reporter-Einheit eine Verbreiterung der Auslenkungsverteilungen dieser Reporter-Einheit in beiden Zuständen einhergeht. Folglich unterscheiden sich die Breiten der Emissionswahrscheinlichkeitsverteilungen von $(\Delta x_{RE}, \Delta y_{RE})$ mitunter in verschiedenen Abschnitten der Messung. Die folgenden Absätze sowie Abbildungen 3.1 und 3.2 zeigen, wie dieser Umstand bei der Implementierung des HMM berücksichtigt wurde.

Segmentweise Betrachtung der Emissionswahrscheinlichkeiten

Die einzelnen Abfolgen von $(\Delta x_{RE}, \Delta y_{RE})$ der mobilen Reporter-Einheiten wurden in bis zu drei aufeinanderfolgende Segmente eingeteilt. Diese Einteilung erfolgte anhand der verschiedenen Wertebereiche der dazugehörigen Abfolge der Median-gefilterten Helligkeit $I_{med,10}$ der mobilen Reporter-Einheit (Abb. 3.1 und 3.2). In jeder Iteration der MLHMM-Prozedur erfolgte für jedes der Segmente eine individuelle Schätzung der Emissionsparameter $\mu_{x,i}, \mu_{y,i}, \sigma_{x,i}, \sigma_{y,i}, i \in \{1, 2\}$. Dementsprechend gingen für jeden Zeitpunkt t in die Berechnung von $\varphi((\Delta x_{RE}(t), \Delta y_{RE}(t)) | \zeta(t) = i)$ nach Gl. (3.4) die Emissionsparameter des Segments ein, in dem der Zeitpunkt t liegt. Die Verteilungen von φ werden innerhalb der einzelnen Segmente als in erster Näherung zeitlich konstant betrachtet.

Zur Einteilung der Segmente wurden je Probe drei Schwellwerte für die Helligkeit der mobilen Reporter-Einheiten $g_1 > g_2 > g_3$ definiert. Das Segment ε_1 umfasst alle Zeitpunkte, zu denen $I_{med,10}$ den Schwellwert g_1 noch nicht dauerhaft unterschritten hat. Das Segment ε_2 umfasst alle Zeitpunkte, zu denen $I_{med,10}$ den Schwellwert g_1 schon dauerhaft unterschritten hat und den Schwellwert g_2 noch nicht dauerhaft unterschritten hat. Das Segment ε_3 ist analog zum Segment ε_2 definiert. Nach dieser Definition liegen Zeitpunkte mit $I_{med,10}(t) < g_k$ noch im Segment ε_k , wenn $I_{med,10}$ zu mindestens einem späteren Zeitpunkt einen größeren Wert als g_k annimmt. Fällt die Median-gefilterte Helligkeit $I_{med,10}$ einer Reporter-Einheit im Verlauf der Messung dauerhaft unter den Wert g_3 , so werden die Auslenkungsabfolgen dieser Reporter-Einheit nur bis zum letzten Zeitpunkt mit $I_{med,10}(t) > g_3$ einer HMM-Analyse unterzogen. Dieser Zeitpunkt entspricht dem Endpunkt des Segments ε_3 . Zeitpunkte mit $I_{med,10}(t) < g_3$ gehen folglich nur dann in die Analyse mit ein, wenn $I_{med,10}$ zu mindestens einem späteren Zeitpunkt einen größeren Wert als g_3 annimmt.

Abbildungen 3.1 und 3.2 zeigen verschiedene Beispiele der Aufteilung in die Segmente. Bei der verfolgten Vorgehensweise kann der Fall eintreten, dass für die Abfolgen

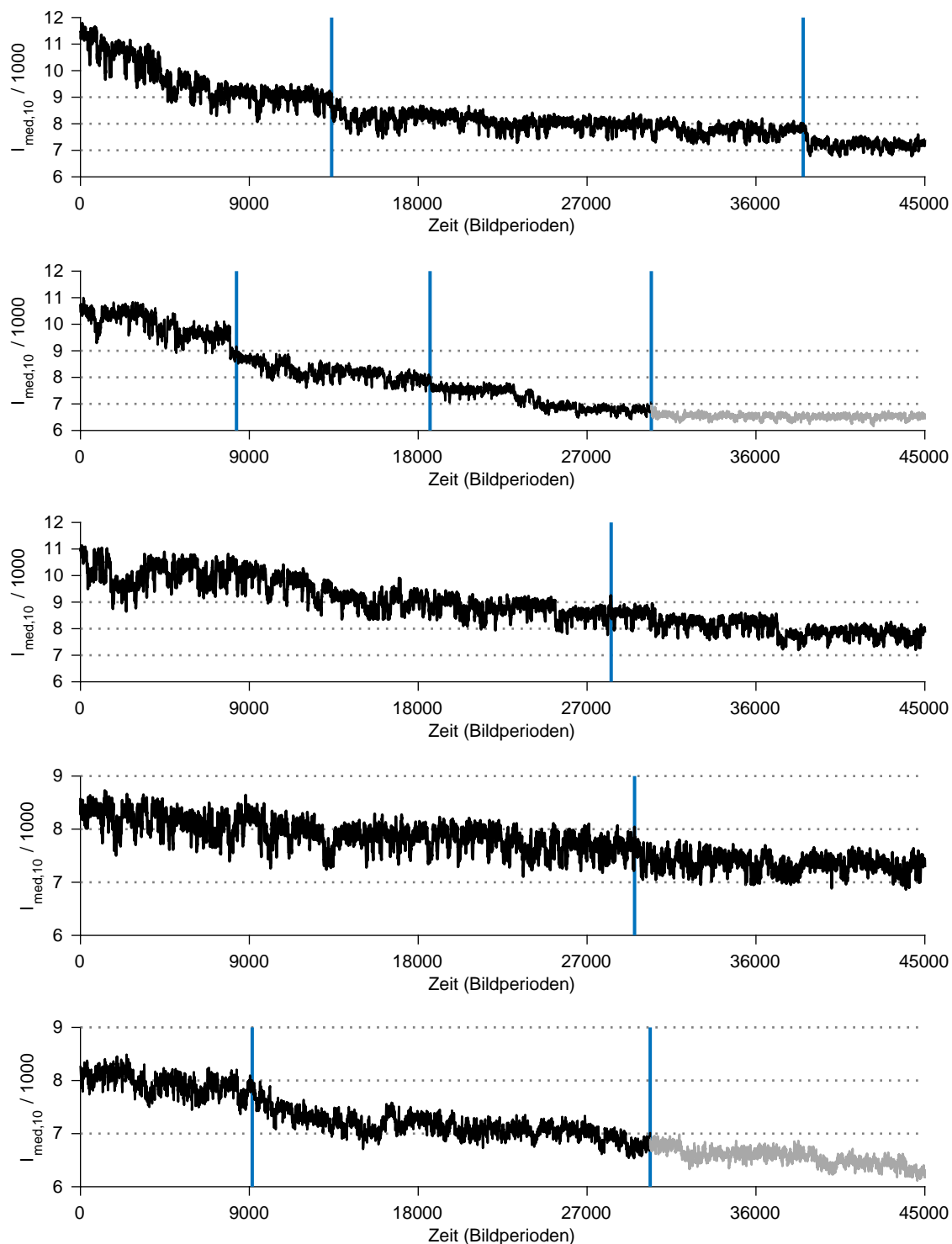


Abbildung 3.1: HMM-Segmente für Objekte mit $L = 546$ bp, distale Verankerungsgeometrie

Schwarze bzw. graue durchgezogene Linien: Abfolgen der Median-gefilterten Helligkeit $I_{\text{med},10}$ der jeweiligen mobilen Reporter-Einheiten. Während grau eingefärbter Abschnitte einer Abfolge liegt $I_{\text{med},10}$ dauerhaft bis zum Ende der Messung unter dem Wert g_3 . Diese Zeitintervalle wurden von der HMM-Analyse der entsprechenden Auslenkungsabfolgen ausgeschlossen. Horizontale gepunktete Linien: Schwellwerte g_1 (9000), g_2 (8000) und g_3 (7000). Vertikale blaue Linien: Segmentgrenzen.

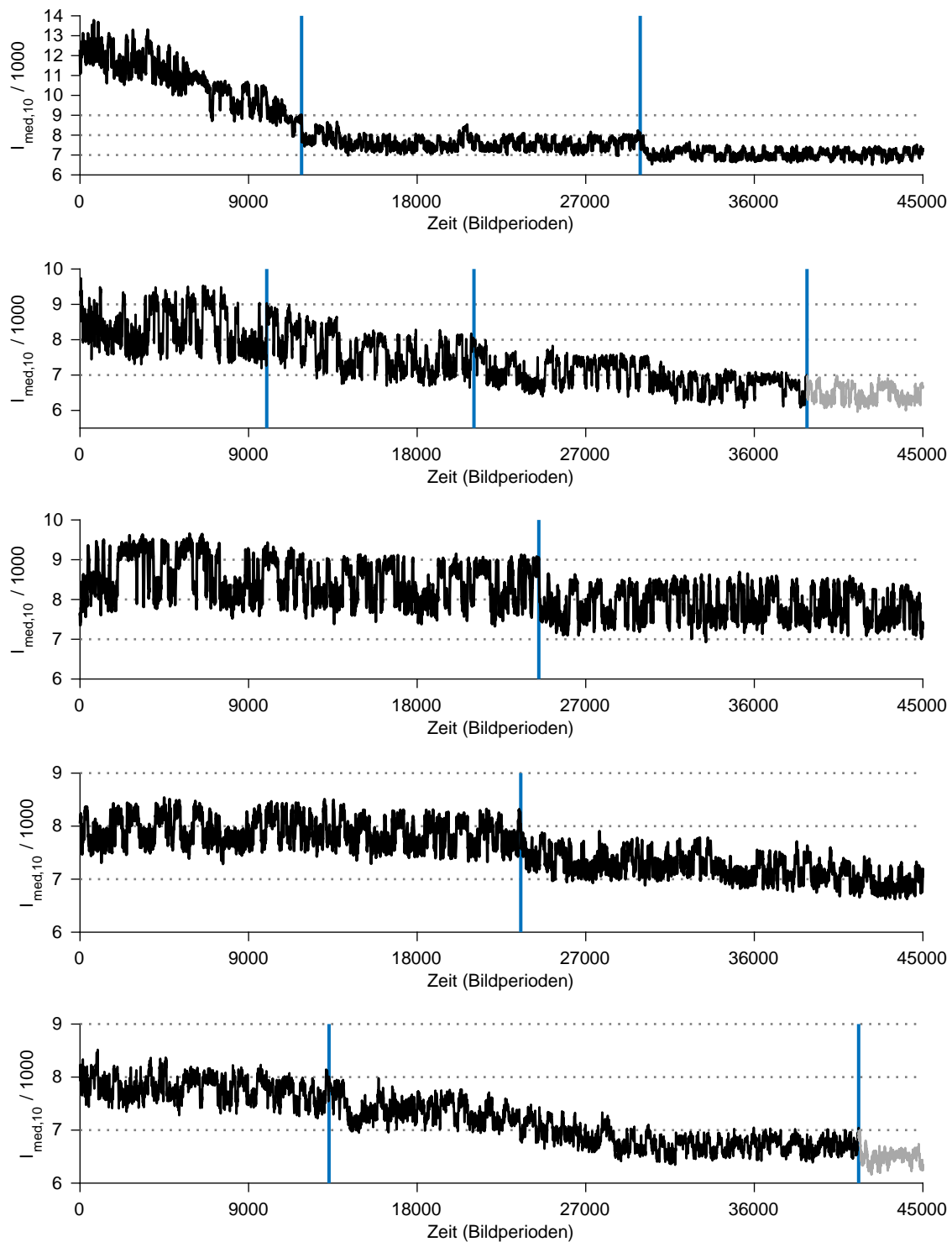


Abbildung 3.2: **HMM-Segmente für Objekte mit $L = 1554$ bp, distale Verankerungsgeometrie**

Schwarze bzw. graue durchgezogene Linien: Abfolgen der Median-gefilterten Helligkeit $I_{\text{med},10}$ der jeweiligen mobilen Reporter-Einheiten. Während grau eingefärbter Abschnitte einer Abfolge liegt $I_{\text{med},10}$ dauerhaft bis zum Ende der Messung unter dem Wert g_3 . Diese Zeitintervalle wurden von der HMM-Analyse der entsprechenden Auslenkungsabfolgen ausgeschlossen. Horizontale gepunktete Linien: Schwellwerte g_1 (9000), g_2 (8000) und g_3 (7000). Vertikale blaue Linien: Segmentgrenzen.

einer Reporter-Einheit nicht alle drei Segmente definiert sind. Ein beliebiges Segment ε_k ist nicht definiert, wenn $I_{\text{med},10}$ während der gesamten Messung unter dem Wert g_k liegt. Ein Segment ε_k mit $k > 1$ ist nicht definiert, wenn der letzte Wert der Abfolge $I_{\text{med},10}$ oberhalb des Werts g_{k-1} liegt.

3.3 Weitere Schritte nach der HMM-Analyse

3.3.1 Extraktion der Verweildauern aus den Zustandsabfolgen

Das für die weitere Auswertung hauptsächlich relevante Ergebnis der HMM-Analyse besteht in der Zustandsabfolge $z(t)$. Sie repräsentiert die Maximum-Likelihood-Schätzung, zu welchen Zeitpunkten das System im Zustand 1 verweilt und zu welchen Zeitpunkten es im Zustand 2 verweilt (Abb. 3.3). Jedes Intervall zwischen zwei Übergängen, innerhalb dessen $z(t)$ konstant ist, wird als fortdauernder Aufenthalt im entsprechenden Zustand betrachtet. Die dazugehörige Verweildauer in Sekunden θ ist das Produkt aus der Dauer des Aufenthalts in Bildperioden mit der Dauer einer Bildperiode, 100 ms. Gilt beispielsweise

$$z(t') = 1 \quad , \quad z(t'+1) = z(t'+2) = \dots = z(t'+50) = 2 \quad \text{und} \quad z(t'+51) = 1,$$

so stellen die Bildperioden $t'+1, \dots, t'+50$ einen fortdauernden Aufenthalt im Zustand 2 mit $\theta = 5$ s dar.

Aussortieren von Aufenthalten mit transienten Klebeereignissen

Nach erfolgter HMM-Analyse war es möglich, transiente Ereignisse unspezifischen Klebens an der Probenoberfläche zu detektieren, die bei der ersten Qualitätskontrolle der aufgenommenen Daten nicht erkannt wurden (Abb. 3.4). Transiente unspezifische Klebeereignisse zeichnen sich dadurch aus, dass innerhalb eines Zeitfensters von einigen Bildperioden

- die Breite der Auslenkungsverteilung in x- und/oder y-Richtung deutlich kleiner ist als die typische Breite der Auslenkungsverteilungen im entsprechenden Zustand
- und/oder der Mittelwert der Auslenkungen in x- und/oder y-Richtung uncharakteristisch stark von Null abweicht.

Die Verweildauern aller Aufenthalte, innerhalb derer ein oder mehrere transiente unspezifische Klebeereignisse detektiert wurden, gingen nicht in die weitere Analyse ein.

3.3.2 Kompensation des Effekts verpasster Aufenthalte

In einer TMFM-Messung sind die Aufenthalte in den beiden Zuständen nur innerhalb gewisser zeitlicher Grenzen detektierbar. Bedingt durch die endliche Beobachtungsdauer der Messung existiert einerseits eine obere Detektionsgrenze θ_{max} . Andererseits

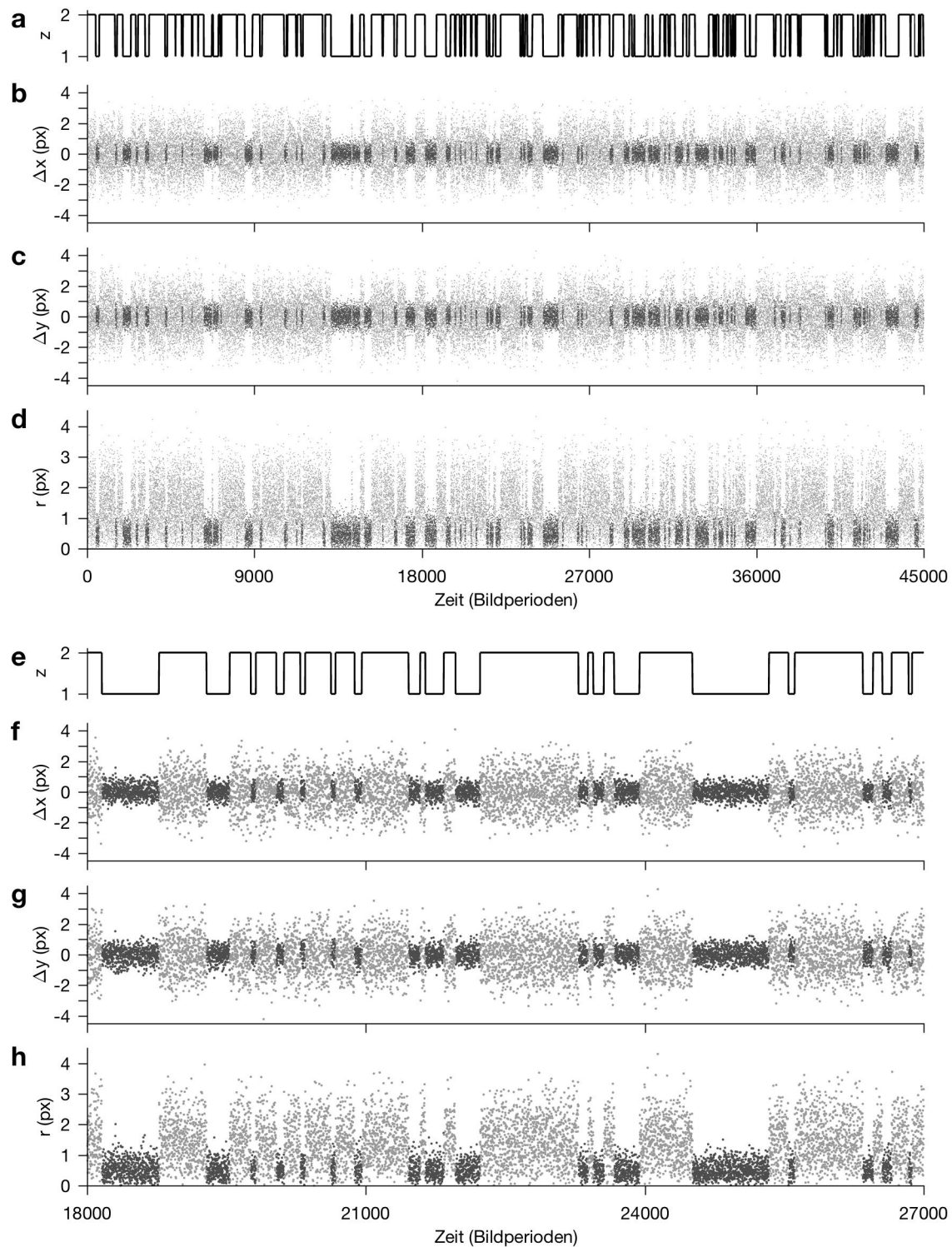


Abbildung 3.3: **Auslenkungsabfolgen mit zugewiesenen Zuständen**

(a) Von der HMM-Prozedur ausgegebene Zustandsabfolge $z(1), \dots, z(T)$. $z=1$: Zustand mit Schlaufe in der Leine. $z=2$: Zustand ohne Schlaufe in der Leine. (b-d) Abfolgen (b) der Auslenkungen in x-Richtung und (c) in y-Richtung sowie (d) der radialen Auslenkung. Dunkelgraue Punkte: Jeweilige Auslenkungen zu Zeitpunkten mit $z=1$. Hellgraue Punkte: Jeweilige Auslenkungen zu Zeitpunkten mit $z=2$. (e-h) Wie a-d, für einen Ausschnitt der Beobachtungsdauer (Bildperioden 18000 bis 27000).

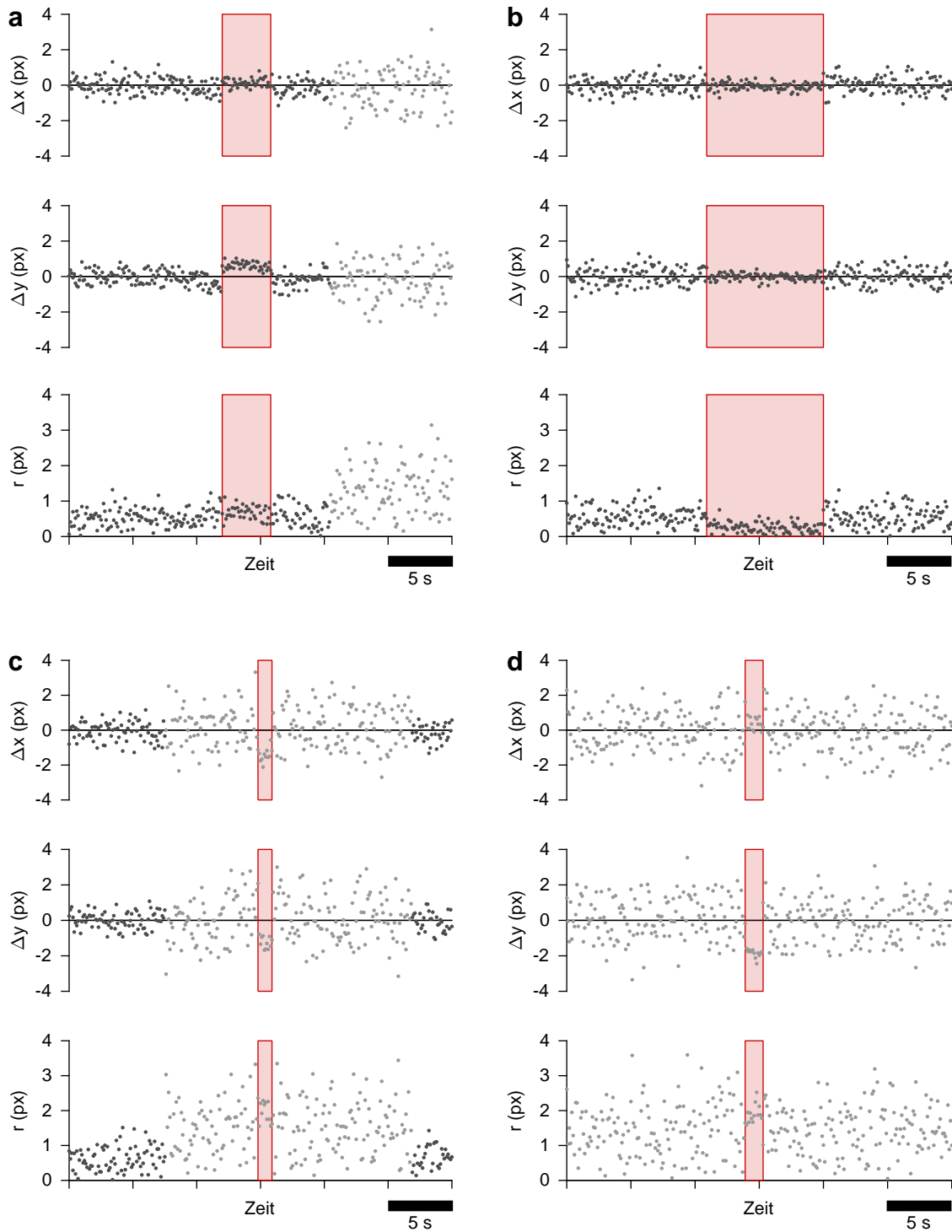


Abbildung 3.4: **Beispiele für kurzlebige unspezifische Klebereignisse**

Jeweils 30 Sekunden (entspricht 300 Bildperioden) lange Ausschnitte der Abfolgen von $\Delta x(t)$, $\Delta y(t)$ und $r(t)$ der mobilen Reporter-Einheiten von vier Objekten mit $L=1554$ bp und Interaktionspartnern in distaler Verankerungsgeometrie. Dunkelgraue Punkte: Jeweilige Auslenkungen zu Zeitpunkten im Zustand mit Schlaufe in der Leine (Zustand 1). Hellgraue Punkte: Jeweilige Auslenkungen zu Zeitpunkten im Zustand ohne Schlaufe in der Leine (Zustand 2). In Rot: Intervalle mit kurzlebigen unspezifischen Klebereignissen (**a,b**) während Aufenthalt im Zustand 1 und (**c,d**) während Aufenthalt im Zustand 2.

besteht wegen der begrenzten zeitlichen Auflösung der Messung eine gewisse Wahrscheinlichkeit, Aufenthalte zu verpassen, deren Verweildauer kürzer ist als der Wert einer unteren Detektionsgrenze θ_{\min} [108, 110–112]. Insbesondere das Verpassen der kurzen Aufenthalte verzerrt die empirischen Verweildauerverteilungen auf zweierlei Weise [108, 110]: Wird ein kurzer Aufenthalt im Zustand 1 (bzw. 2) verpasst, so fehlt die Verweildauer dieses Aufenthalts in der Verweildauerstatistik dieses Zustands. Zum anderen verschmelzen dann die benachbarten Aufenthalte im Zustand 2 (bzw. 1) zu einem einzelnen Aufenthalt mit längerer Verweildauer (vgl. Abb. B.2). Ohne eine entsprechende Berücksichtigung dieses Umstands in der Betrachtung der experimentellen Verweildauern überschätzt man die charakteristischen Verweildauern in den Zuständen gegenüber ihren realen Werten. Der Fehler fällt umso größer aus, je näher die realen Erwartungswerte der Verweildauerverteilungen am Wert von θ_{\min} liegen [108].

Zur Kompensation des Effekts verpasster Aufenthalte in der Analyse der TMFM-Daten wurde die von Stigler et al. [108] beschriebene Prozedur adaptiert. Als obere Detektionsgrenze wurde jeweils die Verweildauer des längsten detektierten Aufenthalts verwendet. Die Werte für die untere Detektionsgrenzen wurden empirisch aus der Betrachtung simulierter Datensätze abgeschätzt. Für nahezu alle TMFM-Datensätze wurde $\theta_{\min,1} = 1,3 \text{ s}$ und $\theta_{\min,2} = 0,7 \text{ s}$ verwendet. Die einzigen Ausnahmen bildeten die Datensätze der Objekte mit der Interaktionssequenz CAGGAACG:CGTTCCTG in distaler Verankerungsgeometrie und $L = 546 \text{ bp}$ ($\theta_{\min,1} = 1,7 \text{ s}$ und $\theta_{\min,2} = 0,9 \text{ s}$) bzw. $L = 2562 \text{ bp}$ ($\theta_{\min,1} = 1,0 \text{ s}$ und $\theta_{\min,2} = 0,2 \text{ s}$).

In der Realität stellen diese Werte keine scharfen Grenzen dar, unterhalb derer alle Aufenthalte im jeweiligen Zustand verpasst werden. Vielmehr nimmt unterhalb einer gewissen Schwelle mit weiter sinkender Verweildauer der erwartete Anteil verpasster Aufenthalte nach und nach zu (vgl. Abb. B.4). Dementsprechend wurden durch den HMM-Algorithmus immer noch einige wenige Aufenthalte mit einer Verweildauer unterhalb der unteren Detektionsgrenze erkannt. Das Modell von Stigler et al. [108] fußt allerdings u. a. auf der Annahme, dass *alle* Aufenthalte mit Verweildauer $\theta < \theta_{\min,i}$ ausnahmslos verpasst werden. Aus Gründen der Konsistenz wurden daher die detektierten Aufenthalte mit Verweildauer $\theta < \theta_{\min,i}$ aus den Zustandsabfolgen entfernt.

Die Verweildauern aus den derart angepassten Zustandsabfolgen einer gesamten Probe wurden zusammengefasst. Anhand der jeweiligen Menge von Verweildauern im Zustand i wurden zunächst durch Maximum-Likelihood-Schätzung die Zeitparameter λ_i der beschränkten Exponentialverteilung mit Verteilungsfunktion

$$F_i(\theta) = \frac{\exp(-\lambda_i \theta) - \exp(-\lambda_i \theta_{\min,i})}{\exp(-\lambda_i \theta_{\max,i}) - \exp(-\lambda_i \theta_{\min,i})}, \quad i \in \{1, 2\} \quad (3.6)$$

und Dichtefunktion

$$f_i(\theta) = \frac{-\lambda_i \cdot \exp(-\lambda_i \theta)}{\exp(-\lambda_i \theta_{\max,i}) - \exp(-\lambda_i \theta_{\min,i})}, \quad i \in \{1, 2\} \quad (3.7)$$

ermittelt. Anschließend wurden die charakteristischen Verweildauern in den zwei Zuständen τ_1 und τ_2 analog zu [108] durch numerische Minimierung des folgenden Feh-

lerterms bestimmt:

$$\left(\exp\left(-\frac{\theta_{\min,2}}{\tau_2}\right) - \lambda_1 \cdot \tau_1 \right)^2 + \left(\exp\left(-\frac{\theta_{\min,1}}{\tau_1}\right) - \lambda_2 \cdot \tau_2 \right)^2 \quad (3.8)$$

Die Assoziationsrate k_{on} und Dissoziationsrate k_{off} berechnen sich aus den inversen charakteristischen Verweildauern:

$$k_{\text{off}} = \frac{1}{\tau_1} \quad , \quad k_{\text{on}} = \frac{1}{\tau_2} \quad (3.9)$$

3.3.3 Detektion und Kompensation molekularer Heterogenität der Probe

DNA-Nano-Objekte derselben Probe zeigten teilweise deutlich unterschiedliche Kinetiken der Hybridisierung und Dissoziation der Interaktionspartner. Die Heterogenität der Kinetik geht aus der Betrachtung der Auslenkungs- und Zustandsabfolgen verschiedener mobiler Reporter-Einheiten (Abb. 3.5) deutlich hervor. Die Ursache der heterogenen Kinetik innerhalb einer Probe könnte in Synthesefehlern einzelner interaktionspartnertragender Oligonukleotide liegen. In Abschnitt 1.4 wurden bereits Ergebnisse aus Ensemble-Experimenten diskutiert, die für das Zutreffen dieser Hypothese sprechen.

Nach der Selbstassemblierung der DNA-Nano-Objekte können Objekte mit vollkommen korrekten Interaktionssequenzen nicht getrennt von DNA-Nano-Objekten mit z. B. Fehlstellen oder Verkürzungen in den Interaktionssequenzen aufgereinigt werden (vgl. Abschnitte 1.4). Daher wurde im Zuge der Analyse der TMFM-Daten ein probabilistischer Ansatz verfolgt, um die DNA-Nano-Objekte nach ihrer individuellen Assoziations- und Dissoziationskinetik in zwei Gruppen zu sortieren. In allen analysierten Proben wurde mit diesem Ansatz die Mehrzahl der Objekte der Probe einer Gruppe mit ähnlicher Kinetik zugeordnet. Diese Gruppe wird mithin als „Hauptpopulation“ bezeichnet. Die Gruppe der „Ausreißer“ umfasst alle DNA-Nano-Objekte, deren Kinetik extrem von der charakteristischen Kinetik der Hauptpopulation abweicht.

Vorgehen zur Klassifizierung in Hauptpopulation und Ausreißer

Die zu Grunde liegende Annahme des Ansatzes lautet, dass die wahren Verweildauern jedes Paares von Interaktionspartnern in den zwei Zuständen unabhängige Zufallszahlen aus Exponentialverteilungen mit Ratenparametern k_{off} bzw. k_{on} sind. Betrachtet man die Summe D aus N unabhängigen Zufallszahlen, welche einer Exponentialverteilung mit Ratenparameter k entstammen, so ist die Wahrscheinlichkeit für das Ereignis „ $D \leq d$ “ durch eine Erlang-Verteilung gegeben:

$$P(D \leq d) = P_{\text{Erlang}}(d; N, k) = 1 - \sum_{n=0}^{N-1} \frac{1}{n!} (kd)^n \exp(-kd) \quad (3.10)$$

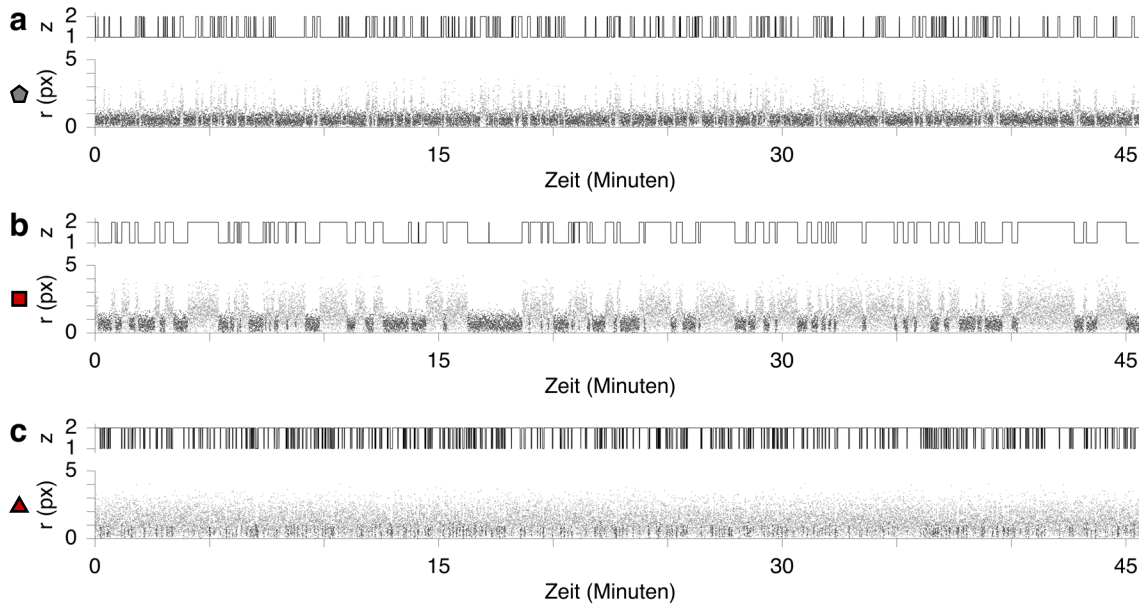


Abbildung 3.5: **Heterogene Kinetik innerhalb einer Probe von DNA-Nano-Objekten – Beispielabfolgen** Interaktionssequenz CCTCCTCC:GGAGGAGG, Konturlänge des Leinenstücks zwischen den Interaktionspartnern $L = 1554$ bp, distale Verankerungsgeometrie. Ausschnitte aus Zustandsabfolgen z und dazugehörigen Abfolgen der radialen Auslenkung r der mobilen Reporter-Einheit **(a)** eines Objekts der Hauptpopulation, **(b)** eines Objekts mit ungewöhnlich hoher mittlerer Verweildauer im dissoziierten Zustand sowie **(c)** eines Objekts mit ungewöhnlich hoher mittlerer Verweildauer im dissoziierten Zustand *und* ungewöhnlich niedriger mittlerer Verweildauer im hybridisierten Zustand. Radiale Auslenkung während Aufenthalt im hybridisierten Zustand in Dunkelgrau, während Aufenthalt im dissoziierten Zustand in Hellgrau. Die Marker links dienen zur Zuordnung des jeweiligen Objekts im Streudiagramm in Abb. 3.6 a.

Aus der Zustandsabfolge jedes DNA-Nano-Objekts erhält man eine Anzahl von Aufenthalten im Zustand 1 (N_1), eine Summe von Verweildauern im Zustand 1 (D_1) sowie die entsprechenden Werte für den Zustand 2 (N_2 und D_2). Allerdings wird davon ausgegangen, dass diese Werte aufgrund von verpassten kurzen Aufenthalten nicht den tatsächlichen Zahlen von Aufenthalten und den tatsächlichen Summen aller Verweildauern in den jeweiligen Zuständen im Beobachtungszeitraum entsprechen. Unter der Annahme exponentialverteilter Verweildauern ist der erwartete Anteil der Aufenthalte mit einer Verweildauer $\theta < \theta_{\min,i}$

$$q_i = 1 - \exp\left(-\frac{\theta_{\min,i}}{\tau_i}\right) \quad , \quad i \in \{1,2\}. \quad (3.11)$$

Die erwartete mittlere Verweildauer dieses Anteils der Aufenthalte ist

$$\tau_{q,i} = \tau_i - \exp\left(-\frac{\theta_{\min,i}}{\tau_i}\right) \cdot (\tau_i + \theta_{\min,i}) \quad , \quad i \in \{1,2\}. \quad (3.12)$$

Damit können Schätzungen \hat{N}_i und \hat{D}_i für die tatsächliche Anzahl der Aufenthalte im Zustand i und die tatsächliche Summe der Verweildauern im Zustand i vorgenommen werden:

$$\hat{N}_i = (1 + \alpha_i) \cdot \frac{N_i + \alpha_j N_j}{1 - \alpha_i \alpha_j} \quad \text{mit} \quad \alpha_i = \frac{q_i}{1 - q_i} \quad , \quad i, j \in \{1, 2\}, j \neq i \quad (3.13)$$

$$\hat{D}_i = D_i + (\hat{N}_i - N_i) \cdot \tau_{q,i} - (\hat{N}_j - N_j) \cdot \tau_{q,j} \quad , \quad i, j \in \{1, 2\}, j \neq i \quad (3.14)$$

Anhand dieser Schätzwerte sowie der Raten k_{off} und k_{on} wurden für jedes DNA-Nano-Objekt die Wahrscheinlichkeiten

$$P_{\text{Erlang}}(\hat{D}_1; \hat{N}_1, k_{\text{off}}) \quad \text{und} \quad P_{\text{Erlang}}(\hat{D}_2; \hat{N}_2, k_{\text{on}})$$

nach Gl. (3.10) berechnet. Ein DNA-Nano-Objekt wurde als Ausreißer klassifiziert, wenn mindestens eine dieser Wahrscheinlichkeiten einen Wert am extremen unteren Rand ($P_{\text{Erlang}} < 10^{-6}$) oder am extremen oberen Rand ($P_{\text{Erlang}} > 1 - 10^{-6}$) der Verteilung annahm.

Die tatsächliche Prozedur zur Klassifizierung der DNA-Nano-Objekte in Hauptpopulation und Ausreißer gestaltete sich wie folgt: Zunächst wurden die Werte von k_{off} und k_{on} auf einem groben zweidimensionalen Gitter variiert, für jede mögliche Kombination der Raten eine Klassifizierung vorgenommen und die Kombination mit der größten Zahl der zur Hauptpopulation zugeordneten Objekte ausgewählt. Im Anschluss wurde die Klassifizierung iterativ optimiert, wobei jede Iteration drei Schritte umfasste:

1. Abschätzung der Raten anhand der Mengen aller einzelnen Verweildauern aus der aktuellen Hauptpopulation gemäß Gleichungen (3.6)-(3.9)
2. Neuberechnung der Wahrscheinlichkeiten P_{Erlang} für beide Zustände gemäß Gleichungen (3.10)-(3.14)
3. Erneute Klassifizierung in Hauptpopulation und Ausreißer auf Basis des Wahrscheinlichkeitskriteriums

Ergebnis der Klassifizierung in Hauptpopulation und Ausreißer

Im Streudiagramm der individuellen mittleren Verweildauern der DNA-Nano-Objekte einer Probe in den beiden Zuständen (Abb. 3.6 a) sind bereits mit bloßem Auge verschiedene Punktcluster erkennbar. Die Klassifizierung in die Hauptpopulation (Cluster oben links, blaue und grauer Marker) und die Ausreißer (rote Marker) ist grundsätzlich gut nachvollziehbar. Allerdings besteht auf den ersten Blick eine scheinbare Widersprüchlichkeit der Klassifikation am Rand des Hauptpopulationsclusters: Objekte werden als Ausreißer aussortiert, obwohl ihre Marker näher am Zentrum des Clusters liegen als die Marker mancher Objekte, die der Hauptpopulation zugeordnet werden. Hierbei handelt es sich jedoch lediglich um einen scheinbaren Widerspruch, denn die Breite der Erlang-Wahrscheinlichkeitsverteilung (Gl. 3.10) korreliert invers mit der Anzahl der einzelnen Verweildauern N . Somit können Objekte, deren mittlere Verweildauern sichtbar vom Cluster abweichen, in die Hauptpopulation aufgenommen werden, sofern die entspre-

chenden Anzahlen der Verweildauern niedrig genug sind. Umgekehrt können Objekte mit hohem N als Ausreißer klassifiziert werden, obwohl ihre mittleren Verweildauern näher am Cluster der Hauptpopulation liegen.

Nach der Klassifizierung wurden die Häufigkeitsverteilungen aller einzelnen Verweildauern der verschiedenen Gruppen von Objekten betrachtet (Abb. 3.6 b,c). Dabei wurden die Häufigkeitsverteilungen (Histogramme) mit Maximum-Likelihood-Schätzungen beschränkter Exponentialverteilungen (gestrichelte Linien) überlagert. Für die Mengen aller einzelnen Verweildauern der Hauptpopulation weisen die Histogramme und die Exponentialverteilungen sehr ähnliche Verläufe auf. Diese Übereinstimmung wird als Indiz für die Homogenität der Kinetik innerhalb der Hauptpopulation interpretiert. Im

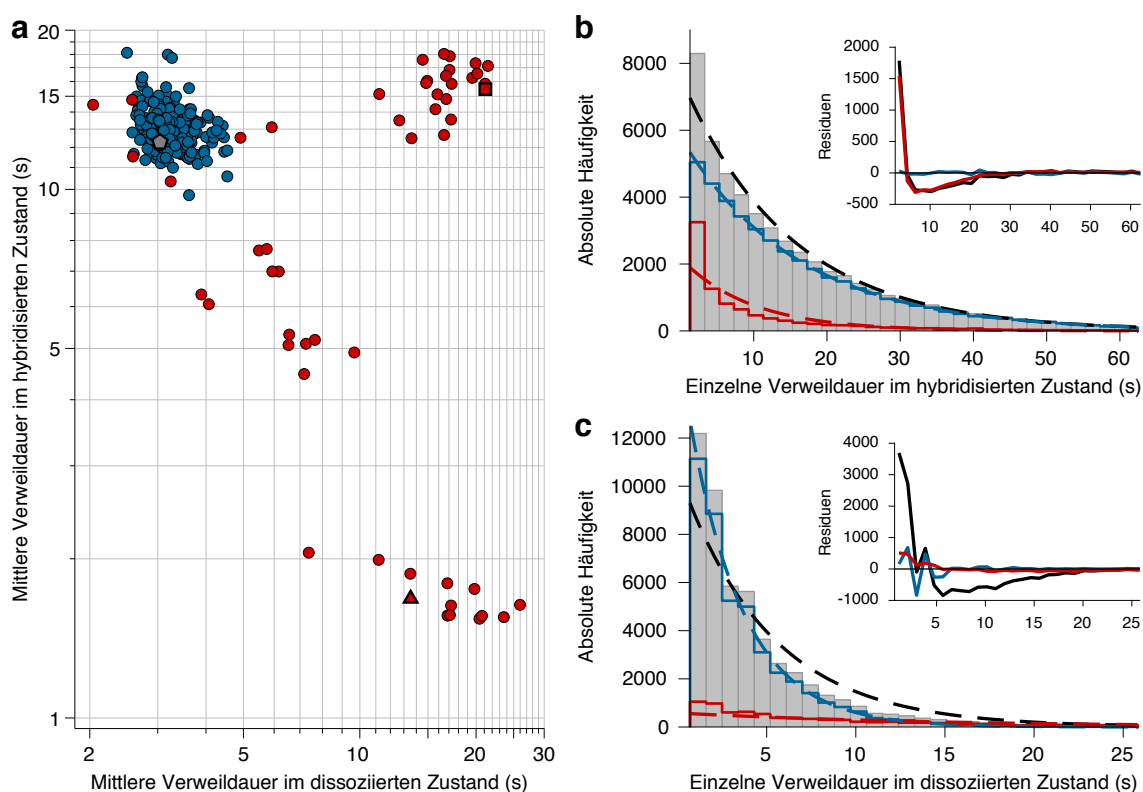


Abbildung 3.6: Heterogene Kinetik in einer TMFM-Messung – Ergebnis der Klassifizierung

Interaktionssequenz CCTCCTCC:GGAGGAGG, Konturlänge des Leinenstücks zwischen den Interaktionspartnern $L = 1554$ bp, distale Verankerungsgeometrie. **(a)** Streudiagramm der mittleren Verweildauern. In jedem Datenpunkt ist für ein DNA-Nano-Objekt die mittlere Verweildauer im hybridisierten Zustand gegen die mittlere Verweildauer im dissoziierten Zustand aufgetragen. Hauptpopulation in Blau/Grau, Ausreißer in Rot. Das graue Pentagon, das rote Quadrat und das rote Dreieck markieren die mittleren Verweildauern jener Objekte, deren Zustands- und Auslenkungsabfolgen in Abb. 3.5 gezeigt sind. **(b)** Verteilungen der einzelnen Verweildauern im hybridisierten Zustand. Graue Balken: einzelne Verweildauern aus den Zustandsabfolgen aller Objekte. Blaue gestufte Linie: einzelne Verweildauern aus den Zustandsabfolgen der Hauptpopulation. Rote gestufte Linie: einzelne Verweildauern aus den Zustandsabfolgen der Ausreißer. Schwarze, blaue und rote gestrichelte Linien: Maximum-Likelihood-Schätzungen von Dichteverteilungen mit der Form von Gl. (3.7), normiert auf die Summe der Fläche aller Klassen des jeweiligen Histogramms. Nebenabbildung: absolute Residuen (Differenz zwischen den Werten der Histogramme und Fit-Funktionen). **(c)** Wie b, für die Verweildauern im dissoziierten Zustand.

Gegensatz dazu folgen die Häufigkeitsverteilungen aller einzelnen Verweildauern der Ausreißer oder der Menge aller Objekte nicht dem Verlauf der Exponentialverteilungen. Diese Verteilungen spiegeln Überlagerungen von Prozessen unterschiedlicher Kinetik wider. Sie können nicht hinreichend durch Wahrscheinlichkeitsverteilungen mit einem einzelnen Exponentialterm beschrieben werden.

Im Fall der in Abb. 3.5 und Abb. 3.6 gezeigten Probe mit der 8-bp-Interaktionssequenz CCTCTCC:GGAGGAGG in distaler Verankerungsgeometrie und $L = 1554$ bp umfasst die Hauptpopulation 248 Objekte. Dagegen wurden 68 Objekte als Ausreißer klassifiziert. Die finalen Werte der charakteristischen Verweildauern der Hauptpopulation ergeben

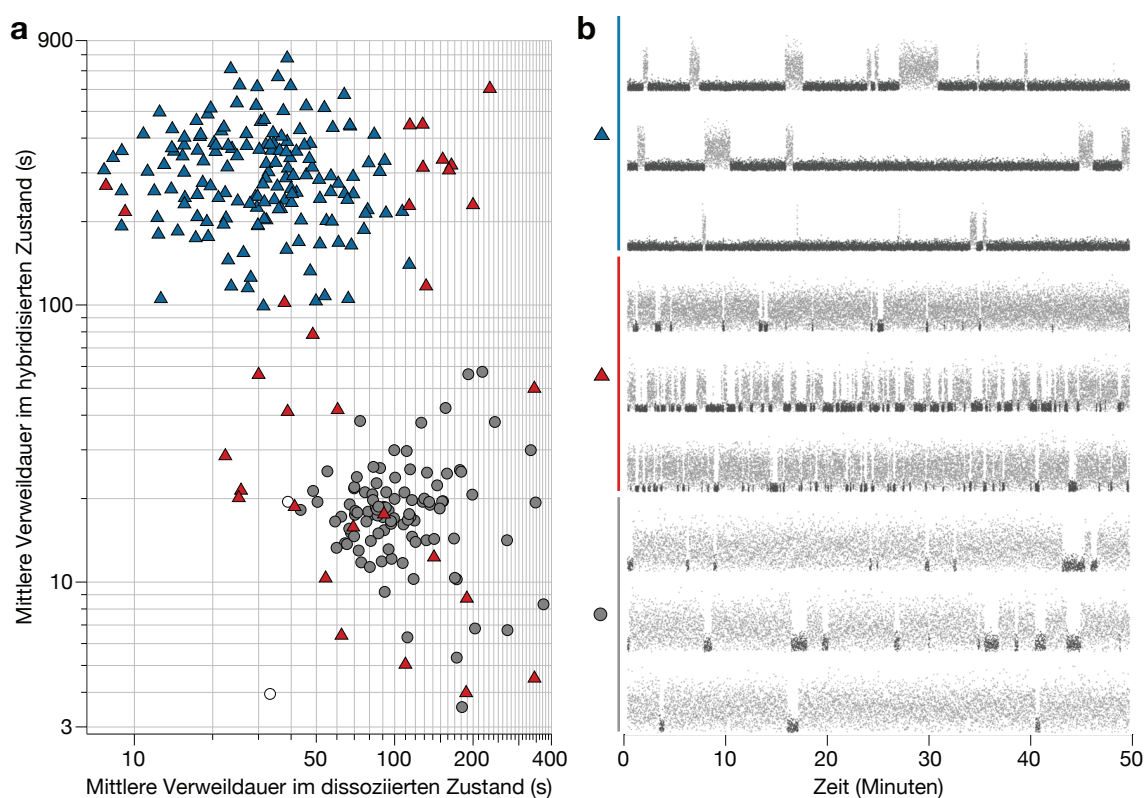


Abbildung 3.7: **Vergleich zwischen 9-bp-Probe und 8-bp-Probe mit ähnlichen Interaktionssequenzen**
 9-bp-Interaktionssequenz: CCAGAGACG:CGTCTCTGG, 8-bp-Interaktionssequenz: CAGGAACG:CGTTCCTG,
 Konturlänge des Leinenstücks zwischen den Interaktionspartnern jeweils $L = 1050$ bp, proximale Verankerungsgeometrie. **(a)** Streudiagramm der mittleren Verweildauern. In jedem Datenpunkt ist für ein DNA-Nano-Objekt die mittlere Verweildauer im hybridisierten Zustand gegen die mittlere Verweildauer im dissoziierten Zustand aufgetragen. Blaue Dreiecke: Hauptpopulation der 9-bp-Probe. Rote Dreiecke: Ausreißer der 9-bp-Probe. Graue Kreise: Hauptpopulation der 8-bp-Probe. Weiße Kreise: Ausreißer der 8-bp-Probe. **(b)** Ausschnitte aus Abfolgen der radialen Auslenkung der mobilen Reporter-Einheiten von neun DNA-Nano-Objekten. Radiale Auslenkung während Aufenthalt im hybridisierten Zustand in Dunkelgrau, während Aufenthalt im dissoziierten Zustand in Hellgrau. Oberes Tripel: Beispielobjekte aus der Hauptpopulation der 9-bp-Probe. Unteres Tripel: Beispielobjekte aus der Hauptpopulation der 8-bp-Probe. Mittleres Tripel: Ausreißer der 9-bp-Probe mit mittlerer Verweildauer im hybridisierten Zustand zwischen 10 und 30 s und mittlerer Verweildauer im dissoziierten Zustand zwischen 20 und 200 s.

sich nach Gleichungen (3.6)-(3.8) zu $\tau_1 \approx 12,4$ s und $\tau_2 \approx 2,7$ s. Bei Berechnung der charakteristischen Verweildauern unter Einbeziehung aller Objekte ohne Aussortieren der Ausreißer würde τ_1 um knapp 3 % und τ_2 um 66 % überschätzt. Für ein weiteres Beispiel, eine Probe von Objekten mit der 9-bp-Interaktionssequenz CCAGAGACG:CGTCTCTGG in proximaler Verankerungsgeometrie und $L = 1050$ bp, würde durch die Einbeziehung der Ausreißer τ_1 um etwa 30 % unterschätzt und τ_2 um etwa 45 % überschätzt.

Interessant ist der Vergleich zwischen dieser 9-bp-Variante und einer 8-bp-Variante mit ähnlicher Interaktionssequenz CAGGAACG:CGTTCCTG (Abb. 3.7). Ausreißer der 9-bp-Probe mit deutlich kürzerer mittlerer Verweildauer im hybridisierten Zustand als die Hauptpopulation der 9-bp-Probe zeigen eine mitunter sehr ähnliche Kinetik zur Hauptpopulation der 8-bp-Probe. Diese Beobachtung ist ein weiteres Indiz für das Zutreffen der Hypothese, dass die heterogene Kinetik verschiedener DNA-Nano-Objekte derselben Probe von Verkürzungen in den interaktionspartnertragenden Staples herrühren könnte.

3.3.4 *Bootstrapping*-Verfahren zur Fehlerabschätzung

Die statistischen Fehler in den charakteristischen Verweildauern und Raten wurden mit einem Bootstrapping-Verfahren abgeschätzt. Bezeichne Θ_i die Menge aller einzelnen Verweildauern der Hauptpopulation der Objekte einer Probe im Zustand i . Bezeichne $N_{\Theta,i}$ die Anzahl der Elemente in Θ_i . Bezeichne B_i eine pseudozufällig und ohne Zurücklegen gezogene Stichprobe mit Umfang $N_{\Theta,i}$ aus der Menge Θ_i . Das Bootstrapping-Verfahren besteht darin, jeweils 10000 derartiger Stichproben B_1 und B_2 zu ziehen und anhand dieser Stichproben die Werte von τ_1 und τ_2 bzw. k_{off} und k_{on} nach Gleichungen (3.6)-(3.9) zu ermitteln. Die Standardabweichungen $\sigma_{\text{bootstrap}}$ der Verteilungen der so erhaltenen 10000 Werte für die charakteristischen Verweildauern in den Zuständen geben einen Anhaltspunkt für den experimentellen Fehler. In der Analyse der TMFM-Daten wurden die Fehlerintervalle der charakteristischen Verweildauern und der Raten stets auf $\pm 3 \cdot \sigma_{\text{bootstrap}}$ abgeschätzt.

3.3.5 Vergleich der Kinetik in verschiedenen Zeitintervallen der TMFM-Messungen

In der Analyse der TMFM-Daten wurde angenommen, dass sich die zu Grunde liegende Assoziations- und Dissoziationskinetiken der beobachteten Einzelstrangpaare im Verlauf einer TMFM-Aufnahme nicht ändern. Grundsätzlich kann jedoch eine mögliche Änderung der Kinetik im Lauf der Zeit nicht vollkommen ausgeschlossen werden. Beispielsweise ist denkbar, dass sich der pH-Wert im Messpuffer durch die Aktivität des Oxygen Scavenging Systems (vgl. Abschnitt 1.5) trotz der Stabilisierung durch 100 mM Tris-HCl mit pH 8 über die Dauer von 75 Minuten leicht absenkt. Darüber hinaus könnten photoinduzierte oder anderweitig hervorgerufen Schäden an den Interaktionssequenzen oder anderen Teilen der DNA-Nano-Objekte auftreten. Die beschädigten Objekte könnten dann ab dem Zeitpunkt der Beschädigung andere charakteristische Verweildauern in den Zuständen aufweisen. Eine weitere mögliche Quelle für eine Ver-

schiebung der gemessenen Verweildauerverteilungen im Verlauf einer Messung könnte in der segmentierten HMM-Analyse (vgl. Abschnitt 3.2.2) liegen.

Um der Frage nach einer möglichen Veränderung der Kinetiken im zeitlichen Verlauf der TMFM-Messungen nachzugehen, wurden die Verteilungen der Verweildauern in den Zuständen in mehreren aufeinanderfolgenden Zeitintervallen miteinander verglichen (Abb. 3.8, Abb. C.12). Der 75 Minuten lange Beobachtungszeitraum jeder Aufnahme wurde in fünf aufeinanderfolgende jeweils 15 Minuten lange Zeitintervalle aufgeteilt. Für jedes dieser Intervalle wurden die Verteilungen der Verweildauern aller Aufenthalte betrachtet, die innerhalb des Intervalls beginnen und enden. Für Messungen mit Stichprobenumfängen von mehr als 1000 einzelnen Verweildauern je Intervall ergab sich eine sehr hohe Ähnlichkeit zwischen den Verläufen der Verweildauerverteilungen aus allen fünf Intervallen (Abb. 3.8, Abb. C.12.1-C.12.2 und C.12.9-C.12.11). Bei Stichprobenumfängen < 1000 Einzelwerte je Intervall zeigten sich Unterschiede zwischen den empirischen Verteilungen aus verschiedenen Intervallen (Abb. C.12.3-C.12.8). In der Ausprägung dieser Unterschiede wurde jedoch keine Regelmäßigkeit festgestellt,

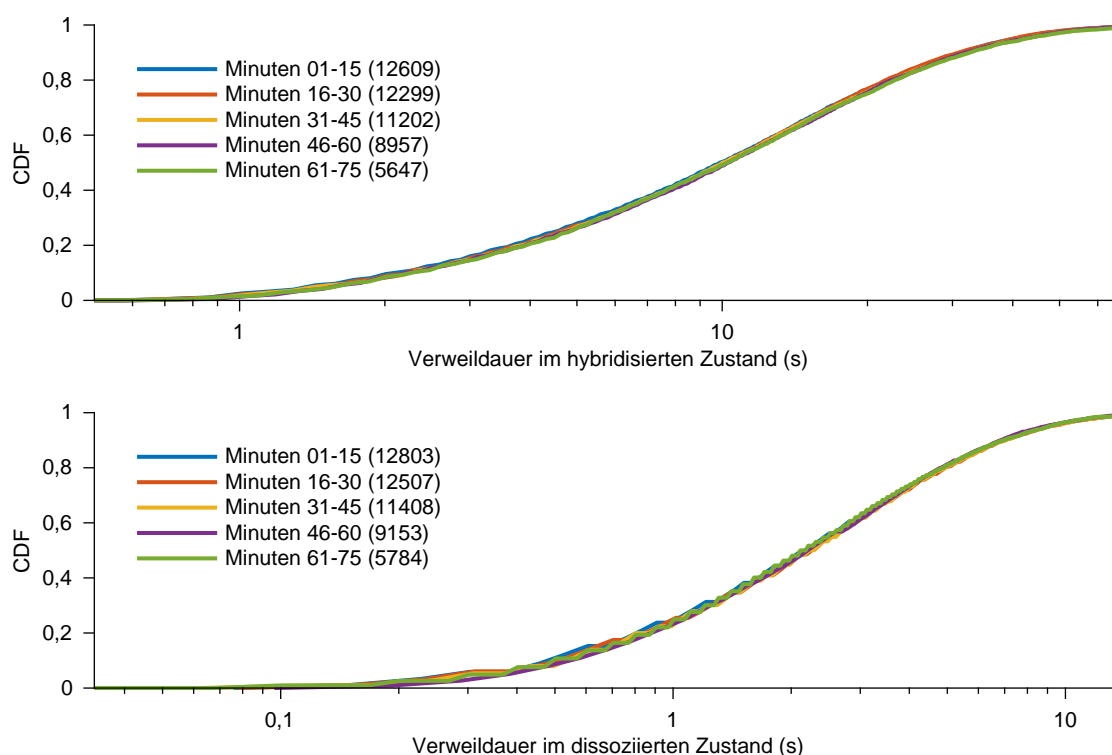


Abbildung 3.8: **Verweildauerverteilungen im zeitlichen Verlauf einer TMFM-Messung**

Interaktionssequenz CCTCTCC:GGAGGAGG, Konturlänge des Leinenstücks zwischen den Interaktionspartnern $L = 1554$ bp, distale Verankerungsgeometrie. Empirische Verteilungen der kumulierten Wahrscheinlichkeit (CDF), ermittelt aus der relativen Häufigkeit der einzelnen Verweildauern. Die Verweildauerachsen sind logarithmisch skaliert. Jedes Diagramm zeigt fünf Verteilungen. Jede dieser Verteilungen wurde nur aus den Verweildauern von Aufenthalten im jeweiligen Zustand generiert, die innerhalb eines gewissen Zeitintervalls beginnen und enden. Legende: Zuordnung der Kurven zu den Zeitintervallen. Zahlen in Klammern: jeweilige Anzahl der einzelnen Verweildauern, die in die Verteilung dieses Zeitintervalls eingehen.

die auf eine systematische Verschiebung der charakteristischen Verweildauern über den Verlauf einer Messung hindeuten würde.

3.4 Kapitelzusammenfassung

Zur Analyse der Auslenkungsabfolgen der mobilen Reporter-Einheiten von Objekten mit Interaktionspartnern wurde die Methode der Hidden-Markov-Modellierung verwendet. Das dafür aufgestellte Hidden-Markov-Modell (HMM) hat zwei Zustände: Der „hybridisierte Zustand“ umfasst alle Mikrozustände des Systems mit teilweiser oder vollständiger Basenpaarung zwischen den Interaktionspartnern. Der „dissoziierte Zustand“ umfasst alle Mikrozustände des Systems ohne jegliche Basenpaarung zwischen den Interaktionspartnern. Die Observable des HMM sind die Auslenkungen der mobilen Reporter-Einheit $(\Delta x_{RE}, \Delta y_{RE})$.

Die Emissionswahrscheinlichkeiten der Observablen wurden als bivariate Normalverteilungen modelliert. Die für viele mobile Reporter-Einheiten mit abnehmender Helligkeit einhergehende Verbreiterung der Auslenkungsverteilungen über den Verlauf einer Messung wurde wie folgt berücksichtigt: Die einzelnen Abfolgen von $(\Delta x_{RE}, \Delta y_{RE})$ wurden anhand der dazugehörigen Helligkeitsabfolgen in bis zu drei Segmente unterteilt. Die Parameter der Emissionswahrscheinlichkeitsverteilungen wurden während der HMM-Prozedur für jedes dieser Segmente individuell optimiert und ausgewertet.

Die HMM-Analyse liefert zu einer Abfolge von Auslenkungen $(\Delta x_{RE}, \Delta y_{RE})$ eine Zustandsabfolge z . Daraus lassen sich die Verweildauern in den beiden Zuständen extrahieren. Aufgrund der beschränkten zeitlichen Auflösung sowie der endlichen Dauer der TMFM-Aufnahmen wird für beide Zustände ein Teil der Aufenthalte – insbesondere sehr kurzlebige Aufenthalte – nicht detektiert. Zur Korrektur des Effekts verpasster Aufenthalte in der Analyse der TMFM-Daten wurde die von Stigler et al. [108] beschriebene Prozedur adaptiert.

DNA-Nano-Objekte derselben Probe zeigten teilweise deutlich unterschiedliche Hybridisierungskinetiken der Interaktionspartner. Die Ursache dieser heterogenen Kinetik könnte in Synthesefehlern einzelner interaktionspartnertragender Staples liegen. Es wurde eine Prozedur implementiert, mit der die Objekte einer Probe basierend auf probabilistischen Überlegungen in zwei Gruppen sortiert werden: die „Hauptpopulation“ mit homogener Kinetik und die Gruppe der „Ausreißer“.

Es besteht Grund zur Annahme, dass die Kinetik der Interaktionspartnerpaare jeweils über den gesamten Verlauf der 75 Minuten langen Messungen gleich ist. Ein Vergleich der Verweildauerverteilungen aus mehreren aufeinanderfolgenden Zeitintervallen ergab keine systematische Veränderung der gemessenen Kinetik im zeitlichen Verlauf der einzelnen Messungen.

4 Hybridisierungskinetik kurzer DNA-Duplexe

Mit der neu entwickelten Methode TMFM wurde die Hybridisierungskinetik kurzer DNA-Doppelstränge in Abhängigkeit von verschiedenen Faktoren untersucht. Dieses Kapitel zeigt die Ergebnisse der drei TMFM-Studien zur Abhängigkeit der Hybridisierungskinetik von der Beschaffenheit des Partikel-Leinen-Systems (Abschnitt 4.1), von der Konzentration von Natriumionen in der umgebenden Lösung (Abschnitt 4.2) und von der Basenpaarsequenz der Interaktionspartner (Abschnitt 4.3).

4.1 Einfluss des umgebenden Partikel-Leinen-Systems

4.1.1 Vorüberlegungen

Die Hybridisierung zweier frei in Lösung vorliegender DNA-Einzelstränge zu einem Doppelstrang kann durch eine Reaktionskinetik zweiter Ordnung beschrieben werden. Die Assoziationsrate dieser bimolekularen Reaktion hat die Einheit $M^{-1} s^{-1}$, die Dissoziationsrate hat die Einheit s^{-1} und die entsprechende Gleichgewichtskonstante $K_{eq,bi}$ hat die Einheit M^{-1} . Im Rahmen von TMFM ist die Hybridisierung von B und B' aufgrund der Verankerung der Einzelstränge im Partikel-Leinen-System ein unimolekularer Prozess, der sich gänzlich durch Reaktionsgleichungen erster Ordnung beschreiben lässt. Sowohl die Assoziationsrate als auch die Dissoziationsrate haben die Einheit s^{-1} , weshalb die entsprechende Gleichgewichtskonstante $K_{eq,uni}$ einheitenlos ist.

Um den unimolekularen Fall mit dem bimolekularen Fall vergleichen zu können, wird in der Literatur der sogenannte *J-Faktor* verwendet [113]. Er errechnet sich aus dem Verhältnis der beiden Gleichgewichtskonstanten:

$$J = \frac{K_{eq,uni}}{K_{eq,bi}} \quad (4.1)$$

Der J-Faktor hat die Einheit Molar und stellt in seiner ursprünglichen Definition die effektive Konzentration des einen Endes eines Fragments doppelsträngiger DNA bezogen auf das andere Ende des Fragments dar [113]. Im Kontext von TMFM steht der J-Faktor für die effektive Konzentration des einen Interaktionspartners im Bezug auf den anderen Interaktionspartner innerhalb des Partikel-Leinen-Systems. Das folgende Gedankenexperiment dient zur Verdeutlichung dieser Betrachtungsweise: Im TMFM-Experiment stelle sich ein bestimmtes Besetzungsgleichgewicht zwischen hybridisiertem und dissoziiertem Zustand der im Partikel-Leinen-System verankerten Interaktionspartner ein. Nun denke man sich den Fall, dass einer der Interaktionspartner weiterhin im Partikel-Leinen-System verankert ist, der andere Interaktionspartner aber mit einer bestimmten

Konzentration frei in Lösung vorliegt. Der J-Faktor ist nun die Konzentration des freien Interaktionspartners, bei der sich dasselbe Besetzungsgleichgewicht einstellen würde. So wie im bimolekularen Fall die Kolokalisationswahrscheinlichkeit der Einzelstränge direkt proportional zu ihren Konzentrationen ist, so ist sie im TMFM-Experiment direkt proportional zum J-Faktor. Dementsprechend gilt: Je größer der J-Faktor, umso stärker verschiebt sich das Besetzungsgleichgewicht hin zum hybridisierten Zustand.

Möglicher Einfluss der Konturlänge L auf die gemessene Kinetik

Der Wert des J-Faktors eines DNA-Fragments hängt von der Konturlänge des Fragments ab. Diese Abhängigkeit wurde bereits in mehreren theoretischen und experimentellen Studien untersucht [114–116]. Der J-Faktor doppelsträngiger DNA ist bei einer bestimmten Konturlänge im Bereich von etwa 500 bp maximal und fällt sowohl zu kleineren als auch zu größeren Konturlängen hin monoton ab. Der Grund dafür liegt darin, dass das Zusammenführen der Enden kurzer und langer DNA-Fragmente energetisch besonders ungünstig ist [113]. Kurze DNA-Fragmente müssen stark gekrümmt werden, um die Enden zusammenzuführen. Mit steigender Konturlänge wird zwar der Biegeenergieanteil immer geringer, dagegen ist jedoch mit dem Zusammenführen der Enden ein immer größerer Entropieverlust verbunden.

Die Rate der Assoziationsreaktion von B und B' zu einem Doppelstrang B:B' ist direkt proportional zur Kolokalisationswahrscheinlichkeit der Einzelstränge. Demnach sind bei gleichbleibender Basenpaarsequenz der Interaktionspartner für DNA-Nano-Objekte mit unterschiedlicher Konturlänge L des Leinenstücks zwischen den Interaktionspartnern (vgl. Abb. 1.5) unterschiedliche Assoziationsraten zu erwarten. Außerdem könnte auch die gemessene Dissoziationsrate von L abhängen: Dem Leinenstück zwischen den Interaktionspartnern wird im hybridisierten Zustand ein energetisch ungünstiger End-zu-End-Abstand aufgezwungen, analog zum makroskopischen Bild einer gestauchten Feder (vgl. Abb. 1.1, rechte Seite). Die entstehende Rückstellkraft könnte den Doppelstrang B:B' destabilisieren und somit die Dissoziation beschleunigen. Diese destabilisierende Kraft sollte sich dann als negative Ableitung der Federenergie zwischen Objekten mit verschiedenem Wert von L ebenfalls unterscheiden.

Möglicher Einfluss der Reporter-Einheiten auf die gemessene Kinetik

Trotz der geringen Größe der Reporter-Einheiten ist es denkbar, dass sie einen störenden Einfluss auf die Hybridisierung der Interaktionspartner haben. Sind B und B' jeweils nahe an den Reportereinheiten verankert, muss sich die mobile Reporter-Einheit in der Nähe der Probenoberfläche und der immobilen Reporter-Einheit befinden, damit sich das Duplex B:B' ausformen kann. Eine sterische Hemmung der Assoziation durch diese erforderliche Nähe ist nicht auszuschließen. Dabei würde der Abstand der Ankerpunkte von B und B' zu den Reporter-Einheiten eine entscheidende Rolle spielen.

Zur Ergründung eines möglichen derartigen Effekts wurden zwei verschiedene Verankerungsgeometrien miteinander verglichen: die distale und die proximale Verankerungsgeometrie. In der distalen Verankerungsgeometrie zweigen die Interaktionspartner

mit den Abständen $D_{\text{imm}} = 252 \text{ bp}$ und $D_{\text{mob}} = 307 \text{ bp}$ zur jeweils näheren Reporter-Einheit ab (vgl. Abb. 1.5 a). Dagegen ragt in der proximalen Verankerungsgeometrie einer der Interaktionspartner aus einem Helixende der immobilen Reporter-Einheit hervor, der andere zweigt mit $D_{\text{mob}} = 55 \text{ bp}$ Abstand zur mobilen Reporter-Einheit von der Leine ab (vgl. Abb. 1.5 c,d). Somit befinden sich im hybridisierten Zustand der Interaktionspartner in der distalen Verankerungsgeometrie 559 bp und in der proximalen Verankerungsgeometrie 55 bp der DNA-Leine zwischen den Reporter-Einheiten.

4.1.2 Ergebnisse der TMFM-Messungen an DNA-Nano-Objekten mit verschiedenen Werten von L und verschiedenen Verankerungsgeometrien

Zur Untersuchung des Einflusses des Werts von L und der Verankerungsgeometrie auf die gemessene Kinetik wurden TMFM-Messungen an unterschiedlichen Varianten von DNA-Nano-Objekten durchgeführt. Für die distale Verankerungsgeometrie wurden drei Varianten von Objekten mit $L = 546 \text{ bp}$, $L = 1554 \text{ bp}$ und $L = 2562 \text{ bp}$ untersucht. Für die proximale Verankerungsgeometrie wurden fünf Varianten von Objekten mit $L = 546 \text{ bp}$, $L = 1050 \text{ bp}$, $L = 1554 \text{ bp}$, $L = 2058 \text{ bp}$ und $L = 2562 \text{ bp}$ untersucht. Alle Varianten hatten dieselbe Interaktionssequenz $\text{CAGGAACG} : \text{CGTTCCTG}$. In den Messpuffer für alle Messungen wurden 500 mM Na^+ eingesetzt.

Die mobilen Reporter-Einheiten aller dieser Varianten zeigten das bereits in den vorangegangenen Kapiteln diskutierte Verhalten: In den Auslenkungsabfolgen wechseln sich Zeitintervalle höherer und geringerer Breite ab. Die abwechselnd unterschiedliche Breite der Auslenkungsabfolgen ergibt sich aus dem wiederholten Wechsel zwischen dem dissoziierten und hybridisierten Zustand der Interaktionspartner B und B' (vgl. Abb. 1.1). Im dissoziierten Zustand der Interaktionspartner (= Zustand ohne Schlaufe in der DNA-Leine) hat die mobile Reporter-Einheit einen höheren Bewegungsradius als im hybridisierten Zustand (= Zustand mit ausgeformter Schlaufe in der DNA-Leine).

Abbildungen 4.1 bis 4.4 stellen die Auslenkungsabfolgen und -verteilungen einzelner Objekte der Hauptpopulationen (vgl. Abschnitt 3.3.3) der verschiedenen Varianten gegenüber. Für Varianten derselben Verankerungsgeometrie ist im hybridisierten Zustand das Ausmaß der Auslenkungen der mobilen Reporter-Einheiten ungefähr gleich. Dagegen nimmt das Ausmaß der Auslenkungen im dissoziierten Zustand bei gleicher Verankerungsgeometrie mit steigendem Wert von L zu. Demnach gilt: Je größer L, umso stärker unterscheidet sich die Beweglichkeit in den zwei Zuständen.

Vergleich der Verweildauern

Aus den Zustandsabfolgen der Objekte mit verschiedenem Wert von L, aber gleicher Verankerungsgeometrie (Abb. 4.1 und 4.3) geht hervor, dass mit steigendem Wert von L im Mittel die Verweildauern im dissoziierten Zustand zunehmen. Der direkte Vergleich von $L = 546 \text{ bp}$ und $L = 2562 \text{ bp}$ (a und c in Abb. 4.1, a und e in Abb. 4.3) zeigt den Zusammenhang besonders deutlich: Beobachtet man bei $L = 546 \text{ bp}$ noch vergleichsweise viele, kurzlebige Aufenthalte im dissoziierten Zustand, so sind es bei $L = 2562 \text{ bp}$ weniger, dafür im Mittel langlebigere Aufenthalte über dieselbe Beobachtungsdauer.

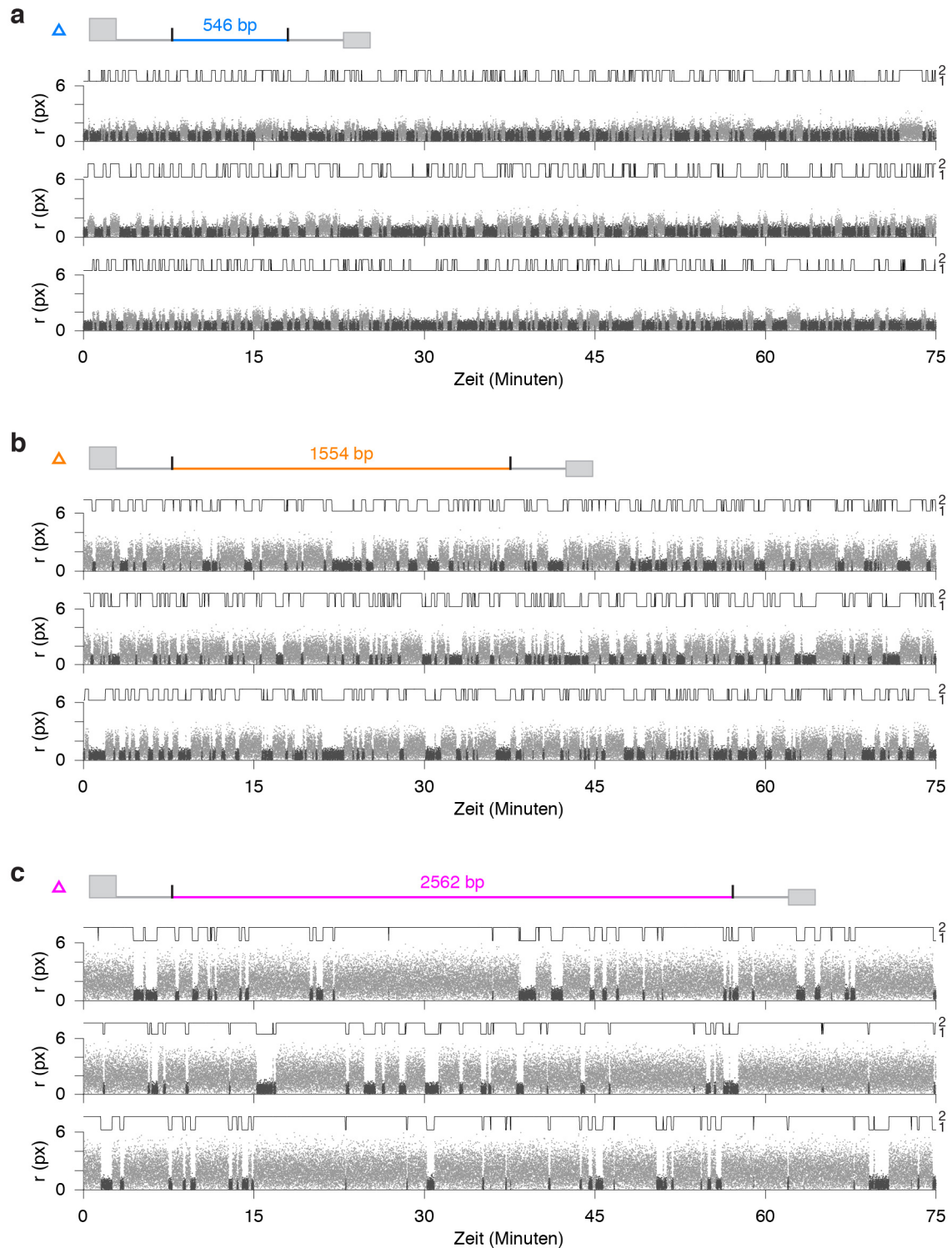


Abbildung 4.1: Zustands- und Auslenkungsabfolgen, distale Verankerungsgeometrie

Interaktionssequenz CAGGAACG : CGTTCCTG. Von a bis c steigende Konturlänge L des Leinenstücks zwischen den Interaktionspartnern: **(a)** 546 bp, **(b)** 1554 bp und **(c)** 2562 bp. Schwarze Linien: Aufenthaltsabfolge aus der HMM-Analyse. 1: hybridisierter Zustand, 2: dissoziierter Zustand. Punkte: Radiale Auslenkungen der mobilen Reporter-Einheiten nach Gl. 2.1 während Aufenthalt im hybridisierten Zustand (dunkelgrau) bzw. Aufenthalt im dissoziierten Zustand (hellgrau). Farben und Marker zur Zuordnung der Objektvarianten in Abb. 4.2, 4.5, 4.7 und 4.9.

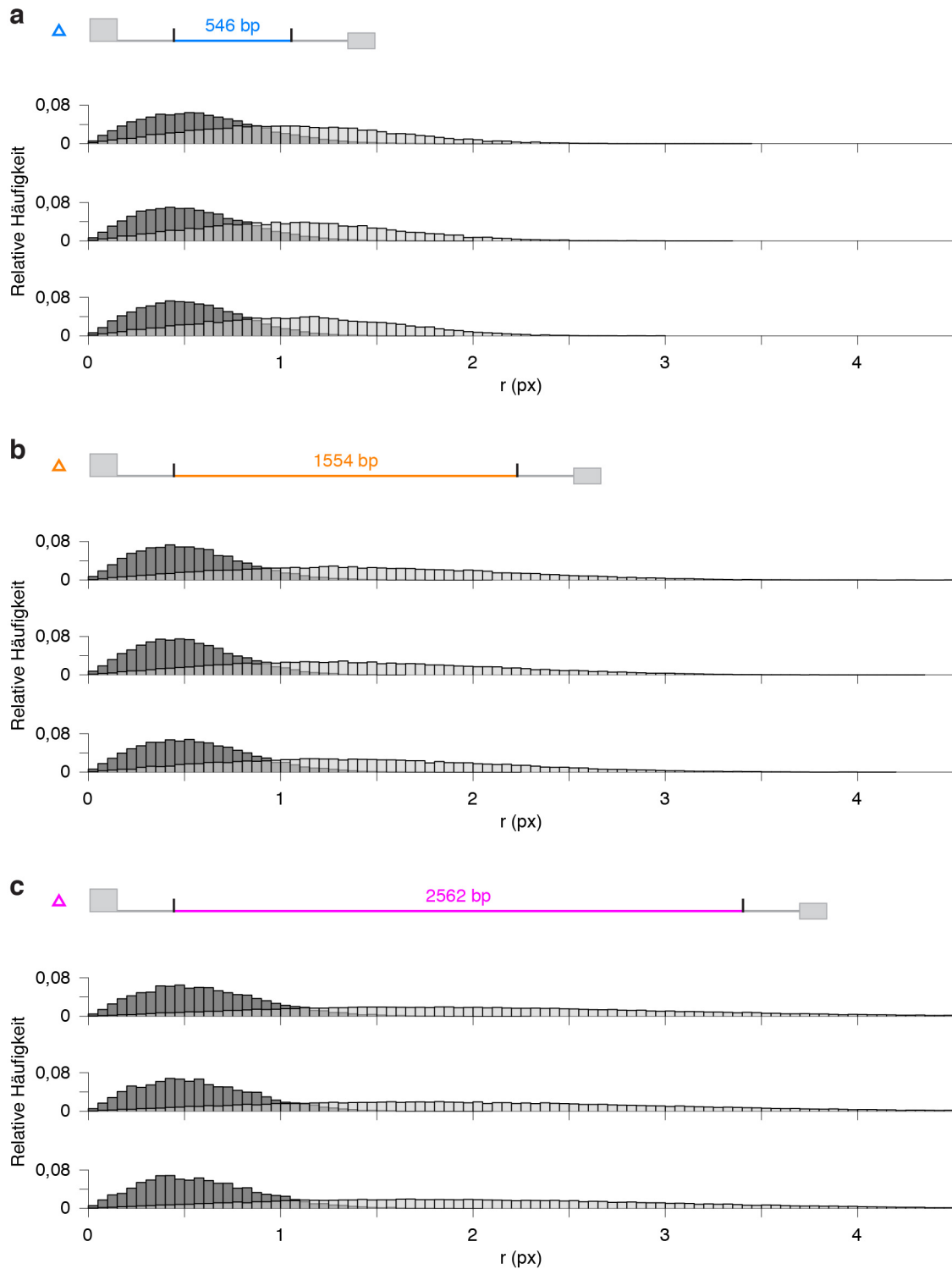


Abbildung 4.2: Verteilungen der radialen Auslenkung r , distale Verankerungsgeometrie

Interaktionssequenz CAGGAACG:CGTTCCTG. Von a bis c steigende Konturlänge L des Leinenstücks zwischen den Interaktionspartnern: **(a)** 546 bp, **(b)** 1554 bp und **(c)** 2562 bp. Histogramme für den hybridisierten Zustand (dunkelgrau) und den dissoziierten Zustand (hellgrau), erstellt aus den entsprechenden Auslenkungsabfolgen in Abb. 4.1. Farben und Marker zur Zuordnung der Objektvarianten in Abb. 4.1, 4.5, 4.7 und 4.9.

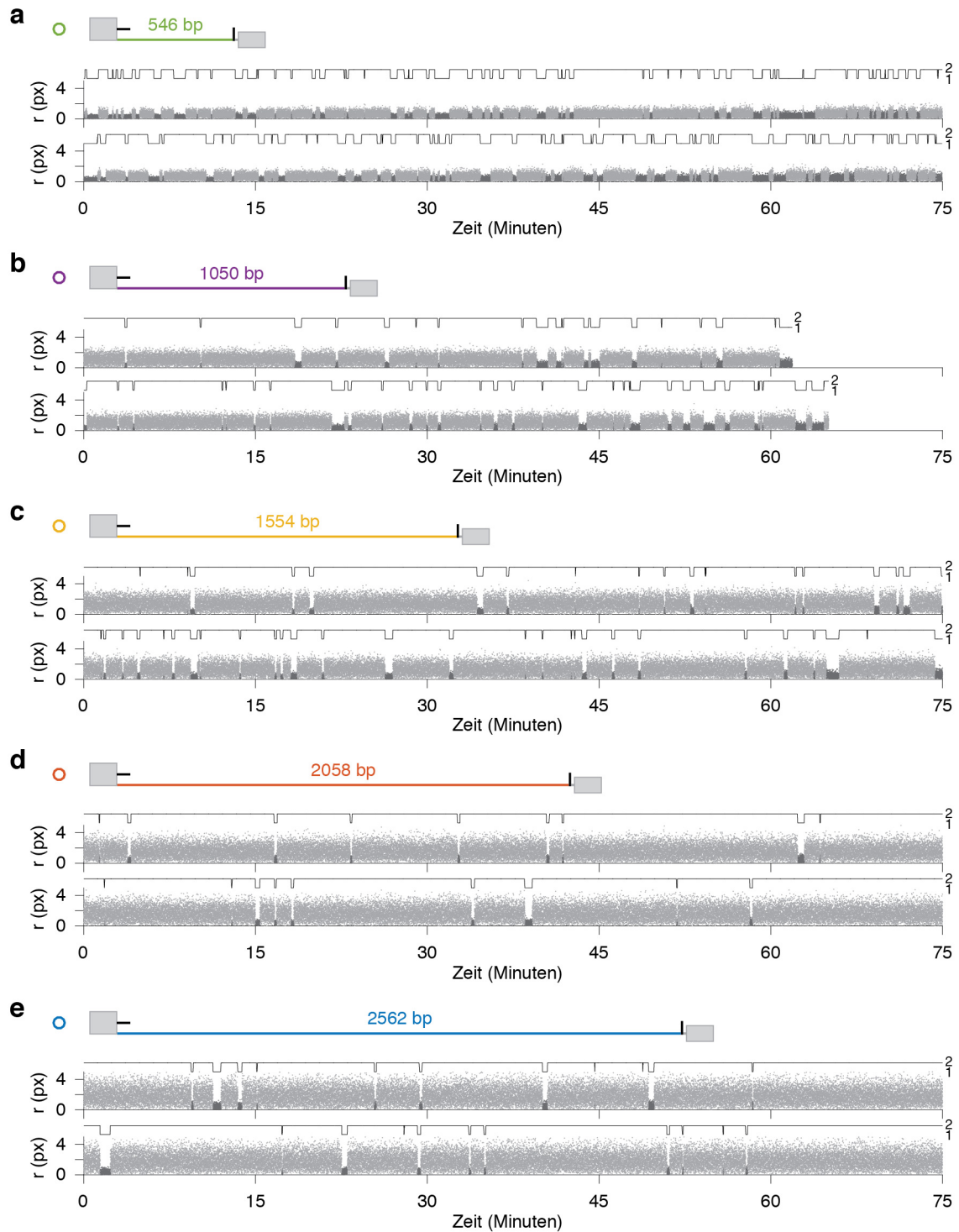


Abbildung 4.3: **Zustands- und Auslenkungsabfolgen, proximale Verankerungsgeometrie**

Interaktionssequenz CAGGAACG:CGTTCCTG. Von a bis e steigende Konturlänge L des Leinenstücks zwischen den Interaktionspartnern: **(a)** 546 bp, **(b)** 1050 bp, **(c)** 1554 bp, **(d)** 2058 bp und **(e)** 2562 bp. Schwarze Linien: Aufenthaltsabfolge aus der HMM-Analyse. 1: hybridisierter Zustand, 2: dissoziierter Zustand. Punkte: Radiale Auslenkungen der mobilen Reporter-Einheiten nach Gl. 2.1 während Aufenthalt im hybridisierten Zustand (dunkelgrau) bzw. Aufenthalt im dissoziierten Zustand (hellgrau). Farben und Marker zur Zuordnung der Objektvarianten in Abb. 4.4, 4.6, 4.8 und 4.9.

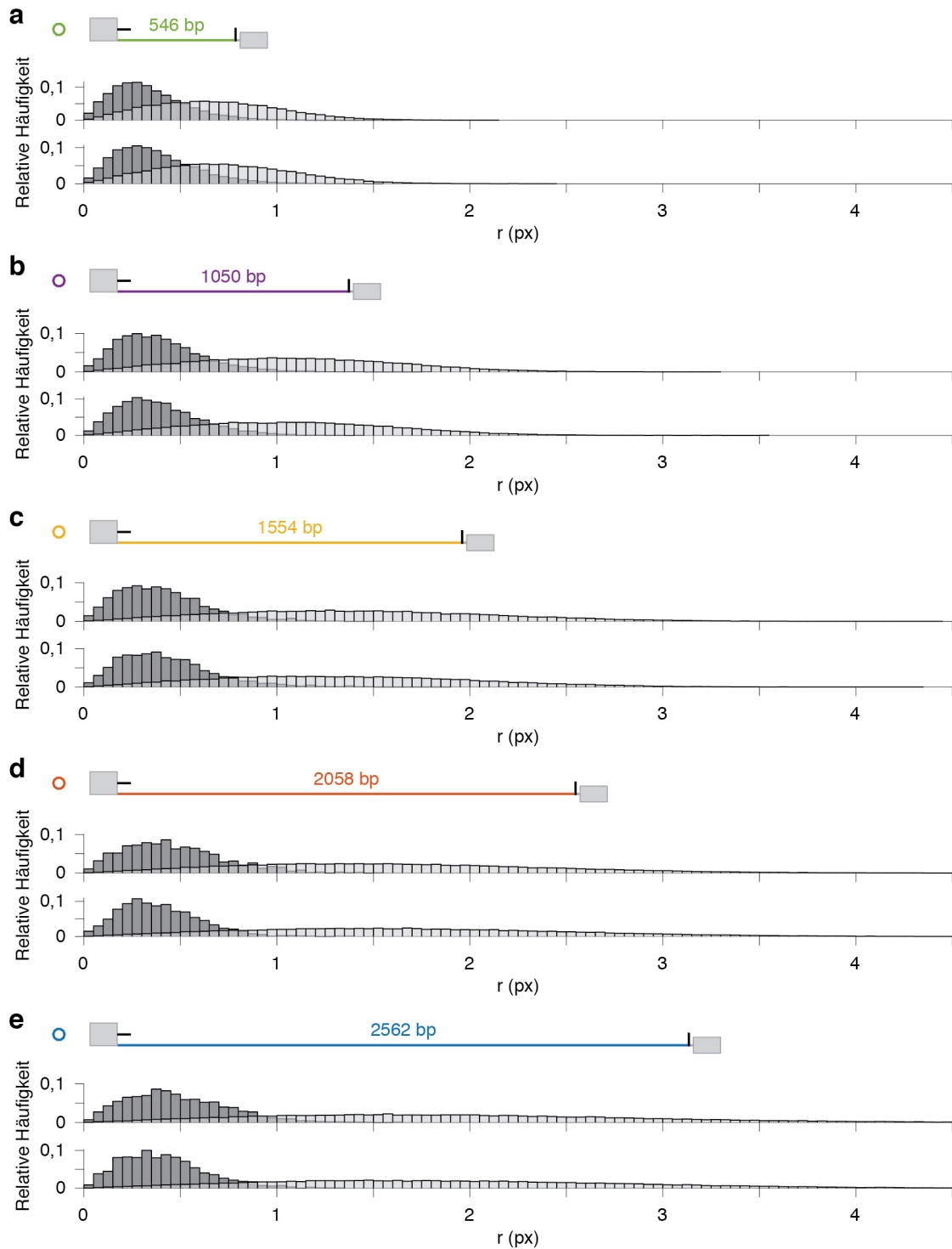


Abbildung 4.4: **Verteilungen der radialen Auslenkung r , proximale Verankerungsgeometrie**

Interaktionssequenz CAGGAACG : CGTTCCTG. Von a bis e steigende Konturlänge L des Leinenstücks zwischen den Interaktionspartnern: **(a)** 546 bp, **(b)** 1050 bp, **(c)** 1554 bp, **(d)** 2058 bp und **(e)** 2562 bp. Histogramme für den hybridisierten Zustand (dunkelgrau) und den dissoziierten Zustand (hellgrau), erstellt aus den entsprechenden Auslenkungsabfolgen in Abb. 4.3. Farben und Marker zur Zuordnung der Objektvarianten in Abb. 4.3, 4.6, 4.8 und 4.9.

Die Verschiebung der Verweildauern im dissoziierten Zustand hin zu größeren Werten mit steigendem Wert von L bei gleicher Verankerungsgeometrie tritt nicht nur für einzelne Objekte auf: Aus den Streudiagrammen der mittleren Verweildauern je Objekt (Abb. 4.5 und 4.6) geht eindeutig hervor, dass dieser Trend für die gesamten Hauptpopulationen der Varianten zutrifft. Die mittleren Verweildauern im hybridisierten Zustand zeigen im Vergleich dazu keine ausgeprägte Abhängigkeit von L . Es zeichnet sich lediglich eine schwache Verschiebung hin zu größeren Verweildauern im hybridisierten Zustand bei kleineren Werten von L ab.

Eine weitere Möglichkeit, die kollektive Abhängigkeit der Kinetik vom Wert von L zu betrachten, besteht darin, für jeden Zustand alle einzelnen Verweildauern der Hauptpopulation jeder Variante zusammenzufassen und die jeweiligen Verweildauerverteilungen gegenüberzustellen (Abb. 4.7 und 4.8). Diese Betrachtung liefert ein mit dem bisherigen Eindruck stimmiges Ergebnis: Die Verteilungen der einzelnen Verweildauern im dissoziierten Zustand aller Objekte der Hauptpopulation verschieben sich mit stei-

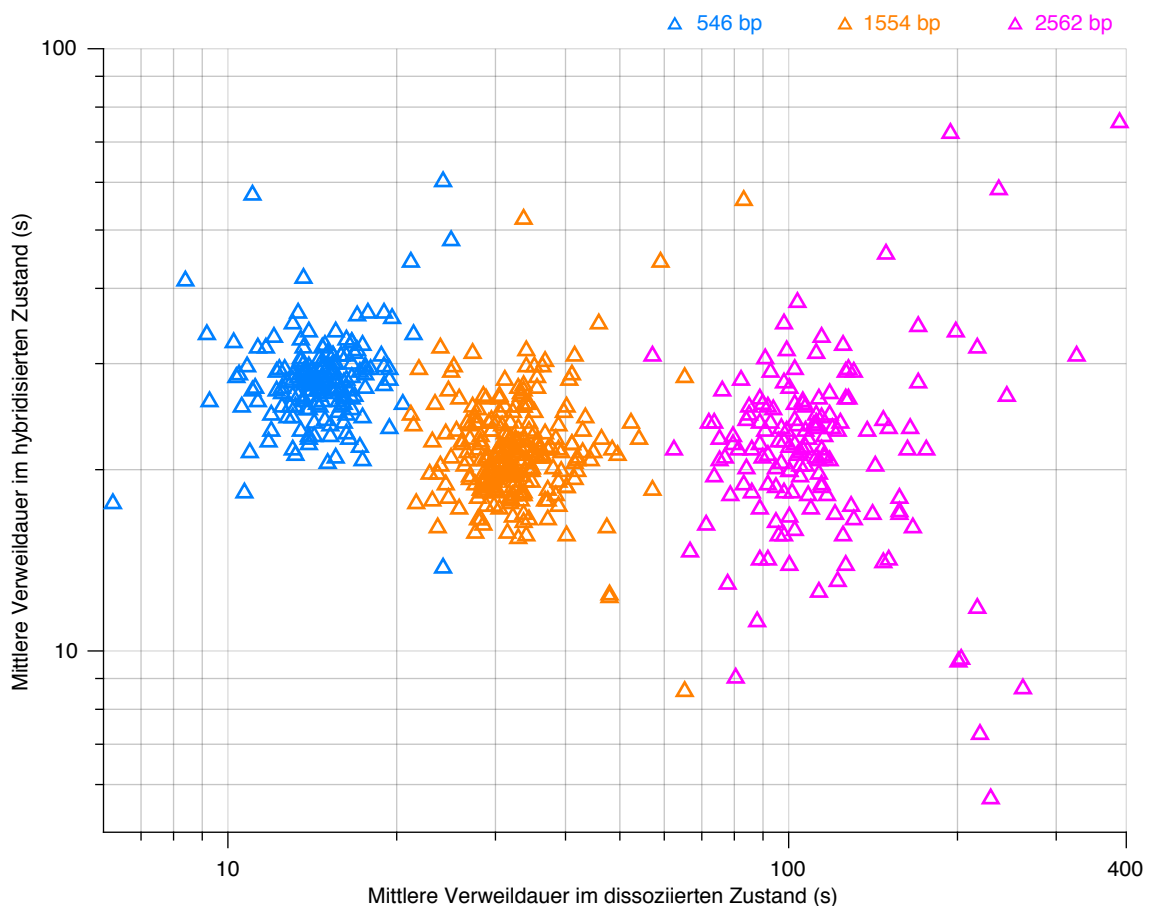


Abbildung 4.5: **Streudiagramm der mittleren Verweildauern, distale Verankerungsgeometrie**

Messdaten von DNA-Nano-Objekten mit der Interaktionssequenz CAGGAACG:CGTTCCTG. In jedem Datenpunkt ist für ein Objekt der jeweiligen Hauptpopulation (vgl. Abschnitt 3.3.3) die mittlere Verweildauer im hybridisierten Zustand gegen die mittlere Verweildauer im dissoziierten Zustand aufgetragen. Die Farben stehen für die Konturlänge L des Leinenstücks zwischen den Interaktionspartnern (siehe Legende oben).

gendem Wert von L immer weiter zu größeren Werten (Abb. 4.7 und 4.8, jeweils rechte Spalte). Dagegen sind die Verteilungen der einzelnen Verweildauern im hybridisierten Zustand aller Objekte der Hauptpopulationen der verschiedenen Varianten sehr ähnlich zueinander (Abb. 4.7 und 4.8, jeweils linke Spalte). Desweiteren wurden aus den Mengen aller einzelnen Verweildauern der jeweiligen Hauptpopulationen die charakteristische Verweildauer im dissoziierten Zustand (τ_2) und die charakteristische Verweildauer im hybridisierten Zustand (τ_1) nach Gleichungen (3.6)-(3.8) ermittelt. Der Wert von τ_2

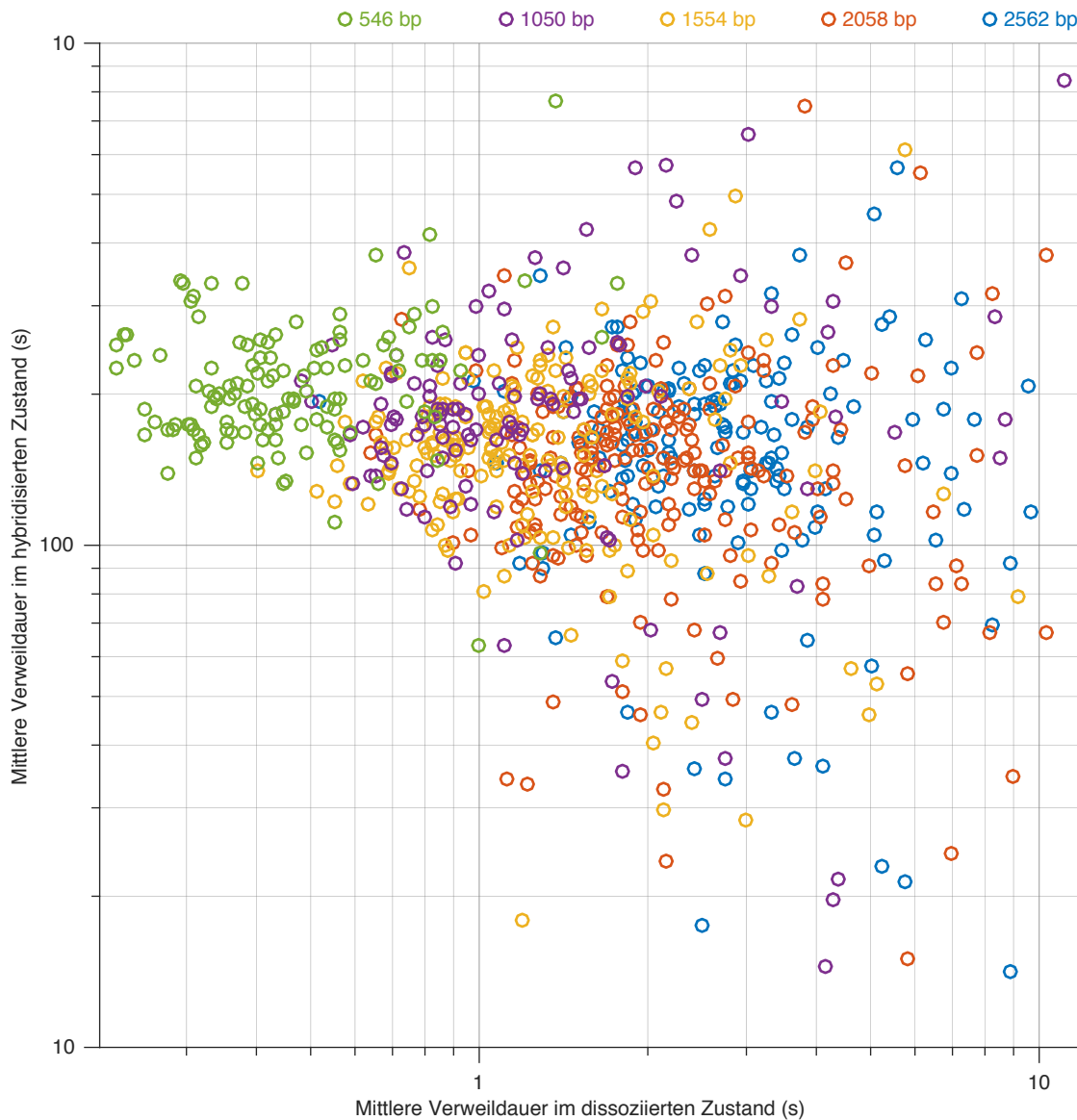


Abbildung 4.6: **Streudiagramm der mittleren Verweildauern, proximale Verankerungsgeometrie**
 Messdaten von DNA-Nano-Objekten mit der Interaktionssequenz CAGGAACG:CGTTCCTG. In jedem Datenpunkt ist für ein Objekt der jeweiligen Hauptpopulation (vgl. Abschnitt 3.3.3) die mittlere Verweildauer im hybridisierten Zustand gegen die mittlere Verweildauer im dissoziierten Zustand aufgetragen. Die Farben stehen für die Konturlänge L des Leinenstücks zwischen den Interaktionspartnern (siehe Legende oben).

ist in der distalen resp. proximalen Verankerungsgeometrie bei $L = 2562$ bp etwa acht- resp. sechsmal so groß wie bei $L = 546$ bp (Abb. 4.7 und 4.8, jeweils rechte Spalte). Der Wert von τ_1 erscheint bei kleineren Werten von L nur leicht erhöht. Zwischen dem größten und dem kleinsten Wert von τ_1 liegt ein Faktor von etwa 1,3 in der distalen Verankerungsgeometrie und etwa 1,4 in der proximalen Verankerungsgeometrie (Abb. 4.7 und 4.8, jeweils linke Spalte).

Die aus den Mengen aller einzelnen Verweildauern der Hauptpopulationen geschätzten Wahrscheinlichkeitsdichteverteilungen in Form von beschränkten Exponentialverteilungen nach Gl. (3.7) stimmen sehr gut mit den Histogrammen der entsprechenden relativen Häufigkeitsdichten überein (Abb. 4.7 und 4.8). Diese Beobachtung steht im Einklang mit der Anschauung des vorliegenden Systems als Zwei-Zustands-System und mit der damit verbundenen Annahme, dass die Lebensdauern etwaiger Zwischenzustände wie z. B. teilweiser Hybridisierung kürzer sind als die zeitliche Auflösungsgrenze der TMFM-Messung (vgl. Abschnitt 3.2.1). Die Übergänge zwischen den zwei Zuständen des Systems werden jeweils hinreichend durch eine einzelne Rate beschrieben.

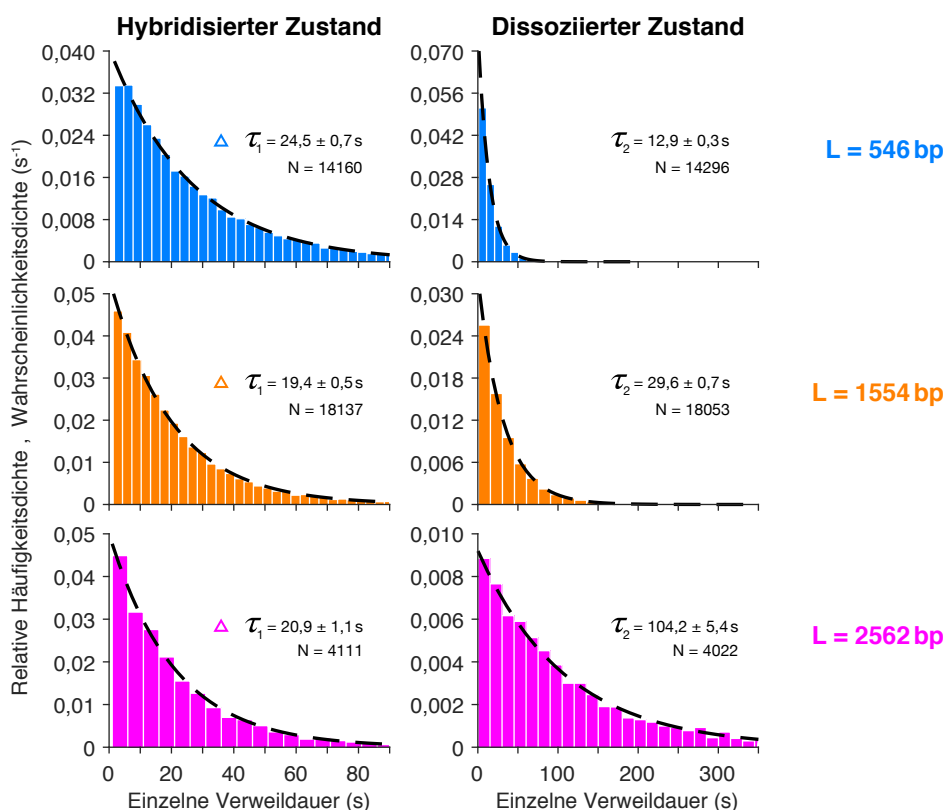


Abbildung 4.7: **Verteilungen der einzelnen Verweildauern, distale Verankerungsgeometrie**

Messdaten derselben DNA-Nano-Objekte wie in Abb. 4.5 (Interaktionssequenz CAGGAACG : CGTTCCTG). Die Konturlänge L des Leinenstücks zwischen den Interaktionspartnern steigt von oben nach unten an. Histogramme: relative Häufigkeitsdichten der einzelnen Verweildauern der jeweiligen Hauptpopulationen (vgl. Abschnitt 3.3.3). Gestrichelte Linien: Maximum-Likelihood-Schätzungen der Wahrscheinlichkeitsdichteverteilungen der jeweiligen Hauptpopulationen nach Gl. (3.7). τ_1, τ_2 : Charakteristische Verweildauern der jeweiligen Hauptpopulationen nach Gl. (3.8). Fehlerbereich $\pm 3 \sigma_{\text{bootstrap}}$ (vgl. Abschnitt 3.3.4). N : jeweilige Anzahl der Einzelwerte.

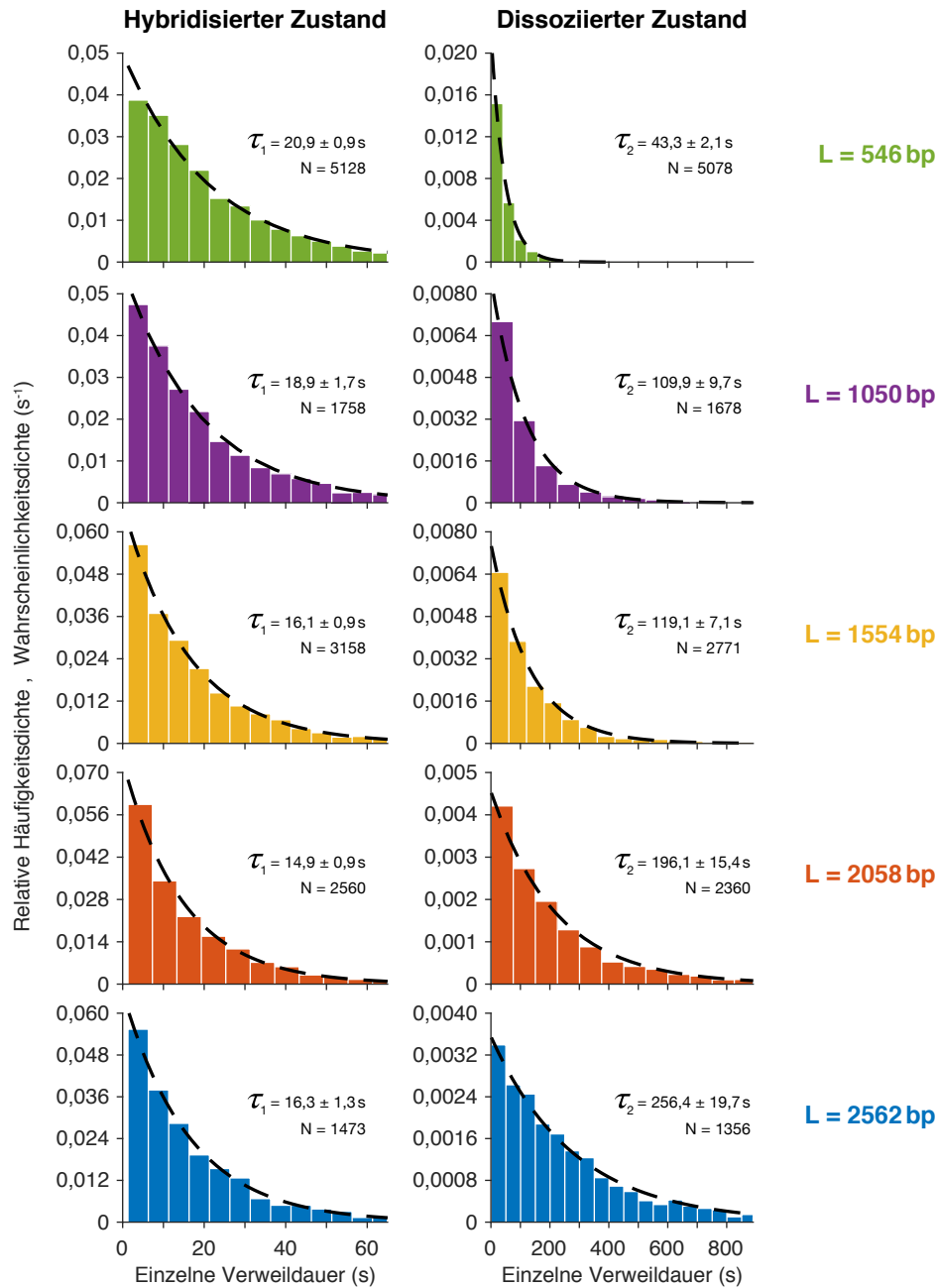


Abbildung 4.8: Verteilungen der einzelnen Verweildauern, proximale Verankerungsgeometrie
 Messdaten derselben DNA-Nano-Objekte wie in Abb. 4.6 (Interaktionssequenz CAGGAACG:CGTTCCTG). Die Konturlänge L des Leinenstücks zwischen den Interaktionspartnern steigt von oben nach unten an. Histogramme: relative Häufigkeitsdichten der einzelnen Verweildauern der jeweiligen Hauptpopulationen (vgl. Abschnitt 3.3.3). Gestrichelte Linien: Maximum-Likelihood-Schätzungen der Wahrscheinlichkeitsdichteverteilungen der jeweiligen Hauptpopulationen nach Gl. (3.7). τ_1, τ_2 : Charakteristische Verweildauern der jeweiligen Hauptpopulationen nach Gl. (3.8). Fehlerbereich $\pm 3 \sigma_{\text{bootstrap}}$ (vgl. Abschnitt 3.3.4). N: jeweilige Anzahl der Einzelwerte.

Raten

Die Assoziationsrate k_{on} und die Dissoziationsrate k_{off} berechnen sich aus den inversen charakteristischen Verweildauern:

$$k_{\text{on}} = \frac{1}{\tau_2} \quad , \quad k_{\text{off}} = \frac{1}{\tau_1} \quad (4.2)$$

Die Assoziationsrate zeigt folglich die inverse Abhängigkeit von L wie τ_2 : bei gleicher Verankerungsgeometrie nimmt k_{on} mit steigendem Wert von L ab (Abb. 4.9 a). Zudem ist k_{on} bei gleichem Wert von L in der distalen Verankerungsgeometrie größer als in der proximalen Verankerungsgeometrie, nämlich um die Faktoren 3,37 / 4,05 / 2,50 bei $L = 546 \text{ bp} / L = 1554 \text{ bp} / L = 2562 \text{ bp}$. Diese Beobachtung lässt sich wie folgt erklären: Zur Kolo-kalisation der Einzelstränge muss in der proximalen Verankerungsgeometrie das Leinenende mit der mobilen Reporter-Einheit nahe an die Probenoberfläche herankommen. Die erforderliche Nähe zur Probenoberfläche ist entropisch ungünstig, da für das Partikel-Leinen-System vergleichsweise weniger Konformationen realisierbar sind als im Fall der Schleifenbildung fernab einer sterisch einschränkenden Oberfläche [117]. Außerdem ist es denkbar, dass trotz gegebener Nähe der Interaktionspartner zueinander die Reporter-Einheiten die Assoziation der Einzelstränge zum Doppelstrang sterisch behindern. Durch die Verankerung der Interaktionspartner in der distalen Verankerungsgeometrie ist die Hybridisierung von B und B' ohne eine zwingende Nähe der mobilen

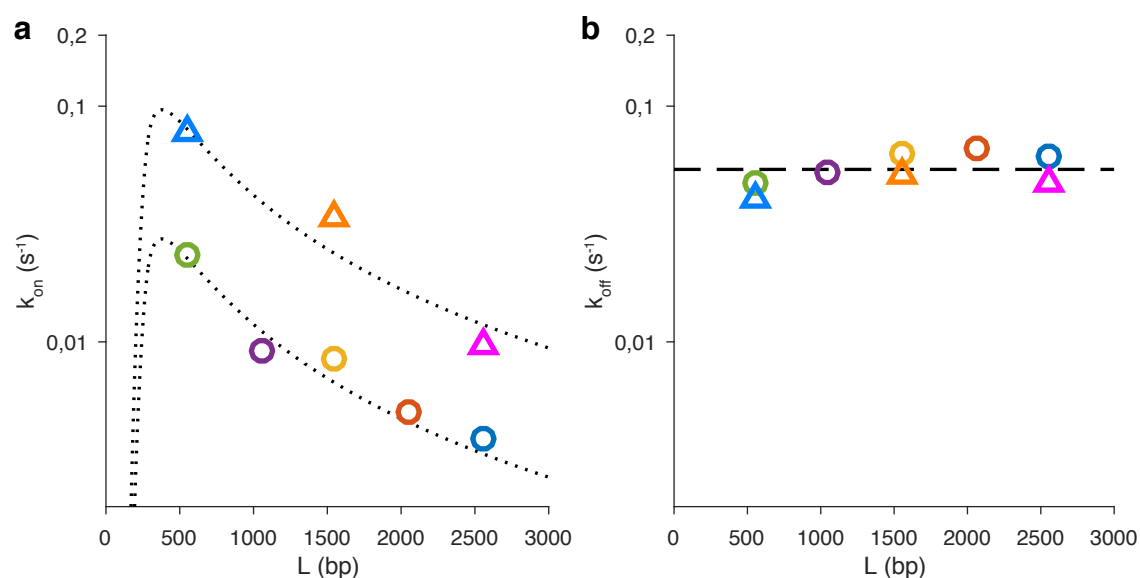


Abbildung 4.9: **Übergangsraten bei proximaler und distaler Verankerungsgeometrie**

Messdaten derselben Hauptpopulationen von DNA-Nano-Objekten wie in Abb. 4.5-4.8. (a) Assoziationsrate k_{on} und (b) Dissoziationsrate k_{off} , jeweils aufgetragen über der Konturlänge L des Leinenstücks zwischen den Interaktionspartnern. Dreiecke: distale Verankerungsgeometrie, Kreise: proximale Verankerungsgeometrie. Die statistischen Fehler ($\pm 3 \sigma_{\text{bootstrap}}$, vgl. Abschnitt 3.3.4) der Raten sind kleiner oder gleich der vertikalen Ausdehnung der Marker. Gepunktete Linien in a: Least-Squares-Fits von Gl. (4.4) an die Messwerte (Details siehe Text). Gestrichelte Linie in b: Mittelwert aus allen acht Werten von k_{off} .

Reporter-Einheit und der Verankerungspunkte von B und B' zur Probenoberfläche und der immobilen Reporter-Einheit möglich, sodass diese Störeffekte schwächer ausgeprägt sind oder gar nicht auftreten. Anders ausgedrückt: Die distale Verankerungsgeometrie ist ähnlicher zum Fall der vollkommenen Abwesenheit einer Oberfläche und der Reporter-Einheiten.

Die Abnahme der Assoziationsrate in beiden Verankerungsgeometrien steht im Einklang zur theoretisch vorhergesagten [114, 116] und experimentell beobachteten [115] Abnahme des J-Faktors mit der Länge eines Fragments doppelsträngiger DNA für Fragmentlängen $\gtrsim 500$ bp. Zur Erstellung der Vergleichskurven in Abb. 4.9 a wurde die Abschätzung für die Kolokalisationswahrscheinlichkeit der Enden eines Fragments doppelsträngiger DNA $G(0; \gamma)$ von Yamakawa & Stockmayer [114] verwendet:

$$G(0; \gamma) = \left(\frac{3}{2\pi} \right)^{\frac{3}{2}} \cdot \gamma^{-\frac{3}{2}} \cdot \left[1 - \frac{5}{8} \gamma^{-1} - \frac{79}{640} \gamma^{-2} \right] \quad (4.3)$$

Dabei steht γ (in der Notation von [114]: t) für die sogenannte reduzierte Konturlänge, den Quotienten aus Konturlänge und dem doppelten Wert der Persistenzlänge. Eine Fit-Funktion $f(\gamma)$ für den Verlauf der mit TMFM gemessenen Assoziationsraten ergibt sich durch Multiplikation von Gl. (4.3) mit einem skalaren Vorfaktor α , welcher die Wahrscheinlichkeit repräsentiert, dass während der Kolokalisation der Einzelstränge B und B' die Assoziation zum Duplex B:B' erfolgt.

$$f(\gamma) = \alpha \cdot G(0; \gamma) \quad (4.4)$$

Diese Funktion wurde durch Variation von α an die gemessenen Assoziationsraten in distaler bzw. proximaler Verankerungsgeometrie angefitet. Dabei wurde für die Persistenzlänge der Wert 50 nm eingesetzt. Im Bereich der untersuchten Werte von L stimmen die Verläufe der Fit-Kurven (gepunktete Linien in Abb. 4.9 a) in etwa mit den Verläufen der Messdaten überein. Diese Übereinstimmung ist jedoch rein qualitativ zu interpretieren. Die Vorhersage erfolgt auf Basis des WLC-Modells und für ein durchgehend doppelsträngiges DNA-Stück, welches sich frei in Lösung befindet. Das DNA-Stück zwischen den Interaktionspartnern in TMFM ist Teil eines Partikel-Leinen-Systems, welches an einer Probenoberfläche verankert ist. Aufgrund von durch die Nähe zur Oberfläche bedingten Volumenausschlusseffekten [117] ist für die DNA-Leine in TMFM eine zur in [114] angenommenen Situation unterschiedliche Dynamik zu erwarten. Darüber hinaus besitzt das Rückgrat des aus Staples bestehenden Einzelstrangs der DNA-Leine sog. *Nicks* (Einzelstrangbrüche) in regelmäßigen Abständen von 42 bp (vgl. Abschnitt 1.2.3, Abb. 1.5 b). In experimentellen Studien wurde für solche „nicked dsDNA“ eine höhere Flexibilität beobachtet als für eine Referenz von dsDNA mit intakten Rückgraten in beiden Einzelsträngen [118, 119]. Allerdings zeigten dieselben Studien, dass sich die Flexibilität der nicked dsDNA jener der intakten dsDNA mit steigender Salzkonzentration in der umgebenden Lösung immer mehr angleicht. Abschnitt 4.2.1 führt diesen Aspekt näher aus.

Im Gegensatz zu den Assoziationsraten ändern sich die gemessenen Dissoziationsraten nur wenig mit L (Abb. 4.9 b). Eine potentielle destabilisierende Wirkung des Partikel-

Leinen-Systems auf das Duplex B:B' sollte bei verschiedenen Werten von L unterschiedlich stark ausgeprägt sein (vgl. Argumentation in Abschnitt 4.1.1). Da sich die Dissoziationsraten verschiedener Objektvarianten kaum unterscheiden, scheint sich das Partikel-Leinen-System höchstens schwach destabilisierend auf das Duplex B:B' auszuwirken. Auffällig ist allerdings, dass die in der proximalen Verankerungsgeometrie gemessenen Dissoziationsraten systematisch um ca. 20 % größer sind als die bei gleichem L in der distalen Verankerungsgeometrie gemessenen Dissoziationsraten. Aufgrund des bereits diskutierten Entropieverlusts bei Nähe der mobilen Reporter-Einheit zur Oberfläche könnte eine destabilisierende Wirkung des Partikel-Leinen-Systems auf das Duplex B:B' in der proximalen Verankerungsgeometrie stärker ausgeprägt sein als in der distalen Verankerungsgeometrie. Jedoch liegt die Abweichung letztlich in derselben Größenordnung wie die Schwankung des Werts der Dissoziationsrate zwischen Messungen bei gleicher Verankerungsgeometrie und verschiedenen Werten von L. Es könnte sich dabei schlichtweg um ein Messartefakt handeln, beispielsweise Temperaturschwankungen auf einer Größenordnung von ± 1 °C zwischen den jeweiligen Messungen.

4.2 Einfluss der Konzentration von Natriumionen

Das Vorliegen von Na^+ in der umgebenden Lösung beeinflusst die Thermodynamik [120,121] und die Kinetik [34,35,45,47,50,122] der DNA-Hybridisierung. Die Abhängigkeit der Hybridisierungskinetik kurzer DNA-Doppelstränge von der Na^+ -Konzentration ist noch nicht abschließend erforscht und ist Gegenstand jüngerer Einzelmolekül-Studien [34,35]. Die Methode TMFM ermöglicht die einfache Realisierung von Einzelmolekül-Messungen der DNA-Hybridisierungskinetik bei verschiedenen Konzentrationen von Na^+ im Messpuffer. Allerdings hängt auch die Flexibilität doppelsträngiger DNA von der Na^+ -Konzentration in der umgebenden Lösung ab [118,119,123–133], weshalb eine Veränderung von $[\text{Na}^+]$ auch eine Veränderung der mechanischen Eigenschaften der DNA-Leine bewirken könnte.

4.2.1 Eine kurze Zusammenfassung relevanter Literatur

$[\text{Na}^+]$ -Abhängigkeit der Thermodynamik und Kinetik der DNA-Hybridisierung

Aus den Ergebnissen kalorimetrischer Experimente entwickelte SantaLucia [120] folgende empirische Korrekturformeln für die freie Energie ΔG° und die Entropie ΔS° der Duplexbindung von DNA-Oligonukleotiden:

$$\Delta G^\circ([\text{Na}^+]) = \Delta G^\circ(1 \text{ M NaCl}) - 0,114 \cdot \frac{N}{2} \cdot \ln([\text{Na}^+]) \cdot 1 \frac{\text{kcal}}{\text{mol}} \quad (4.5)$$

$$\Delta S^\circ([\text{Na}^+]) = \Delta S^\circ(1 \text{ M NaCl}) + 0,368 \cdot \frac{N}{2} \cdot \ln([\text{Na}^+]) \cdot 1 \frac{\text{cal}}{\text{mol} \cdot \text{K}} \quad (4.6)$$

In beiden Formeln steht N für die Gesamtzahl der im jeweiligen Doppelstrang vorhandenen Phosphatgruppen. Die Korrektur gilt im Konzentrationsbereich von Na^+ zwi-

schen 0,05 M und 1,1 M. Dabei wird angenommen, dass sich die Enthalpie ΔH° nicht mit $[\text{Na}^+]$ ändert und dass ΔH° und ΔS° näherungsweise temperaturunabhängig sind [121].

In einigen Ensemble-Studien [45,47,50,122] der Kinetik der DNA-Hybridisierung wurden die Assoziations- und Dissoziationsraten von DNA-Duplexen bei verschiedenen Konzentrationen von Na^+ im Bereich von einigen mM bis zu 3,2 M ermittelt. All diese Studien berichteten über eine starke Zunahme der Assoziationsraten mit steigender Konzentration von Na^+ . Insbesondere beobachteten Braunlin et al. [122] einen Anstieg der Assoziationsrate von etwa einer Größenordnung über den $[\text{Na}^+]$ -Bereich von 46 bis 267 mM. Dieser Anstieg verlief näherungsweise proportional zum Logarithmus von $[\text{Na}^+]$. Die Dissoziationsraten dagegen zeigten in all diesen Ensemble-Studien nur eine schwache bis gar keine Abhängigkeit von $[\text{Na}^+]$.

Bisher wurden nur wenige Einzelmolekül-Studien zur Abhängigkeit der DNA-Hybridisierungskinetik von der Kationenkonzentration veröffentlicht. Cisse et al. [34] berichteten im Jahr 2012 über eine auf Förster-Resonanz-Energie-Transfer (FRET) basierende Untersuchung an zueinander revers-komplementären, paarweise in Vesikeln isolierten Einzelsträngen. Die Hybridisierungskinetik der Interaktionspartner wurde bei verschiedenen Na^+ -Konzentrationen im Bereich von 5 bis 50 mM gemessen. Über diesen Konzentrationsbereich wurde eine Zunahme der Assoziationsrate mit $[\text{Na}^+]$ um etwa eine Größenordnung beobachtet. Dagegen veränderte sich die Dissoziationsrate im Rahmen des Messfehlers kaum. In der ebenfalls auf Einzelmolekül-FRET basierenden Studie von Dupuis et al. [35] aus dem Jahr 2013 wurde ein Einzelstrang an einer Oberfläche immobilisiert. Der dazu revers-komplementäre Einzelstrang befand sich frei in Lösung. Die Hybridisierungskinetik der Interaktionspartner wurde bei verschiedenen Na^+ -Konzentrationen im Bereich von 25 mM bis 1 M gemessen. Der Logarithmus der Assoziationsrate stieg linear mit dem Logarithmus von $[\text{Na}^+]$ an. Zwischen den Werten der Assoziationsrate bei der höchsten und bei der niedrigsten Konzentration von Na^+ lag ein Faktor von etwa 40. Die Dissoziationsrate nahm über denselben Konzentrationsbereich von Na^+ um einen Faktor von etwa 4 ab.

$[\text{Na}^+]$ -Abhängigkeit der Persistenzlänge doppelsträngiger DNA

Es steht fest, dass sich die Persistenzlänge doppelsträngiger DNA durch das Vorliegen von Na^+ in der umgebenden Lösung verringert [123–133]. Allerdings finden sich in der Literatur widersprüchliche Aussagen darüber, ob sich die Abnahme der Persistenzlänge nur bis zu einer Konzentration von Na^+ von etwa 0,1 M vollzieht, oder ob die Persistenzlänge für $[\text{Na}^+] \gtrsim 0,1 \text{ M}$ mit steigender Konzentration von Na^+ weiter abnimmt [123].

Im Kontext von TMFM sind diesbezüglich insbesondere die Ergebnisse der TPM-Messungen von Brunet et al. [132] und Guilbaud et al. [133] interessant. Darin wurde die Persistenzlänge doppelsträngiger DNA aus der mittleren quadratischen radialen Auslenkung von Partikeln am Ende von DNA-Stücken mit 1201 oder 2060 bp Konturlänge bei einer Reihe von Konzentrationen von Na^+ bis zu 3 M ermittelt. Es zeigte sich für $[\text{Na}^+] > 0,1 \text{ M}$ eine schwache, aber dennoch messbare Abnahme der Persistenzlänge mit steigender Konzentration von Na^+ . Zwischen $[\text{Na}^+] = 0,1 \text{ M}$ und $[\text{Na}^+] = 1 \text{ M}$ verringerte sich die Persistenzlänge um etwa 10 % [133].

Salzabhängigkeit der Flexibilität von dsDNA mit Nicks

Die Flexibilität eines DNA-Doppelstrangs kann sich durch ein Aufklappen an den Positionen von Nicks (Einzelstrangbrüchen) lokal deutlich erhöhen [118,119]. Daher zeigt dsDNA mit Nicks im direkten Vergleich eine deutlich höhere Flexibilität als eine entsprechende Referenz ohne Nicks [118]. Jedoch kann sich durch die Stapelwechselwirkung zwischen den benachbarten Basen auch an den Nicks die helikale Form des Doppelstrangs ausformen. Diese Konformation unterscheidet sich von einer vollkommen intakten Doppelhelix dann lediglich durch die Lücke im Rückgrat eines der Einzelstränge, aber nicht durch die relative Orientierung der benachbarten Basen. Mit zunehmender Salzkonzentration in der umgebenden Lösung verschiebt sich das Gleichgewicht der Stapelwechselwirkung immer mehr hin zum gestapelten Zustand [119]. Dementsprechend wurde beobachtet, dass sich die Flexibilität von dsDNA mit Nicks bei zunehmender Salzkonzentration verringert und sich dadurch der Flexibilität der Referenz ohne Nicks annähert [118].

4.2.2 Mögliche Änderung des J-Faktors mit der Konzentration von Na^+

Der J-Faktor skaliert mit der Kolokalisationswahrscheinlichkeit der Interaktionspartner B und B'. Für diese Wahrscheinlichkeit spielt die Flexibilität der DNA-Leine eine maßgebliche Rolle.

Das Rückgrat des aus Staples bestehenden Einzelstrangs der Leine besitzt Nicks in regelmäßigen Abständen von 42 bp. Den Ergebnissen von [118] und [119] nach zu urteilen sollte sich mit steigender Konzentration von Na^+ immer häufiger der gestapelte Zustand der benachbarten Basen an den Nicks einstellen. Dadurch sollte die Persistenzlänge der DNA-Leine immer mehr gegen die Persistenzlänge des Doppelstrangs in der idealen Konformation mit dauerhafter Stapelung aller benachbarten Basen streben. Dagegen geben insbesondere die Ergebnisse von [132] und [133] Anlass zur Annahme, dass sich dieser obere Grenzwert mit steigender Konzentration von Na^+ immer weiter verringert. Es ist denkbar, dass sich die beiden Effekte zumindest teilweise ausgleichen. Eine nicht verschwindende $[\text{Na}^+]$ -Abhängigkeit der Flexibilität der DNA-Leine, und damit des J-Faktors, würde jedoch die Messung der $[\text{Na}^+]$ -Abhängigkeit der Hybridisierungskinetik des Interaktionspartnerpaares verfälschen.

4.2.3 TMFM-Messungen bei einer Reihe von Na^+ -Konzentrationen

Die Hybridisierungskinetik des 8-bp-Duplex CAGGAACG:CGTTCCTG wurde bei drei verschiedenen Konzentrationen von Na^+ im Messpuffer und mit drei verschiedenen Konturlängen L des Leinenstücks zwischen den Interaktionspartnern gemessen. An Proben von DNA-Nano-Objekten mit $L = 546$ bp und $L = 1554$ bp wurden jeweils TMFM-Messungen bei $[\text{Na}^+] = 150$ mM, $[\text{Na}^+] = 500$ mM und $[\text{Na}^+] = 1$ M durchgeführt. An Proben von DNA-Nano-Objekten mit $L = 2562$ bp wurden TMFM-Messungen bei $[\text{Na}^+] = 500$ mM und $[\text{Na}^+] = 1$ M durchgeführt. Details zur Versuchsdurchführung finden sich in Anhang A.2.4.

[Na⁺]-Abhängigkeit der Verweildauern im dissoziierten und hybridisierten Zustand

Aus den TMFM-Daten jeder Messung wurden Zustandsabfolgen und Verweildauerstatistiken der einzelnen DNA-Nano-Objekte extrahiert. Die Streudiagramme der mittleren Verweildauern je Objekt (Abb. 4.10 a) zeigen, dass sich die mittleren Verweildauern im dissoziierten Zustand mit zunehmender Konzentration von Na⁺ deutlich verkürzen. Im Gegensatz dazu ändern sich die mittleren Verweildauern im hybridisierten Zustand kaum mit [Na⁺]. Diese beiden Beobachtungen gelten für alle Werte von L und spie-

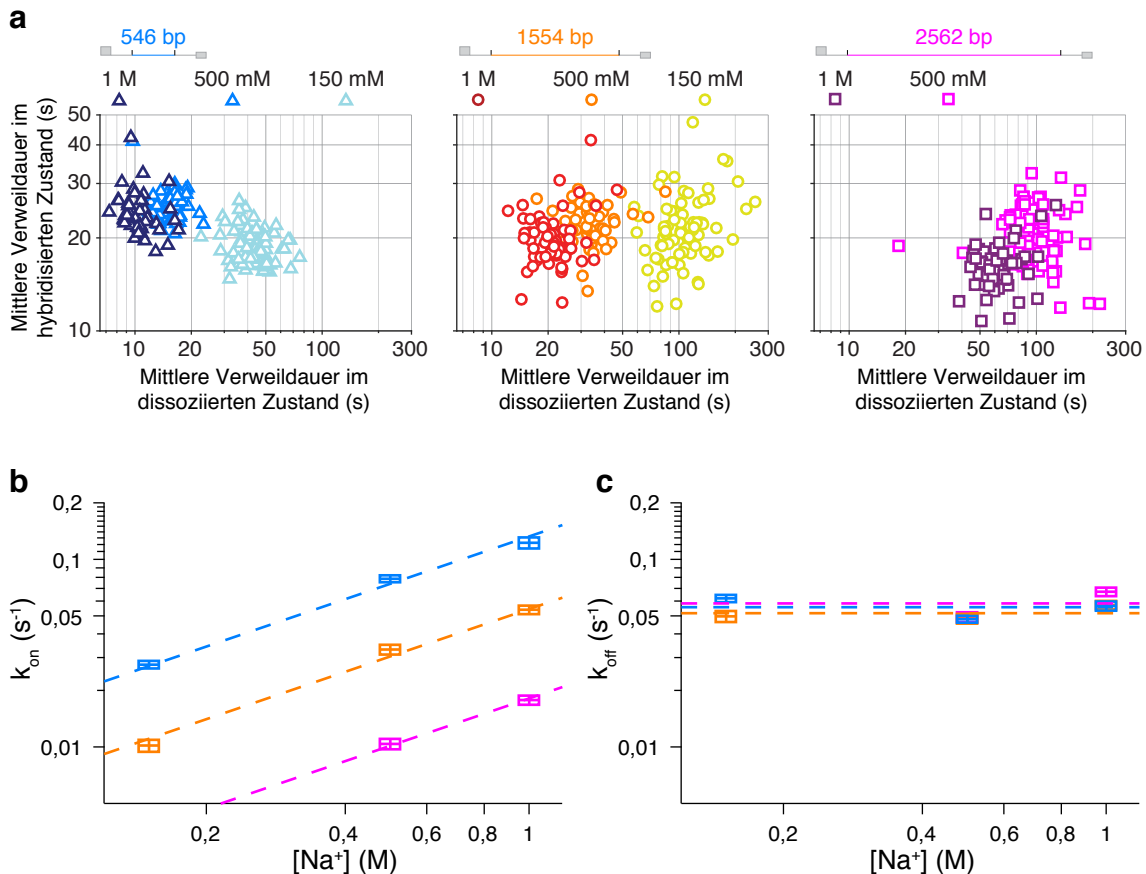


Abbildung 4.10: **[Na⁺]-Abhängigkeit der mittleren Verweildauern je Objekt sowie der globalen Raten** Messdaten von Objekten mit der Interaktionssequenz CAGGAACG:CGTTCCTG, mit unterschiedlichen Werten der Konturlänge L zwischen den Interaktionspartnern und mit distaler Verankerungsgeometrie. **(a)** Streudiagramme der mittleren Verweildauern im hybridisierten Zustand gegen die mittleren Verweildauern im dissoziierten Zustand der einzelnen Objekte der Hauptpopulation (vgl. Abschnitt 3.3.3) für L = 546 bp (links), L = 1554 bp (Mitte) und L = 2562 bp (rechts), jeweils bei verschiedenen Konzentrationen von Na⁺ (beachte Legenden oberhalb der Diagramme). **(b)** Assoziationsrate k_{on} gegen die Konzentration von Na⁺ für L = 546 bp (blau), L = 1554 bp (orange) und L = 2562 bp (violett). Doppelt-logarithmische Achsenskalierung. Breite der Marker: $\pm 5\%$ Fehler in [Na⁺] wegen Pipettierfehlern und möglicher Verdunstung der Probe. Höhe der Marker: statistischer Fehler der Rate ($\pm 3 \sigma_{bootstrap}$, vgl. Abschnitt 3.3.4). Gestrichelte Linien: Schar von Fit-Geraden an eine lineare Beziehung zwischen $\ln(k_{on} \cdot 1 \text{ s})$ und $\ln([Na^+]/1 \text{ M})$ unter der Annahme eines für jeden Wert von L konstanten J-Faktors (siehe Text). **(c)** Dissoziationsrate k_{off} gegen die Konzentration von Na⁺. Achsenskalierung, Farben und Marker wie in b. Gestrichelte Linien: Mittelwerte der bei den verschiedenen Konzentrationen von Na⁺ gemessenen Dissoziationsraten jeder Objektvariante.

geln sich auch in den Verläufen der Assoziations- und Dissoziationsraten mit $[\text{Na}^+]$ wider¹. Die Assoziationsrate wächst monoton mit $[\text{Na}^+]$ (Abb. 4.10 b). Sie erhöht sich von $[\text{Na}^+] = 150 \text{ mM}$ bis $[\text{Na}^+] = 1 \text{ M}$ um einen Faktor von etwa 5,3 (bei $L = 1554 \text{ bp}$) bzw. 4,5 (bei $L = 546 \text{ bp}$). Der Verlauf der Dissoziationsrate mit $[\text{Na}^+]$ zeigt keinen klaren Trend (Abb. 4.10 c), und zwischen dem höchsten und niedrigsten gemessenen Wert dieser Rate liegt lediglich ein Faktor von etwa 1,4.

Auf Gl. (4.1) aufbauend kann man die beiden Raten, den J-Faktor und ΔG° wie folgt zueinander in Bezug setzen:

$$\frac{k_{\text{on}}}{k_{\text{off}}} = K_{\text{eq,uni}} = J \cdot K_{\text{eq,bi}} \quad (4.7a)$$

$$\Rightarrow \ln(k_{\text{on}}) - \ln(k_{\text{off}}) = \ln\left(\frac{J}{1 \text{ M}}\right) + \ln(K_{\text{eq,bi}} \cdot 1 \text{ M}) \quad (4.7b)$$

$$\Leftrightarrow \ln(k_{\text{on}}) - \ln(k_{\text{off}}) = \ln\left(\frac{J}{1 \text{ M}}\right) - \frac{1}{k_{\text{B}}T} \cdot \Delta G^\circ \quad (4.7c)$$

Da sich die in den TMFM-Experimenten gemessene Dissoziationsrate fast nicht mit $[\text{Na}^+]$ ändert, dominiert die $[\text{Na}^+]$ -Abhängigkeit von $\ln(k_{\text{on}})$ die $[\text{Na}^+]$ -Abhängigkeit der linken Seite von Gl. (4.7c). Nähme man an, dass sich der J-Faktor im untersuchten Konzentrationsbereich nicht maßgeblich mit $[\text{Na}^+]$ ändert, so bestünde laut der Korrekturformel von SantaLucia [120] (Gl. (4.5)) eine lineare Beziehung zwischen $\ln(k_{\text{on}})$ und $\ln([\text{Na}^+])$. Die Steigung dieser linearen Beziehung müsste für jeden Wert von L gleich sein. Der J-Faktor ginge lediglich in den Achsenabschnitt von $\ln(k_{\text{on}}(\ln([\text{Na}^+])))$ ein.

Eine Schar von drei Fit-Geraden derselben Steigung wurde per Least-Squares-Fit an die Verläufe von $\ln(k_{\text{on}})$ über $\ln([\text{Na}^+])$ für die drei Objektvarianten mit $L = 546 \text{ bp}$, $L = 1554 \text{ bp}$ und $L = 2562 \text{ bp}$ angepasst (Abb. 4.10 b). Die Fit-Geraden liegen im Rahmen der experimentellen Fehler nahe an den jeweiligen Verläufen von $\ln(k_{\text{on}})$ über $\ln([\text{Na}^+])$. Allerdings entsteht für die Datensätze mit $L = 546 \text{ bp}$ und $L = 1554 \text{ bp}$ der Eindruck, dass leicht rechtsgekrümmte Kurven den jeweiligen Verläufen von $\ln(k_{\text{on}})$ über $\ln([\text{Na}^+])$ noch besser entsprechen würden als die hier erstellten Fit-Geraden.

Beweglichkeit der mobilen Reporter-Einheiten im dissoziierten Zustand bei verschiedenen Konzentrationen von Na^+

Die Beweglichkeit der mobilen Reporter-Einheiten im dissoziierten Zustand wurde stichprobenartig untersucht. Die Stichproben umfassten jeweils $N = 3000$ Bildperioden. Als Maß für die Beweglichkeit wurde die mittlere quadratische radiale Auslenkung $\langle R^2 \rangle$ der gesamten Stichprobe betrachtet:

$$\langle R^2 \rangle = \frac{1}{N} \sum_{n=1}^N (\Delta x_{\text{RE}}(n)^2 + \Delta y_{\text{RE}}(n)^2) \quad (4.8)$$

¹Auch für die Messdaten der $[\text{Na}^+]$ -Reihe wiesen die Verweildauerverteilungen in beiden Zuständen eine hohe Ähnlichkeit zu Exponentialverteilungen auf. Die Systeme lassen sich demnach hinreichend als Zwei-Zustands-Systeme mit einer einzelnen Rate je Übergang beschreiben.

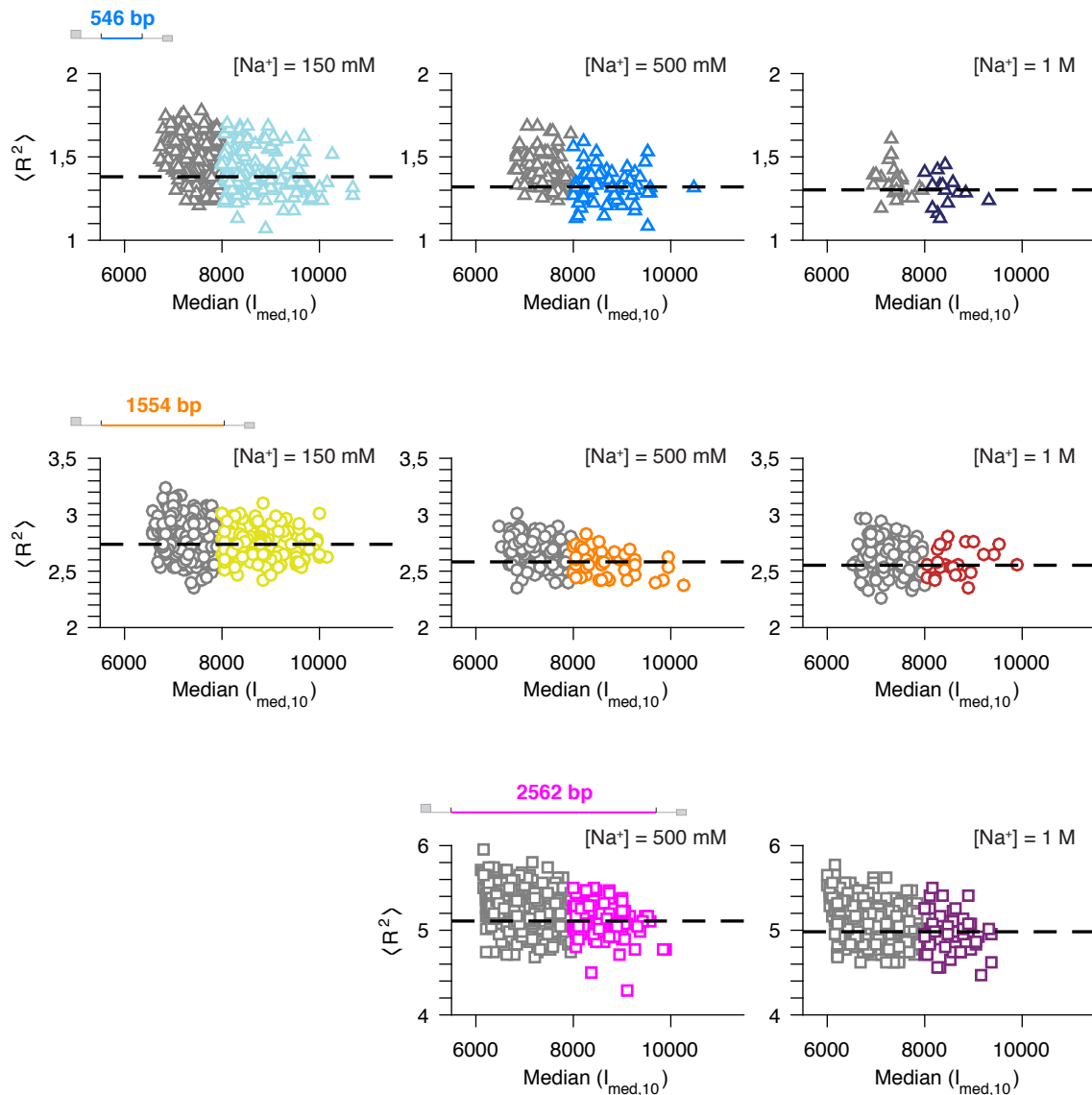


Abbildung 4.11: **Mittlere quadratische Auslenkung der mobilen Reporter-Einheiten im dissoziierten Zustand in Abhängigkeit von $[Na^+]$** für dieselben DNA-Nano-Objekte wie in Abb. 4.10, mit der Interaktionssequenz CAGGAACG:CGTTCCTG, mit unterschiedlichen Werten der Konturlänge L zwischen den Interaktionspartnern und mit distaler Verankerungsgeometrie. Jedes Streudiagramm zeigt die Daten für einen Wert von L (obere Zeile: $L = 546$ bp, mittlere Zeile: $L = 1554$ bp, untere Zeile: $L = 2562$ bp) und eine Konzentration von Na^+ (in jedem Diagramm oben rechts angegeben). Jeder Marker repräsentiert eine Stichprobe von 3000 Bildperioden aus den Auslenkungs- und Helligkeitsabfolgen einer mobilen Reporter-Einheit im dissoziierten Zustand. Aufgetragen ist jeweils die mittlere quadratische radiale Auslenkung $\langle R^2 \rangle$ nach Gl. (4.8) der Stichprobe gegen den Median der entsprechenden 3000 Werte von $l_{med,10}$ (siehe Gl. (2.3)) der Stichprobe. Graue Marker: Stichproben mit $Median(l_{med,10}) < 8000$. Farbige Marker: Stichproben mit $Median(l_{med,10}) \geq 8000$. Gestrichelte Linien: Mittelwerte $\langle R^2 \rangle$ von $\langle R^2 \rangle$ aller Stichproben mit $Median(l_{med,10}) \geq 8000$ bei den entsprechenden Werten von L und $[Na^+]$.

Innerhalb jedes einzelnen Datensatzes für einen Wert von L und eine Konzentration von Na^+ weist $\langle R^2 \rangle$ eine breite Streuung auf (Abb. 4.11). Insbesondere zeigt sich auch hier der bereits in Kapitel 2 diskutierte Zusammenhang zwischen der Helligkeit und der gemessenen Beweglichkeit der Reporter-Einheiten: Bei geringerer mittlerer Helligkeit und damit schlechterem Signal-zu-Rausch-Verhältnis ist die Positionsabschätzung weniger genau. Infolgedessen wirkt die Beweglichkeit höher, $\langle R^2 \rangle$ nimmt zu.

Im Vergleich der Streudiagramme für die Datensätze mit gleichem Wert von L und verschiedenen Konzentrationen von Na^+ (Zeilen in Abb. 4.11) zeigt sich ein deutlicher Trend: Im Mittel nimmt $\langle R^2 \rangle$ mit steigender Konzentration von Na^+ ab. Für einen quantitativen Vergleich wurde je Datensatz aus allen Stichproben mit einer mittleren Helligkeit von mindestens 8000 das arithmetische Mittel von $\langle R^2 \rangle$ gebildet. Dieser Mittelwert wird im Folgenden als $\overline{\langle R^2 \rangle}$ bezeichnet. Zwischen $[\text{Na}^+] = 150 \text{ mM}$ und $[\text{Na}^+] = 500 \text{ mM}$ nimmt $\overline{\langle R^2 \rangle}$ für $L = 546 \text{ bp}$ um etwa 4,4 % und für $L = 1554 \text{ bp}$ um etwa 5,7 % ab. Zwischen $[\text{Na}^+] = 500 \text{ mM}$ und $[\text{Na}^+] = 1 \text{ M}$ nimmt $\overline{\langle R^2 \rangle}$ für $L = 546 \text{ bp}$ um etwa 1,4 %, für $L = 1554 \text{ bp}$ um etwa 1,2 % und für $L = 2562 \text{ bp}$ um etwa 2,5 % ab.

4.2.4 Diskussion und Fazit

Die Abnahme von $\overline{\langle R^2 \rangle}$ mit $[\text{Na}^+]$ deutet auf eine Zunahme der Flexibilität der DNA-Leine mit $[\text{Na}^+]$ hin. Mit einer Zunahme der Leinenflexibilität könnte eine Zunahme des J-Faktors einhergehen. Der gemessene Verlauf der Assoziationsrate (Abb. 4.10 b) würde dann eine Überlagerung der $[\text{Na}^+]$ -Abhängigkeit der Assoziationskinetik des Duplex B:B' mit der $[\text{Na}^+]$ -Abhängigkeit des J-Faktors darstellen. Durch Korrektur um den J-Faktor-bedingten Anteil der Zunahme von k_{on} erhielte man die reine Abhängigkeit der bimolekularen Duplex-Assoziationsrate $k_{\text{on,bi}}$ von $[\text{Na}^+]$. Der Verlauf von $k_{\text{on,bi}}$ mit $[\text{Na}^+]$ wäre dann flacher als der in Abb. 4.10 b gezeigte Verlauf von k_{on} mit $[\text{Na}^+]$. Je nach Form der $[\text{Na}^+]$ -Abhängigkeit des J-Faktors ergäbe sich möglicherweise keine lineare Beziehung zwischen $\ln(k_{\text{on,bi}})$ und $\ln([\text{Na}^+])$.

Ebenso ist grundsätzlich denkbar, dass der gemessene Verlauf der Dissoziationsrate (Abb. 4.10 c) eine Überlagerung der $[\text{Na}^+]$ -Abhängigkeit der Dissoziationskinetik des Duplex B:B' mit der $[\text{Na}^+]$ -Abhängigkeit der Flexibilität des Leinenstücks zwischen den Interaktionspartnern darstellt. Eine mögliche destabilisierende Kraft der Schlaufe auf das Duplex B:B' könnte mit zunehmender Flexibilität der DNA-Leine immer schwächer ausgeprägt sein. Die Dissoziation würde folglich bei niedrigen Konzentrationen von Na^+ durch die Schlaufe im Mittel stärker beschleunigt als bei hohen Konzentrationen von Na^+ . In diesem Fall müsste die „J-Faktor-bereinigte“ Dissoziationsrate des Duplex B:B' mit $[\text{Na}^+]$ zunehmen, damit sich in Summe die beobachtete annähernde Unabhängigkeit der TMFM-Dissoziationsrate k_{off} von $[\text{Na}^+]$ ergäbe. Eine Zunahme der Duplex-Dissoziationsrate mit $[\text{Na}^+]$ stünde jedoch im Widerspruch zu bisherigen experimentellen Beobachtungen [34, 35, 45, 47, 50, 122].

Um die Frage nach einer möglichen $[\text{Na}^+]$ -Abhängigkeit des J-Faktors in der TMFM-Methode eindeutig zu beantworten, sind weitere Experimente notwendig. Das relativ geringe Ausmaß der Änderung von $\overline{\langle R^2 \rangle}$ mit $[\text{Na}^+]$ erlaubt die Annahme, dass sich der J-Faktor im Bereich zwischen $[\text{Na}^+] = 150 \text{ mM}$ und $[\text{Na}^+] = 1 \text{ M}$ höchstens schwach

ändert. Diese ist die einfachste Annahme, unter welcher die hier präsentierten Messergebnisse der $[\text{Na}^+]$ -Abhängigkeit der Duplex-Kinetik im Einklang zur etablierten Vorhersage der $[\text{Na}^+]$ -Abhängigkeit der Duplex-Stabilität nach SantaLucia [120, 121] stehen.

Zum Abschluss dieses Abschnitts sei darauf hingewiesen, dass die zwei beteiligten Einzelstränge aufgrund ihrer speziellen Basensequenzen schwache Sekundärstrukturen ausbilden können. In Abschnitt 4.3 wird näher darauf eingegangen, wie sich das Auftreten solcher Sekundärstrukturen auf die gemessene Duplexkinetik auswirken kann. Inwieweit die $[\text{Na}^+]$ -Abhängigkeit der Stabilität von Sekundärstrukturen mit der $[\text{Na}^+]$ -Abhängigkeit der Duplexkinetik zusammenspielen könnte, wurde im Rahmen der vorliegenden Arbeit nicht untersucht.

4.3 Einfluss der Basenparsequenz

Der Zusammenhang zwischen Basenparsequenz und DNA-Hybridisierungskinetik war bereits Gegenstand mehrerer Ensemble-Studien [24, 28–30, 33, 42, 134]. Einige dieser Arbeiten untersuchten speziell den Einfluss der Bildung von Sekundärstrukturen auf die Duplex-Assoziationsrate [28–30, 42, 134]. Sekundärstrukturen können durch Basenpaarung zwischen revers-komplementären Abschnitten innerhalb von Einzelsträngen entstehen, wie die linke Hälfte des Schemas in Abb. 4.12 veranschaulicht. Aufgrund ihrer haarnadelähnlichen Form werden solche Sekundärstrukturen auch als *Hairpins* bezeichnet. Die Bildung von Hairpins konkurriert mit der Ausformung des Doppelstrangs (siehe Abb. 4.12). Der Zustand mit dissoziierten Einzelsträngen ohne jegliche Sekundärstrukturen sei als S_0 bezeichnet. Aus diesem Zustand können sich entweder Hairpins in den Einzelsträngen formieren (H1, H2), oder die Hairpin-freien Einzelstränge assoziieren über einen Nukleationszustand mit wenigen Basenpaaren (N) zu einem vollständigen Duplex (D). In den Hairpin-behafteten Zuständen sind nicht alle Basen der Einzelstränge sterisch zugänglich, sodass die Ausformung von inter-Strang-Basenpaaren gehemmt wird. Erst durch Aufschmelzen aller Hairpins würde der Zustand S_0 und mit ihm die volle Fähigkeit zur Nukleation wiederhergestellt. Es sind jedoch auch Reaktionspfade zwischen den Hairpin-behafteten Zuständen und dem Duplex-Zustand denkbar, die kein zwischenzeitliches Verweilen im Zustand S_0 erfordern (Abb. 4.12, in Grau angedeutet). Aus den sterisch noch zugänglichen ungepaarten Basen der Reaktionspartner könnte sich rein theoretisch ein Nukleus bilden, woraufhin die verbleibenden intra-Strang-Basenpaare durch Bildung weiterer inter-Strang-Basenpaare verdrängt werden könnten [30, 42, 135]. Allerdings ist anzunehmen, dass eine so geartete Duplex-Assoziation im Mittel langsamer vonstatten gehen würde als die Duplex-Assoziation aus dem Zustand S_0 .

Verschiedene quantitative Studien [24, 28, 29, 33, 42] offenbarten die starke Abhängigkeit der Duplex-Assoziationskinetik von der Basenparsequenz. In Vergleichsstudien von Duplexvarianten ähnlicher Schmelztemperatur, aber unterschiedlicher Basenparsequenz, umspannte der Bereich der gemessenen Assoziationsraten über zwei Größenordnungen [28, 29, 42]. Dabei sind mit der Duplex-Assoziation konkurrierende Hairpins offenbar nicht die einzige Ursache heterogener Assoziationskinetik. Auch für Duplexe aus

Einzelsträngen mit einer vernachlässigbaren Neigung zur Bildung von Hairpins sind Unterschiede von über einer Größenordnung in den Assoziationsraten möglich [24, 33]. Demnach scheint auch die Duplex-Assoziationsrate aus dem sekundärstrukturfreien Zustand S_0 von der Basenpaarsequenz abzuhängen. Die zu Grunde liegenden Regelmäßigkeiten und Mechanismen sind allerdings noch nicht vollkommen verstanden [33, 42]. Zu ihrer Ergründung bedarf es weiterer systematischer Studien. Eine Schwierigkeit besteht jedoch darin, Duplexvarianten zu finden, die den Sequenzraum möglichst vollständig abdecken, deren Einzelstränge aber gleichzeitig möglichst frei von Sekundärstrukturen sind. Auf Basis der bisher vorhandenen experimentellen Daten wurden bereits empirische Ansätze zur quantitativen Vorhersage der Duplex-Assoziationsrate in Abhängigkeit von der Basenpaarsequenz entwickelt [24, 42], welche in Abschnitt 4.3.2 näher thematisiert werden.

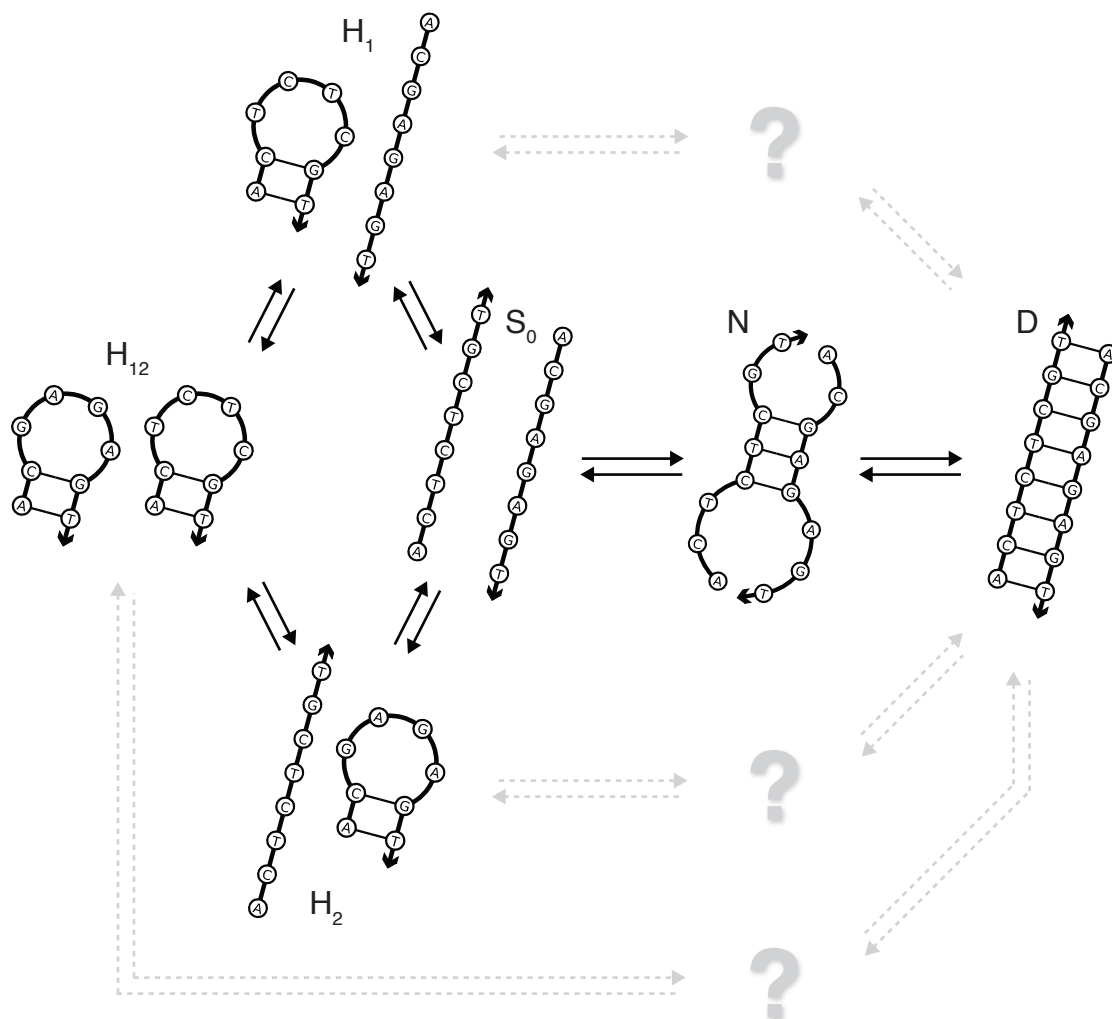


Abbildung 4.12: **Konkurrenz zwischen Hairpins und Duplexbindung**

S_0 : Zustand ohne jegliche Basenpaarung. N: Nukleationszustand. D: Komplett hybridisierter Doppelstrang. H1, H2: Zustände mit einem Hairpin in einem der Einzelstränge. H12: Zustand mit Hairpins in beiden Einzelsträngen. In Grau angedeutet: Theoretisch denkbare Reaktionspfade der Duplex-Assoziation aus den Zuständen H1, H2 oder H12 ohne zwischenzeitliches Verweilen im Zustand S_0 .

Im Gegensatz zur Abhängigkeit der Assoziationsraten von der Basenpaarsequenz liegen bis dato kaum experimentelle Daten zur Abhängigkeit der Dissoziationsraten von der Basenpaarsequenz vor. Ergebnisse von Simulationen deuten auf einen Mechanismus hin, durch den Hairpins die Duplex-Dissoziation begünstigen könnten [38]: Zwischenzustände eines bereits partiell aufgeschmolzenen Duplex könnten durch Bildung von intra-Strang-Basenpaaren stabilisiert werden. Dadurch würde eine erneute Bildung von inter-Strang-Basenpaaren verhindert und der Übergang zur vollständigen Dissoziation des Duplex begünstigt. Allerdings ist dem Verfasser dieser Dissertation keine Veröffentlichung bekannt, die von der experimentellen Beobachtung dieses Effekts berichtet.

4.3.1 Die sieben Sequenzvarianten

Teil der hier präsentierten Arbeit war eine TMFM-Studie der Assoziations- und Dissoziationskinetik von sieben verschiedenen Duplex-Sequenzvarianten (siehe Tabelle 4.1). Alle diese Varianten haben eine Länge von acht Basenpaaren und wurden nach dem Kriterium ausgewählt, dass sich die freien Energien ihrer Duplexbindung laut Vorhersage mit dem Nearest-Neighbor-Modell ΔG_{NN} kaum unterscheiden². Dafür unterscheiden sich die Varianten in ihrem relativen Anteil von G:C-Basenpaarungen, in der Anzahl möglicher Hairpins und in den jeweiligen freien Energien der Hairpin-Bildung². Als Maßzahl für die Neigung zur Ausbildung von Hairpins wird hier die Größe p_{Hairpin} als die Wahrscheinlichkeit definiert, dass in mindestens einem der Einzelstränge ein mit der Duplex-Assoziation konkurrierendes Hairpin vorliegt. Bezeichne s_0 den Zustand frei von jeglichen Hairpins für Strang B und s'_0 den Zustand frei von jeglichen Hairpins für Strang B'. Setzt man die freien Energien der Zustände s_0 und s'_0 in den energetischen Nullpunkt für die jeweiligen Einzelstränge, so folgt:

$$\begin{aligned} p_{\text{Hairpin}} &= 1 - p(s_0) \cdot p(s'_0) \\ &= 1 - \left(1 + \sum_{m=1}^M \exp\left(\frac{\Delta G_{NN,m}}{k_B T}\right) \right)^{-1} \cdot \left(1 + \sum_{m'=1}^{M'} \exp\left(\frac{\Delta G_{NN,m'}}{k_B T}\right) \right)^{-1} \end{aligned} \quad (4.9)$$

In Gleichung (4.9) bezeichnet

- M die Anzahl der berücksichtigten Hairpins in Strang B,
- M' die Anzahl der berücksichtigten Hairpins in Strang B',
- $\Delta G_{NN,m}$ die freie Energie der Bildung des m-ten Hairpins² in Strang B und
- $\Delta G_{NN,m'}$ die freie Energie der Bildung des m'-ten Hairpins² in Strang B'.

Bei der Berechnung von p_{Hairpin} wurden nur Hairpins berücksichtigt, bei denen mindestens zwei Basen aus der acht Basen langen Interaktionssequenz Teil des Hairpin-Stammes sind. Weitere mögliche Hairpins wurden vernachlässigt, weil erstens ihr Auftreten sehr unwahrscheinlich ist und sie zweitens nicht zwingend eine für die Duplex-

²Die ΔG_{NN} sowie die freien Energien der Hairpin-Bildung wurden mithilfe des DINAmelt Servers [95] vorhergesagt, für eine Reaktionstemperatur von 23 °C, $[\text{Na}^+] = 0,5 \text{ M}$ und $[\text{Mg}^{++}] = 0 \text{ M}$. In all diesen Vorhersagen wurden die Sequenzen inklusive T₁₂-Verbinder eingegeben.

Assoziation hinderliche Sekundärstruktur darstellen. Abbildung C.13 illustriert dieses Auswahlkriterium am Beispiel der Hairpins von Strang B' der Variante v3.

Die Varianten lassen sich anhand von p_{Hairpin} grob in zwei Gruppen aufteilen: Die Varianten v1 bis v4 bilden eher schwache Hairpins mit geringer Besetzungswahrscheinlichkeit aus. Für diese Varianten ist p_{Hairpin} somit niedrig (vgl. Tabelle 4.1). In den Varianten v5 bis v7 können sich deutlich stabilere Hairpins ausformen, weshalb p_{Hairpin} für diese Varianten deutlich höher ausfällt. Besonders stabil sind dabei das GAGA-Tetraloop-Motiv [121,136] in der Variante v6 und das GCA-Triloop-Motiv [121,137] in der Variante v7. Die Variante v5 weist bezüglich der Hairpins eine Besonderheit auf, die sie von allen anderen Varianten unterscheidet: Die Stämme der möglichen Hairpins in Strang B dieser Variante entstehen ausschließlich unter Beteiligung des T_{12} -Verbinders, indem die zwei aufeinanderfolgenden Adenine der Interaktionssequenz Basenpaarungen mit Thyminen an verschiedenen Stellen des Verbinders eingehen (siehe Abb. C.14).

Tabelle 4.1: **Überblick über die Eigenschaften der Sequenzvarianten**

Die Sequenzen sind in Richtung 5'-Ende nach 3'-Ende angegeben. Beim prozentualen GC-Gehalt werden die 8 bp langen Interaktionssequenzen, nicht jedoch die T_{12} -Verbinder berücksichtigt. Die unterstrichenen Sequenzabschnitte in Varianten v2, v5, v6 und v7 gehen intra-Strang-Basenpaarungen ein. ΔG_{NN} : Nach dem Nearest-Neighbor-Modell mithilfe des DINAmelt Servers [95] vorhergesagte Gibbs'sche freie Energie der Hybridisierung für eine Reaktionstemperatur von 23 °C, $[\text{Na}^+] = 0,5 \text{ M}$ und $[\text{Mg}^{++}] = 0 \text{ M}$. p_{Hairpin} : Wahrscheinlichkeit, dass in mindestens einem der Einzelstränge ein mit der Duplex-Assoziation konkurrierendes Hairpin vorliegt (Details zur Berechnung: siehe Text und Gl. (4.9)).

Variante	Sequenz B	Sequenz B'	% G:C	ΔG_{NN} (kcal/mol)	p_{Hairpin}
v1	$T_{12}\text{CCTCCTCC}$	$T_{12}\text{GGAGGAGG}$	75	-11,8	0,027
v2	$T_{12}\text{TACGCCAT}$	$T_{12}\text{ATGGCGTA}$	50	-11,7	0,160
v3	$T_{12}\text{ACGCATCA}$	$T_{12}\text{TGATGCGT}$	50	-11,7	0,047
v4	$T_{12}\text{CGCGATAT}$	$T_{12}\text{ATATCGCG}$	50	-11,6	0,003
v5	$T_{12}\text{CAGGAACG}$	$T_{12}\text{CGTTCCTG}$	62,5	-11,7	0,330
v6	$T_{12}\text{ACTCTCGT}$	$T_{12}\text{ACGAGAGT}$	50	-11,6	0,954
v7	$T_{12}\text{ACGCAGTA}$	$T_{12}\text{TACTGCGT}$	50	-11,7	0,963

4.3.2 Hybridisierungskinetik der Sequenzvarianten

Vorüberlegungen

Vor der Betrachtung der TMFM-Messergebnisse sollen zunächst Hypothesen diskutiert werden, wie sich die Formation von Hairpins in den Einzelsträngen quantitativ auf die gemessene Duplex-Hybridisierungskinetik auswirken könnte. In den meisten möglichen Hairpins befindet sich der Großteil der acht Basen der Interaktionssequenz entweder im Stamm des Hairpins oder innerhalb einer kurzen, stark gekrümmten Schleife. Die Annahme liegt nahe, dass eine solche Konformation eines Einzelstrangs die Assoziation beider Stränge zu einem stabilen Nukleus sterisch nahezu unmöglich macht. Trifft diese Annahme zu, so wäre die Assoziation nur dann möglich, wenn sich das Sys-

tem im vollkommen sekundärstrukturfreien Zustand S_0 (vgl. Abb. 4.12) befindet. Die Duplex-Assoziationsrate $k_{\text{on},0}$ aus dem Zustand S_0 lässt sich dann folgendermaßen mit der im Experiment gemessenen Assoziationsrate k_{on} und p_{Hairpin} verknüpfen:

$$\begin{aligned} k_{\text{on}} &= p(S_0) \cdot k_{\text{on},0} = p(s_0) \cdot p(s'_0) \cdot k_{\text{on},0} = (1 - p_{\text{Hairpin}}) \cdot k_{\text{on},0} \\ \Leftrightarrow k_{\text{on},0} &= \frac{k_{\text{on}}}{1 - p_{\text{Hairpin}}} \end{aligned} \quad (4.10)$$

Analog lässt sich argumentieren, dass die Ausformung der meisten Hairpins rein sterisch erst dann möglich ist, wenn das entsprechende Duplex bereits vollständig dissoziiert ist. Der weiter oben erwähnte mögliche Einfluss von Hairpins auf die Duplex-Dissoziation wäre dann für die hier untersuchten kurzen Duplexe vernachlässigbar. Die im TMFM-Experiment gemessene Dissoziationsrate k_{off} wäre unabhängig von Hairpins in den Einzelsträngen.

In der Vorhersage der freien Energien der Duplexbindung ΔG_{NN} wird die Konkurrenz zwischen Hairpins in den Einzelsträngen und der Duplexbindung nicht berücksichtigt. Die Gleichgewichtskonstante

$$K_{\text{eq,NN}} = \exp\left(-\frac{\Delta G_{\text{NN}}}{k_{\text{B}}T}\right) \quad (4.11)$$

beschreibt demnach das Gleichgewicht zwischen dem Duplex-Zustand und dem Zustand S_0 . Zudem nimmt die Vorhersage die Situation zweier sich frei in Lösung befindlicher Moleküle an. Die Gleichgewichtskonstante

$$K_{\text{eq,exp}} = \frac{k_{\text{on},0}}{k_{\text{off}}} \quad (4.12)$$

stellt das experimentelle, unimolekulare Pendant zu $K_{\text{eq,NN}}$ dar. Mit dem J-Faktor und Gl. (4.10) ergibt sich die folgende Beziehung zwischen den beiden Gleichgewichtskonstanten, den im TMFM-Experiment gemessenen Raten und p_{Hairpin} :

$$\begin{aligned} K_{\text{eq,exp}} &= \frac{k_{\text{on}}}{k_{\text{off}}} \cdot \frac{1}{1 - p_{\text{Hairpin}}} = J \cdot K_{\text{eq,NN}} \\ \Rightarrow \frac{k_{\text{on}}}{1 - p_{\text{Hairpin}}} &= J \cdot K_{\text{eq,NN}} \cdot k_{\text{off}} \end{aligned} \quad (4.13)$$

Die TMFM-Messungen für alle Sequenzvarianten wurden mit derselben Art von DNA-Nano-Objekten durchgeführt, mit Konturlänge des Leinenstücks zwischen den Interaktionspartnern $L = 1554$ bp und distaler Verankerungsgeometrie. Bei allen Messungen wurde die gleiche Konzentration von Na^+ (500 mM) eingesetzt. Der J-Faktor sollte folglich in allen Messungen denselben Wert haben. Letztlich münden die bisherigen Überlegungen in der Hypothese, dass eine annähernd lineare Beziehung zwischen $k_{\text{on}}/(1 - p_{\text{Hairpin}})$ und $(K_{\text{eq,NN}} \cdot k_{\text{off}})$ bestehen sollte. Für die Sequenzvarianten v1 bis v4 mit $p_{\text{Hairpin}} \ll 1$ gilt $\frac{1}{1 - p_{\text{Hairpin}}} \approx 1$. Für diese Varianten sollte dann eine annähernd lineare Beziehung zwischen k_{on} und $(K_{\text{eq,NN}} \cdot k_{\text{off}})$ bestehen.

Messergebnisse und Diskussion

Aus den TMFM-Daten der verschiedenen Messungen wurden die Verweildauern der einzelnen DNA-Nano-Objekte in den beiden Zuständen extrahiert. Ein Streudiagramm der mittleren Verweildauern im hybridisierten Zustand gegen die mittleren Verweildauern im dissoziierten Zustand der einzelnen DNA-Nano-Objekte offenbart die starken Unterschiede in den gemessenen Hybridisierungskinetiken der verschiedenen Sequenzvarianten (Abb. 4.13). Während die Punktwolken der Varianten v1 bis v4 nahe beieinander liegen, weichen die Punktwolken der Varianten v5 bis v7 deutlich ab. Sie sind um bis zu zwei Größenordnungen zu höheren Verweildauern im dissoziierten Zustand verschoben. Diese Verschiebung ist mit steigendem p_{Hairpin} immer stärker ausgeprägt. Interessanterweise fallen die mittleren Verweildauern im dissoziierten Zustand der Variante v7 um eine ganze Größenordnung höher aus als die der Variante v6, obwohl sich die p_{Hairpin} dieser zwei Varianten nur noch um rund 0,01 unterscheiden.

Auch die mittleren Verweildauern der Objekte der verschiedenen Sequenzvarianten im hybridisierten Zustand weisen Unterschiede auf: Bei den Varianten v6 und v7 mit sehr hohem p_{Hairpin} sind sie sichtbar geringer als bei den Varianten mit niedrigerem p_{Hairpin} . Zwischen den niedrigsten Werten (für die Variante v7) und den höchsten Werten (für die Varianten v3 und v4) einzelner DNA-Nano-Objekte liegt etwa eine Größenordnung. Die Variante v5 ähnelt hinsichtlich der mittleren Verweildauern im hybridisierten Zustand eher der Gruppe der Varianten v1 bis v4 mit niedrigem p_{Hairpin} als den zwei anderen Varianten mit hohem p_{Hairpin} , v6 und v7.

Die Streuung der Punktwolken für die Varianten v5, v6 und besonders v7 ist deutlich größer als die Streuung der Punktwolken für die Varianten v1 bis v4. Die Ursachen für diesen Effekt sind zweierlei Natur: Erstens ist es eine grundlegende Eigenschaft der Exponentialverteilung, dass die Varianz des arithmetischen Mittels aus exponentialverteilten Zufallsvariablen mit dem Erwartungswert der Verteilung anwächst. Zweitens war die Dauer der Datenaufnahme für alle Sequenzvarianten gleich, weshalb man bei den Varianten v5 bis v7 je einem DNA-Nano-Objekt deutlich weniger Übergänge zwischen den Zuständen beobachtet als bei den Varianten v1 bis v4. Dieser Unterschied ist zwischen den Varianten v1 und v7 am stärksten ausgeprägt (siehe Nebenabbildung im gestrichelten Kasten unten links in Abb. 4.13). Je weniger Übergänge für ein DNA-Nano-Objekt beobachtet werden, umso weniger Einzelwerte gehen in die Berechnung der mittleren Verweildauern ein. Dadurch wachsen die statistischen Fehler der Mittelwerte, womit sich die Verteilungen der Mittelwerte verbreitern.

Die einzelnen Verweildauern aller Objekte der Hauptpopulationen wurden für jede Variante zusammengefasst und die Verweildauerverteilungen betrachtet (Abb. 4.14). Die entsprechenden charakteristischen Verweildauern τ_1 und τ_2 wurden nach Gleichungen (3.6)-(3.8) ermittelt. Ein Vergleich der Werte von τ_1 und τ_2 der verschiedenen Sequenzvarianten (siehe Angaben innerhalb der Diagramme in Abb. 4.14) dient zur Quantifizierung der bereits festgestellten Unterschiede. Die charakteristischen Verweildauern im hybridisierten Zustand reichen von $\tau_1 \approx 6,4$ s für die Variante v6 bis $\tau_1 \approx 27,5$ s für die Variante v3. Die charakteristischen Verweildauern im dissoziierten Zustand reichen von $\tau_2 \approx 2,7$ s für die Variante v1 bis $\tau_2 \approx 454,6$ s für die Variante v7.

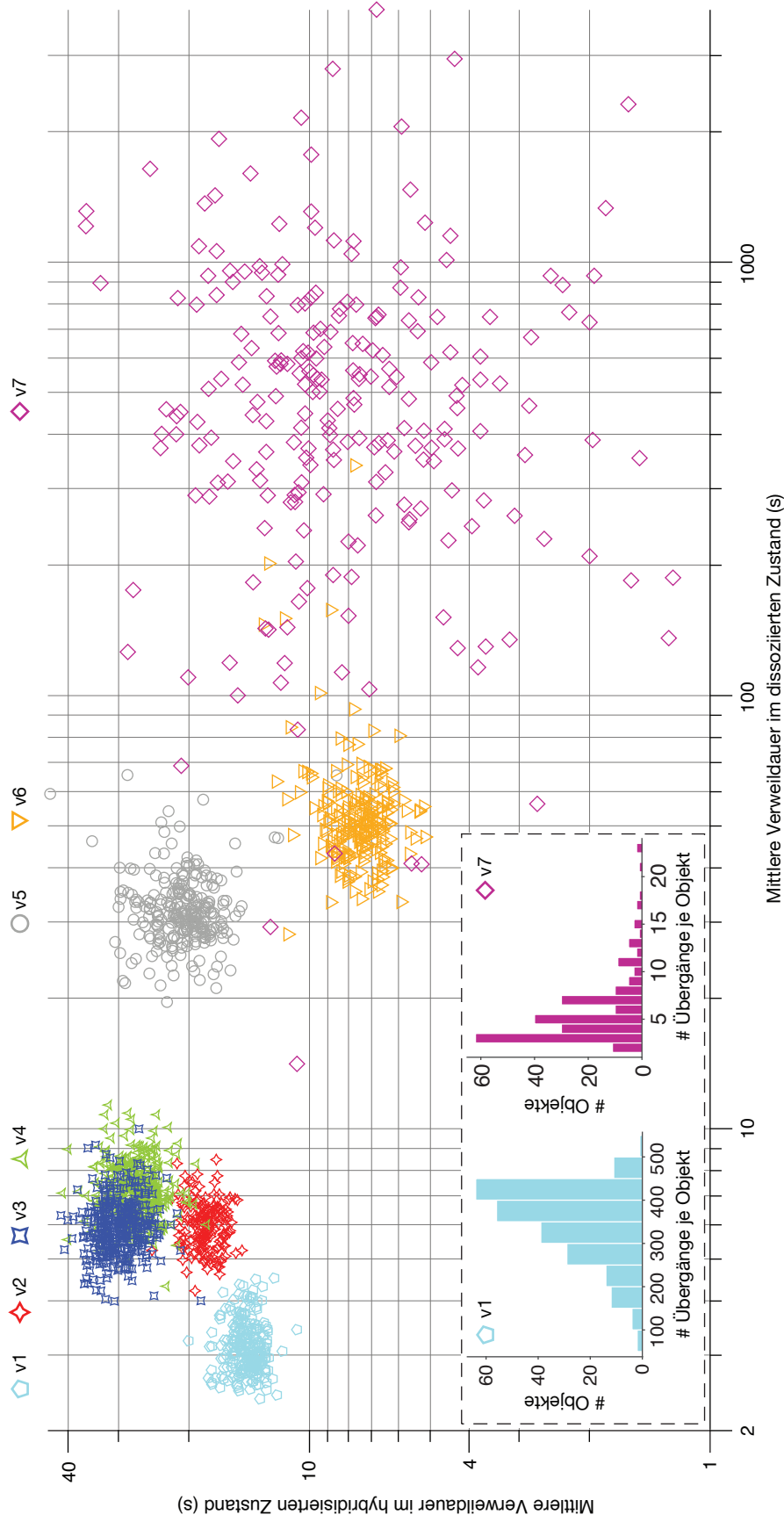


Abbildung 4.13: Streudiagramm der mittleren Verweildauern einzelner DNA-Nano-Objekte für die sieben Sequenzvarianten

Messdaten von DNA-Nano-Objekten mit Konturlängen des Leinenstücks zwischen den Interaktionspartnern $L = 1554$ bp und mit distaler Verankerungsgeometrie. In jedem Datenpunkt ist für ein DNA-Nano-Objekt der Hauptpopulation (vgl. Abschnitt 3.3.3) der jeweiligen Probe die mittlere Verweildauer im hybridisierten Zustand gegen die mittlere Verweildauer im dissoziierten Zustand aufgetragen. Farben und Marker stehen für die jeweilige Sequenzvariante (vgl. Legende oberhalb des Diagramms). Nebenabbildung im gestrichelten Kasten: Häufigkeitsverteilungen der Anzahl der Übergänge je DNA-Nano-Objekt für die Varianten v1 (links) und v7 (rechts). Man beachte die unterschiedliche Skalierung der Abszissenachse.

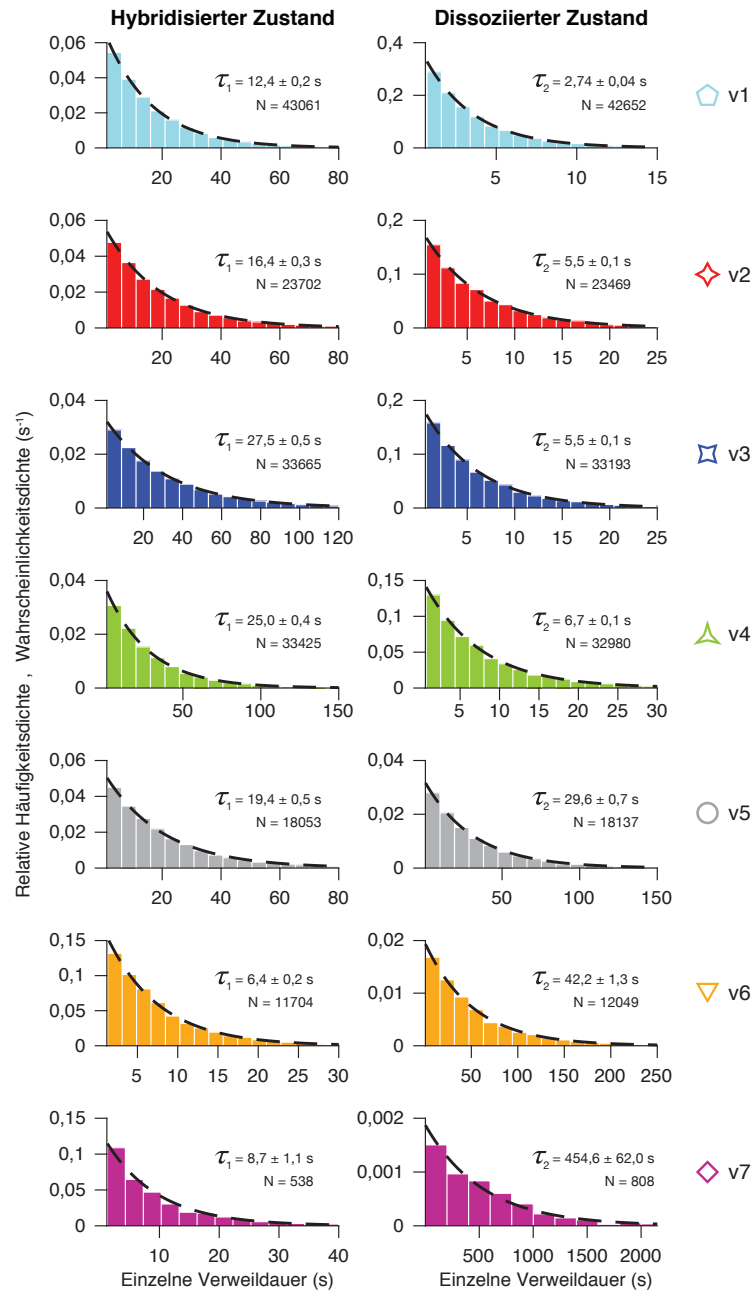


Abbildung 4.14: **Verteilungen der einzelnen Verweildauern für die sieben Sequenzvarianten**

Messdaten derselben DNA-Nano-Objekte wie in Abb. 4.13 (Konturlänge des Leinenstücks zwischen den Interaktionspartnern $L = 1554$ bp, distale Verankerungsgeometrie). Histogramme: Relative Häufigkeitsdichten der einzelnen Verweildauern der jeweiligen Hauptpopulationen (vgl. Abschnitt 3.3.3). Gestrichelte Linien: Maximum-Likelihood-Schätzungen der Wahrscheinlichkeitsdichteverteilungen der jeweiligen Hauptpopulationen nach Gl. (3.7). τ_1, τ_2 : Charakteristische Verweildauern der jeweiligen Hauptpopulationen nach Gl. (3.8). Fehlerbereich $\pm 3 \sigma_{\text{bootstrap}}$ (vgl. Abschnitt 3.3.4). N: jeweilige Anzahl der Einzelwerte.

Die aus den Mengen aller einzelnen Verweildauern geschätzten Wahrscheinlichkeitsdichteverteilungen in Form von beschränkten Exponentialverteilungen nach Gl. (3.7) wurden mit den Histogrammen der relativen Häufigkeitsdichten überlagert (Abb. 4.14). Allgemein herrscht eine sehr gute Übereinstimmung zwischen den Verläufen der geschätzten exponentialverteilten Wahrscheinlichkeitsdichten und den Histogrammen. Lediglich für die Variante v7 gibt es sichtbare leichte Abweichungen, welche auf den geringen Stichprobenumfang für diese Variante zurückgeführt werden können. Wurden für die Varianten v1 bis v6 jeweils mindestens 10.000 bzw. im Fall der Variante v1 sogar über 40.000 einzelne Verweildauern je Zustand erfasst, so waren es für die Variante v7 weniger als 1000 Verweildauern je Zustand. Selbst für die Variante v7 stimmen jedoch die geschätzten exponentialverteilten Wahrscheinlichkeitsdichten annehmbar mit den Histogrammen überein (Abb. 4.14, unterste Zeile). Somit lässt sich die Hybridisierung für jede der sieben Duplexvarianten als Zwei-Zustands-System beschreiben – einschließlich der Varianten v5 bis v7, welche eine starke Neigung zur Ausformung von Hairpins aufweisen.

Die Assoziationsrate k_{on} und die Dissoziationsrate k_{off} für jede einzelne Variante wurden gemäß Gl. (4.2) aus den inversen charakteristischen Verweildauern ermittelt. Darauf folgte eine Prüfung der eingangs formulierten Hypothesen bezüglich der Beziehung zwischen k_{on} , k_{off} , $K_{\text{eq,NN}}$ und p_{Hairpin} (siehe pp. 108-109). Die Auftragung von k_{on} gegen $(K_{\text{eq,NN}} \cdot k_{\text{off}})$ lässt für die Varianten v1 bis v4 mit $p_{\text{Hairpin}} \ll 1$ eine annähernd lineare Beziehung erkennen, von der die Varianten v5 bis v7 stark abweichen (Abb. 4.15 a). Bei Berücksichtigung von p_{Hairpin} und Auftragung von $k_{\text{on}}/(1 - p_{\text{Hairpin}})$ gegen $(K_{\text{eq,NN}} \cdot k_{\text{off}})$ zeigt sich insgesamt eine größere Übereinstimmung mit einer linearen Beziehung (Abb. 4.15 b, c). Obgleich sich auch die Varianten v5 und v7 durch das Einbeziehen von p_{Hairpin} in die Assoziationsrate etwas an diese lineare Beziehung annähern, weichen sie immer noch deutlich davon ab. Der Einfluss konkurrierender Hairpins auf die Hybridisierungskinetik der Varianten v5 und v7 scheint größer zu sein als auf Basis der oben angestellten Überlegungen erwartet. Im Folgenden werden mögliche Gründe dafür diskutiert.

Die Varianten v5 und v7 unterscheiden sich von der Variante v6 in der Anzahl der möglichen Hairpins je Einzelstrang. Beide Einzelstränge der Variante v6 können jeweils nur ein einziges mögliches Hairpin ausformen. Für Strang B' der Varianten v5 und v7 gilt dies ebenfalls, aber in Strang B der Variante v5 resp. v7 sind zwei resp. fünf verschiedene Hairpins möglich (vgl. Abb. C.14 resp. C.15). In der Berechnung von p_{Hairpin} werden die verschiedenen Hairpins als einzelne, diskrete energetische Zustände behandelt. Womöglich ist diese Anschauung zu stark vereinfacht und bildet die tatsächlichen mikroskopischen Prozesse nicht adäquat ab. Die Gesamtwahrscheinlichkeit, dass irgendeine hemmende Sekundärstruktur vorliegt, könnte insgesamt größer sein als die Summe der Besetzungswahrscheinlichkeiten der isoliert betrachteten Zustände vollständig ausgeformter Hairpins.

Vermutlich kann selbst für kurze Duplex-Sequenzen nicht vollkommen ausgeschlossen werden, dass sich die Neigung zur Formation von Hairpins auch auf die Duplex-Dissoziation auswirkt. Es ist denkbar, dass ein 8-bp-Duplex, welches bereits auf wenige Basenpaare aufgeschmolzen ist, mit höherer Wahrscheinlichkeit vollständig dissoziiert,

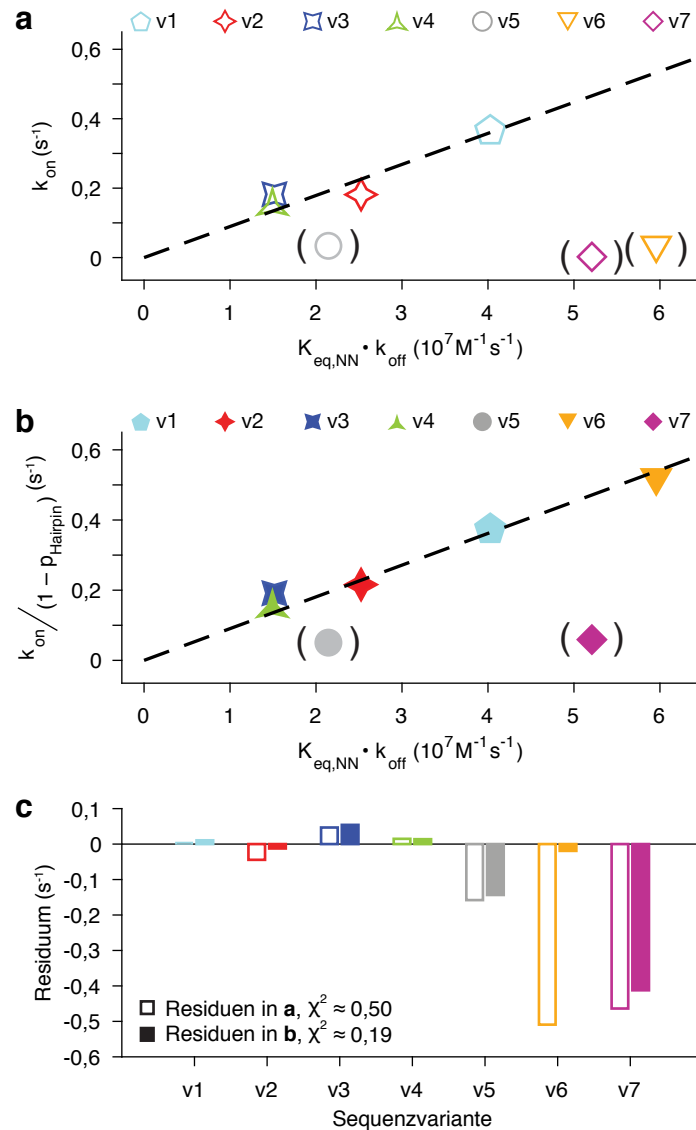


Abbildung 4.15: **Die Beziehung zwischen k_{on} , k_{off} , $K_{eq,NN}$ und $p_{Hairpin}$ für die sieben Sequenzvarianten** Messdaten derselben DNA-Nano-Objekte wie in Abb. 4.13 und Abb. 4.14 (Konturlänge des Leinenstücks zwischen den Interaktionspartnern $L = 1554$ bp, distale Verankerungsgeometrie). Farben und Marker stehen für die jeweilige Sequenzvariante (vgl. Legenden oberhalb von a und b). Die Klammern um einzelne Marker in a und b kennzeichnen die nicht in die jeweiligen Fits einbezogenen Varianten. **(a)** Auftragung von k_{on} gegen $(K_{eq,NN} \cdot k_{off})$. Linie: Least-Squares-Fit einer linearen Beziehung $k_{on} = j \cdot (K_{eq,NN} \cdot k_{off})$ an die Daten der Varianten v1 bis v4 (Steigung $j_a \approx 8,95 \cdot 10^{-9} M$). **(b)** Auftragung von $k_{on}/(1 - p_{Hairpin})$ gegen k_{off} . Linie: Least-Squares-Fit einer linearen Beziehung $k_{on}/(1 - p_{Hairpin}) = j \cdot (K_{eq,NN} \cdot k_{off})$ an die Daten der Varianten v1 bis v4 sowie v6 (Steigung $j_b \approx 9,05 \cdot 10^{-9} M$). **(c)** Hohle Balken: Residuen von k_{on} zur Fit-Geraden in a. Gefüllte Balken: Residuen von $k_{on}/(1 - p_{Hairpin})$ zur Fit-Geraden in b. χ^2 : Jeweils Summe der quadrierten Residuen aller Varianten.

wenn transiente intra-Strang-Interaktionen möglich sind und mit der erneuten Bildung von inter-Strang-Basenpaaren konkurrieren. Insbesondere im Fall von Strang B der Variante v5 ist es sterisch durchaus möglich, dass sich Paarungen zwischen Basen der Interaktionssequenz und des T₁₂-Verbinders ausformen, während z. B. noch der vier Basenpaare lange Abschnitt CAGG:CCTG intakt ist (vgl. Tabelle 4.1 und Abb. C.14). Auch bei einer etwaigen Destabilisierung der Duplex-Variante v7 könnte der T₁₂-Verbinder involviert sein: Im Stamm eines der möglichen Hairpins von Strang B besteht ein Basenpaar aus dem Adenin am 3'-Ende des Stranges und dem an die Interaktionssequenz angrenzenden Thymin des T₁₂-Verbinders (vgl. Tabelle 4.1 und Abb. C.15 a). In einem partiell aufgeschmolzenen Zwischenzustand des Duplex könnte die Formierung dieses intra-Strang-Basenpaares den Startpunkt für eine anschließende Verdrängung noch bestehender benachbarter inter-Strang-Basenpaare und die vollständige Ausformung eines stabilen Hairpin-Stammes darstellen.

Nichtsdestotrotz stützen die experimentellen Beobachtungen für fünf der sieben Sequenzvarianten die Hypothese einer linearen Beziehung zwischen $k_{\text{on}}/(1 - p_{\text{Hairpin}})$ und $(K_{\text{eq,NN}} \cdot k_{\text{off}})$ in der Form von Gl. (4.13). Darin beziffert die Steigung dieser linearen Beziehung den J-Faktor. Die Steigungen der Fit-Geraden in Abb. 4.15 betragen beide etwa 9 nM. Dieser Wert für den J-Faktor liegt in einem ähnlichen Bereich wie bereits publizierte theoretische [116] oder experimentelle [115] Werte für den J-Faktor von DNA-Fragmenten vergleichbarer Konturlänge.

Stapelung von Purin-Basen in den Einzelsträngen

Die Studie der Hybridisierungskinetik der sieben Sequenzvarianten fand in Kooperation mit dem Lehrstuhl für Biomolekulare Dynamik am Physik Department der TUM statt. Prof. Dr. Martin Zacharias führte Molekulardynamik-Simulationen der Interaktion zwischen den jeweiligen Einzelsträngen der sieben Sequenzvarianten durch. An dieser Stelle sei Prof. Zacharias erneut für die gute und erfolgreiche Zusammenarbeit gedankt!

In der gemeinsamen Publikation [138] werden die Simulationen näher beschrieben und die gesammelten Ergebnisse diskutiert. An dieser Stelle soll eine aus den Simulationen gewonnene Erkenntnis hervorgehoben werden: Durch stapelförmige Anordnung benachbarter Purin-Basen können sich DNA-Einzelstränge bereits teilweise helikal orientieren. Ein solches Arrangement könnte als Schablone für die Anlagerung des Komplementärstrangs dienen und somit die Bildung eines stabilen Nukleus sowie die anschließende Vervollständigung des Duplex erleichtern. Die Interaktionssequenz von Strang B' der Variante v1, GGAGGAGG, besteht ausschließlich aus Purin-Basen. Im TMFM-Experiment wurde für diese Variante die mit Abstand höchste Assoziationsrate gemessen. Die Ergebnisse der Simulationen deuten darauf hin, dass Basenstapelung im Einzelstrang B' dieser besonders schnellen Assoziation zu Grunde liegen könnte. Eine verstärkte Tendenz zur Stapelung von Purin-Basen findet sich auch in Strang B' der Variante v6. Dies könnte zumindest teilweise erklären, warum sich die Assoziationsraten der Varianten v6 und v7 trotz der sehr ähnlichen vorhergesagten Hairpin-Stabilitäten dieser zwei Varianten um mehr als eine Größenordnung unterscheiden.

4.3.3 Vergleich mit Modellen zur Vorhersage von Duplex-Assoziationsraten

Hata et al. [42] und Zhang et al. [24] veröffentlichten in den Jahren 2017 und 2018 jeweils eine Studie zur Sequenzabhängigkeit von Duplex-Assoziationsraten. In beiden Studien wurden Assoziationsraten verschiedener Duplexe aus Nichtgleichgewichts-Ensemble-Messungen bestimmt. Basierend auf den experimentellen Daten wurden dabei Formalismen zur Vorhersage der Duplex-Assoziationsrate aus der Basenpaarsequenz entwickelt. Die folgenden Abschnitte gehen näher auf die beiden Studien ein und stellen die jeweiligen Vorhersagen für die Assoziationsraten der Duplex-Varianten v1 bis v7 den entsprechenden Messwerten aus den TMFM-Experimenten gegenüber.

Vorhersage nach Hata et al.

Hata et al. [42] untersuchten die Assoziationskinetik von 46 verschiedenen Duplexen mit einer Länge von je 23 bp bei einer Temperatur von 25 °C und einer Na⁺-Konzentration von 0,195 M. Die Basenpaarsequenzen wurden derart gewählt, dass alle 46 Duplexe eine sehr ähnliche thermodynamische Stabilität besaßen: Die Schmelztemperaturen lagen bei den gewählten Bedingungen alle im Bereich zwischen 61 und 66 °C. In einer Stopped-Flow-Apparatur wurden die einzelsträngigen Interaktionspartner schnell vermischt und anschließend der Anstieg der Konzentration doppelsträngiger DNA durch Auslesen der Fluoreszenz des dsDNA-Farbstoffs Pico Green[®] verfolgt. Die aus den Messdaten bestimmten Assoziationsraten umspannten über zwei Größenordnungen.

Die Autoren führten die starken Unterschiede in den Assoziationsraten auf zwei Effekte zurück: zum einen auf die unterschiedliche Fähigkeit der verschiedenen Sequenzen zur Bildung von Nuklei für die Duplex-Formation und zum anderen auf die unterschiedliche Neigung der Einzelstränge zur Ausformung von Hairpins. Sie formulierten ein Modell zur Vorhersage der Assoziationsrate der DNA-Hybridisierung, dem die Annahme zu Grunde liegt, dass Abschnitte von drei aufeinanderfolgenden Basenpaaren einen ausreichend stabilen Nukleus zur Ausformung eines vollständigen Duplex bilden können. Dabei sind nicht alle möglichen Basenpaar-Tripel gleichermaßen zur Bildung eines stabilen Nukleus geeignet. Der Parameter δ_{seq} gibt die Fähigkeit zur Nukleation eines bestimmten Tripels an. Eine weitere wichtige Annahme ist, dass das Vorliegen von Hairpins nicht zwingend die Duplex-Assoziation verhindert. Solange Tripel von Basen noch zugänglich sind, können sie mit dazu passenden Tripeln des Komplementärstrangs Nuklei ausformen. Ein Basentripel gilt dann als zugänglich, wenn es mehr als zwei Nukleotide Abstand von Hairpins im selben Einzelstrang hat.

In die Berechnung der vorhergesagten Assoziationsrate k_{calc} für ein beliebiges Duplex B:B' gehen die δ_{seq} aller in B und B' enthaltenen Basentripel sowie die freien Energien der Bildung aller berücksichtigten Hairpins ein. Die Assoziationsrate der Hybridisierung zweier Einzelstränge ohne Hairpins und mit $\delta_{\text{seq}} = 1$ für alle Basentripel wird mit k_0 bezeichnet und ist ein Faktor in der Formel für k_{calc} . Der Wert von k_0 ist unbekannt, aber für Duplexe derselben Länge gleich. Folglich lassen sich mit dem Modell von Hata et al. zwar Vorhersagen über das Verhältnis der Assoziationsraten verschiedener Duplexe gleicher Länge treffen, jedoch keine absoluten Assoziationsraten vorhersagen. Die

Autoren optimierten die Werte von δ_{seq} für alle 64 möglichen Basentripel auf das Ziel einer größtmöglichen Korrelation zwischen k_{calc}/k_0 und der experimentell bestimmten Assoziationsrate (in [42] als k_{app} bezeichnet). Eine Randbedingung der Optimierung war die Gleichheit der δ_{seq} für zueinander revers-komplementäre Tripel. Der optimale Korrelationskoeffizient zwischen k_{calc}/k_0 und k_{app} betrug $r = 0,92$. Die erhaltenen Werte für δ_{seq} lagen im Bereich von 0 bis 3,5. Für 16 verschiedene Tripel ergab sich der Minimalwert von $\delta_{\text{seq}} = 0$, für die Tripel AGG und CCT der Maximalwert von $\delta_{\text{seq}} = 3,5$.

Für die sieben Sequenzvarianten der vorliegenden Arbeit wurden die Werte von k_{calc}/k_0 nach dem Vorhersageformalismus von Hata et al. berechnet und mit den aus den TMFM-Daten ermittelten Assoziationsraten k_{on} verglichen (Abb. 4.16). Es zeigt sich eine positive Korrelation zwischen den Raten aus dem TMFM-Experiment und der Vorhersage nach Hata et al. ($r = 0,86$). Die Vorhersage liefert für die Variante v1 mit der experimentell größten Assoziationsrate auch den größten Wert von k_{calc}/k_0 . Für die Varianten v5 bis v7 mit den experimentell kleinsten Assoziationsraten liefert die Vorhersage vergleichsweise kleine Werte von k_{calc}/k_0 . Darüber hinaus passen die Verhältnisse der k_{calc}/k_0 für die Varianten v1, v5, v6 und v7 in etwa zu den Verhältnissen der experimentellen Assoziationsraten für diese Varianten.

Bei idealer Übereinstimmung von Vorhersage und TMFM-Experiment würden alle Datenpunkte in Abb. 4.16 auf einer Ursprungsgerade liegen. Dies ist jedoch nicht der Fall. Das Modell scheint entweder die Werte von k_{calc}/k_0 für die Varianten v2 bis v4 zu unterschätzen oder den Wert von k_{calc}/k_0 für die Variante v1 zu überschätzen. Das Ergebnis der Vorhersage für die Varianten v2 bis v4 passt in einem weiteren Aspekt nicht zum Ergebnis der TMFM-Messung: Zwischen dem größten und kleinsten Wert von k_{calc}/k_0 innerhalb dieser Gruppe liegt ein Faktor von über 5, zwischen dem größten und kleinsten Wert von k_{on} dagegen nur ein Faktor von etwa 1,23.

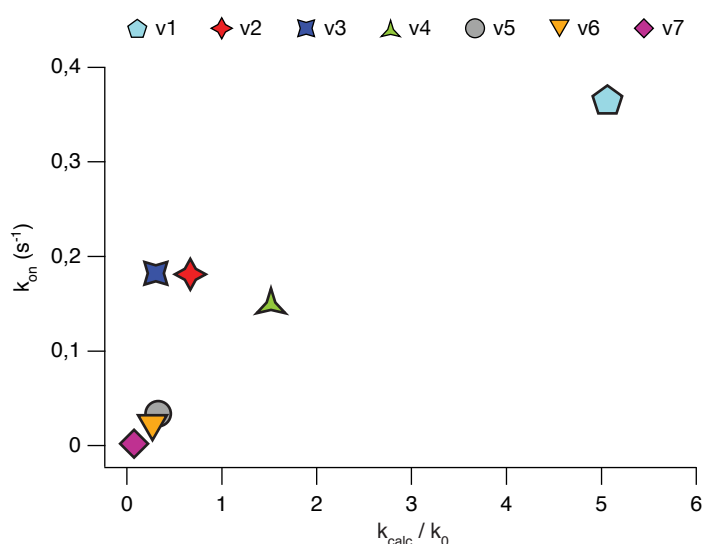


Abbildung 4.16: Vergleich der TMFM-Assoziationsrate k_{on} mit der Vorhersage nach Hata et al.

Farben und Marker stehen für die jeweilige Sequenzvariante (vgl. Legende). k_{calc}/k_0 wurde nach dem Formalismus von Hata et al. berechnet (siehe insbesondere pp. 7-9 in [42]).

Vorhersage nach Zhang et al.

Zhang et al. [24] untersuchten die Assoziationskinetik von 100 verschiedenen Duplexen mit einer Länge von je 36 bp bei verschiedenen Reaktionstemperaturen von 28 bis 55 °C. Die insgesamt 210 experimentell bestimmten Assoziationsraten wurden alle in fünffach konzentrierter PBS-Lösung (diese enthält u. a. 0,685 M Na⁺ und 0,0135 M K⁺) gemessen. Die Einzelstränge der verschiedenen Duplexe wiesen eine unterschiedlich stark ausgeprägte Neigung zur Bildung von Hairpins auf. Anders als die sieben Sequenzvarianten in der hier präsentierten Arbeit oder die von Hata et al. untersuchten Duplexe wurden die von Zhang et al. untersuchten Duplexe nicht gezielt auf ähnliche thermodynamische Stabilität der Duplexbindung hin selektiert. Zhang et al. bezeichnen jeweils einen der Einzelstränge eines Paares als *Target*, den anderen Einzelstrang als *Probe*. Im Experiment wurde zunächst einer der Einzelstränge an ein Fluorophor, der andere an ein fluoreszenzlöschendes Molekül gekoppelt. Nach anschließender Mischung der so modifizierten Interaktionspartner wurde das Fluoreszenzsignal des Reaktionsgemischs verfolgt, welches mit der zunehmenden Assoziation von Target und Probe abnimmt. Aus dieser Abnahme wurde die jeweilige Assoziationsrate extrahiert.

Im Gegensatz zum Vorhersageformalismus nach Hata et al. fließen in den Vorhersageformalismus nach Zhang et al. keinerlei modellhafte Annahmen über den Mechanismus der Duplex-Assoziation ein. Vielmehr basiert die Vorhersage nach Zhang et al. auf einem Algorithmus namens *Weighted Neighbor Voting* (WNV). Dabei wird jede Hybridisierungsreaktion anhand sogenannter *Features* (bioinformatischer Merkmale) charakterisiert. Ein solches Feature ist beispielsweise die gemittelte freie Energie der Bildung aller Komplexe von Target und Probe, die im Gleichgewicht mit einer Wahrscheinlichkeit von mehr als 0,5% auftreten, wobei in der Mittelung der Beitrag jedes Komplexes mit seiner jeweiligen Auftrittswahrscheinlichkeit gewichtet wird. Ein weiteres Beispiel für ein Feature ist schlichtweg die Reaktionstemperatur in °C. Final verwendeten Zhang et al. sechs verschiedene, unterschiedlich stark gewichtete Features. Eine Hybridisierungsreaktion mit Index m wird dementsprechend auf einen Punkt $\vec{f}(m)$ in einem sechsdimensionalen Raum abgebildet. Die vorhergesagte Assoziationsrate einer beliebigen Hybridisierungsreaktion $\hat{k}_{\text{hyb}}(j)$ ergibt sich aus dem gewichteten Mittelwert aller bekannten Assoziationsraten $k_{\text{hyb}}(m)$. Die Wichtungsfaktoren für die $k_{\text{hyb}}(m)$ fallen dabei exponentiell mit dem Abstand der Punkte $\vec{f}(j)$ und $\vec{f}(m)$ im Feature-Raum ab.

Die Autoren stellten ein Online-Tool für die Vorhersage von \hat{k}_{hyb} zur Verfügung³. In der Eingabemaske werden die Reaktionstemperatur, die Salinität⁴ in mol/l und die Basensequenz des Target abgefragt. Die Vorhersage erfolgt für die Hybridisierungsreaktion zwischen einem Einzelstrang dieser Sequenz und dem dazu vollständig reverskomplementären Einzelstrang. Gäbe man die Sequenzen der Varianten v1 bis v7 inklusive T₁₂-Verbinder ein, so würde das Tool Assoziationsraten für jeweils 20 bp lange Duplexe vorhersagen, z. B. im Fall von Variante v5 für das Duplex der Basenpaarsequenz T₁₂CAGGAACG:CGTTCCTGA₁₂, und nicht für das Duplex der Basenpaarsequenz CAGGAACG:CGTTCCTG. Daher wurden zum Zweck der Vorhersage von \hat{k}_{hyb} für die Va-

³<http://nablab.rice.edu/nabtools/kinetics>

⁴Aus dem Artikel von Zhang et al. geht nicht eindeutig hervor, welche Größe damit gemeint ist. Aus dem Kontext wurde geschlossen, dass es sich dabei um die Konzentration monovalenter Ionen handelt.

rianten v1 bis v7 jeweils nur die acht Basenpaare langen 3'-terminalen Interaktionssequenzen eingegeben. Infolgedessen kann jedoch die Vorhersage mögliche Einflüsse des T₁₂-Verbinders, insbesondere die Beteiligung des T₁₂-Verbinders an Hairpins, nicht mit einbeziehen. Die Stämme nahezu aller in dieser Arbeit berücksichtigten Hairpins der Varianten v1, v4 und v5 beinhalten aber Basen des T₁₂-Verbinders. Für die Varianten v1 und v4 ist die Auftrittswahrscheinlichkeit dieser Hairpins in erster Näherung vernachlässigbar, jedoch nicht für Variante v5. Von vornherein würde man dementsprechend für Variante v5 keine gute Übereinstimmung der Vorhersage von \hat{k}_{hyb} mit der experimentellen Assoziationsrate k_{on} erwarten.

Für die sieben Sequenzvarianten der vorliegenden Arbeit wurden die Werte von \hat{k}_{hyb} mit dem Online-Tool von Zhang et al. berechnet und mit den aus den TMFM-Daten ermittelten Assoziationsraten k_{on} verglichen (Abb. 4.17). Abhängig davon, ob in der Ein-

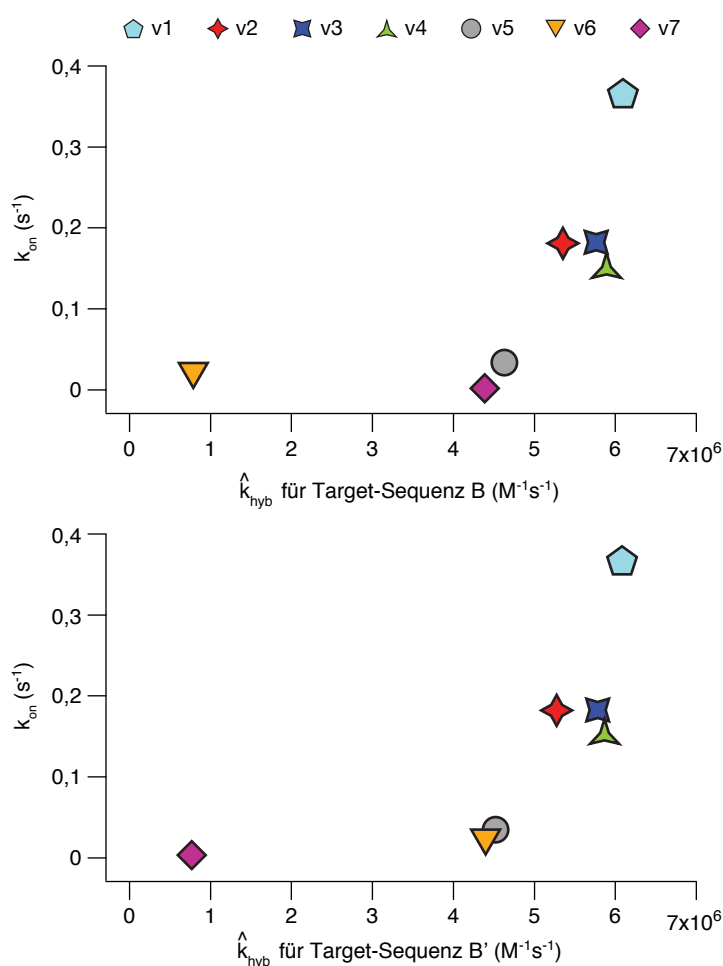


Abbildung 4.17: **Vergleich der TMFM-Assoziationsrate k_{on} mit der Vorhersage nach Zhang et al.**

Farben und Marker stehen für die jeweilige Sequenzvariante (vgl. Legende). Oben: Vorhersage von \hat{k}_{hyb} nach Eingabe der Interaktionssequenz von Strang B als Target-Sequenz. Unten: Vorhersage von \hat{k}_{hyb} nach Eingabe der Interaktionssequenz von Strang B' als Target-Sequenz. Für beide Vorhersagen wurde als Reaktionstemperatur 23 °C und als Salinität⁴ 0,5 mol/l eingegeben.

gabemaske als Target-Sequenz die Interaktionssequenz von Strang B (Abb. 4.17, oben) oder die Interaktionssequenz von Strang B' (Abb. 4.17, unten) eingegeben wird, liefert die Vorhersage für dieselbe Duplexvariante unterschiedliche Werte von \hat{k}_{hyb} . Diese Asymmetrie ergibt sich für alle Varianten außer v1 und ist besonders ausgeprägt für die Varianten v6 und v7 mit der größten Neigung zur Ausbildung von Hairpins. Die Ursache dafür liegt in der Auswahl der Features für den WNV-Algorithmus. Eines der finalen sechs Features repräsentiert die Neigung des Probe-Stranges zur Ausformung von Sekundärstrukturen. Dagegen wurde das Feature, welches die Neigung des Target-Stranges zur Ausformung von Sekundärstrukturen repräsentiert, nicht in die finale Auswahl der Features aufgenommen. Mögliche Hairpins in der eingegebenen Sequenz spielen demnach für die Vorhersage mit WNV keine Rolle, wohl aber mögliche Hairpins in der dazu revers-komplementären Sequenz. Dies führt offensichtlich dazu, dass \hat{k}_{hyb} für Variante v6 resp. Variante v7 sehr viel höher ausfällt, wenn als Target-Sequenz die Interaktionssequenz von Strang B' resp. Strang B eingegeben wird, als andersherum.

Für die Variante v1 mit der größten experimentellen Assoziationsrate liefert die Vorhersage den größten Wert für \hat{k}_{hyb} , für die Varianten v6 und v7 mit den kleinsten experimentellen Assoziationsraten liefert die Vorhersage die kleinsten Werte für \hat{k}_{hyb} . Für die Variante v6 sind sowohl der experimentelle Wert als auch der kleinere der zwei vorhergesagten Werte der Assoziationsrate um etwa eine Größenordnung kleiner als die entsprechenden Werte für Variante v1. Während jedoch die experimentelle Assoziationsrate für Variante v7 um eine weitere Größenordnung kleiner ist als die experimentelle Assoziationsrate für Variante v6, sind für diese zwei Varianten sowohl die jeweils kleineren Werte als auch die jeweils größeren Werte von \hat{k}_{hyb} nahezu identisch.

Die Streuung der vorhergesagten Assoziationsrate \hat{k}_{hyb} für die drei Varianten v2 bis v4 entspricht ungefähr der Streuung der experimentellen Assoziationsrate k_{on} für diese Untergruppe. Allerdings stimmt die Sortierung $\hat{k}_{\text{hyb}}(v4) > \hat{k}_{\text{hyb}}(v3) > \hat{k}_{\text{hyb}}(v2)$ nicht mit der Sortierung $k_{\text{on}}(v3) \approx k_{\text{on}}(v2) > k_{\text{on}}(v4)$ überein. Desweiteren sind die \hat{k}_{hyb} für die Varianten v2 bis v4 im Mittel nur um weniger als 10% kleiner als für Variante v1, während die experimentellen Assoziationsraten für diese drei Varianten im Mittel um einen Faktor von über 2 kleiner sind als die experimentelle Assoziationsrate von Variante v1. Es hat somit den Anschein, als würde die Vorhersage nach Zhang et al. entweder die \hat{k}_{hyb} für Variante v1 unter- oder die \hat{k}_{hyb} für Varianten v2 bis v4 überschätzen.

Bei idealer Übereinstimmung von Vorhersage und TMFM-Experiment würden die Datenpunkte in der Auftragung von k_{on} gegen \hat{k}_{hyb} auf einer Ursprungsgerade liegen. Selbst bei Verwendung des jeweils kleineren Werts von \hat{k}_{hyb} für die Varianten v6 und v7, sowie mit Vernachlässigung von Variante v5, ist dies nicht der Fall.

Fazit

In einigen Aspekten herrscht Übereinstimmung zwischen den Vorhersagen von Hata et al. bzw. Zhang et al. und den Daten aus dem TMFM-Experiment, im Detail bestehen dazwischen aber teils deutliche Diskrepanzen. Die Identifikation besonders langsam und besonders schnell assoziierender Sequenzvarianten scheint verlässlicher zu funktionieren als die genaue quantitative Vorhersage der Assoziationsraten.

Die Vorhersageformalismen von Hata et al. bzw. Zhang et al. wurden mit Daten aus Messungen an Duplexen mit 23 bp bzw. 36 bp Länge kalibriert. Bei der Duplex-Assoziation der mit 8 bp deutlich kürzeren Sequenzvarianten v1 bis v7 könnten Effekte zum Tragen kommen, welche bei der Assoziation längerer Duplexe nur eine untergeordnete bis gar keine Rolle spielen. Außerdem handelt es sich bei den Experimenten sowohl von Hata et al., als auch von Zhang et al., um Ensemble-Messungen, die nicht um den von fehlerhaften Kopien der interagierenden Moleküle herrührenden Anteil des Messsignals bereinigt werden können – im Gegensatz zu den TMFM-Messungen (vgl. Abschnitt 3.3.3). Bei den Messungen beider Gruppen lagen die Interaktionspartner frei in Lösung vor, wogegen die Interaktionspartner in den TMFM-Messungen im Partikel-Leinen-System verankert sind. Allein die Unterschiede in der Gestaltung der verschiedenen Experimente könnten die Unterschiede zwischen den beiden Vorhersagen und den TMFM-Daten erklären.

Die beobachteten Übereinstimmungen zwischen den Vorhersagen und den TMFM-Daten lassen jedoch schließen, dass die Modelle von Hata et al. und Zhang et al. bereits einen Schritt in die richtige Richtung darstellen. Es besteht Grund zur Zuversicht, dass mit der Durchführung weiterer Messungen und der Weiterentwicklung von Vorhersagemodellen in Zukunft die Assoziationsraten von Duplexen beliebiger Sequenz mit immer höherer Verlässlichkeit vorhergesagt werden können.

4.4 Kapitelzusammenfassung

Die neu entwickelte Methode TMFM wurde zur Messung der Hybridisierungskinetik 8 bp langer DNA-Duplexe in Abhängigkeit von der Beschaffenheit des Partikel-Leinen-Systems, von der Na^+ -Konzentration in der umgebenden Lösung und von der Basenpaarsequenz der Interaktionspartner verwendet. Bei der Interpretation der Messdaten spielt der sogenannte J-Faktor eine bedeutende Rolle. Er beziffert die effektive Konzentration der im Partikel-Leinen-System verankerten Interaktionspartner zueinander.

Die Ergebnisse aller Messungen weisen eine wichtige Gemeinsamkeit auf: Die Verweildauerverteilungen der DNA-Nano-Objekte der Hauptpopulation im hybridisierten und im dissoziierten Zustand besitzen eine große Ähnlichkeit zu Exponentialverteilungen. Alle untersuchten Systeme lassen sich somit hinreichend als Zwei-Zustands-Systeme mit jeweils einer einzelnen Rate für die Übergänge zwischen den zwei Zuständen beschreiben.

Bei Messungen mit verschiedenen Konturlängen L des Leinenstücks zwischen den Interaktionspartnern im Bereich von 546 bp bis 2562 bp zeigte die Dissoziationsrate k_{off} nur eine schwache bis gar keine Abhängigkeit von L . Dagegen nahm die Assoziationsrate k_{on} mit zunehmendem Wert von L deutlich ab. Diese Abnahme von k_{on} ist auf eine Abnahme des J-Faktors zurückzuführen. Die Form des Verlaufs von k_{on} über L entspricht qualitativ der Form des Verlaufs der theoretischen Vorhersage des J-Faktors in Abhängigkeit der Konturlänge eines Fragments doppelsträngiger DNA nach [114].

Die gemessenen Raten hängen vom Abstand der Ankerpunkte der Interaktionspartner an der DNA-Leine zu den Reporter-Einheiten ab. Die Hybridisierungskinetik wurde

für drei Werte von L jeweils in der distalen Verankerungsgeometrie mit großem Abstand zwischen den Ankerpunkten und den Reporter-Einheiten sowie in der proximalen Verankerungsgeometrie mit kleinem Abstand zwischen den Ankerpunkten und den Reporter-Einheiten gemessen. In der distalen Verankerungsgeometrie waren bei jeweils gleichem Wert von L die gemessenen Dissoziationsraten etwas kleiner und die gemessenen Assoziationsraten deutlich größer als in der proximalen Verankerungsgeometrie. Die distale Verankerungsgeometrie verringert demnach den störenden Einfluss des umgebenden Partikel-Leinen-Systems auf die Hybridisierung des Duplex $B:B'$, was einer Erhöhung des J -Faktors entspricht.

TMFM-Messungen bei einer Reihe von Na^+ -Konzentrationen offenbarten eine deutliche Zunahme von k_{on} mit $[\text{Na}^+]$. Von $[\text{Na}^+] = 150 \text{ mM}$ bis $[\text{Na}^+] = 1 \text{ M}$ stieg k_{on} um einen Faktor von etwa 5 an. Dabei wurde eine annähernd lineare Beziehung zwischen $\ln(k_{\text{on}})$ und $\ln([\text{Na}^+])$ beobachtet. Dagegen änderte sich k_{off} nur geringfügig mit $[\text{Na}^+]$. Eine gewisse Verfälschung der gemessenen Raten durch eine mögliche $[\text{Na}^+]$ -Abhängigkeit der Flexibilität der DNA-Leine und somit des J -Faktors kann allerdings nicht vollkommen ausgeschlossen werden. Für eine zumindest leichte Abhängigkeit des J -Faktors von $[\text{Na}^+]$ spricht die leichte Abnahme der mittleren quadratischen Auslenkungen der mobilen Reporter-Einheiten mit steigender Konzentration von Na^+ .

Desweiteren wurde die Hybridisierungskinetik von sieben Duplex-Sequenzvarianten untersucht, die laut theoretischer Vorhersage eine fast identische thermodynamische Stabilität, jedoch eine stark unterschiedliche Neigung zur Bildung von Hairpins in den jeweiligen Einzelsträngen aufweisen. Die Messdaten illustrieren, welchen starken verzerrenden Effekt Hairpins auf die Duplex-Hybridisierungskinetik haben können. Zum einen sind die Assoziationsraten der Varianten mit großer Neigung zur Bildung von Hairpins um bis zu zwei Größenordnungen kleiner als die Assoziationsraten der Varianten mit geringer Neigung zur Bildung von Hairpins. Zum anderen tendieren die Varianten mit größerer Neigung zur Bildung von Hairpins zu höheren Werten der Dissoziationsrate im Vergleich zu den Varianten mit geringer Neigung zur Bildung von Hairpins.

Eine auf einfachen Annahmen basierende Korrektur von k_{on} konnte den Einfluss von Hairpins nicht für alle Varianten zufriedenstellend kompensieren. Für fünf der sieben Sequenzvarianten zeigt sich allerdings eine lineare Beziehung zwischen der korrigierten Assoziationsrate $k_{\text{on}}/(1 - p_{\text{Hairpin}})$ und dem Produkt der theoretisch vorhergesagten Gleichgewichtskonstante mit der gemessenen Dissoziationsrate ($K_{\text{eq,NN}} \cdot k_{\text{off}}$). Aus dieser

linearen Beziehung konnte der J -Faktor bei $L = 1554 \text{ bp}$ in der distalen Verankerungsgeometrie abgeschätzt werden. Der erhaltene Wert von 9 nM liegt in der gleichen Größenordnung wie bisherige theoretische [116] oder experimentelle [115] Literaturwerte für Fragmente doppelsträngiger DNA ähnlicher Konturlänge.

Die aus den TMFM-Messungen ermittelten Assoziationsraten für die Sequenzvarianten v1 bis v7 wurden mit zwei verschiedenen Vorhersagen der Duplex-Assoziationsrate anhand der Basenpaarsequenz [24, 42] verglichen. Diese Vergleiche ergaben teilweise Übereinstimmungen zwischen den Vorhersagen und den TMFM-Daten, offenbarten jedoch auch deutliche Unterschiede.

5 Ensemble-Kompetitions-Experimente zur Bestimmung des J-Faktors in TMFM

Durch die Verankerung der Interaktionspartner B und B' in einem Partikel-Leinen-System wird in der TMFM-Messung keine bimolekulare Reaktion, sondern eine unimolekulare Reaktion beobachtet. Um die Ergebnisse aus den TMFM-Messungen zu dem Fall in Bezug setzen zu können, dass beide Bindepartner frei in Lösung vorliegen, ist die Kenntnis des J-Faktors der DNA-Nano-Objekte notwendig. In Abschnitt 4.3 wurde basierend auf den TMFM-Messungen mit Duplexen unterschiedlicher Sequenz bereits eine Schätzung des J-Faktors vorgenommen. Diese Schätzung ergab für Objekte mit einer Konturlänge des Leinenstücks zwischen den Interaktionspartnern von $L = 1554$ bp und mit der distalen Verankerungsgeometrie einen Wert von etwa 9 nM.

Dieses Kapitel handelt von ergänzenden Ensemble-Experimenten zur Bestimmung des J-Faktors der DNA-Nano-Objekte. Abschnitt 5.1 erklärt das Konzept dieser Experimente. Abschnitt 5.2 beschreibt die Durchführung und Analyse der Experimente. Abschnitt 5.3 legt dar, wie der J-Faktor aus den Messdaten extrahiert wurde, und diskutiert die Ergebnisse. In Abschnitt 5.4 wird ein Resümee zu den Kompetitions-Experimenten gezogen.

5.1 Konzept der Ensemble-Kompetitions-Experimente

Die Ensemble-Kompetitions-Experimente beruhen auf der Einbringung eines zusätzlichen einzelsträngigen Bindepartners mit der Sequenz B in die Probenlösung (siehe Abb. 5.1). Dieser zusätzliche Bindepartner wird in dieser Arbeit als „Kompetitor“ bezeichnet. Befindet sich ein DNA-Nano-Objekt im dissoziierten Zustand, so kann der Kompetitor an den Interaktionspartner B' dieses Objekts binden, bevor eine erneute Bindung zwischen den zwei im Objekt verankerten Interaktionspartnern zustande kommen kann (Abb. 5.1 a, rechte Seite des Schemas). Erst wenn sich der Kompetitor wieder vom Interaktionspartner B' gelöst hat, besteht die Möglichkeit zur Bindung zwischen den verankerten Interaktionspartnern B und B' (Abb. 5.1 a, linke Seite des Schemas).

Bezeichne C den Zustand des Systems mit dem DNA-Nano-Objekt im hybridisierten Zustand und freiem Kompetitor, bezeichne U den Zustand des Systems mit dem DNA-Nano-Objekt im dissoziierten Zustand und freiem Kompetitor und bezeichne D den Zustand des Systems mit Bindung des Kompetitors an den Interaktionspartner B' (vgl. Abb. 5.1 a). Fügt man Kompetitor zu einer equilibrierten Probe von DNA-Nano-Objekten hinzu, sind zunächst nur die Zustände C und U besetzt. Mit der Zeit stellt sich ein neues Besetzungsgleichgewicht zwischen allen drei Zuständen ein.

Der Anteil der Objekte im Zustand C kann per Gel-Elektrophorese bestimmt werden, da Objekte mit Schlaufe in der DNA-Leine eine höhere Mobilität im Gel besitzen als Objekte ohne Schlaufe in der Leine (vgl. Abschnitt 1.3). Für Objekte mit der proximalen Verankerungsgeometrie ist die Auftrennung dieser Populationen im Gel deutlicher ausgeprägt als für Objekte mit der distalen Verankerungsgeometrie. Zudem korreliert das Ausmaß der Auftrennung positiv mit der Länge der DNA-Leine (vgl. Abb. 1.11). Deshalb wurden für die Ensemble-Kompetitions-Experimente DNA-Nano-Objekte mit der

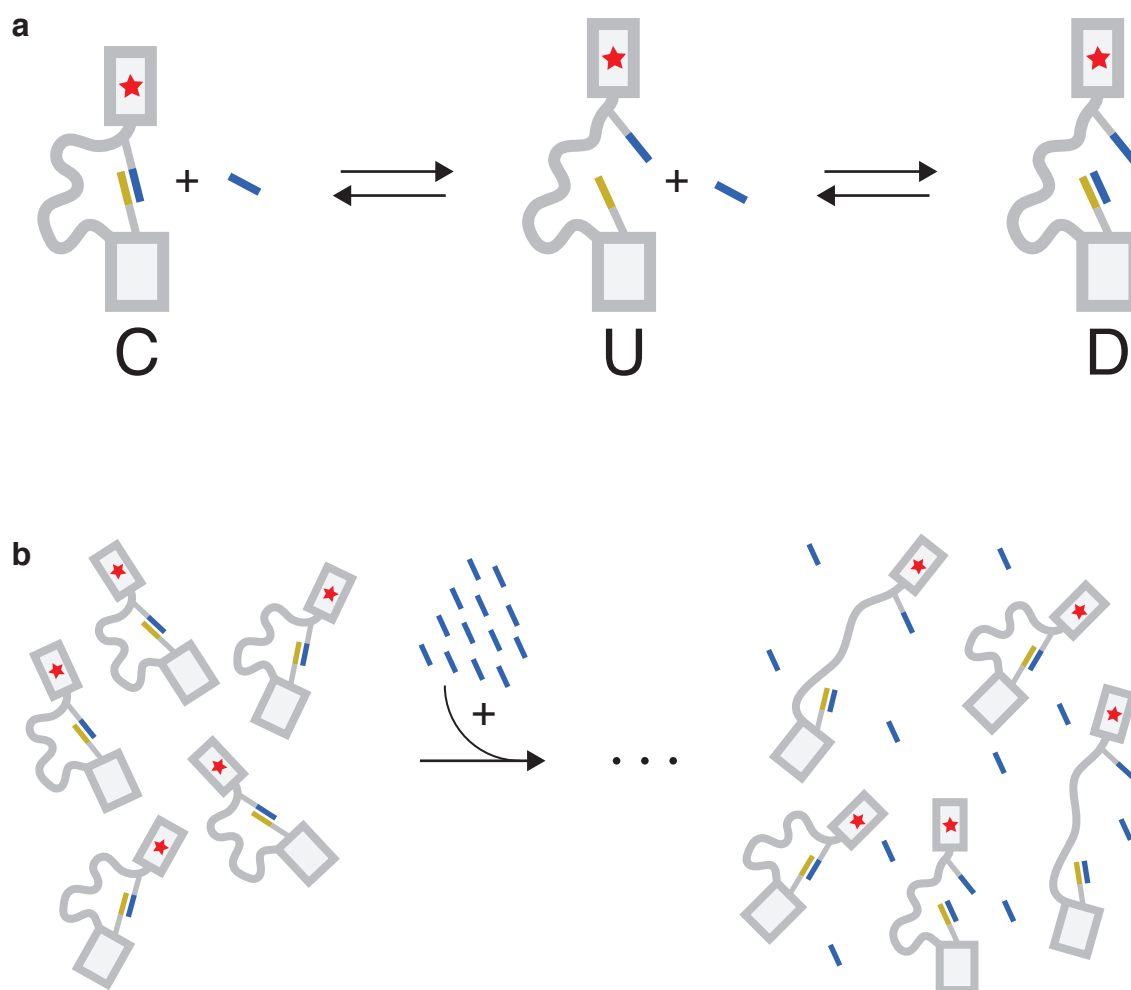


Abbildung 5.1: **Schemata der Reaktion zwischen dem Kompetitor und den DNA-Nano-Objekten**

DNA-Nano-Objekte in Grau, Sequenz B in Blau, Sequenz B' in Gold. Die roten Sterne deuten die Cy5-Farbstoffe an der mobilen Reporter-Einheit an. **(a)** Reaktionsschema mit den drei Zuständen des Systems aus einem DNA-Nano-Objekt und einem Kompetitor-Strang. Links: Zustand C mit dem DNA-Nano-Objekt im hybridisierten Zustand und freiem Kompetitor. Mitte: Zustand U mit dem DNA-Nano-Objekt im dissoziierten Zustand und freiem Kompetitor. Rechts: Zustand D mit Bindung des Kompetitors an den Interaktionspartner B' im Objekt. **(b)** Links: Vor dem Hinzufügen des Kompetitors liegen die DNA-Nano-Objekte mit 12 bp langer Interaktionssequenz fast dauerhaft im hybridisierten Zustand vor. Rechts: Nach dem Hinzufügen der Kompetitor-Stränge verschiebt sich die Besetzungsverteilung nach und nach vom Zustand C zum Zustand D, bis sich ein neues Besetzungsgleichgewicht eingestellt hat.

proximalen Verankerungsgeometrie und einer Konturlänge des Leinenstücks zwischen den Interaktionspartnern von $L = 2058$ bp verwendet.

Damit das Bandenprofil nach der zur Auftrennung der Probe notwendigen mehrstündigen Laufzeit des Gels immer noch die Verteilung zum Zeitpunkt der Auftragung der Proben auf das Gel widerspiegelt, muss die Bindung zwischen dem Kompetitor und dem Interaktionspartner B' ausreichend stabil sein. Andernfalls könnte sich der Kompetitor im Gel vom DNA-Nano-Objekt lösen, was eine Verschiebung des Bandenprofils von der langsameren zur schnelleren Bande nach sich zieht (vgl. Abb 1.13 in Abschnitt 1.4). Aus diesem Grund wurde für die Ensemble-Kompetitions-Experimente die 12 bp lange Interaktionssequenz CCAAGAGACACG:CGTGTCTCTTGG ausgewählt. Dementsprechend lautet die Sequenz des Kompetitors CCAAGAGACACG.

Die Vorhersage mit dem Nearest-Neighbor-Modell ergibt für die freie Energie der Bildung des 12-bp-Duplex etwa $-18,1$ kcal/mol¹. Die analog dazu vorhergesagten freien Energien der in Kapitel 4 untersuchten 8-bp-Duplexe lagen im Bereich zwischen $-10,8$ und $-10,6$ kcal/mol (vgl. Tabelle 4.1). Demnach sollte rein rechnerisch für die 12-bp-Duplexe die relative Besetzung des dissoziierten Zustands um vier bis fünf Größenordnungen geringer sein als für die 8-bp-Duplexe. Es kann folglich davon ausgegangen werden, dass im Kompetitions-Experiment die Besetzungswahrscheinlichkeit des Zustands U im Vergleich zu den Besetzungswahrscheinlichkeiten der Zustände C und D nur sehr gering ist. Da das Verweilen des Systems im Zustand U einen notwendigen Zwischenschritt des Übergangs vom Zustand C in den Zustand D darstellt, ist anzunehmen, dass sich nach dem Hinzufügen des Kompetitors zu einer Probe das neue Gleichgewicht nur sehr langsam einstellen wird.

Die Entwicklung der Besetzungsverteilung zwischen den Zuständen C und D nach Hinzufügen des Kompetitors wurde in Zeitreihen untersucht. Diese Zeitreihen werden im folgenden Abschnitt näher beschrieben.

5.2 Durchführung und Analyse der Kompetitions-Zeitreihen

Die DNA-Nano-Objekte wurden wie in Protokoll A.2.1 beschrieben hergestellt. Die intakte Spezies wurde per SEC aufgereinigt (vgl. Abschnitt 1.4 und Anhang A.2.3). Die aufgereinigten Objekte wurden auf mehrere PCR-Reaktionsgefäße verteilt (Abb. 5.2 a). Den einzelnen Proben einer Reihe wurde zu verschiedenen Zeitpunkten Kompetitor in einer definierten finalen Konzentration (je nach Zeitreihe 20, 40 oder 100 nM) hinzugefügt (Abb. 5.2 b). Die finale Konzentration der DNA-Nano-Objekte nach allen Schritten der Probenpräparation liegt schätzungsweise einem Bereich von etwa bis 0,1 bis 1,0 nM.

Vor und nach dem Hinzufügen des Kompetitors wurden die Proben in einer geschlossenen PCR-Apparatur bei einer konstanten Temperatur von 23 °C aufbewahrt. Zur Vermeidung von Adsorption der DNA-Nano-Objekte oder der Kompetitor-Stränge an den Wänden der Reaktionsgefäße enthielt die Pufferlösung aller Proben jeweils 0,05 % (v/v)

¹Die Vorhersage erfolgte mithilfe des DINAmelt Servers [95] für die Interaktionssequenzen inkl. T₁₂-Verbinder, eine Reaktionstemperatur von 23 °C, $[\text{Na}^+] = 0,5$ M und $[\text{Mg}^{++}] = 0$ M.

der Detergenzien Triton X-100 und TWEEN 20. Das Gesamtvolumen der Probe betrug nach Hinzufügen aller Bestandteile in jedem Reaktionsgefäß 100 μl .

Nach Aufbewahrungszeiten von mehreren Tagen bis Wochen wurden die Proben in Agarose-Gelen untersucht (Abb. 5.2c). Die Referenzprobe ohne Kompetitor zeigt im Gel eine einzelne Hauptbande. Die Proben mit Kompetitor zeigen neben einer Bande auf derselben Höhe eine zusätzliche, langsamere Bande. Mit zunehmender Zeitspanne zwischen dem Hinzufügen des Kompetitors zur Probe und dem Auftragen der Pro-

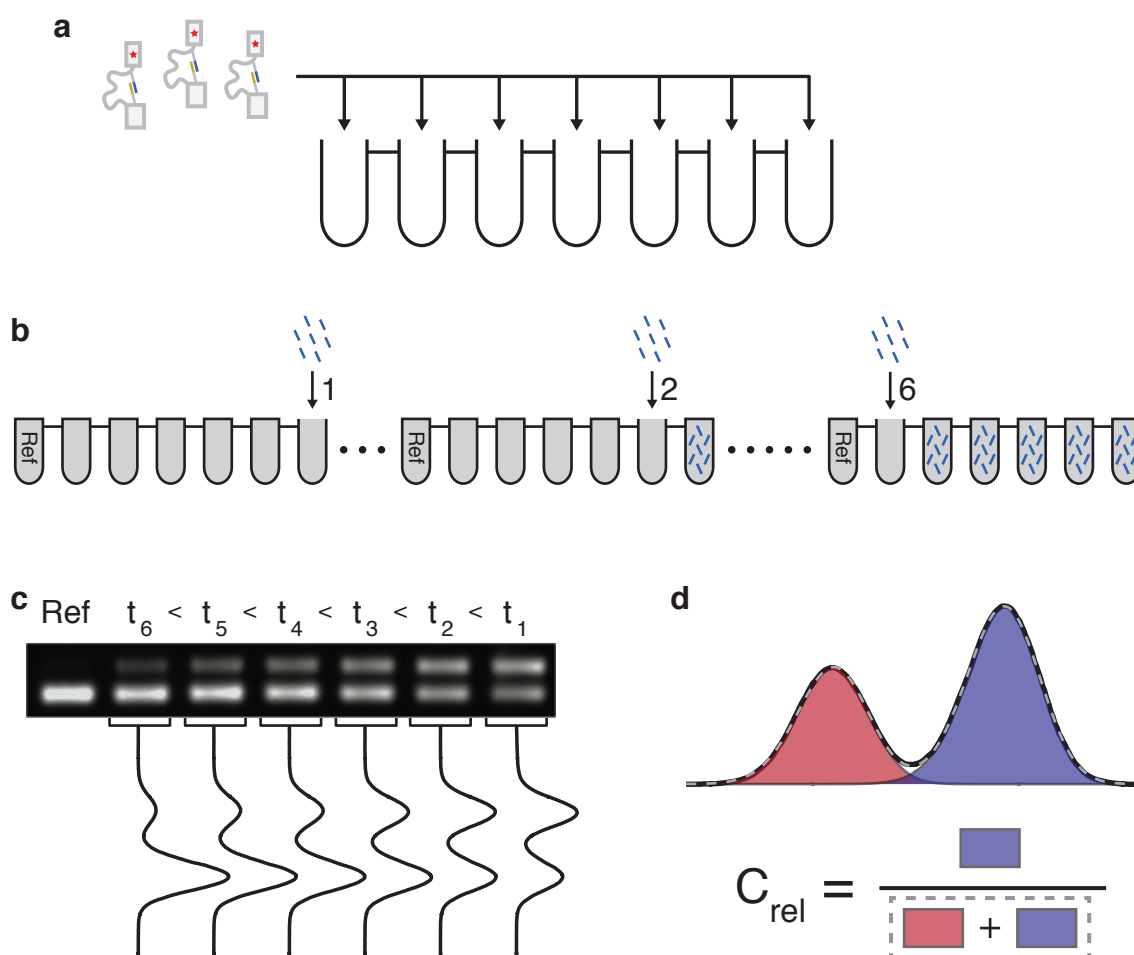


Abbildung 5.2: **Illustration der Durchführung und der Analyse der Ensemble-Kompetitions-Zeitreihen** (a) Aufteilung der aufgereinigten DNA-Nano-Objekte auf mehrere Reaktionsgefäße. (b) Hinzufügen der Kompetitor-Stränge zu verschiedenen Zeitpunkten. Ref: Referenzprobe ohne Kompetitor. (c) Scan eines Agarose-Gels zur Analyse der Proben. t_i bezeichnet die Zeit zwischen dem Hinzufügen des Kompetitors und dem Auftragen der Proben auf das Gel. Für das gezeigte Beispiel war $t_6 = 24$ h, $t_5 = 48$ h, $t_4 = 72$ h, $t_3 = 108$ h, $t_2 = 164$ h und $t_1 = 212$ h, mit einer Kompetitor-konzentration von 100 nM in der Probenlösung. Schwarze Linien unterhalb des Gel-Scans: Helligkeitsprofile, erstellt durch Summierung der Pixelhelligkeiten über die Breite der entsprechenden Gel-Bahnen. (d) Ermittlung von C_{rel} (relativer Anteil der Objekte in Zustand C). Schwarze Linie: Helligkeitsprofil. Blau: Fläche unter der Fit-Kurve der Normalverteilung für die schnellere Bande. Rot: Fläche unter der Fit-Kurve der Normalverteilung für die langsamere Bande. Graue gestrichelte Linie: Summe der zwei Normalverteilungen. C_{rel} berechnet sich aus dem Quotienten der blauen Fläche und der Summe aus der blauen und der roten Fläche.

be auf das Agarose-Gel ist die schnellere Bande immer schwächer und die langsamere Bande immer stärker ausgeprägt. Die schnellere Bande enthält den Anteil der Objekte im Zustand C, die langsamere Bande den Anteil der Objekte in den Zuständen D und U. Für jede Gel-Bahn wurde durch Summierung der Pixelhelligkeiten über die Breite der Bahn ein Helligkeitsprofil erstellt (Abb. 5.2 c). An das Helligkeitsprofil jeder Doppelbande wurde eine Summe aus zwei Normalverteilungen angefitzt (Abb. 5.2 d). Der Quotient aus der Fläche unter der Normalverteilung für die schnellere Bande und der Summe der Flächen unter beiden Normalverteilungen ergibt den Messwert für C_{rel} .

5.3 Berechnung des J-Faktors aus dem zeitlichen Verlauf von C_{rel}

Die Form des zeitlichen Verlaufs von C_{rel} ähnelt für jede der Zeitreihen mit 20, 40 und 100 nM Kompetitor einem exponentiellen Abfall (Abb. 5.3). Ein Besetzungsgleichgewicht scheint sich in jeder Reihe erst nach einer Zeit von mehreren Wochen einzustellen.

Zur Berechnung des J-Faktors aus dem zeitlichen Verlauf von C_{rel} wurden zwei Modelle formuliert. Der maßgebliche Unterschied zwischen den zwei Modellen liegt in der Grundannahme über eine mögliche Heterogenität der Kinetik der aufgereinigten DNA-Nano-Objekte. Während Modell 1 die Kinetik aller Objekte als gleich betrachtet, berücksichtigt Modell 2 die Möglichkeit verschiedener Populationen von Objekten mit unterschiedlicher Kinetik.

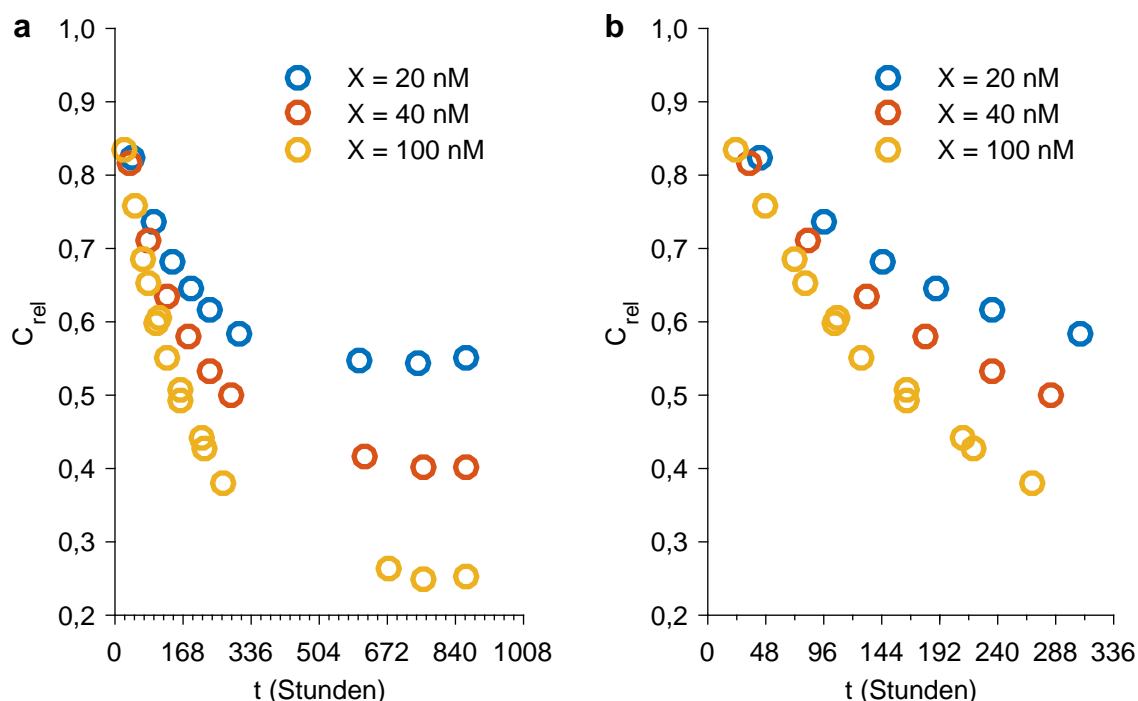


Abbildung 5.3: Verlauf der Messwerte von C_{rel} aus den drei Wettbewerbs-Zeitreihen

(a) Messwerte von C_{rel} über t (Zeitspanne zwischen dem Hinzufügen des Kompetitors und dem Auftragen der Proben auf das Agarose-Gel). Drei Zeitreihen für die Kompetitor-Konzentrationen $X = 20$ nM, $X = 40$ nM und $X = 100$ nM (siehe Legende). (b) Vergrößerte Darstellung des Bereichs bis $t = 336$ h.

Modell 1

Bezeichnen C_{rel} , U_{rel} und D_{rel} die relativen Besetzungen der Zustände C, U und D. Bezeichne X die Konzentration des Kompetitors in der Probenlösung. Bezeichne $k_{\text{on,bi}}$ die bimolekulare Duplex-Assoziationsrate, d. h. die Assoziationsrate für den Fall, dass sich beide Interaktionspartner frei in Lösung befinden. Sei die Duplex-Dissoziationsrate k_{off} unabhängig davon, ob die Dissoziation aus dem Zustand C oder aus dem Zustand D erfolgt. Die zeitliche Entwicklung der relativen Besetzungen der Zustände C, U und D wird dann durch das folgende Differentialgleichungssystem beschrieben:

$$\frac{d}{dt}C_{\text{rel}} = k_{\text{on,bi}} \cdot J \cdot U_{\text{rel}} - k_{\text{off}} \cdot C_{\text{rel}} \quad (5.1a)$$

$$\frac{d}{dt}U_{\text{rel}} = -k_{\text{on,bi}} \cdot (J + X) \cdot U_{\text{rel}} + k_{\text{off}} \cdot (C_{\text{rel}} + D_{\text{rel}}) \quad (5.1b)$$

$$\frac{d}{dt}D_{\text{rel}} = k_{\text{on,bi}} \cdot X \cdot U_{\text{rel}} - k_{\text{off}} \cdot D_{\text{rel}} \quad (5.1c)$$

Da X in allen Wettbewerbs-Zeitreihen deutlich größer ist als die Gesamtkonzentration der DNA-Nano-Objekte, wird X über den Verlauf eines Wettbewerbs-Experiments als annähernd konstant betrachtet. Aufgrund der weiter oben diskutierten geringen Besetzungswahrscheinlichkeit des Zustands U im Vergleich zu den Besetzungswahrscheinlichkeiten der Zustände C und D kann folgende Näherung angewendet werden:

$$U_{\text{rel}} \ll 1 \quad \Leftrightarrow \quad (C_{\text{rel}} + D_{\text{rel}}) \approx 1 \quad (5.2)$$

Außerdem wird angenommen, dass jedes Objekt, welches in den Zustand U übergeht, sehr schnell wieder in einen der anderen beiden Zustände C oder D übergeht. Deshalb kann neben U_{rel} auch die zeitliche Änderung von U_{rel} vernachlässigt werden:

$$\frac{d}{dt}U_{\text{rel}} \approx 0 \quad (5.3)$$

Unter allen getroffenen Annahmen und mit der Randbedingung $C_{\text{rel}}(0) = 1$ ergibt sich folgende Lösung für den zeitlichen Verlauf von C_{rel} nach Modell 1 (Anhang D.1 zeigt eine ausführlichere Herleitung):

$$C_{\text{rel}}(t) = \frac{1}{1 + \frac{X}{J}} + \frac{1}{1 + \frac{J}{X}} \cdot \exp(-k_{\text{off}} \cdot t) \quad (5.4)$$

Modell 2

Die Aufreinigung per SEC separiert die DNA-Nano-Objekte nach dem Anteil der Zeit, den sie im Zustand mit Schlaufe in der Leine verbringen. So werden Objekte aus der Probe entfernt, die z. B. aufgrund von Synthesefehlern in den Interaktionssequenzen nur selten oder nie eine Schlaufe in der DNA-Leine ausformen (in Abschnitt 1.4 als die „defekte Spezies“ bezeichnet). Es ist jedoch denkbar, dass in einem Teil der Objekte für die Wettbewerbs-Experimente Defekte auftreten, die die Bindung zwischen den

Interaktionspartnern zwar abschwächen, aber nicht dazu führen, dass diese Objekte hauptsächlich ohne Schlaufe vorliegen und in der SEC-Aufreinigung entfernt werden.

Man betrachte beispielsweise den Fall von Objekten mit nur ein bis zwei fehlenden Nukleotiden in der Interaktionssequenz B und korrekter Interaktionssequenz B'. Solche Objekte verbrächten mehr Zeit in Zustand U und böten somit häufiger eine Gelegenheit für die Bindung eines Kompetitors mit korrekter Sequenz B an den Interaktionspartner B'. Je nach Ausprägung der Defekte könnte sogar eine Strangverdrängung durch den Kompetitor und damit ein direkter Übergang vom Zustand C in den Zustand D möglich sein. Die Bindung zwischen dem Kompetitor und dem Interaktionspartner korrekter Sequenz wäre zudem energetisch günstiger als die Bindung zwischen den zwei verankerten Interaktionspartnern.

Als einfache Form der Berücksichtigung derartiger Szenarien wird in Modell 2 ein Anteil Γ von Objekten mit fehlerhaften Interaktionssequenzen eingeführt, die sich von den Objekten mit völlig intakten Interaktionssequenzen in zwei Punkten unterscheiden:

1. Die Abnahme von C_{rel} erfolgt für diesen Anteil der Objekte mit der Rate $k_{\Gamma} > k_{off}$.
2. Dieser Anteil der Objekte geht wegen der höheren Affinität des Kompetitors zum Interaktionspartner B' vollständig in den Zustand D über.

Es ergibt sich folgender Ausdruck für den zeitlichen Verlauf von C_{rel} nach Modell 2:

$$C_{rel}(t) = (1 - \Gamma) \cdot \left(\frac{1}{1 + \frac{X}{J}} + \frac{1}{1 + \frac{J}{X}} \cdot \exp(-k_{off} \cdot t) \right) + \Gamma \cdot \exp(-k_{\Gamma} \cdot t) \quad (5.5)$$

Fits der Modelle an die Messdaten

Beide Modelle wurden in individuellen und globalen Least-Squares-Fits an die Messdaten aus den Kompetitions-Zeitreihen gefittet (Abb. 5.4 und 5.5). In den individuellen Fits wurden die Modellparameter auf die bestmögliche Übereinstimmung der Funktionswerte mit den Daten aus einer einzelnen Kompetitor-Zeitreihe hin optimiert. In den globalen Fits wurden die Modellparameter auf die bestmögliche Übereinstimmung der Funktionswerte mit den Daten aus allen drei Kompetitor-Zeitreihen hin optimiert.

Die Kurven aus den individuellen Fits von Modell 1 zeigen eine gewisse Übereinstimmung mit den Messdaten aus den jeweiligen Zeitreihen (Abb. 5.4 a, b). Jedoch weichen die Verläufe der Fit-Kurven im Bereich $t < 336$ h sichtbar von den Datenpunkten der Messung von C_{rel} ab (Abb. 5.4 b). Die Kurven aus dem globalen Fit von Modell 1 weichen allgemein deutlich von den Verläufen der Messwerte ab, insbesondere für die Kompetitorkonzentrationen 20 nM und 100 nM (Abb. 5.4 c, d).

Die Fits von Modell 2 an die experimentellen Daten zeigen im Vergleich zu den Fits von Modell 1 eine deutlich größere Übereinstimmung mit den Verläufen der Messwerte von C_{rel} (Abb. 5.5). Insbesondere stimmen die Kurven aus dem globalen Fit von Modell 2 sehr viel besser mit den Daten überein als die Kurven aus dem globalen Fit von Modell 1 (Abb. 5.4 und 5.5, jeweils c und d).

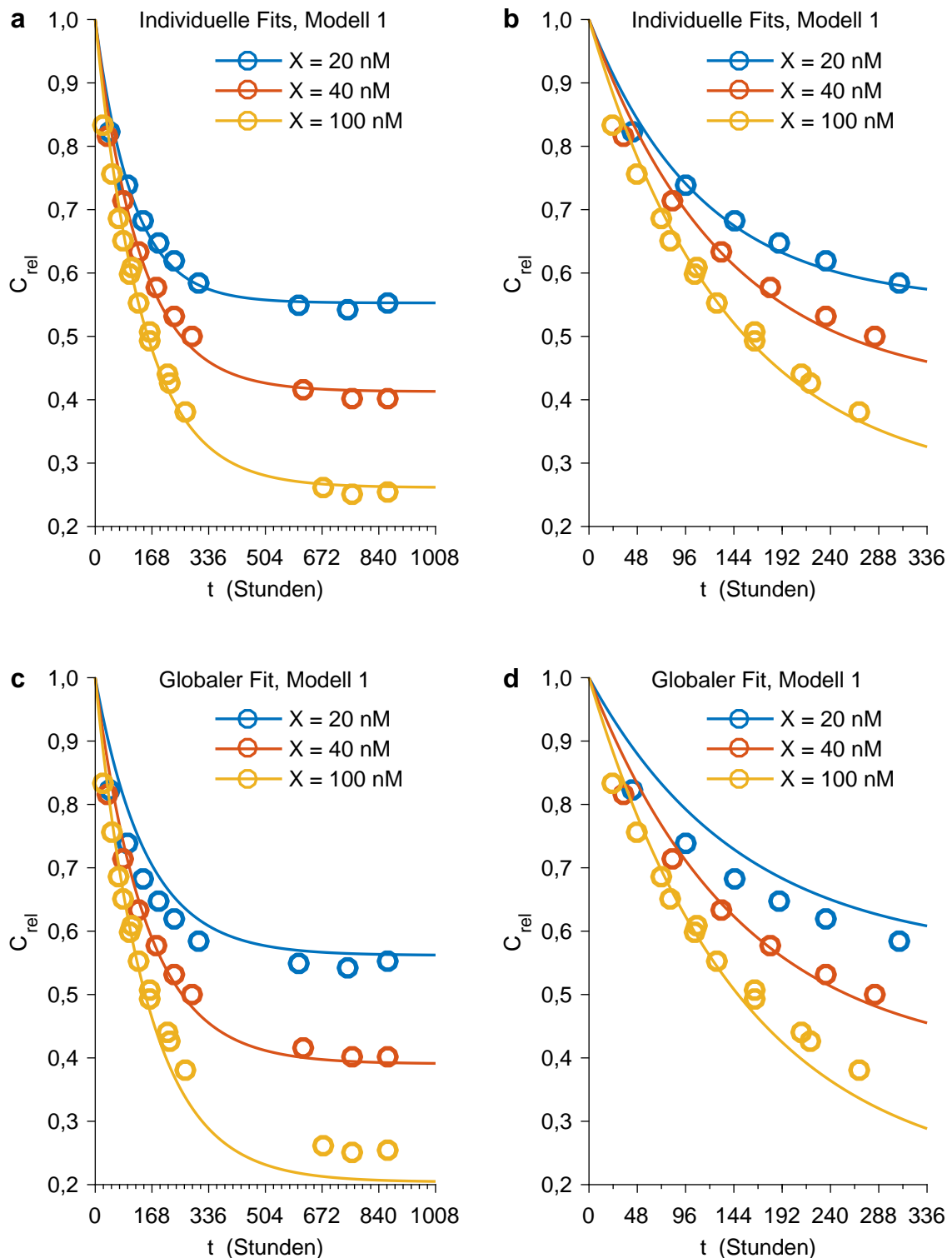


Abbildung 5.4: **Least-Squares-Fits von Modell 1 an die Messwerte von C_{rel}**

Marker: Messwerte von C_{rel} aus den drei verschiedenen Zeitreihen (siehe Legende). **(a)** Linien: Fit-Kurven aus den individuellen Fits von Modell 1 an die Daten aus den einzelnen Zeitreihen. **(b)** Wie a, mit vergrößerter Darstellung des Bereichs bis $t = 336$ h. **(c)** Linien: Fit-Kurven aus dem globalen Fit von Modell 1 an die Daten aus allen drei Zeitreihen. **(d)** Wie c, mit vergrößerter Darstellung des Bereichs bis $t = 336$ h.

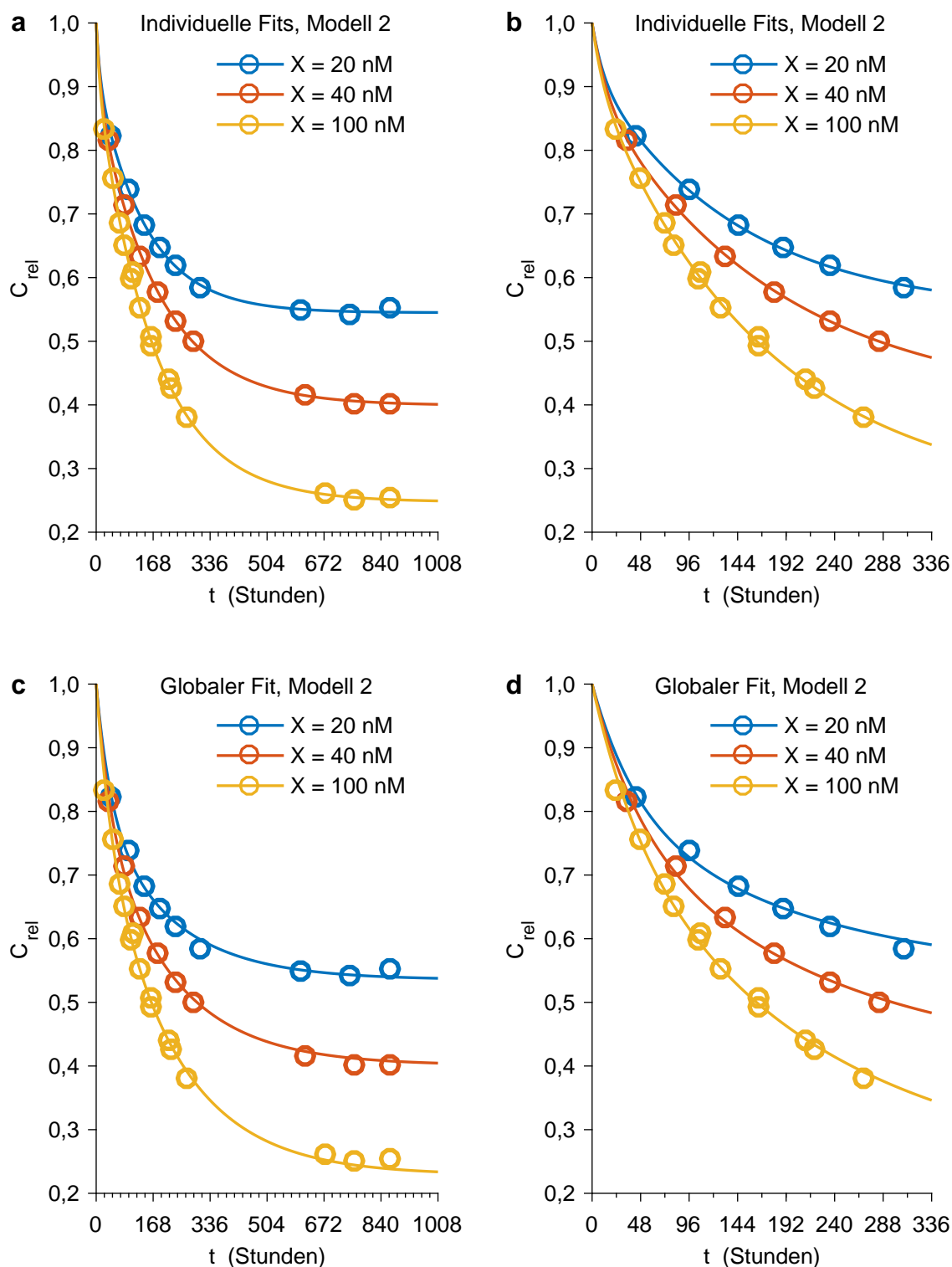


Abbildung 5.5: **Least-Squares-Fits von Modell 2 an die Messwerte von C_{rel}**

Marker: Messwerte von C_{rel} aus den drei verschiedenen Zeitreihen (siehe Legende). **(a)** Linien: Fit-Kurven aus den individuellen Fits von Modell 2 an die Daten aus den einzelnen Zeitreihen. **(b)** Wie a, mit vergrößerter Darstellung des Bereichs bis $t = 336 \text{ h}$. **(c)** Linien: Fit-Kurven aus dem globalen Fit von Modell 2 an die Daten aus allen drei Zeitreihen. **(d)** Wie c, mit vergrößerter Darstellung des Bereichs bis $t = 336 \text{ h}$.

Tabelle 5.1: **Werte der Fit-Parameter aus den Least-Squares-Fits von Modell 1 an die Messdaten**
 In Klammern: 95 %-Konfidenzintervalle der Parameterwerte.

Fit an Zeitreihe	J (nM)	k_{off} (10^{-6}s^{-1})
X=20 nM	24,7 (22,8 26,6)	2,50 (2,13 2,87)
X=40 nM	28,1 (24,4 31,8)	2,08 (1,73 2,44)
X=100 nM	35,4 (30,7 40,1)	2,01 (1,85 2,19)
Global	25,6 (23,0 28,2)	1,86 (1,69 2,02)

Tabelle 5.2: **Werte der Fit-Parameter aus den Least-Squares-Fits von Modell 2 an die Messdaten**
 In Klammern: 95 %-Konfidenzintervalle der Parameterwerte.

Fit an Zeitreihe	J (nM)	k_{off} (10^{-6}s^{-1})	Γ (%)	k_{Γ} (10^{-6}s^{-1})
X=20 nM	28,9 (26,1 31,7)	1,95 (1,72 2,18)	7,8 (3,8 11,7)	25,1 (-67,7 118)
X=40 nM	32,2 (31,3 33,2)	1,56 (1,49 1,63)	10,6 (9,0 12,2)	18,7 (10,4 27,0)
X=100 nM	37,1 (34,9 39,3)	1,66 (1,54 1,78)	8,5 (5,7 11,3)	24,7 (-3,4 52,9)
Global	39,7 (36,6 42,8)	1,32 (1,25 1,39)	19,5 (17,2 21,8)	6,40 (5,25 7,54)

Die Ergebnisse der individuellen Fits von Modell 2 sind allerdings mit einer gewissen Skepsis zu betrachten: Die Streuung der Werte der Parameter Γ und k_{Γ} sowie die großen 95 %-Konfidenzintervalle dieser Parameterwerte (vgl. Tabelle 5.2) deuten auf ein Überfitten der Daten durch das Modell hin. Der Grund hierfür könnte die geringe Zahl von Messwerten der einzelnen Zeitreihen bei kurzen Zeiten sein.

Anders verhält es sich beim globalen Fit von Modell 2: Die Konfidenzintervalle der Werte von Γ und k_{Γ} sind hier deutlich kleiner. Außerdem unterscheiden sich die Werte von Γ und k_{Γ} aus dem globalen Fit deutlich von den Werten von Γ und k_{Γ} aus den individuellen Fits. Ein Anteil fehlerhafter Objekte von $\Gamma \approx 20\%$ wirkt zwar auf den ersten Blick überraschend hoch, ist jedoch in Anbetracht bereits an anderer Stelle diskutierter Beobachtungen (vgl. Abschnitte 1.4 und 3.3.3) nicht unrealistisch.

Wertebereich des J-Faktors aus den Fits der Modelle an die Daten

Für beide Modelle liefern die individuellen Fits verschiedene Werte für J. Der globale Fit von Modell 2 passt am besten zu den Daten, weist allerdings mit 39,7 nM den größten

Wert von J aller Fits auf. Einerseits deutet die Heterogenität der erhaltenen Werte für J darauf hin, dass beide Modelle die Entwicklung der Probe nach Hinzufügen des Kompetitors nicht vollkommen realitätsgetreu beschreiben. Andererseits sind die Unterschiede nicht exorbitant und der Wertebereich von $25 \text{ nM} \lesssim J \lesssim 40 \text{ nM}$ liegt in derselben Größenordnung wie bereits publizierte theoretische [116] oder experimentelle [115] Werte für den J -Faktor von DNA-Fragmenten vergleichbarer Konturlänge.

5.4 Resümee zu den Ensemble-Kompetitions-Experimenten

Die Ergebnisse der Ensemble-Kompetitions-Experimente dienen zur groben Einordnung des J -Faktors der DNA-Nano-Objekte, jedoch in der hier präsentierten Form nicht zu seiner exakten Bestimmung. Mögliche Ansatzpunkte zur Verbesserung liegen in der Erhebung größerer Datenmengen bei mehreren Konkurrentorkonzentrationen und in feineren Zeitabständen sowie in der Entwicklung ausgefeilterer Modelle.

Die Populationsanalyse per Gel-Elektrophorese verlangt eine langsame Dissoziationskinetik der Interaktionspartner, die zu einer notwendigen langen Dauer der Zeitreihen führt. Die langsame Equilibrierung der Proben reduziert zwar einerseits mögliche Fehler in der zeitlichen Verortung der Messpunkte und vereinfacht die theoretische Beschreibung des Systems, andererseits stellt sie einen klaren Nachteil der Methode dar, wenn Zeit eine knappe Ressource ist. Außerdem eignet sich die Populationsanalyse in Gelen schlechter für Objekte mit kürzerer Leinenlänge und der distalen Verankerungsgeometrie, da die Auftrennung im Gel für diese Objekte weniger stark ausgeprägt ist.

In Abschnitt 4.3 wurde aus den Messergebnissen von TMFM-Experimenten mit verschiedenen Sequenzvarianten der J -Faktor von DNA-Nano-Objekten mit $L = 1554 \text{ bp}$ und der distalen Verankerungsgeometrie auf etwa 9 nM geschätzt. In Abschnitt 4.1 wurde festgestellt, dass der J -Faktor in der proximalen Verankerungsgeometrie bei gleichem Wert von L geringer ist als in der distalen Verankerungsgeometrie und darüber hinaus mit steigendem Wert von L abfällt (vgl. Abb. 4.9). Es ist daher interessant, dass sich aus den Ensemble-Kompetitions-Experimenten für Objekte mit $L = 2058 \text{ bp}$ und der proximalen Verankerungsgeometrie deutlich höhere Werte als 9 nM für den J -Faktor ergeben. Diese Beobachtung untermauert zum einen den störenden Einfluss der Probenoberfläche im TMFM-Experiment auf die Bindung zwischen den verankerten Interaktionspartnern. Zum anderen kann daraus gefolgert werden, dass die Ergebnisse aus den Ensemble-Kompetitions-Experimenten mit Objekten fernab einer Probenoberfläche nicht repräsentativ sind für die Situation oberflächengebundener Objekte im TMFM-Experiment.

Ein geeigneterer Weg zur Bestimmung des J -Faktors der DNA-Nano-Objekte könnte in der direkten TMFM-Messung der Hybridisierungskinetik bei verschiedenen Konzentrationen von Konkurrentor in der umgebenden Lösung liegen. Diese Messungen könnten mit einer Auswahl der 8-bp -Interaktionssequenzen aus Kapitel 4 und entsprechender Sequenz des Konkurrentors erfolgen. Die theoretischen Überlegungen in Anhang D.2 geben einen Anhaltspunkt zur Berechnung des J -Faktors aus den Ergebnissen solcher Einzelmolekül-Kompetitions-Experimente.

Fazit und Ausblick

Diese Dissertation beschreibt die Entwicklung, Charakterisierung und Anwendung der neuartigen Einzelmolekül-Methode Tethered Multi-Fluorophore Motion (TMFM). Mit Hilfe des Konstruktionsprinzips Scaffolded DNA-Origami wurden DNA-Nano-Objekte für TMFM entworfen und hergestellt. Die Dynamik der hergestellten Objekte wurde experimentell charakterisiert. Darüber hinaus wurde eine Prozedur für die Analyse von TMFM-Daten mit zwei Beweglichkeitszuständen der mobilen Reporter-Einheiten implementiert. Mit der neu entwickelten Methode wurden drei Studien zur Hybridisierungskinetik kurzer DNA-Doppelstränge in Abhängigkeit von verschiedenen Parametern durchgeführt.

Die folgenden Absätze heben einige Besonderheiten von TMFM hervor und zeigen Ansatzpunkte für die Weiterentwicklung der Methode auf.

Gleichgewichts-Einzelmolekül-Messung ohne Austausch der Interaktionspartner

TMFM ermöglicht die Beobachtung der Interaktion desselben Paares von Interaktionspartnern im Gleichgewicht über einen Zeitraum von mehr als einer Stunde. In den hier durchgeführten Experimenten wurden bis zu ≈ 500 Assoziations- und Dissoziationsereignisse ein- und desselben Interaktionspartnerpaares registriert. Die einzige dem Autor bekannte weitere Einzelmolekül-Studie der DNA-Hybridisierungskinetik ohne Austausch der Interaktionspartner ist die Arbeit von Cisse et al. [34] aus dem Jahr 2012. Dabei wurden die interagierenden DNA-Einzelstränge paarweise in Vesikeln isoliert. Die Messung der Hybridisierungskinetik basiert auf Förster-Resonanz-Energie-Transfer (FRET) zwischen einzelnen Fluoreszenzfarbstoffmolekülen an den Interaktionspartnern. Die Beobachtungsdauer einzelner Interaktionspartnerpaare ist in diesem Ansatz somit durch die Lebenszeit der einzelnen Farbstoffmoleküle begrenzt und beträgt bestenfalls wenige Minuten.

Somit stellt die in der hier vorgelegten Arbeit erfolgte Einzelmolekül-Messung der DNA-Hybridisierungskinetik im Gleichgewicht, ohne Austausch der Interaktionspartner und über einen ausgedehnten Zeitraum nach dem Kenntnisstand des Autors ein absolutes Novum dar.

Detektion molekularer Heterogenität

Die Langzeitbeobachtung ohne Austausch der Interaktionspartner legt den Grundstein für die verlässliche Identifikation von Subpopulationen unterschiedlicher Kinetik innerhalb einer Probe auf der Basis probabilistischer Überlegungen. Im Fall der in dieser Arbeit präsentierten TMFM-Messdaten erfolgte für jede Probe eine Einteilung der

Objekte in eine Hauptpopulation homogener Kinetik und in Ausreißer-Objekte mit einer davon deutlich abweichenden Kinetik. Eine derartige Einteilung ist für Daten aus Einzelmolekül-Messungen mit Austausch der Interaktionspartner nur begrenzt möglich und für Daten aus Ensemble-Messungen vollkommen unmöglich.

Modulation der gemessenen Kinetik

Die Verankerung der molekularen Interaktionspartner kann an nahezu beliebigen Positionen des Partikel-Leinen-Systems erfolgen. Die Separation der Interaktionspartner entlang der Kontur der DNA-Leine kann dabei gezielt zur Modulation des J-Faktors und somit der im TMFM-Experiment gemessenen Assoziationskinetik genutzt werden. Es ist beispielsweise denkbar, schnell assoziierende Interaktionspartner durch ein möglichst langes Leinenstück zu trennen, um die Bindekinetik für die Beobachtung mit der begrenzten Bildrate einer Kamera-Aufnahme zugänglich zu machen.

Messung molekularer Kinetiken in Abhängigkeit innerer oder äußerer Parameter

Die Studie der Hybridisierungskinetik der sieben Sequenzvarianten in Abschnitt 4.3 demonstriert die gute Eignung von TMFM zur Messung bimolekularer Kinetiken in Abhängigkeit innerer Parameter. Damit sind spezifische Eigenschaften der Interaktionspartner – in diesem Fall die spezifische Basenpaarsequenz – gemeint. Die äußeren Bedingungen wie der Aufbau der Partikel-Leinen-Systeme und insbesondere die Zusammensetzung des Messpuffers waren in allen Experimenten gleich. Unterschiede in den gemessenen Kinetiken spiegeln ausschließlich die Unterschiede zwischen den verschiedenen Interaktionspartnerpaaren wider.

Die Studie der $[\text{Na}^+]$ -Abhängigkeit der Hybridisierungskinetik in Abschnitt 4.2 ist ein Beispiel für die Messung von Kinetiken in Abhängigkeit äußerer Parameter. Die Ergebnisse dieser Untersuchung offenbaren eine Schwachstelle von TMFM bzw. DNA-Leinen-basierter Methoden im Allgemeinen: Eine Veränderung äußerer Parameter beeinflusst unter Umständen auch die mechanischen Eigenschaften der DNA-Leine. Unterschiede in den gemessenen Kinetiken rühren dann nicht zwingend allein von Änderungen des Verhaltens der Interaktionspartner her. Diese Problematik bestünde sehr wahrscheinlich auch bei der Messung molekularer Kinetiken in Abhängigkeit anderer äußerer Parameter wie der Temperatur, des pH-Werts, der Viskosität oder DNA-bindender Moleküle. In einigen dieser Fälle könnten womöglich durch entsprechende Kontrollexperimente geeignete Korrekturen der Messungen erarbeitet werden.

TMFM-Messungen mit Proteinen als Interaktionspartner

Der mögliche Anwendungsbereich von TMFM beschränkt sich nicht auf Messungen der DNA-Hybridisierungskinetik. Durch punktuelle Modifikationen des Systems sollten beispielsweise TMFM-Studien von Protein-DNA-Interaktionen und Protein-Protein-Interaktionen realisierbar sein. Moderne chemische Konjugationsmethoden [139, 140] ermöglichen die kovalente Verbindung von Oligonukleotiden mit spezifischen Ami-

nosäureresten von Proteinen. In der Literatur finden sich zahlreiche Beispiele für die Verankerung von Protein-DNA-Konjugaten an DNA-Nano-Objekten [16,141–153]. Doppelsträngige Erkennungsmotive für DNA-bindende Proteine könnten entweder ähnlich wie die hier verwendeten einzelsträngigen Interaktionspartner an der DNA-Leine verankert oder durch gezielte Anpassung der Scaffold-Sequenz direkt in die DNA-Leine eingefügt werden.

Weiterentwicklung der DNA-Nano-Objekte mit spezifisch designten Scaffold-Sequenzen

In den vergangenen Jahren wurden verschiedene Methodiken zur Konzeption und Herstellung der Scaffold-Stränge für Scaffolded DNA-Origami vorgestellt [154–161]. Ein für TMFM relevanter Aspekt ist die direkte Produktion linearer statt zirkulärer Scaffold-Stränge [154,156–158,161]. Insbesondere stellt die Einbringung autokatalytischer Deoxyribozym-Sequenzmotive in die Scaffold-Sequenz eine mit biotechnologischen Produktionsverfahren vereinbare Vorgehensweise zur Linearisierung des Scaffolds dar [161,162].

Ein weiterer Ansatzpunkt besteht in der Konzeption verschiedener Varianten von Scaffold-Strängen mit konstanten Abschnitten für die Reporter-Einheiten sowie die Verankerungspunkte der Interaktionspartner und zwischengelagerten variablen Abschnitten für die verschiedenen Teile der DNA-Leine. Mit derartigen Scaffold-Strängen wäre der in Abschnitt 1.2.4 beschriebene Aufroll-Mechanismus obsolet, was eine deutlich geringere Variation zwischen den Staple-Sätzen in den Selbstassemblierungsreaktionen verschiedener Objektvarianten bedeuten würde. Zum einen wäre die Herstellung der Objekte dadurch weniger fehleranfällig, zum anderen würde die Vergleichbarkeit zwischen verschiedenen Leinenlängen-Varianten der Objekte erhöht.

Ansätze zur Stabilisierung des Fluoreszenzsignals der mobilen Reporter-Einheiten

Die Verbreiterung der Auslenkungsabfolgen der Reporter-Einheiten im Verlauf der Messungen wegen der schleichenden Abnahme des Signal-zu-Rausch-Verhältnisses (siehe Abschnitte 2.3 und 2.5.1) verkompliziert die HMM-Analyse. Ohne eine entsprechende Berücksichtigung des Messartefakts in der HMM-Prozedur würde sich die Qualität der Zustandszuweisung verringern. Die in ihrer Umsetzung einfache segmentweise Betrachtung der Emissionswahrscheinlichkeiten (siehe Abschnitt 3.2.2) kompensiert das Messartefakt zum Teil, aber nicht vollständig. Innerhalb der einzelnen Segmente einer Abfolge tritt im Allgemeinen immer noch eine Verbreiterung der Auslenkungsverteilungen auf, obgleich diese deutlich weniger stark ausgeprägt ist als die Verbreiterung über die Gesamtdauer der Abfolge. Einerseits könnte eine ausgefeiltere Prozedur das Messartefakt womöglich noch besser kompensieren. Andererseits könnte die Ursache des Artefakts durch eine weitere Stabilisierung des Fluoreszenzsignals der mobilen Reporter-Einheiten eingedämmt werden.

Eine offensichtliche Verbesserungsmöglichkeit besteht in der Verwendung besser zur Langzeitbeobachtung geeigneterer Fluoreszenzfarbstoffe. In den hier durchgeführten TMFM-Experimenten besaßen die mobilen Reporter-Einheiten 20 Cy5-Farbstoffe, die immobilen Reporter-Einheiten 20 Cy3-Farbstoffe. In der Regel zeigten die immobilen

Reporter-Einheiten ein deutlich helleres und stabileres Fluoreszenzsignal als die mobilen Reporter-Einheiten. Allein durch Vertauschen der Farbstoffsorten der Reporter-Einheiten ließe sich folglich bereits eine Verbesserung des Signals der mobilen Reporter-Einheiten erreichen. Unter der Vielzahl kommerziell verfügbarer Fluoreszenzfarbstoffe befinden sich vermutlich noch besser geeignete Alternativen, wie z. B. der in DNA-PAINT-Messungen [21, 163, 164] standardmäßig verwendete Fluoreszenzfarbstoff Cy3B.

Zusätzlich könnte in einem neuen strukturellen Design der mobilen Reporter-Einheit die Anzahl der Fluoreszenzfarbstoffmoleküle erhöht werden. Dabei gilt nach wie vor die Maßgabe, einerseits das Volumen der Reporter-Einheit möglichst klein zu halten, und andererseits zur Vermeidung von Fluoreszenzauslöschung einen gewissen Mindestabstand zwischen den einzelnen Farbstoffmolekülen zu wahren. Eine Möglichkeit zur optimalen räumlichen Anordnung der Fluoreszenzfarbstoffe könnte in der Neukonzeption der mobilen Reporter-Einheit als drahtgitterartiges Gebilde aus DNA-Doppelhelizes [165–171] statt der bisherigen Helixbündelform liegen.

Vor Kurzem wurden Ansätze zur Erhöhung des Fluoreszenzsignals oder zur Verlängerung der Beobachtungsdauer fluoreszenzmarkierter Objekte durch kontinuierlichen Austausch farbstofftragender DNA-Stränge vorgestellt [164, 172]. Die Bindung der farbstofftragenden DNA-Stränge an die Objekte ist dabei schwach genug, dass sie sich nach kurzer Zeit wieder von den Objekten lösen und durch überschüssige Stränge aus der Probenlösung ersetzt werden können. Grundsätzlich käme ein derartiges System der Fluoreszenzmarkierung auch für die mobilen Reporter-Einheiten der TMFM-Objekte in Frage. Falls es sich bei den mit TMFM untersuchten Interaktionspartnern um DNA-Duplexe handelt, könnten die überschüssigen farbstofftragenden Einzelstränge jedoch einen störenden Einfluss auf die Hybridisierung der Interaktionspartner haben. Außerdem könnte das dauerhaft hohe Hintergrundsignal in der Probenlösung das Tracking der Reporter-Einheiten erschweren, falls die farbstofftragenden DNA-Stränge nicht derart gestaltet sind, dass sie im ungebundenen Zustand eine vernachlässigbare Emission aufweisen [173].

Globale Optimierung der Übergangsmatrix in der HMM-Prozedur

Im Rahmen der hier vorgestellten HMM-Prozedur (siehe Kapitel 3) erfolgten die Optimierung der Übergangsmatrix \mathcal{T} und die Zustandszuweisung anhand der Auslenkungsabfolgen individuell für jede mobile Reporter-Einheit. Dieses Vorgehen ist aufgrund der beobachteten Heterogenität der Kinetik innerhalb einzelner Proben sinnvoll. Nach der Identifikation der Hauptpopulation einer Probe könnte allerdings eine globale Optimierung der Übergangsmatrix anhand der Auslenkungsabfolgen aller Objekte der Hauptpopulation vorgenommen werden. Eine anschließende erneute Zustandszuweisung auf Basis dieser global optimierten Übergangsmatrix stünde auf einem solideren statistischen Fundament als die Zustandszuweisung auf Basis der individuell optimierten Übergangsmatrizen, insbesondere im Fall von Proben mit wenigen Zustandsübergängen je einzelner Abfolge. Es wäre sogar eine iterative Prozedur denkbar, in der jede Iteration die globale Optimierung der Übergangsmatrix, die erneute Zustandszuweisung je Aus-

lenkungsabfolge und eine erneute Klassifizierung der Objekte in Hauptpopulation und Ausreißer beinhaltet.

Weiterentwicklung vom MLHMM zum BHMM

Eine weitere Möglichkeit zur Verbesserung der HMM-Prozedur besteht in der Umstellung vom hier verwendeten Maximum Likelihood HMM (MLHMM) zu einem *Bayesian* HMM (BHMM), wie z. B. in [107] beschrieben. Aufgrund der meist unvollständigen Statistik experimenteller Daten können die MLHMM-Schätzungen der Modellparameter mit sehr großen Unsicherheiten behaftet sein, deren verlässliche Bestimmung nicht ohne Weiteres möglich ist. Ein Bayesian HMM arbeitet von vornherein mit Wahrscheinlichkeitsverteilungen der Modellparameter. Das Ergebnis der BHMM-Analyse beinhaltet die statistischen Unsicherheiten sowohl der Modellparameter als auch der Zuweisungen der verborgenen Zustände. [107] Durch die Implementierung einer BHMM-Prozedur wäre somit eine differenziertere Interpretation der TMFM-Messdaten möglich.

Beitrag der Ergebnisse dieses Promotionsprojekts zur Sequenzoptimierung für DNA-PAINT

Ein Teil der in dieser Arbeit präsentierten Ergebnisse zur Sequenzabhängigkeit der DNA-Hybridisierungskinetik (siehe Abschnitt 4.3) diente als Ausgangspunkt für eine weiterführende Studie [163] der Arbeitsgruppe von Prof. Ralf Jungmann². Das Ziel dieser Studie war die weitere Optimierung der Super-Resolution-Mikroskopiemethode *DNA Points Accumulation in Nanoscale Topography* (DNA-PAINT) [20,21]. Die Funktionsweise dieser Methode beruht auf der wiederholten kurzzeitigen Bindung von fluoreszenzmarkierten *Imager*-Strängen an dazu revers-komplementäre, an definierten Punkten der abzubildenden Probe verankerte *Docking*-Stränge. Für einen möglichst hohen zeitlichen Durchsatz der DNA-PAINT-Messungen und für die Anwendung der Methode zur Abbildung zellulärer Systeme ist eine Maximierung der Anzahl detektierter Photonen je markierter Position pro Zeiteinheit notwendig. Die Rate der erneuten Assoziation der *Imager*-Stränge an unbesetzte *Docking*-Stränge spielt dabei eine maßgebliche Rolle.

In der Veröffentlichung von Schueder et al. [163] berichtet die Gruppe u. a. über die Optimierung der Sequenzen der *Imager*- und *Docking*-Stränge. Ausgehend von der Sequenz des am schnellsten assoziierenden 8-bp-Duplex CCTCCTCC:GGAGGAGG der hier präsentierten Arbeit (Sequenzvariante v1 in Abschnitt 4.3) testeten Schueder et al. verschiedene 7-bp-Duplexe mit dazu ähnlicher Basenpaarsequenz. Dabei zeigte die Kombination des *Imager*-Strangs mit der Sequenz TCCTCCC und dem *Docking*-Strang mit der Sequenz GGGAGGA die schnellste Assoziationskinetik. Mit diesen Sequenzen erfolgte die Neubesetzung unbesetzter *Docking*-Stränge im Mittel um einen Faktor von etwa 5 schneller als mit bisher verwendeten Kombinationen aus *Imager*- und *Docking*-Strängen.

Das Beispiel der Sequenzoptimierung für DNA-PAINT veranschaulicht den möglichen Nutzen der Charakterisierung biomolekularer Kinetiken für bionanotechnologische Anwendungen. Die hier vorgestellte Methode TMFM birgt das Potential, einen entscheidenden Beitrag zum fundierten Verständnis biomolekularer Systeme zu leisten.

²Max-Planck-Institut für Biochemie, Martinsried

A Materialien und Methoden

Sofern nicht anders angegeben, wurden Chemikalien von der Firma Carl Roth bezogen.

A.1 Reagenzien

A.1.1 DNA

Für alle in dieser Arbeit hergestellten DNA-Nano-Objekte wurde Scaffold vom Typ p8064 verwendet. Dieser wurde nach der in [92] beschriebenen Methodik von anderen Mitarbeiter*innen der AG Dietz hergestellt. Herzlichen Dank an das „Team Scaffold“!

Alle für diese Arbeit verwendeten Staple-Oligonukleotide wurden von der Firma Eurofins Genomics bezogen. Verschiedene Staple-Gruppen wurden in verschiedenen Reinheitsgraden bestellt:

- HPLC Die interaktionspartnertragenden Staples für die Einzelmolekül-Experimente und alle Staples mit Fluoreszenzfarbstoffmodifikationen
- HYPUR Das Kompetitor-Oligonukleotid und die interaktionspartnertragenden Staples für die Ensemble-Kompetitions-Studie
- HPSF Alle anderen Staples

Listen mit den Staple-Sequenzen finden sich in den zusätzlichen Informationen zu [138].

A.1.2 Puffer

CutSmart

Standardpuffer der Firma New England Biolabs (NEB) für Enzymreaktionen.
50 mM Kaliumacetat | 20 mM Tris-Acetate | 10 mM Magnesiumacetat | 100 µg/ml BSA

FoB500N

5 mM Tris-HCl (pH 8) | 1 mM EDTA | 500 mM NaCl

0,5xTBE

44,5 mM Tris | 44,5 mM Borsäure | 1 mM EDTA

6xGel-Ladepuffer

5 mM Tris | 15 % (w/v) Ficoll-400 | 0,1 % (w/v) Bromphenolblau

T50

10 mM Tris | 50 mM NaCl

10xTE

100 mM Tris | 10 mM EDTA

Fluoreszenzmikroskopie-Aufnahmepuffer (FM-AB)

100 mM Tris-HCl (pH 8) | 2 mM Trolox | 0,8 % (v/v) D-Glucose |
442 U/ml Glucose-Oxidase | 2170 U/ml Catalase | 150 mM oder 500 mM oder 1 M NaCl
(Trolox, Glucose, Glucose-Oxidase und Catalase von Sigma-Aldrich)

Fluoreszenzmikroskopie-Waschpuffer (FM-WB)

Wie FM-AB, jedoch ohne Glucose-Oxidase und Catalase.

A.2 Laborprotokolle

A.2.1 Herstellung der DNA-Nano-Objekte

Selbstassemblierungsreaktionen

Die Reaktionsgemische für die Selbstassemblierung der DNA-Nano-Objekte hatten standardmäßig ein Volumen von 50 μ l. Als Faltungspuffer wurde standardmäßig einfacher CutSmart-Puffer verwendet. Scaffold und Staples wurden standardmäßig in folgenden finalen Konzentrationen eingesetzt:

10 nM	Scaffold
12,5 nM	Interaktionspartnertragende Staples
50 nM	Zum auszuschneidenden Scaffold-Teil komplementäre Staples
100 nM	Alle anderen Staples

Die Reaktionsgemische wurden in einer PCR-Apparatur vom Typ Tetrad (MJ Research, früher: Bio-Rad) standardmäßig der folgenden Temperaturabfolge unterzogen: 65 °C für 15 Min., dann schrittweise Abkühlung (1 h pro °C) über den Bereich von 60 bis 40 °C. Die Proben wurden entweder direkt weiterverarbeitet oder bis zur Weiterverarbeitung bei 4 °C gelagert.

Verdau mit Restriktionsenzym Cla I

Jedes Reaktionsgemisch für den enzymatischen Verdau enthielt 10 µl des Reaktionsgemischs einer Selbstassemblierungsreaktion und 10 U Cla I (New England Biolabs), bei einer finalen Konzentration von 1xCutSmart-Puffer in einem Gesamtvolumen von 50 µl. Dieses neue Gemisch wurde für 2 h bei 37 °C inkubiert und im Anschluss entweder direkt weiterverarbeitet oder bis zur Weiterverarbeitung bei 4 °C gelagert.

A.2.2 Gel-Elektrophorese

Die Gele zur Auftrennung von Proben der hergestellten DNA-Nano-Objekte enthielten standardmäßig 0,8 % (w/w) UltraPure Agarose (Invitrogen) und 4,4 mM MgCl₂ in 0,5xTBE-Puffer. Der Laufpuffer war standardmäßig 0,5xTBE-Puffer mit 4,4 mM MgCl₂. Vor dem Auftragen auf das Gel wurden die Proben der hergestellten DNA-Nano-Objekte im Verhältnis 5:1 mit 6xGel-Ladepuffer vermischt. Die Auftrennung im Gel erfolgte über eine Laufzeit von mindestens 90 Minuten bei einer Spannung von 75 V. Während der Laufzeit wurde die Gelkammer zur Kühlung in ein Eiswasserbad gestellt. Bei Laufzeiten von mehr als 90 Minuten wurden der Laufpuffer und das Eiswasserbad standardmäßig alle 90 Minuten erneuert. Die Gele wurden mit einem Laser-Scanner (Modell: Typhoon FLA 9500, GE Healthcare) gescannt. Die Abbildung der DNA-Nano-Objekte im Gel erfolgte dabei über passende Anregung der Fluoreszenzfarbstoffe Cy3 bzw. Cy5 und Detektion des entsprechenden Fluoreszenzlichts.

Aufreinigung der DNA-Nano-Objekte aus Agarose-Gelen

Für die TMFM-Experimente wurden die DNA-Nano-Objekte aus Agarose-Gelen aufgereinigt. Vor dem Scan wurden die Ränder des jeweiligen Aufreinigungsgels an mehreren Stellen eingekerbt. Auf einem Ausdruck des Gel-Scans wurden die Objekt-Banden markiert. Der Ausdruck wurde unter eine transparente Platte gelegt. Das Gel wurde auf die Platte gelegt und über die Kerben an den Rändern mit dem Ausdruck bündig ausgerichtet. An den Positionen der markierten Banden wurden Stücke aus dem Gel ausgeschnitten. Die DNA-Nano-Objekte wurden per Zentrifugation durch Zelluloseacetat-Filter mit einer Porengröße von 0,45 µm (Bio-Rad oder Costar) aus den ausgeschnittenen Gel-Stücken extrahiert. Der Filter-Durchfluss wurde entweder direkt weiterverarbeitet oder bis zur Weiterverarbeitung bei 4 °C gelagert.

A.2.3 Size Exclusion Chromatography (SEC)

Für SEC wurde eine Apparatur vom Typ 1260 Infinity (Agilent Technologies) mit einer Säule vom Typ Bio SEC-5 (Agilent Technologies) verwendet. Proben der verdauten DNA-Nano-Objekte wurden im Verhältnis 1:1 mit FoB500N-Puffer vermischt. Die Mischung wurde bei Raumtemperatur über Nacht equilibriert und dann in den Autosampler der Apparatur geladen. Die Probe wurde mit FoB500N-Puffer bei einer Flussrate von 2,0 ml/min durch die Säule gespült. Die Zeit je Fraktion betrug 0,17 Minuten.

Auswahl der Fraktionen für die Ensemble-Kompetitions-Experimente

Proben aus den SEC-Fraktionen wurden in Agarose-Gelen analysiert. Jene Fraktionen, die im Gel die Bande der intakten Spezies, aber keine Bande der defekten Spezies zeigten, wurden zur Weiterverwendung in den Ensemble-Kompetitions-Experimenten zu einem großen Probenvolumen kombiniert. Zur Vermeidung von Adsorption der DNA-Nano-Objekte an den Reaktionsgefäßwänden wurden Triton X-100 und TWEEN 20 (beide Sigma-Aldrich) in einer finalen Konzentration von jeweils 0,05 % (v/v) hinzugefügt. Bis zur Weiterverwendung wurde die Probe bei 4 °C gelagert.

A.2.4 Einzelmolekül-Fluoreszenzmikroskopie

Herstellung der Probenkammern

Zunächst wurden Deckgläschen für die Böden der Probenkammern einseitig mit einer Mischung im Verhältnis 1:1000 oder 1:100 von bioPEG-Silan zu mPEG-Silan (Laysan Bio) nach der in [174] beschriebenen Prozedur funktionalisiert. Die Probenkammern wurden durch Aufsetzen eines gründlich gereinigten Plexiglasblocks mit Laser-geschnittenen zylinderförmigen Aussparungen auf die bioPEG/mPEG-funktionalisierte Oberfläche hergestellt. Dabei wurden alle Grenzflächen zwischen Deckglas und Plexiglasblock mit Vakuumfett abgedichtet. Die offene Oberseite der Probenkammern wurde mit einem gründlich gereinigten, sonst unbehandelten Deckgläschen abgedeckt, das nur zum pipettieren in die Kammer oder aus der Kammer abgenommen wurde.

Probenpräparation für die TMFM-Versuche

Eine Lösung von 0,05 mg/ml NeutrAvidin (Invitrogen) in T50-Puffer wurde in die Probenkammer pipettiert und dort für mindestens 30 Minuten belassen. Überschüssiges NeutrAvidin wurde mit T50-Puffer abgewaschen. Anschließend wurde die Kammer mit 10xTE-Puffer gespült, der zusätzlich 500 mM NaCl enthielt.

Die Gel-aufgereinigten DNA-Nano-Objekte wurden vor der Einbringung in die Probenkammer je nach Probe fünf- bis zwanzigfach verdünnt. Zur Verdünnung wurde 10xTE-Puffer verwendet, der zusätzlich dieselbe NaCl-Konzentration enthielt wie der FM-AB für die erste Aufnahme der jeweiligen Probe (150 mM oder 500 mM). Die verdünnte Probe von DNA-Nano-Objekten wurde in die Probenkammer pipettiert und dort in der Regel für wenige Minuten belassen, bis eine geeignete Oberflächendichte von DNA-Nano-Objekten (50 bis 200 Objekte pro Mikroskopie-Sichtfeld) erreicht war. Überschüssige DNA-Nano-Objekte wurden mit FM-WB abgewaschen.

Um einen vorzeitigen Start der enzymatischen Reaktionen von Glucose-Oxidase und Catalase zu verhindern, wurde der FM-AB erst kurz vor dem Start jeder Aufnahme angemischt. Der Puffer in der Probenkammer wurde gegen den frisch angemischten FM-AB ausgetauscht. Dann wurde eine geeignete Stelle der Probenoberfläche ausgewählt und die Aufnahme gestartet. Vor jeder weiteren Aufnahme wurde gründlich mit FM-WB gewaschen und die Kammer mit frisch angemischtem FM-AB gefüllt.

Besonderheit bei den Na⁺-Konzentrations-Reihen

Für die [Na⁺]-Reihen wurden jeweils an derselben Probe von DNA-Nano-Objekten mehrere Messungen bei verschiedenen Na⁺-Konzentrationen vorgenommen. Dabei wurde mit der niedrigsten Na⁺-Konzentration der jeweiligen Reihe begonnen (150 mM für Proben von Objekten mit L = 546 bp und L = 1554 bp, 500 mM für Proben von Objekten mit L = 2562 bp). Bei jeder Na⁺-Konzentration wurden TMFM-Aufnahmen von je 75 Minuten Dauer an zwei verschiedenen Orten der Probenoberfläche aufgenommen, bevor die Na⁺-Konzentration auf den nächsthöheren Wert erhöht wurde.

Zwischen zwei Messungen bei derselben Na⁺-Konzentration wurde die Probenkammer mit FM-WB mit derselben Na⁺-Konzentration gespült und anschließend mit frisch angemischtem FM-AB mit derselben Na⁺-Konzentration gefüllt. Bei jeder Erhöhung der Na⁺-Konzentration wurde die Probenkammer gründlich mit FM-WB mit der neuen Na⁺-Konzentration gespült und anschließend mit frisch angemischtem FM-AB mit der neuen Na⁺-Konzentration gefüllt.

A.3 Software

Entwurf der DNA-Nano-Objekte und Interaktionspartnerpaare

Die DNA-Nano-Objekte wurden mithilfe von caDNAno v0.2, caDNAnoSQ [93] und CanDo [91, 94] entworfen. Die Positionen der Erkennungsmotive für die Restriktionsendonuklease Cla I im Scaffold wurden mit dem Programm EnzymeX (Nucleobytes) bestimmt.

Auswertung der Fluoreszenzmikroskopiedaten

Die Auswertung der Fluoreszenzmikroskopiedaten erfolgte mit MATLAB (MathWorks). Der Programmcode ist über das folgende GitHub-Repositorium frei zugänglich:
<https://github.com/mschickinger/TMFM>

B Ergänzende Informationen zu Kapitel 2

B.1 Kompensation des Drifts der TMFM-Probe in der xy-Ebene

Der Tracking-Algorithmus liefert letztendlich für jede Reporter-Einheit Abfolgen von x- und y-Koordinaten zu insgesamt T äquidistanten Zeitpunkten

$$(x(t), y(t)), \quad t = 1, 2, 3, \dots, T$$

Die Koordinatenpaare repräsentieren die Projektion der Position der entsprechenden Reporter-Einheit auf die xy-Ebene. Abbildung B.1 a zeigt die Abfolgen $(x(t), y(t))$ der mobilen Reporter-Einheit eines DNA-Nano-Objekts ohne Interaktionspartner. Da in nahezu jeder Messung Drift der Probe in der xy-Ebene auftrat, erscheint ein Streudiagramm der getrackten xy-Positionen in der Regel asymmetrisch und in die Richtung des Drifts elongiert, wie Abb. B.1 b verdeutlicht. Zur Drift-Korrektur in xy-Richtung wurden jeweils die mittleren xy-Positionen per Median-Bildung über ein Intervall von 101 Einzelbildern abgeschätzt (vgl. Abb. B.1 a und b):

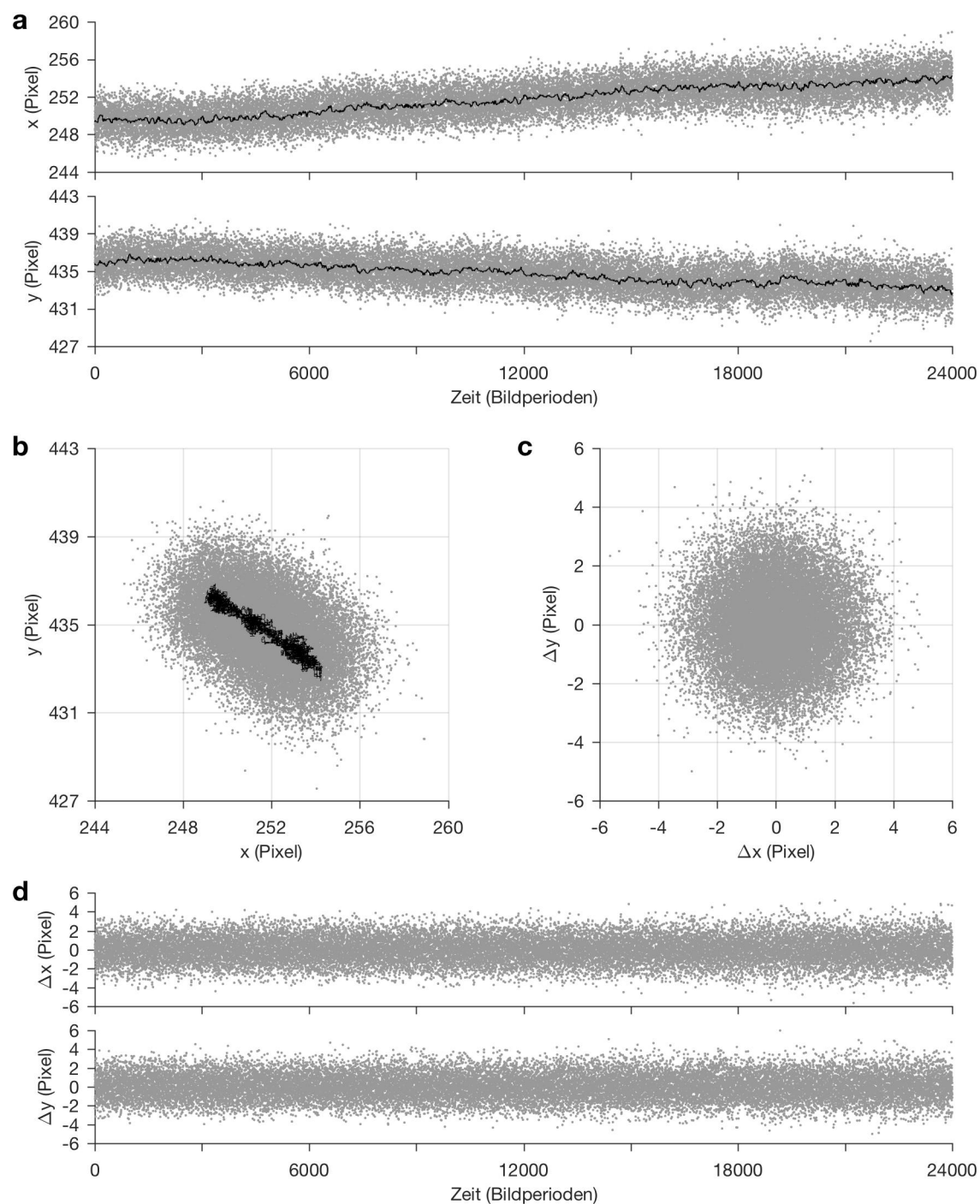
$$\hat{x}(t) = \begin{cases} \text{med}(x(t-50), \dots, x(t+50)) & \text{für } t \in \{51, \dots, T-50\} \\ \hat{x}(51) & \text{für } t \in \{1, \dots, 50\} \\ \hat{x}(T-50) & \text{für } t \in \{T-49, \dots, T\} \end{cases} \quad (\text{B.1})$$

(analog für $\hat{y}(t)$)

Anschließend wurden die xy-Auslenkungen der Reporter-Einheiten aus diesen mittleren Positionen berechnet (vgl. Abb. B.1 c und d):

$$(\Delta x(t), \Delta y(t)) = (x(t), y(t)) - (\hat{x}(t), \hat{y}(t)) \quad (\text{B.2})$$

Bei einer Bildrate der Aufnahme von 10 Bildern pro Sekunde entfernt diese Form der Drift-Korrektur Komponenten der Bewegung der Reporter-Einheiten auf einer Zeitskala von mehr als etwa 10 Sekunden bzw. mit einer Frequenz von weniger als etwa 0,1 Hz.

Abbildung B.1: **Driftkorrektur**

(a) Punkte: Abfolgen der Koordinaten der getrackten Position $x(t)$ und $y(t)$. Schwarze Linien: Abfolgen von $\hat{x}(t)$ und $\hat{y}(t)$ nach Gl. (B.1). **(b)** Punkte: Streudiagramm der getrackten xy-Position. Schwarze Linie: Bahn der geschätzten mittleren Position $(\hat{x}(t), \hat{y}(t))$. **(c)** Streudiagramm der Auslenkung aus der geschätzten mittleren Position $(\Delta x(t), \Delta y(t))$ nach Gl. (B.2). **(d)** Abfolgen von $\Delta x(t)$ und $\Delta y(t)$.

B.2 Erprobung der HMM-Analyse an simulierten Datensätzen

Um die Verlässlichkeit der Ermittlung von Zustandsabfolgen allein basierend auf der unterschiedlichen Breite der Auslenkungsverteilungen zu überprüfen, wurde die implementierte HMM-Prozedur (vgl. Abschnitt 3.2) an simulierten Datensätzen erprobt.

Zur Erstellung einer simulierten verborgenen Zustandsabfolge werden exponentialverteilte Pseudozufallszahlen mit den Erwartungswerten $\tau_{\text{SIM},1}$ bzw. $\tau_{\text{SIM},2}$ generiert. Diese Zufallszahlen stellen die simulierten Verweildauern des Systems in den verborgenen Zuständen 1 und 2 dar. Abwechselnde Aneinanderreihung einzelner Verweildauern in den zwei Zuständen und Auslesen des aktuellen simulierten Zustands zu den äquidistanten Zeitpunkten $1, \dots, T$ liefert eine simulierte verborgene Zustandsabfolge

$$\zeta_{\text{SIM}}(1), \dots, \zeta_{\text{SIM}}(T).$$

Für die dazugehörige simulierte Observablenabfolge bzw. Auslenkungsabfolge werden Pseudozufallszahlen aus Normalverteilungen mit Erwartungswert 0 und Standardabweichung $\sigma_{\text{SIM},1}$ bzw. $\sigma_{\text{SIM},2}$ gezogen, wobei stets $\sigma_{\text{SIM},2} > \sigma_{\text{SIM},1}$ gilt. Jedem Zeitpunkt wird ein Wertepaar aus der dem Zustand $\zeta_{\text{SIM}}(t)$ entsprechenden Verteilung zugeordnet. So entsteht eine simulierte Auslenkungsabfolge

$$(\Delta x_{\text{SIM}}(1), \Delta y_{\text{SIM}}(1)), \dots, (\Delta x_{\text{SIM}}(T), \Delta y_{\text{SIM}}(T))$$

mit abwechselnden Intervallen niedriger bzw. hoher Streuung der Auslenkungen.

Mit der HMM-Analyse (vgl. Abschnitte 3.1 und 3.2) kann aus einer simulierten Auslenkungsabfolge eine Schätzung für die wahrscheinlichste verborgene Zustandsabfolge

$$z(1), \dots, z(T)$$

gewonnen werden¹. Eine Möglichkeit zur Qualitätskontrolle der HMM-Analyse besteht im Vergleich zwischen der simulierten verborgenen Zustandsabfolge (Abb. B.2 a) und der durch die HMM-Prozedur ausgegebenen Schätzung (Abb. B.2 b). In der Regel ist die Zustandszuweisung durch den HMM-Algorithmus nicht zu 100 % korrekt, sprich: Einem gewissen Anteil der Zeitpunkte mit $\zeta_{\text{SIM}}(t) = i$ wird der Zustand $z(t) \neq i$ zugewiesen (Abb. B.2 c). Eine fehlerhafte Zustandszuweisung kann einerseits zur Folge haben, dass einzelne Übergänge zwischen den Zuständen in der Abfolge z im Vergleich zur Abfolge ζ_{SIM} zeitlich verschoben sind, wodurch einzelne Verweildauern in den Zuständen in der Abfolge z im Vergleich zur Abfolge ζ_{SIM} kürzer oder länger ausfallen. Andererseits kann es vorkommen, dass kurze Aufenthalte überhaupt nicht registriert, d. h. sozusagen „verpasst“ werden (vgl. durch Pfeile gekennzeichnete Beispiele in Abb. B.2, nahe den Indices 130 und 3080).

¹In der HMM-Analyse der simulierten Daten wurden die Abfolgen zur Betrachtung der Emissionswahrscheinlichkeiten nicht wie in Abschnitt 3.2.2 beschrieben in mehrere Segmente unterteilt. Die Verteilung der Emissionswahrscheinlichkeiten wurde für jede simulierte Abfolge mit einer einzigen bivariaten Normalverteilung modelliert.

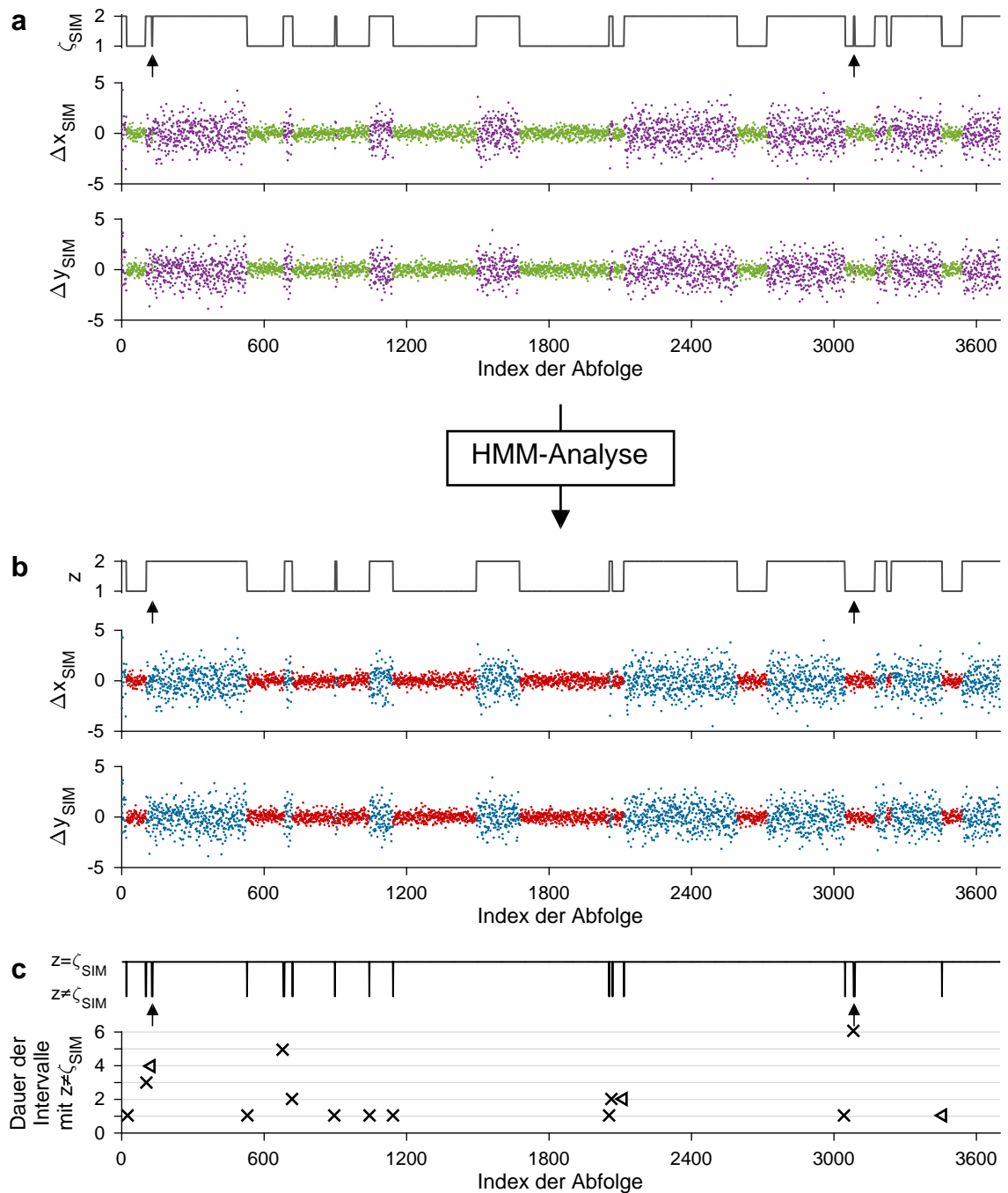


Abbildung B.2: **Vergleich der simulierten Zustandsabfolge mit der Schätzung durch die HMM-Prozedur**
(a) Simulierte Abfolge von Zuständen ζ_{SIM} und dazugehörige simulierte Abfolgen von x - und y -Auslenkungen. Auslenkungen während Aufenthalts mit $\zeta_{SIM} = 1$ in Grün, während Aufenthalts mit $\zeta_{SIM} = 2$ in Violett. Pfeile kennzeichnen jeweils einen durch die HMM-Analyse verpassten kurzen Aufenthalt in Zustand 1 bzw. Zustand 2. **(b)** Durch den HMM-Algorithmus bestimmte Zustandsabfolge z und entsprechend eingefärbte Auslenkungsabfolgen ($z = 1$ in Rot, $z = 2$ in Blau). Pfeile: wie in (a). **(c)** Vergleich der Zustandsabfolgen z und ζ_{SIM} . Das Niveau der schwarzen Linie zeigt eine korrekte Zuweisung (oben) bzw. falsche Zuweisung (unten) des Zustands in z an. Darunter: Jeweilige Dauer der Intervalle mit $\zeta_{SIM} = 1$ und $z = 2$ (\triangleleft) bzw. $\zeta_{SIM} = 2$ und $z = 1$ (\times). Pfeile: wie in (a).

Es wurden simulierte Datensätze mit verschiedenen Kombinationen von Erwartungswerten $\tau_{\text{SIM},1}$ und $\tau_{\text{SIM},2}$ für die Verweildauern sowie verschiedenen Werten der Standardabweichungen $\sigma_{\text{SIM},1}$ und $\sigma_{\text{SIM},2}$ generiert und mit der HMM-Prozedur analysiert. In diesem Abschnitt werden exemplarisch die Ergebnisse der Analyse von simulierten Datensätzen mit allen möglichen Kombinationen von $\tau_{\text{SIM},1}, \tau_{\text{SIM},2} \in \{50, 150, 500\}$ gezeigt. Die Werte der Standardabweichungen der simulierten Auslenkungsverteilungen betragen dabei $\sigma_{\text{SIM},1} = 0,5$ und $\sigma_{\text{SIM},2} = 1,2$. Das Verhältnis dieser beiden Werte entspricht etwa dem Verhältnis der bei Objekten mit $L = 1554 \text{ bp}$ und distaler Verankerungsgeometrie im Experiment beobachteten Standardabweichungen der Auslenkungen im hybridisierten bzw. dissoziierten Zustand. Jeder simulierte Datensatz umfasste 100 Zustandsabfolgen sowie die dazugehörigen Auslenkungsabfolgen, mit einer jeweiligen Abfolgenlänge von $T = 45000$. Dementsprechend kann jeder simulierte Datensatz als Nachahmung einer TMFM-Messung einer Probe von 100 Objekten über einen Zeitraum von 75 Minuten mit einer Bildrate von 10 Bildern pro Sekunde angesehen werden, wobei die wahren charakteristischen Verweildauer in den Zuständen $0,1 \text{ s} \cdot \tau_{\text{SIM},1}$ und $0,1 \text{ s} \cdot \tau_{\text{SIM},2}$ betragen.

Zur Qualitätskontrolle der HMM-Prozedur wurden für jeden Datensatz die folgenden Merkmale betrachtet:

- Die Quote von Zeitpunkten mit korrekter Zustandszuweisung $z(t) = \zeta_{\text{SIM}}(t)$
- Das Ausmaß verpasster Aufenthalte in Abhängigkeit ihrer Verweildauer
- Die Abweichung der Verweildauerverteilungen aus den vom HMM-Algorithmus bestimmten z -Abfolgen von den der Simulation zu Grunde liegenden Exponentialverteilungen
- Die Abweichungen der Maximum-Likelihood-Schätzer für die charakteristischen Verweildauern in den Zuständen ohne Berücksichtigung verpasster Aufenthalte von den Erwartungswerten der ursprünglichen Verteilungen $\tau_{\text{SIM},1}$ und $\tau_{\text{SIM},2}$

Verlässlichkeit der Zustandszuweisung

Die Quote korrekter Zustandszuweisung durch den HMM-Algorithmus ist allgemein gesprochen sehr hoch und korreliert mit den Werten von $\tau_{\text{SIM},1}$ und $\tau_{\text{SIM},2}$ (Abb. B.3). Im Fall $\tau_{\text{SIM},1} = \tau_{\text{SIM},2} = 500$ weist der HMM-Algorithmus in jeder der 100 Abfolgen des simulierten Datensatzes mindestens 99,7% der Zeitpunkte den korrekten Zustand zu. Für $\tau_{\text{SIM},1} = \tau_{\text{SIM},2} = 50$ ist die Zustandszuweisung immer noch in mindestens 97,4% der Zeitpunkte jeder Abfolge korrekt. Die mit rund 95,4% geringste Quote der korrekten Erkennung des Zustands 1 trat in einer Abfolge des Datensatzes mit $\tau_{\text{SIM},1} = 50$ und $\tau_{\text{SIM},2} = 500$ auf. Die mit rund 94,6% geringste Quote der korrekten Erkennung des Zustands 2 trat in einer Abfolge des Datensatzes mit $\tau_{\text{SIM},1} = 500$ und $\tau_{\text{SIM},2} = 50$ auf. Desweiteren fällt auf, dass für alle Datensätze mit $\tau_{\text{SIM},1} = \tau_{\text{SIM},2}$ die Quote korrekt erkannter Zeitpunkte im Zustand 1 im Mittel höher ist als die Quote korrekt erkannter Zeitpunkte im Zustand 2.

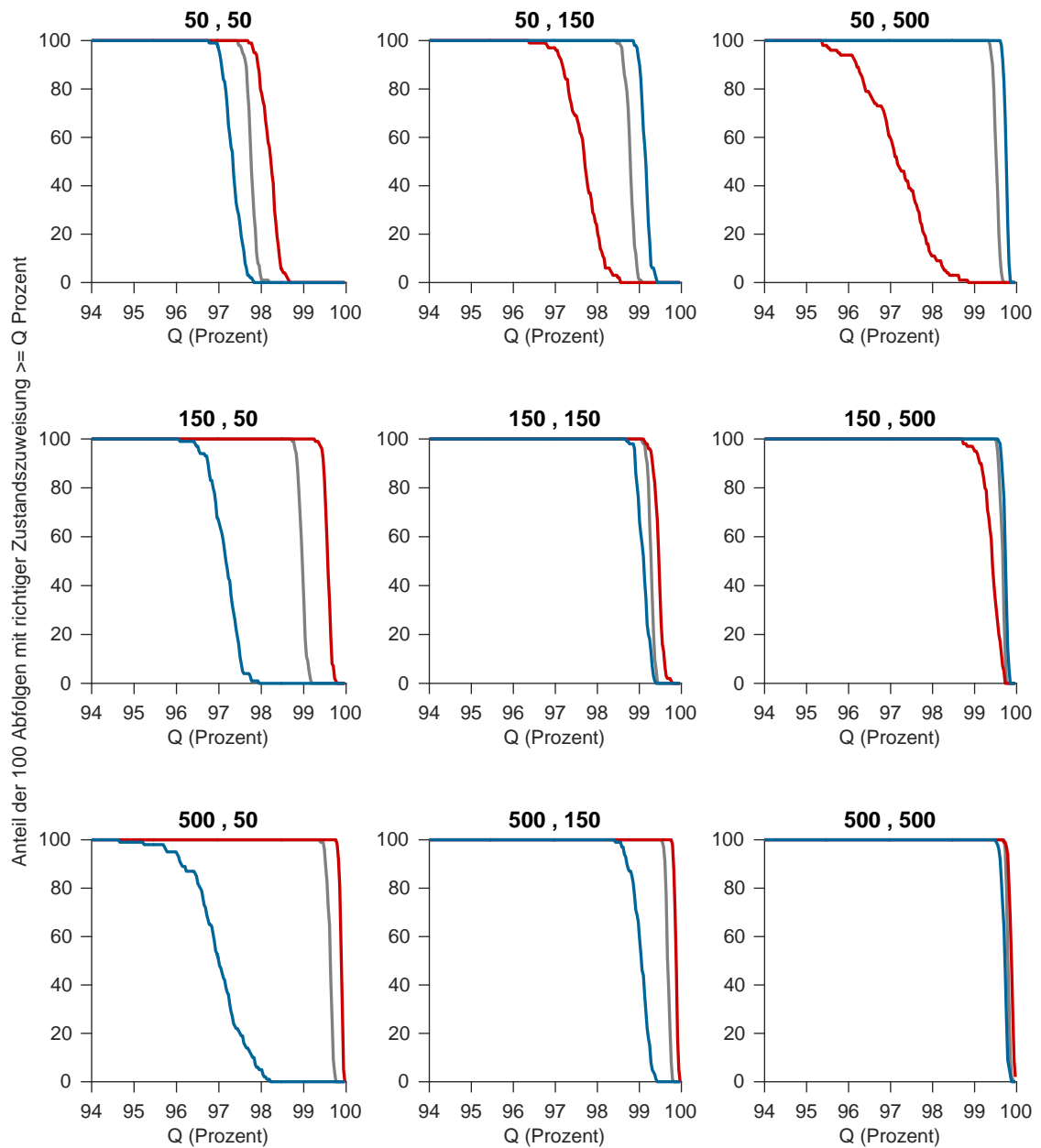


Abbildung B.3: **Quote korrekter Zustandszuweisung durch den HMM-Algorithmus**

Statistik über die korrekte Zustandszuweisung für neun verschiedene simulierte, mit dem HMM-Algorithmus analysierte Datensätze. Anzahlen der Zustandsabfolgen von insgesamt Hundert pro Datensatz mit einer korrekten Zuweisung zu Q oder mehr Prozent der Zeitpunkte mit $\zeta_{SIM}(t) = 1$ (rot), der Zeitpunkte mit $\zeta_{SIM}(t) = 2$ (blau) und insgesamt (grau). Oberhalb der Diagramme sind jeweils die Erwartungswerte $\tau_{SIM,1}$, $\tau_{SIM,2}$ der Exponentialverteilungen für die zufallsgenerierten Verweildauern angegeben.

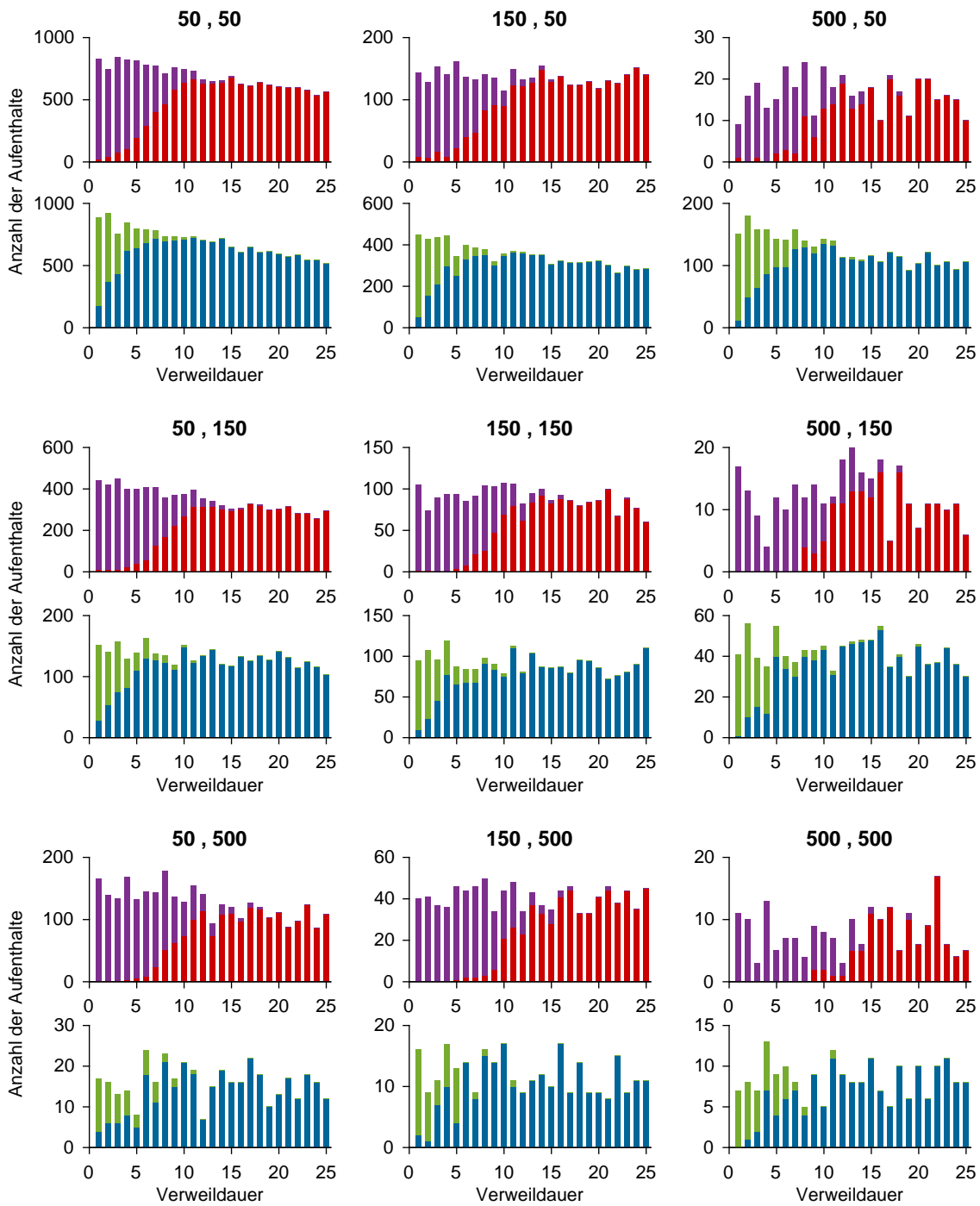


Abbildung B.4: **Verpasste und detektierte Aufenthalte mit Verweildauern bis zu zwei Sekunden**

Gestapelte Balkendiagramme für neun verschiedene simulierte, mit dem HMM-Algorithmus analysierte Datensätze. Anzahlen der verpassten Aufenthalte mit einer bestimmten Verweildauer in Violett (Zustand 1) bzw. Grün (Zustand 2). Anzahlen der detektierten Aufenthalte mit einer bestimmten Verweildauer in Rot (Zustand 1) bzw. Blau (Zustand 2). Oberhalb der Diagramme sind jeweils die Erwartungswerte $\tau_{SIM,1}$, $\tau_{SIM,2}$ der Exponentialverteilungen für die zufallsgenerierten Verweildauern angegeben.

Verpasste Aufenthalte

In der Analyse jedes der simulierten Datensätze wurden kurze Aufenthalte verpasst. Der Anteil verpasster Aufenthalte korreliert dabei invers mit der Höhe der Werte von $\tau_{\text{SIM},1}$ und $\tau_{\text{SIM},2}$: Für den Datensatz mit $\tau_{\text{SIM},1} = \tau_{\text{SIM},2} = 500$ werden rund 2 % resp. 1 % der Aufenthalte in Zustand 1 resp. 2 nicht detektiert. Für alle Datensätze mit $\tau_{\text{SIM},1} = 50$ werden über 10 % der Aufenthalte im Zustand 1 nicht erkannt, mit einem Maximalwert von mehr als 17 % nicht erkannter Aufenthalte im Zustand 1 für den Datensatz mit $\tau_{\text{SIM},1} = 50$ und $\tau_{\text{SIM},2} = 500$.

Die maximale Verweildauer verpasster Aufenthalte liegt bei 23. Alle Aufenthalte mit einer Dauer von mindestens 24 in den simulierten Zustandsabfolgen erscheinen durch vereinzelte fehlerhafte Zustandszuweisung womöglich verkürzt, werden aber dennoch detektiert. Die Wahrscheinlichkeit, dass ein Aufenthalt in der HMM-Analyse verloren geht, steigt mit abnehmender Verweildauer der Aufenthalte (Abb. B.4). Für die Mehrzahl der simulierten Datensätze wird unterhalb einer Verweildauer von etwa 10 resp. 5 ein beträchtlicher Anteil der Aufenthalte im Zustand 1 resp. 2 verpasst. Während in den Datensätzen mit $\tau_{\text{SIM},1} = \tau_{\text{SIM},2}$ der Zustand 1 bei der Quote der korrekten Zustandszuweisung besser abschneidet, verhält es sich beim Verpassen kurzer Aufenthalte umgekehrt: Für jeden Datensatz mit $\tau_{\text{SIM},1} = \tau_{\text{SIM},2}$ ist der Anteil verpasster Aufenthalte im Zustand 1 höher als im Zustand 2.

Verfälschung der Verweildauerverteilungen durch verpasste Aufenthalte

Das Verpassen von Aufenthalten verzerrt die empirischen Verweildauerstatistiken der Zustände auf zweierlei Weise: Wird z. B. ein kurzer Aufenthalt im Zustand 1 verpasst, geht dieser Aufenthalt erstens nicht in die Statistik der Verweildauern des Systems im Zustand 1 ein. Zweitens erscheinen der vorherige und der nachfolgende Aufenthalt im Zustand 2 nicht als separate Aufenthalte, sondern verschmelzen zu einem einzigen Aufenthalt mit entsprechend längerer Verweildauer. Im Vergleich der empirischen, aus den Zustandsabfolgen z gewonnenen Verweildauerverteilungen mit den Exponentialverteilungen zur Erzeugung der simulierten Verweildauern ist die Verzerrung nachvollziehbar (Abb. B.5). Bei kürzeren bis mittleren Verweildauern weichen die empirischen Verteilungen deutlich von den theoretischen, der Simulation zu Grunde liegenden Verteilungen ab.

Schätzung der charakteristischen Verweildauern ohne Berücksichtigung verpasster Aufenthalte

Wird eine Stichprobe von Zufallszahlen aus einer zu Grunde liegenden Exponentialverteilung gezogen, so ist der Maximum-Likelihood-Schätzer für den Erwartungswert dieser Exponentialverteilung das arithmetische Mittel der Stichprobenwerte. Bei Vernachlässigung des Verpassens kurzer Aufenthalte durch die HMM-Analyse ist der naive Schätzer für die charakteristische Verweildauer im Zustand i das arithmetische Mittel $\overline{\theta_{\text{HMM},i}}$ aller aus den Zustandsabfolgen z extrahierten Verweildauern im Zustand i . Die-

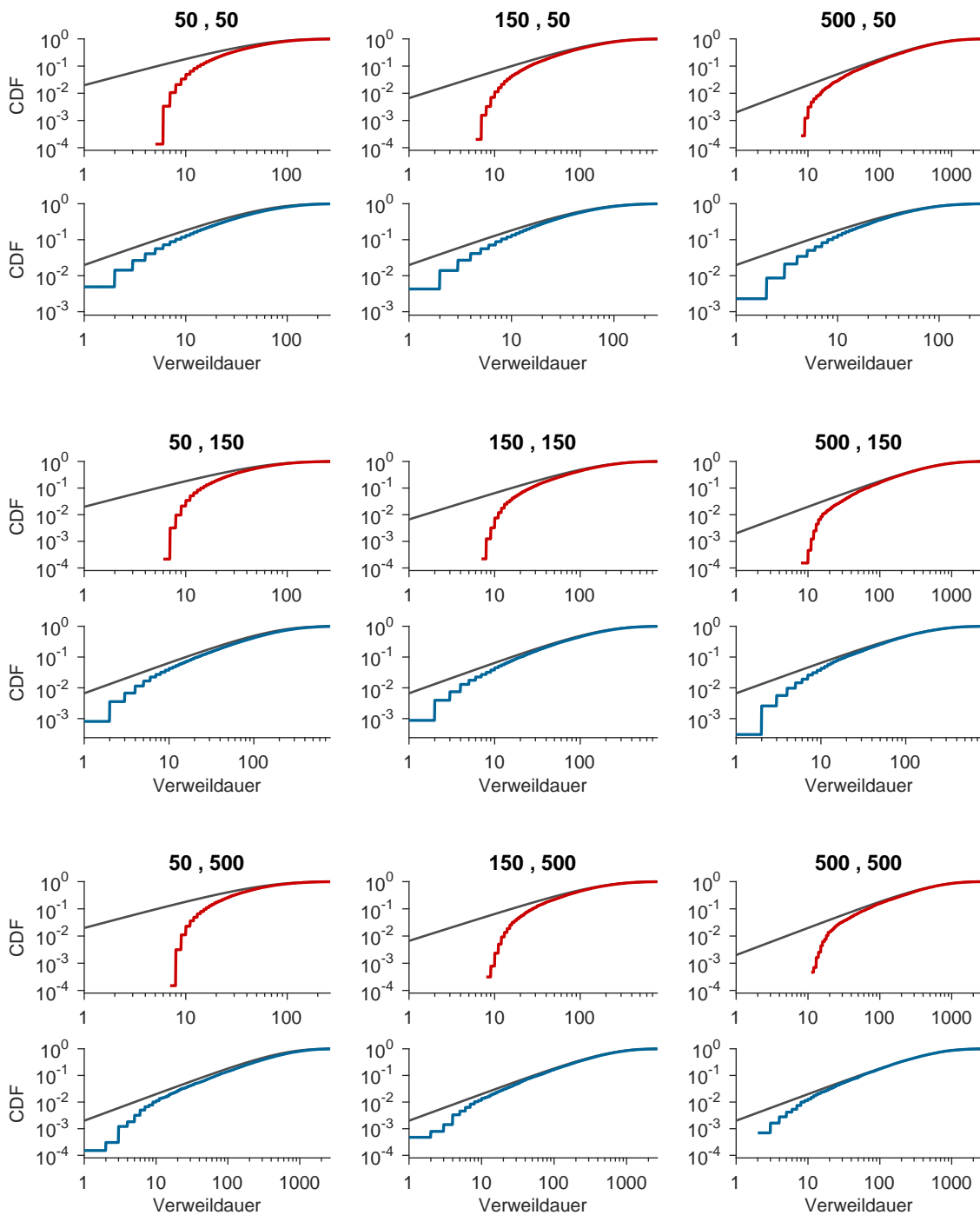


Abbildung B.5: **Vergleich der empirischen und theoretischen Verweildauerverteilungen**

Rote resp. blaue Linien: Empirische kumulierte relative Häufigkeiten der Verweildauern in Zustand 1 (jeweils oberes Diagramm) resp. Zustand 2 (jeweils unteres Diagramm) aus den Zustandsabfolgen z von neun simulierten, mit dem HMM-Algorithmus analysierten Datensätzen. Graue Kurven: Exponentialverteilungen mit Erwartungswert $\tau_{\text{SIM},1}$ resp. $\tau_{\text{SIM},2}$ für die zufallsgenerierten Verweildauern des Datensatzes. Die jeweiligen Werte $\tau_{\text{SIM},1}$, $\tau_{\text{SIM},2}$ sind oberhalb der Diagramme angegeben.

Tabelle B.1: **Geschätzte Verweildauern ohne Berücksichtigung verpasster Aufenthalte**

$\tau_{SIM,1}$ resp. $\tau_{SIM,2}$: Erwartungswerte der Exponentialverteilungen für die simulierten Verweildauern im Zustand 1 resp. 2. $\overline{\theta_{SIM,1}}$ resp. $\overline{\theta_{SIM,2}}$: Arithmetische Mittel der simulierten Verweildauern im Zustand 1 resp. 2. $\overline{\theta_{HMM,1}}$ resp. $\overline{\theta_{HMM,2}}$: Arithmetische Mittel der empirischen, aus den Zustandsabfolgen z extrahierten Verweildauern im Zustand 1 resp. 2. $N_{SIM,1}$, $N_{SIM,2}$, $N_{HMM,1}$, $N_{HMM,2}$: Jeweilige Gesamtzahl der in den Mittelwert eingehenden einzelnen Verweildauern.

$\tau_{SIM,1}$	$\tau_{SIM,2}$	$\overline{\theta_{SIM,1}}$ ($N_{SIM,1}$)	$\overline{\theta_{SIM,2}}$ ($N_{SIM,2}$)	$\overline{\theta_{HMM,1}}$ ($N_{HMM,1}$)	$\overline{\theta_{HMM,2}}$ ($N_{HMM,2}$)
50	50	50,1 (45015)	49,6 (45015)	61,5 (36988)	59,8 (36980)
50	150	49,6 (22454)	150,4 (22403)	60,1 (18566)	181,4 (18509)
50	500	50,1 (8185)	495,9 (8110)	61,1 (6681)	604,1 (6602)
150	50	149,8 (22414)	50,0 (22475)	169,8 (19858)	55,5 (19918)
150	150	150,2 (14944)	149,4 (14942)	164,5 (13687)	162,4 (13684)
150	500	150,6 (6893)	495,8 (6837)	163,1 (6383)	534,6 (6326)
500	50	493,3 (8147)	50,0 (8228)	547,3 (7337)	54,4 (7420)
500	150	496,3 (6851)	149,9 (6902)	521,2 (6530)	156,3 (6581)
500	500	504,7 (4453)	490,0 (4448)	521,2 (4318)	504,6 (4310)

ser Mittelwert überschätzt jedoch für beide Zustände die charakteristische Verweildauer gegenüber dem Erwartungswert der zu Grunde liegenden Exponentialverteilung sowie gegenüber dem Mittelwert $\overline{\theta_{SIM,i}}$ der simulierten Verweildauern im Zustand i (Tab. B.1). Der relative Fehler des Schätzers ist umso größer, je kleiner die tatsächlichen Erwartungswerte $\tau_{SIM,1}$ und $\tau_{SIM,2}$ sind. Bei allen Datensätzen mit $\tau_{SIM,1} = 50$ werden bei naiver Mittelung ohne Berücksichtigung verpasster Aufenthalte die charakteristischen Verweildauern um mehr als 20 % überschätzt.

Um eine derartige Überschätzung der charakteristischen Verweildauern in der Auswertung der experimentellen TMFM-Daten zu umgehen oder zumindest bestmöglich abzuschwächen, wurde eine Prozedur zur Kompensation des Effekts verpasster Aufenthalte basierend auf der Methode von Stigler et al. [108] angewandt. Abschnitt 3.3.2 beschreibt diese Prozedur.

C Weitere ergänzende Abbildungen und Tabellen

Tabelle C.1: **Überblick über die charakteristischen Lebensdauern und Raten sowie die Umfänge der Datensätze der verschiedenen TMFM-Messungen an kurzen Duplexen**

L: Konturlänge des Leinenstücks zwischen den Interaktionspartnern. τ_1 resp. τ_2 : Charakteristische Verweildauern der jeweiligen Hauptpopulationen (vgl. Abschnitt 3.3.3) im hybridisierten resp. dissoziierten Zustand nach Gl. (3.8). k_{off} : Dissoziationsrate ($1/\tau_1$). k_{on} : Assoziationsrate ($1/\tau_2$). Fehler in τ_1 , τ_2 , k_{off} und k_{on} jeweils $\pm 3 \sigma_{\text{bootstrap}}$ (vgl. Abschnitt 3.3.4). N_1 resp. N_2 : Anzahl der einzelnen Verweildauern im hybridisierten resp. dissoziierten Zustand. N_{HP} : Anzahl der DNA-Nano-Objekte in der jeweiligen Hauptpopulation.

Verankerungs- geometrie	Inter- aktions- länge (bp)	L (bp)	Basen- sequenz Strang B	[NaCl] (M)	τ_1 (s)	τ_2 (s)	k_{off} (s ⁻¹)	k_{on} (s ⁻¹)	N_1	N_2	N_{HP}	
Distal	8	546			24,51 ± 0,66	12,85 ± 0,33	0,0408 ± 0,0011	0,0778 ± 0,0020	14296	14160	170	
		1554		0,5	19,42 ± 0,45	29,59 ± 0,70	0,0515 ± 0,0012	0,0338 ± 0,0008	18053	18137	256	
		2562			20,92 ± 1,05	104,17 ± 5,43	0,0478 ± 0,0024	0,0096 ± 0,0005	4022	4111	154	
				0,15	16,10 ± 0,73	36,36 ± 1,72	0,0621 ± 0,0028	0,0275 ± 0,0013	4436	4486	77	
		546		0,5	20,79 ± 0,86	12,66 ± 0,53	0,0481 ± 0,0020	0,0790 ± 0,0033	5159	5104	62	
				CAGGAACG	1,0	17,67 ± 1,00	8,14 ± 0,54	0,0566 ± 0,0032	0,1229 ± 0,0081	3359	3298	47
				0,15	20,08 ± 1,37	98,04 ± 7,69	0,0498 ± 0,0034	0,0102 ± 0,0008	1835	1978	80	
		1554		0,5	20,62 ± 1,23	30,12 ± 1,81	0,0485 ± 0,0029	0,0332 ± 0,0020	2621	2624	56	
				1,0	17,48 ± 0,92	18,55 ± 0,96	0,0572 ± 0,0030	0,0539 ± 0,0028	3528	3454	66	
		2562		0,5	20,24 ± 1,23	96,15 ± 6,47	0,0494 ± 0,0030	0,0104 ± 0,0007	2487	2515	84	
			1,0	14,86 ± 0,79	56,18 ± 3,16	0,0673 ± 0,0036	0,0178 ± 0,0010	3425	3463	71		
			1554	CCTCCTCC		12,38 ± 0,20	2,74 ± 0,04	0,0808 ± 0,0013	0,3649 ± 0,0058	43061	42652	248
			1554	TACGCCAT		16,42 ± 0,32	5,51 ± 0,11	0,0609 ± 0,0012	0,1814 ± 0,0037	23702	23469	167
			1554	ACGCATCA	0,5	27,47 ± 0,45	5,48 ± 0,10	0,0364 ± 0,0006	0,1824 ± 0,0032	33665	33193	366
			1554	CGGATAT		25,06 ± 0,44	6,74 ± 0,11	0,0399 ± 0,0007	0,1484 ± 0,0025	33425	32980	336
			1554	ACTCTCGT		6,41 ± 0,18	42,19 ± 1,25	0,1561 ± 0,0044	0,0237 ± 0,0007	11704	12049	212
			1554	ACGCAGTA		8,70 ± 1,06	454,55 ± 61,98	0,1150 ± 0,0140	0,0022 ± 0,0003	538	808	230
	Proximal	8	546			20,92 ± 0,92	43,29 ± 2,06	0,0478 ± 0,0021	0,0231 ± 0,0011	5078	5128	128
			1050			18,94 ± 1,65	109,89 ± 9,66	0,0528 ± 0,0046	0,0091 ± 0,0008	1678	1758	123
			1554	CAGGAACG	0,5	16,08 ± 0,90	119,05 ± 7,09	0,0622 ± 0,0035	0,0084 ± 0,0005	2771	3158	186
2058					14,93 ± 0,89	196,08 ± 15,38	0,0670 ± 0,0040	0,0051 ± 0,0004	2360	2560	202	
2562					16,29 ± 1,33	256,41 ± 19,72	0,0614 ± 0,0050	0,0039 ± 0,0003	1356	1473	159	
9		1050	CCAGAGACG	0,5	303,03 ± 27,55	35,46 ± 3,14	0,0033 ± 0,0003	0,0282 ± 0,0025	2165	1881	223	

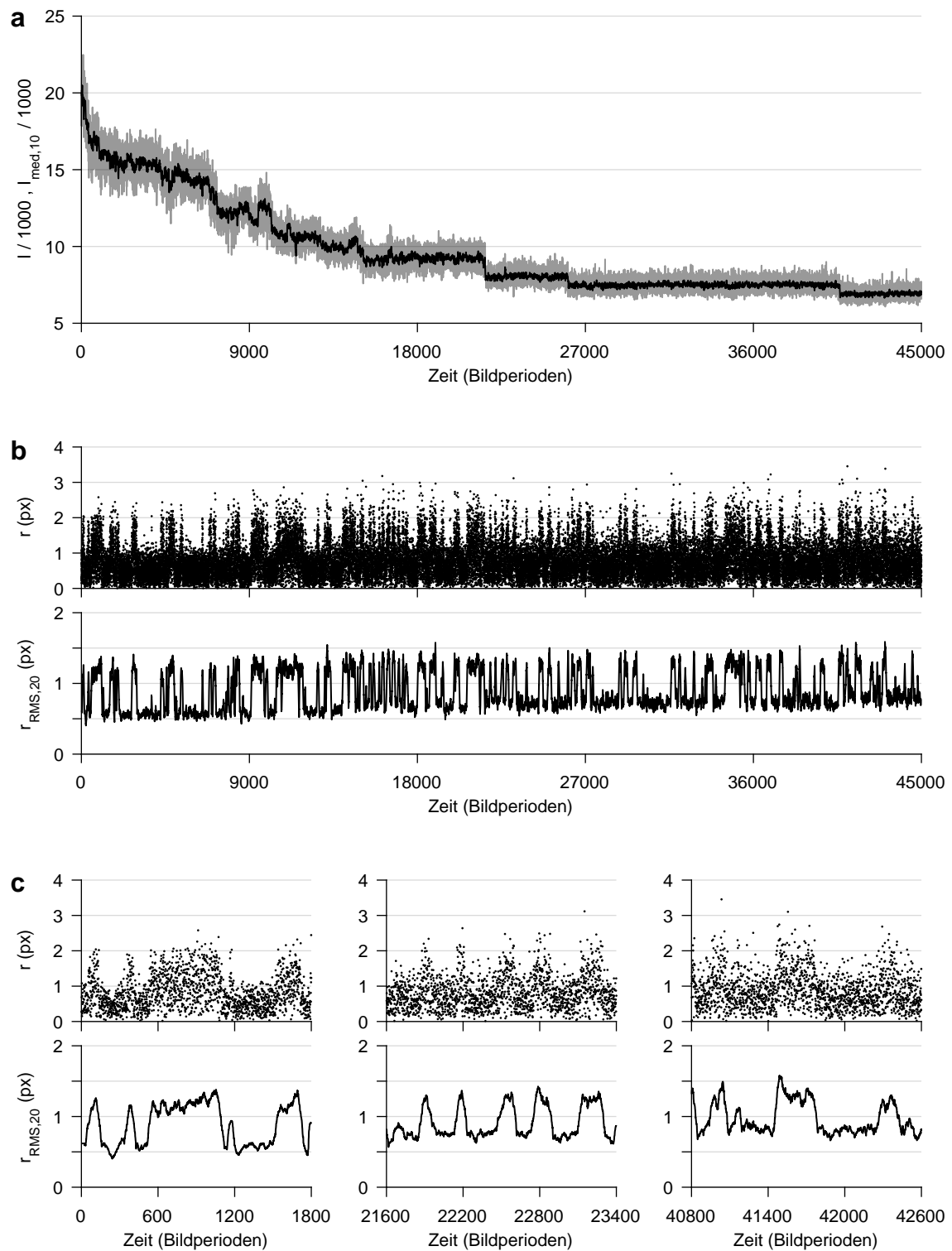


Abbildung C.1: **Helligkeit und gemessene Beweglichkeit einer mobilen Reporter-Einheit, Bsp. 3**

Objekt mit Interaktionspartnern in distaler Verankerungsgeometrie, Konturlänge $L = 546$ bp. **(a)** Grau: Helligkeitsabfolge $I(t)$. Schwarz: Median-gefilterte Helligkeitsabfolge $I_{\text{med},10}$ nach Gl. (2.3). **(b)** Abfolgen der radialen Auslenkung r nach Gleichung (2.1) und von $r_{\text{RMS},20}$ nach Gleichung (2.2) über die gesamte Dauer der Messung. **(c)** Drei verschiedene jeweils 3 Minuten lange Ausschnitte der Abfolgen von r und $r_{\text{RMS},20}$.

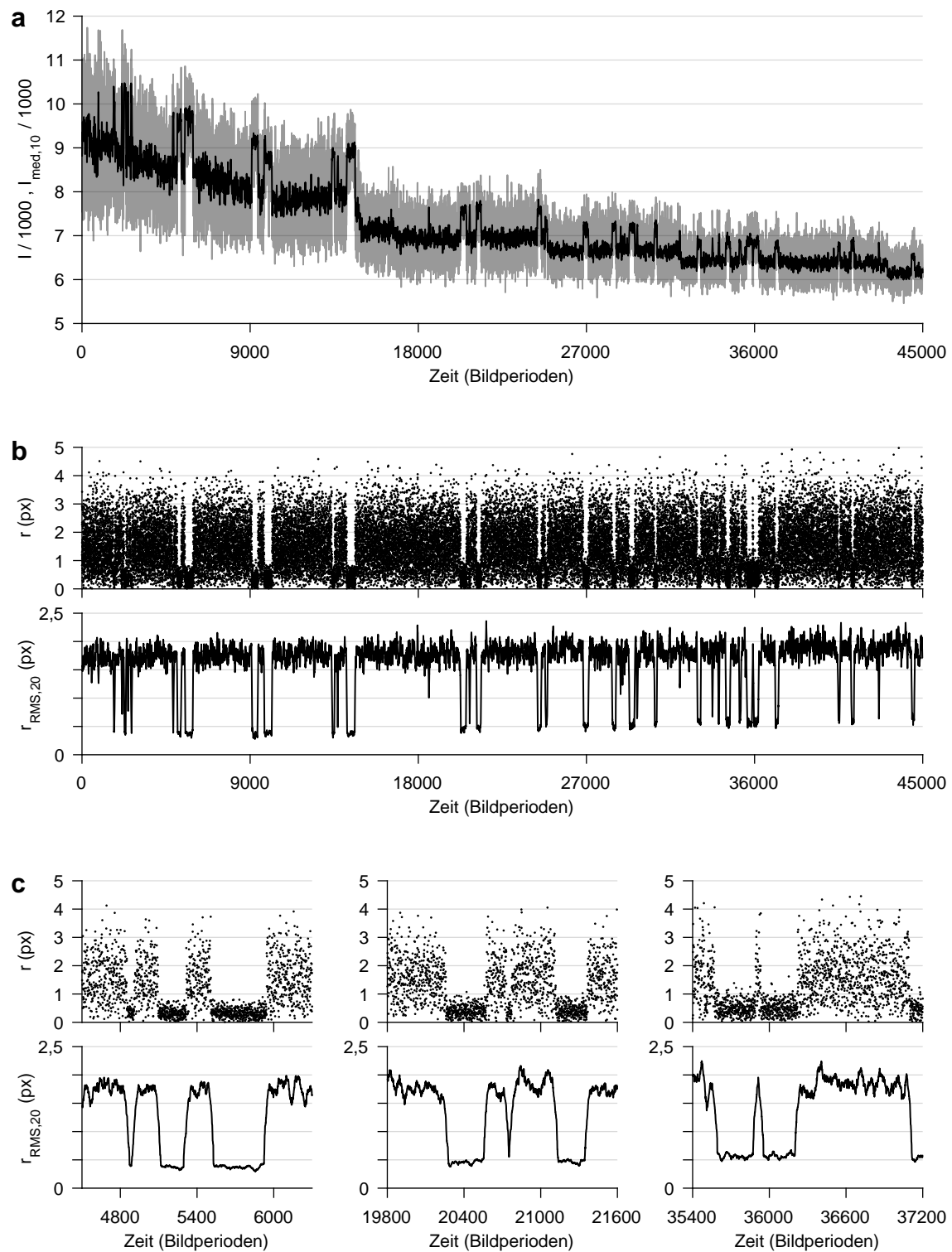


Abbildung C.2: Helligkeit und gemessene Beweglichkeit einer mobilen Reporter-Einheit, Bsp. 4

Objekt mit Interaktionspartnern in proximaler Verankerungsgeometrie, Konturlänge $L = 2058$ bp. (a) Grau: Helligkeitsabfolge $I(t)$. Schwarz: Median-gefilterte Helligkeitsabfolge $I_{med,10}$ nach Gl. (2.3). (b) Abfolgen der radialen Auslenkung r nach Gleichung (2.1) und von $r_{RMS,20}$ nach Gleichung (2.2) über die gesamte Dauer der Messung. (c) Drei verschiedene jeweils 3 Minuten lange Ausschnitte der Abfolgen von r und $r_{RMS,20}$.

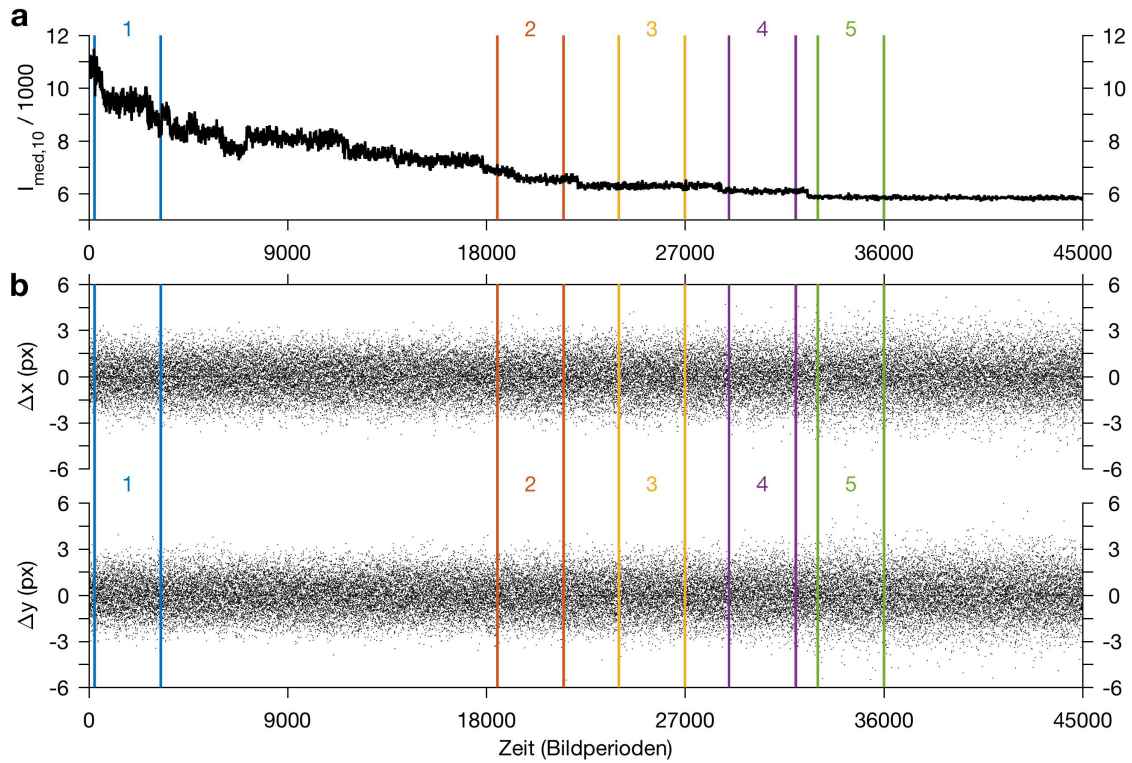


Abbildung C.3: **Helligkeits- und der Auslenkungsabfolgen einer mobilen Reporter-Einheit, Ref. B0-L6**
(a) Schwarze Linie: $I_{med,10}$, Median-gefilterte Abfolge der summierten Pixelhelligkeit um die Position der Reporter-Einheit nach Gl. (2.3). **(b)** Punkte: Δx und Δy , Abfolgen der Auslenkungen in x- und y-Richtung nach Gl. (2.2). In Farbe: Grenzen und Nummerierung der fünf Zeitintervalle, denen die Kennzahlen und Verteilungen in Tab. C.2 und Abb. C.4 entstammen.

Tabelle C.2: **Kennzahlen der Helligkeit und der Auslenkungen für die in Abb. C.3 markierten Intervalle**
k: Nummer des Intervalls. $I_{med,10}$: Median-gefilterte summierte Pixelhelligkeit nach Gl. (2.3). Jeweils gerundete Werte des Maximums, Medians und Minimums im k-ten Intervall. $m_{x,k} / m_{y,k}$: Arithmetisches Mittel der Stichprobe von $\Delta x / \Delta y$ im k-ten Intervall. $s_{x,k} / s_{y,k}$: Standardabweichung der Stichprobe von $\Delta x / \Delta y$ im k-ten Intervall.

k	$I_{med,10}$			$m_{x,k}$	$m_{y,k}$	$s_{x,k}$	$s_{y,k}$
	Maximum	Median	Minimum				
1	11209	9464	8117	0,003	0,010	1,127	1,126
2	7121	6575	6285	-0,012	0,020	1,197	1,167
3	6519	6288	6086	-0,008	-0,009	1,250	1,223
4	6278	6091	5920	-0,006	0,002	1,259	1,254
5	6042	5856	5701	-0,024	0,041	1,366	1,364

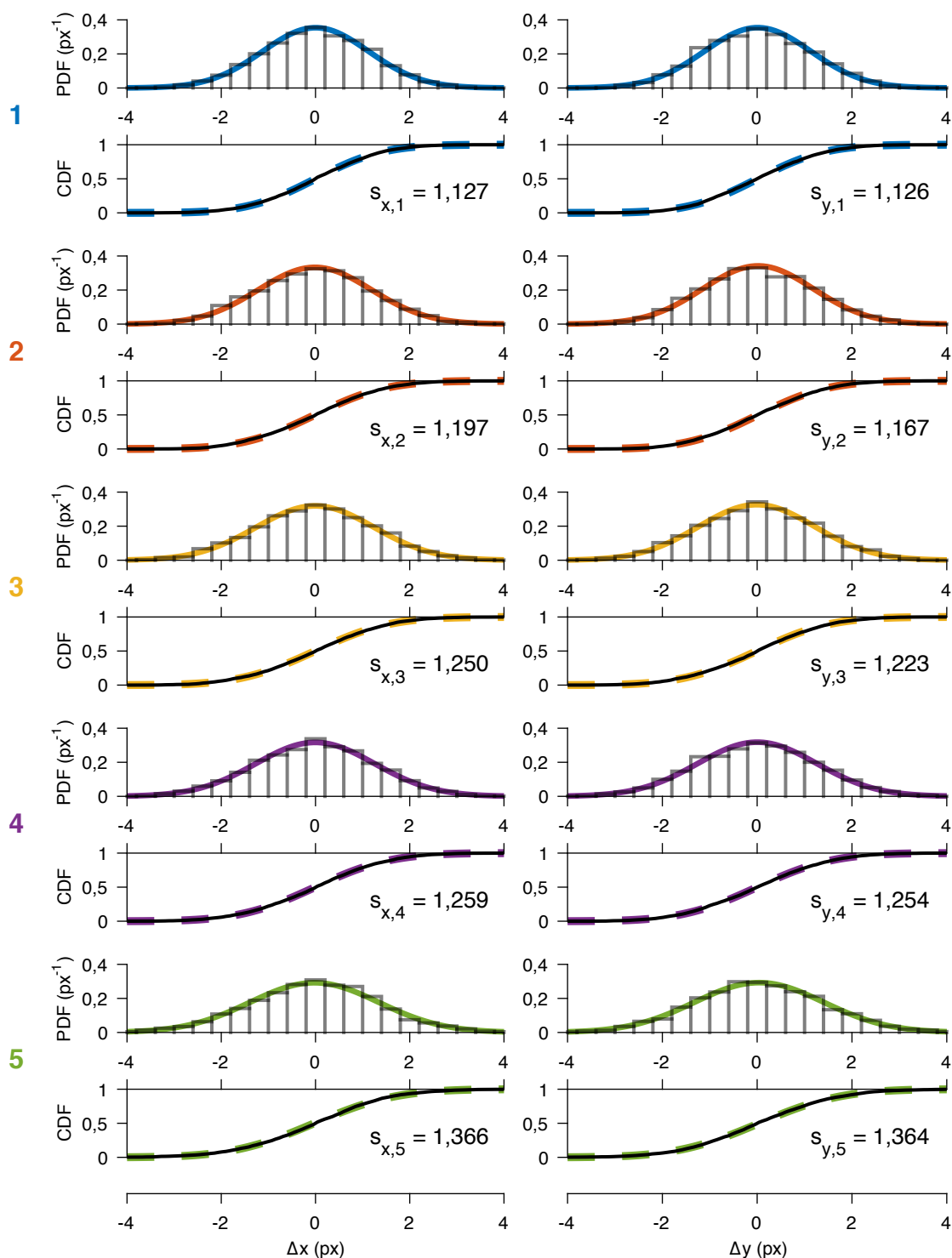


Abbildung C.4: **Verteilungen der Auslenkungsstichproben aus den fünf Zeitintervallen**

Verteilungen von Δx (linke Spalte) und Δy (rechte Spalte). Histogramme: Aus der relativen Häufigkeit ermittelte Wahrscheinlichkeitsdichte. Schwarze Linien: Kumulierte relative Häufigkeit. Farbige durchgezogene resp. gestrichelte Linien: Normalverteilte Wahrscheinlichkeitsdichte resp. Verteilungsfunktion mit Erwartungswert $m_{x,k} / m_{y,k}$ (arithmetisches Mittel der Stichprobe) und Standardabweichung $s_{x,k} / s_{y,k}$ (Standardabweichung der Stichprobe), $k \in \{1, \dots, 5\}$.

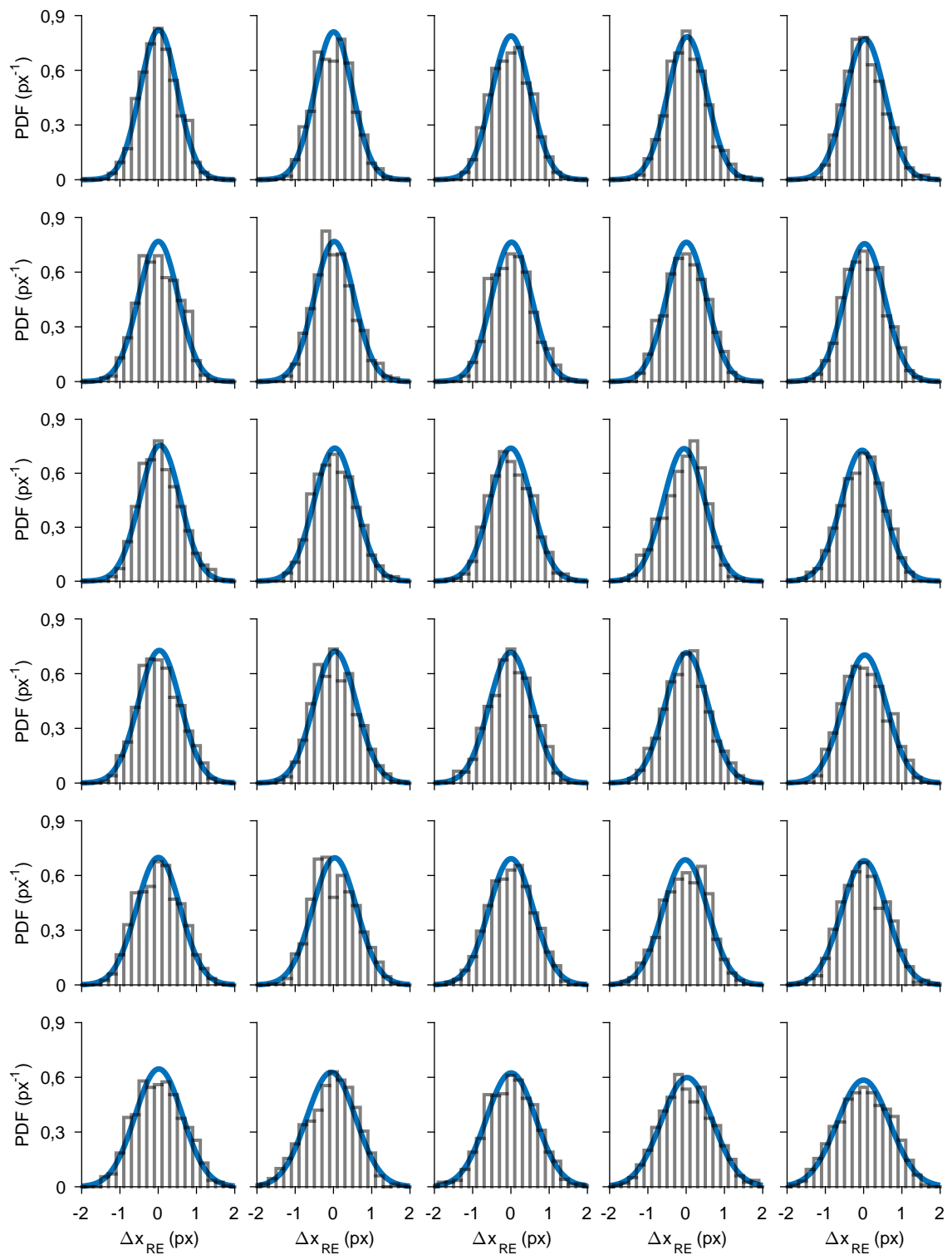


Abbildung C.5: **Wahrscheinlichkeitsdichteverteilungen von Stichproben aus Δx_{RE} von 30 mobilen Reporter-Einheiten der Referenz B0-L2** Histogramme: Aus der relativen Häufigkeit geschätzte Wahrscheinlichkeitsdichte. Blaue Linien: Normalverteilte Wahrscheinlichkeitsdichte nach Gl. (2.5). Die Stichproben sind nach dem Wert der Standardabweichung sortiert. Stichprobenumfang: je 1000 Bildperioden.

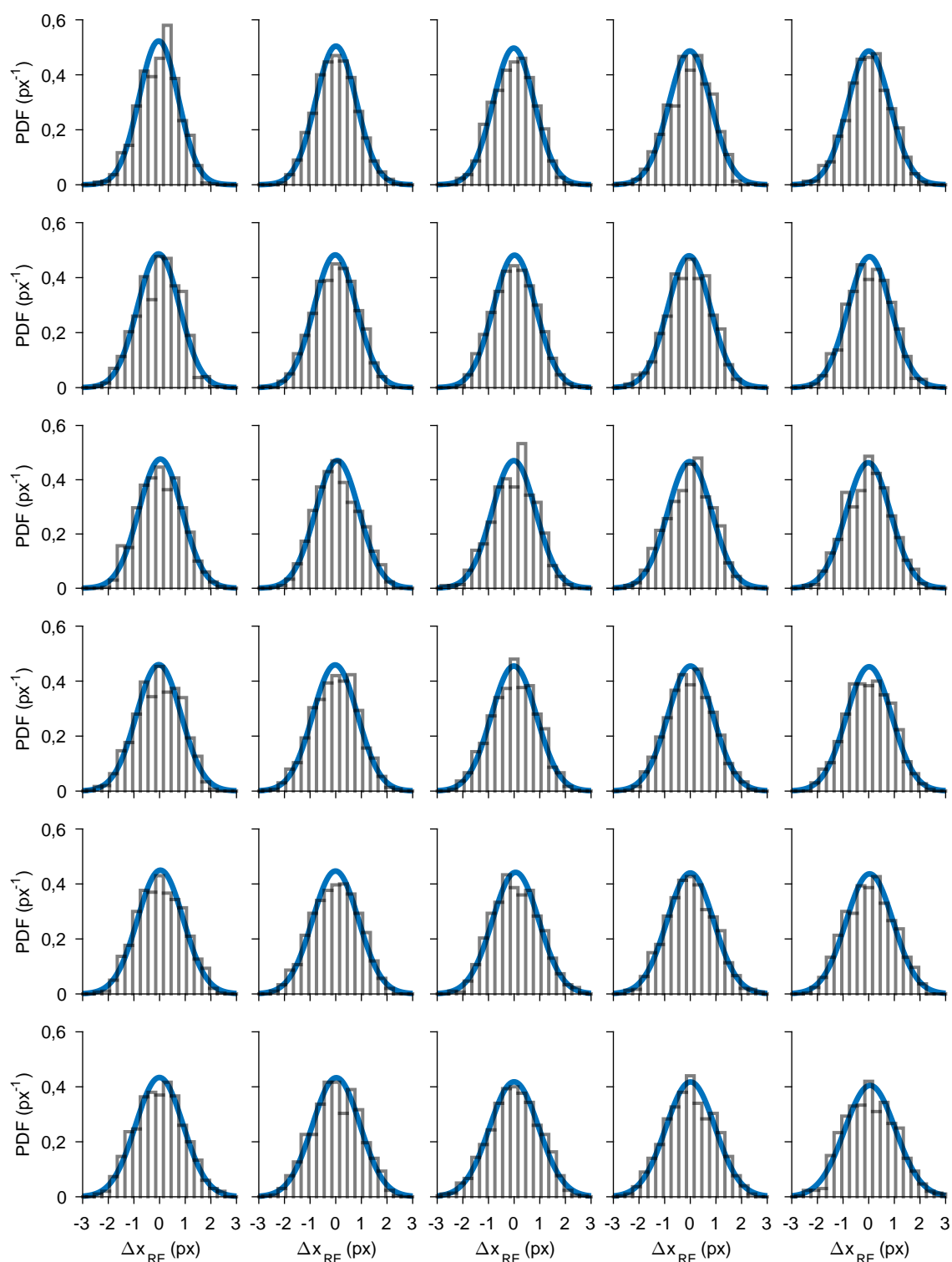


Abbildung C.6: **Wahrscheinlichkeitsdichteverteilungen von Stichproben aus Δx_{RE} von 30 mobilen Reporter-Einheiten der Referenz B0-L4** Histogramme: Aus der relativen Häufigkeit geschätzte Wahrscheinlichkeitsdichte. Blaue Linien: Normalverteilte Wahrscheinlichkeitsdichte nach Gl. (2.5). Die Stichproben sind nach dem Wert der Standardabweichung sortiert. Stichprobenumfang: je 1000 Bildperioden.

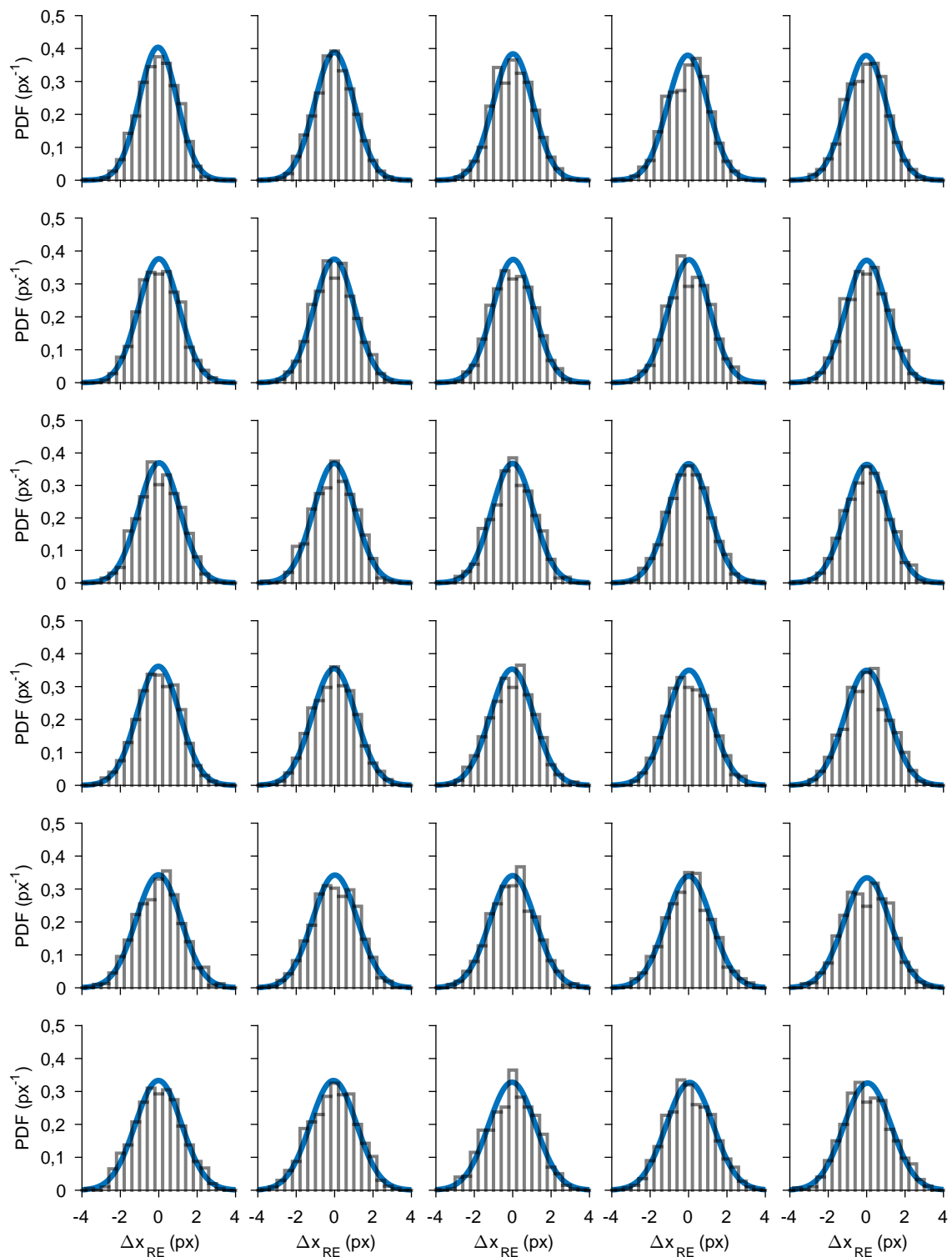


Abbildung C.7: **Wahrscheinlichkeitsdichteverteilungen von Stichproben aus Δx_{RE} von 30 mobilen Reporter-Einheiten der Referenz B0-L6** Histogramme: Aus der relativen Häufigkeit geschätzte Wahrscheinlichkeitsdichte. Blaue Linien: Normalverteilte Wahrscheinlichkeitsdichte nach Gl. (2.5). Die Stichproben sind nach dem Wert der Standardabweichung sortiert. Stichprobenumfang: je 1000 Bildperioden.

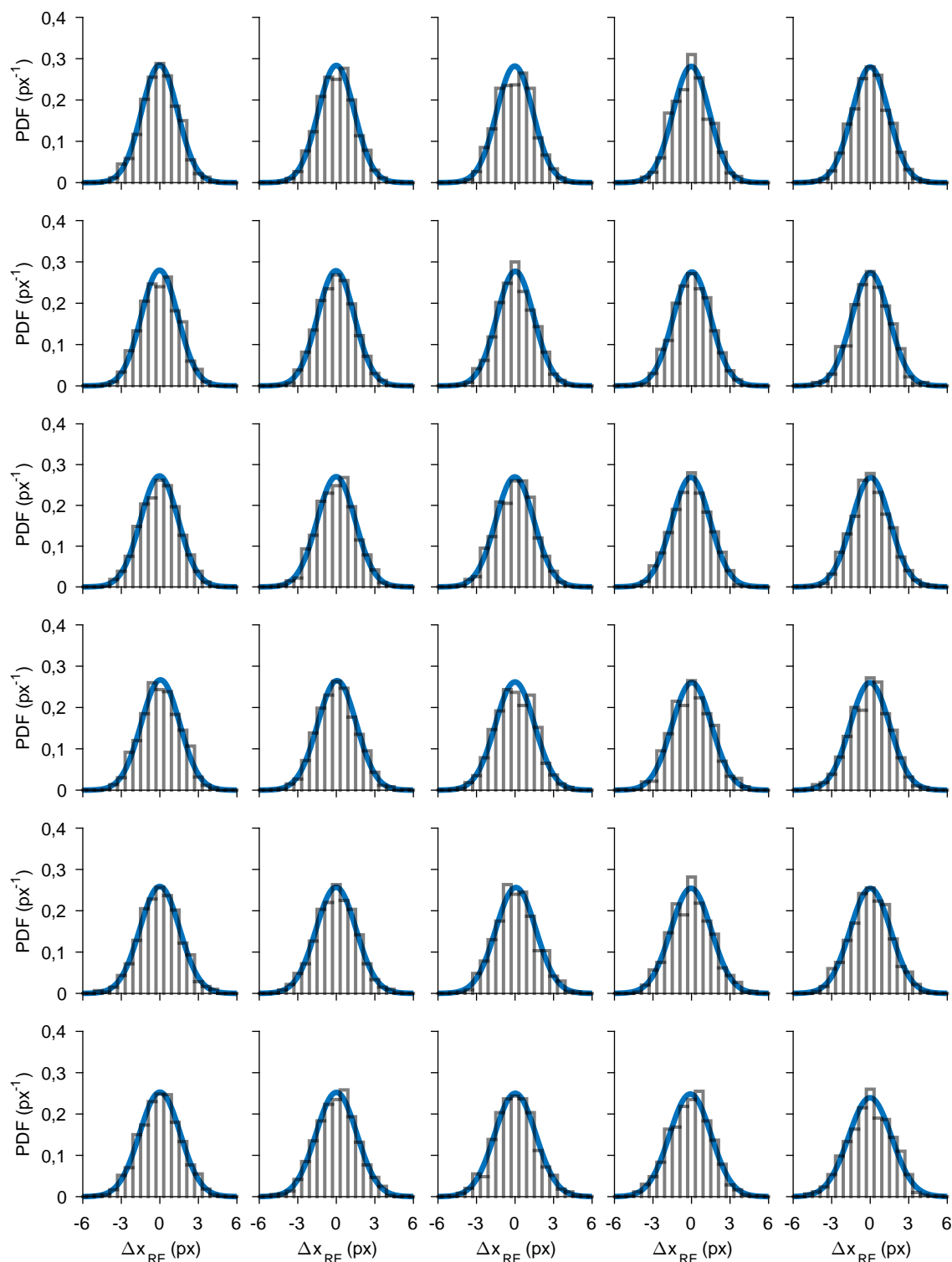


Abbildung C.8: **Wahrscheinlichkeitsdichteverteilungen von Stichproben aus Δx_{RE} von 30 mobilen Reporter-Einheiten der Referenz B0-L10** Histogramme: Aus der relativen Häufigkeit geschätzte Wahrscheinlichkeitsdichte. Blaue Linien: Normalverteilte Wahrscheinlichkeitsdichte nach Gl. (2.5). Die Stichproben sind nach dem Wert der Standardabweichung sortiert. Stichprobenumfang: je 1000 Bildperioden.

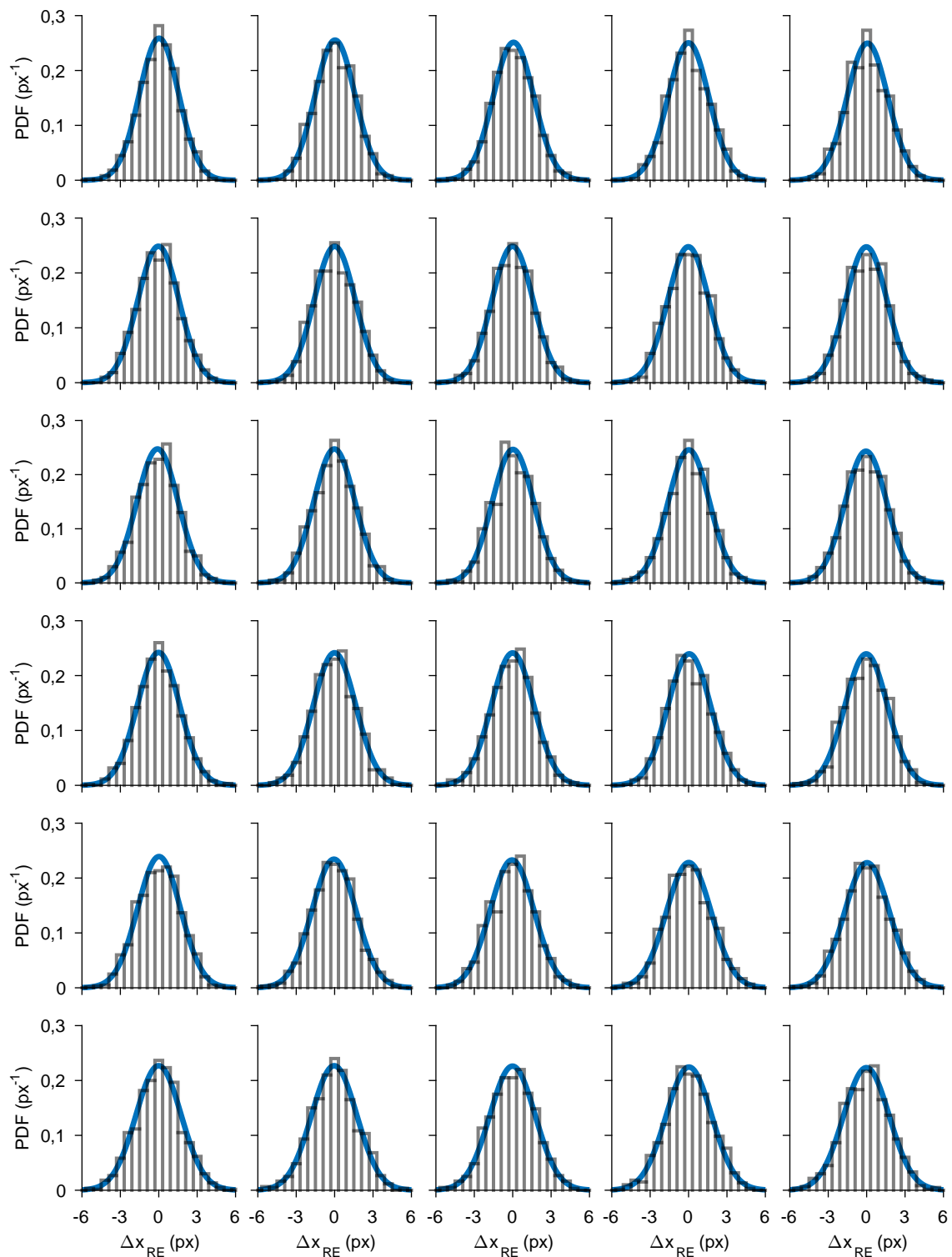


Abbildung C.9: **Wahrscheinlichkeitsdichteverteilungen von Stichproben aus Δx_{RE} von 30 mobilen Reporter-Einheiten der Referenz 0-L12** Histogramme: Aus der relativen Häufigkeit geschätzte Wahrscheinlichkeitsdichte. Blaue Linien: Normalverteilte Wahrscheinlichkeitsdichte nach Gl. (2.5). Die Stichproben sind nach dem Wert der Standardabweichung sortiert. Stichprobenumfang: je 1000 Bildperioden.

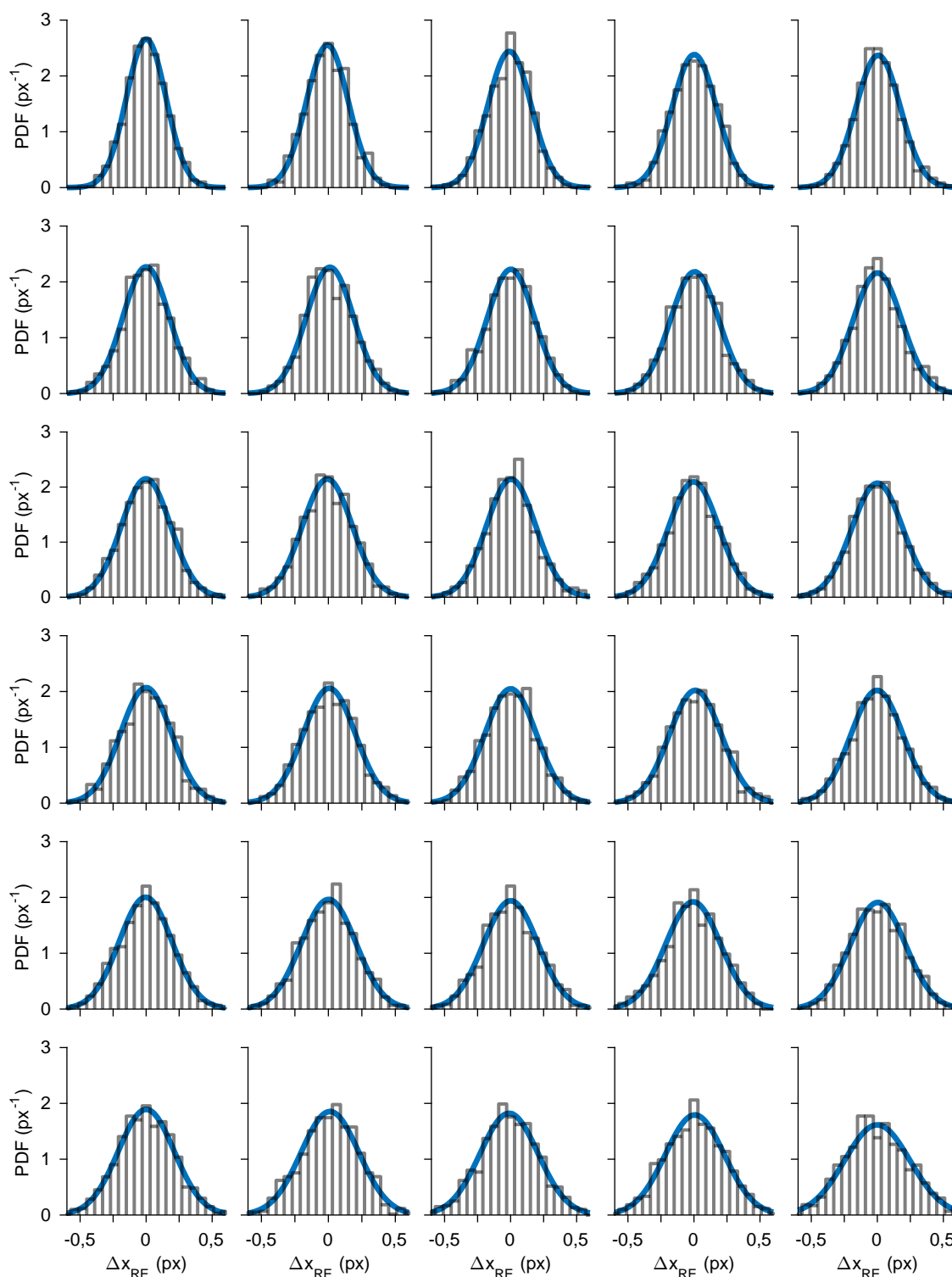


Abbildung C.10: **Wahrscheinlichkeitsdichteverteilungen von Stichproben aus Δx_{RE} von 30 mobilen Reporter-Einheiten der Referenz i9p-L4** Histogramme: Aus der relativen Häufigkeit geschätzte Wahrscheinlichkeitsdichte. Blaue Linien: Normalverteilte Wahrscheinlichkeitsdichte nach Gl. (2.5). Die Stichproben sind nach dem Wert der Standardabweichung sortiert. Stichprobenumfang: je 1000 Bildperioden.

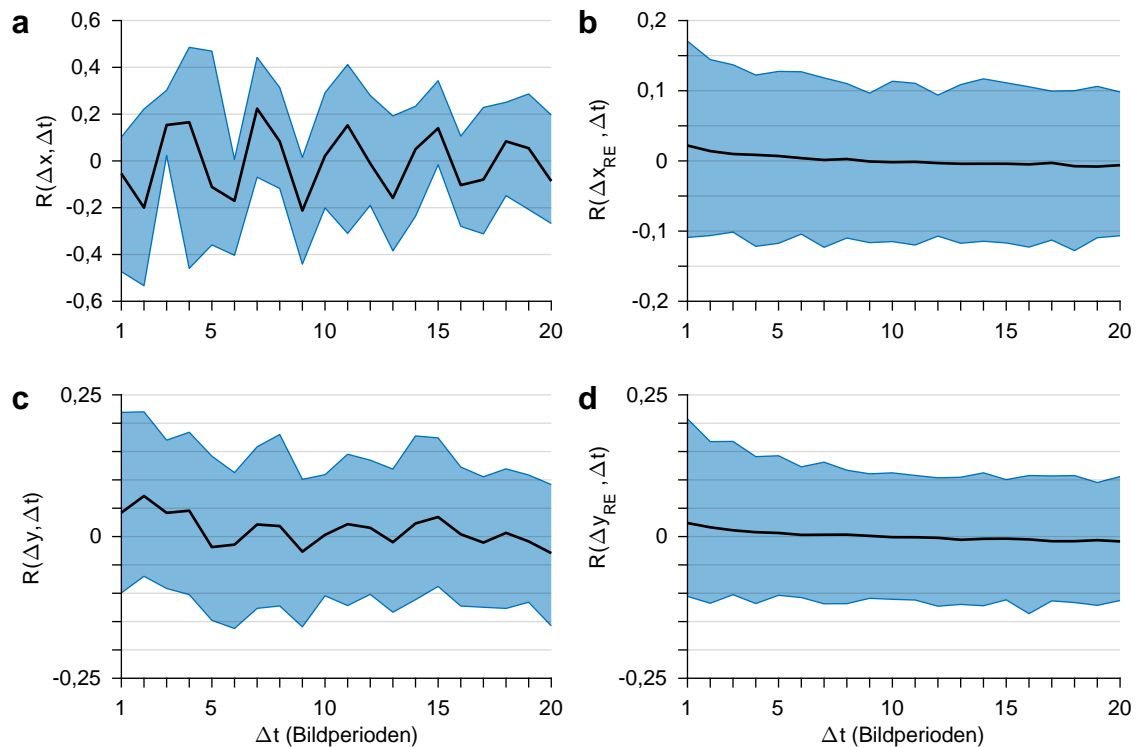


Abbildung C.11: **Normierte Autokorrelationen der Abfolgen Δx , Δx_{RE} , Δy und Δy_{RE} mobiler Reporter-Einheiten der Referenz i9p-L4 in vielen ($N = 1154$) jeweils 500 Bildperioden langen Ausschnitten** (a) $R(\Delta x, \Delta t)$, (b) $R(\Delta x_{RE}, \Delta t)$, (c) $R(\Delta y, \Delta t)$ und (d) $R(\Delta y_{RE}, \Delta t)$ nach Gl. (2.6) für zeitliche Verschiebungen Δt um 1 bis 20 Bildperioden (0,1 bis 2 s bei einer Bildrate von 10 Bildern pro Sekunde). Die blauen Flächen werden begrenzt durch die empirischen 0,005- und 0,995-Quantile der Menge der Werte aus allen 1154 Intervallen. Schwarze Kurven: jeweils arithmetisches Mittel der Werte aus allen 1154 Intervallen.

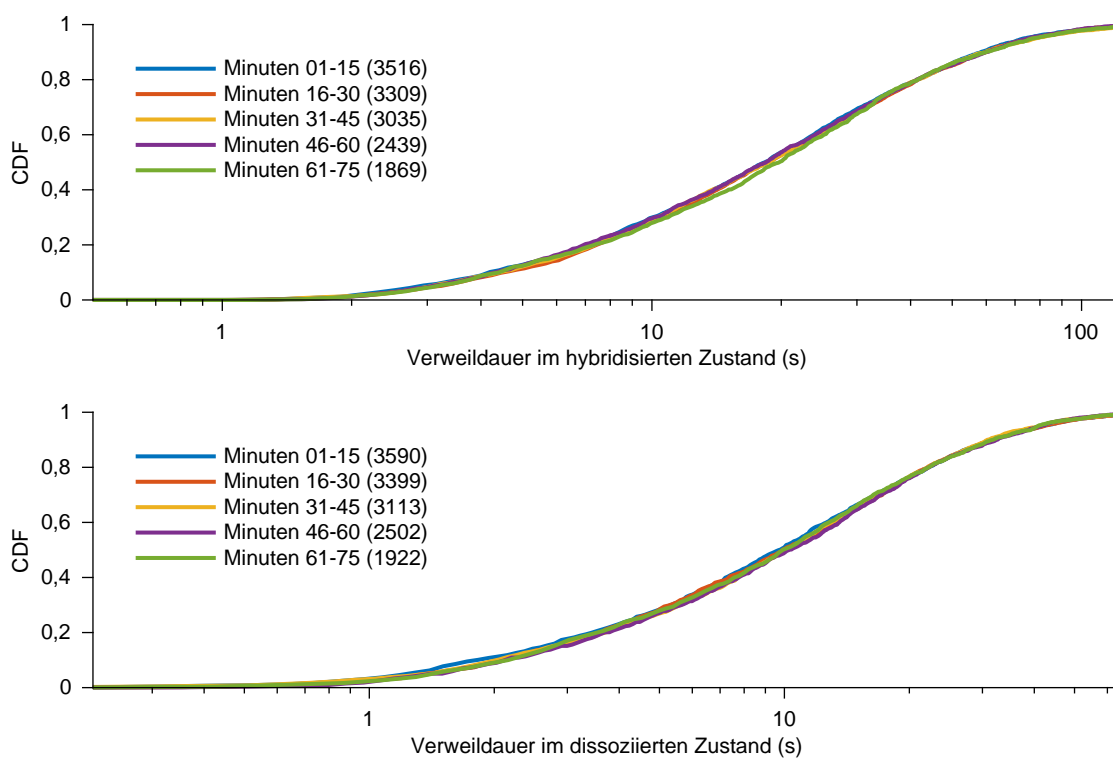


Abb. C.12.1: **CAGGAACG:CGTTCCTG, L = 546 bp, distale Verankerungsgeometrie**

Abbildung C.12: Verweildauerverteilungen im zeitlichen Verlauf der TMFM-Messungen

Empirische Verteilungen der kumulierten Wahrscheinlichkeit (CDF), ermittelt aus der relativen Häufigkeit der einzelnen Verweildauern. Die Verweildauerachsen sind logarithmisch skaliert. Jedes Diagramm zeigt fünf Verteilungen. Jede dieser Verteilungen wurde nur aus den Verweildauern von Aufenthalten im jeweiligen Zustand generiert, die innerhalb eines gewissen Zeitintervalls beginnen und enden. Legende: Zuordnung der Kurven zu den Zeitintervallen. Zahlen in Klammern: jeweilige Anzahl der einzelnen Verweildauern, die in die Verteilung dieses Zeitintervalls eingehen. Die Unterabbildungen C.12.1-C.12.13 zeigen eine derartige Betrachtung für Proben von DNA-Nano-Objekten mit verschiedenen Interaktionssequenzen, Konturlängen L des Leinenstücks zwischen den Interaktionspartnern und Verankerungsgeometrien (für die Betrachtung der Probe mit Interaktionssequenz CCTCCTCC:GGAGGAGG, $L = 1554$ bp und distaler Verankerungsgeometrie siehe Abb. 3.8). Alle TMFM-Messungen dieser Proben fanden bei $[\text{Na}^+] = 0,5$ M im Messpuffer statt.

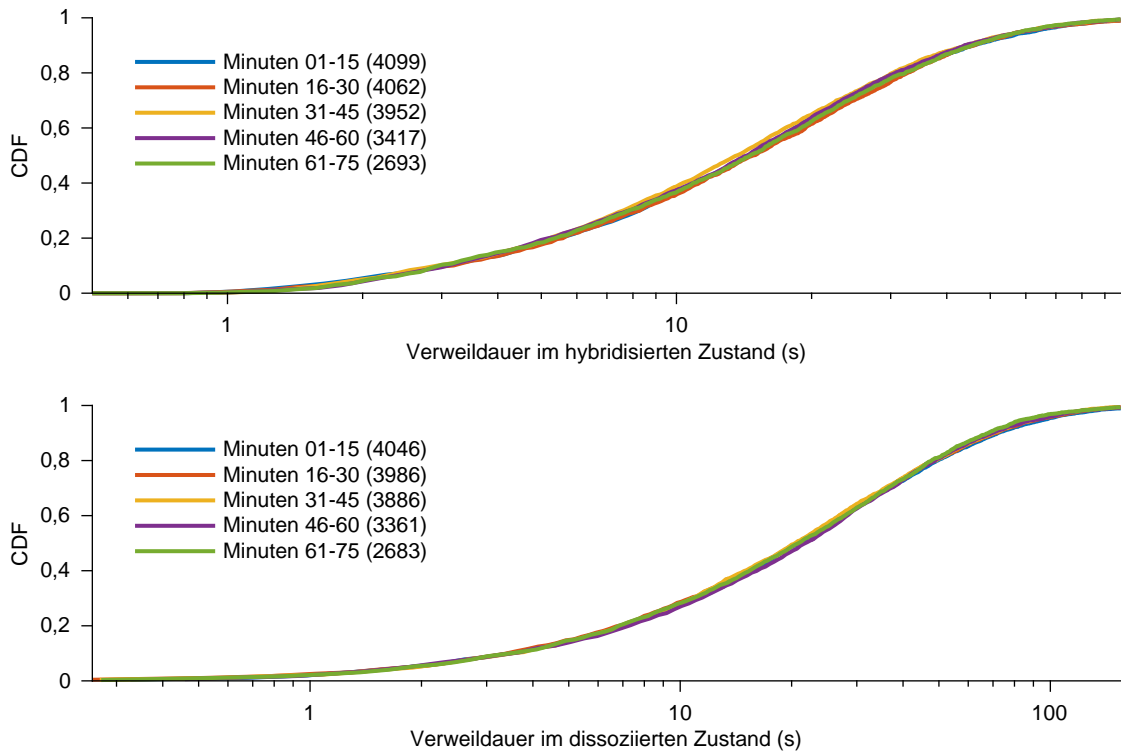


Abb. C.12.2: **CAGGAACG:CGTTCCTG, L = 1554 bp, distale Verankerungsgeometrie**

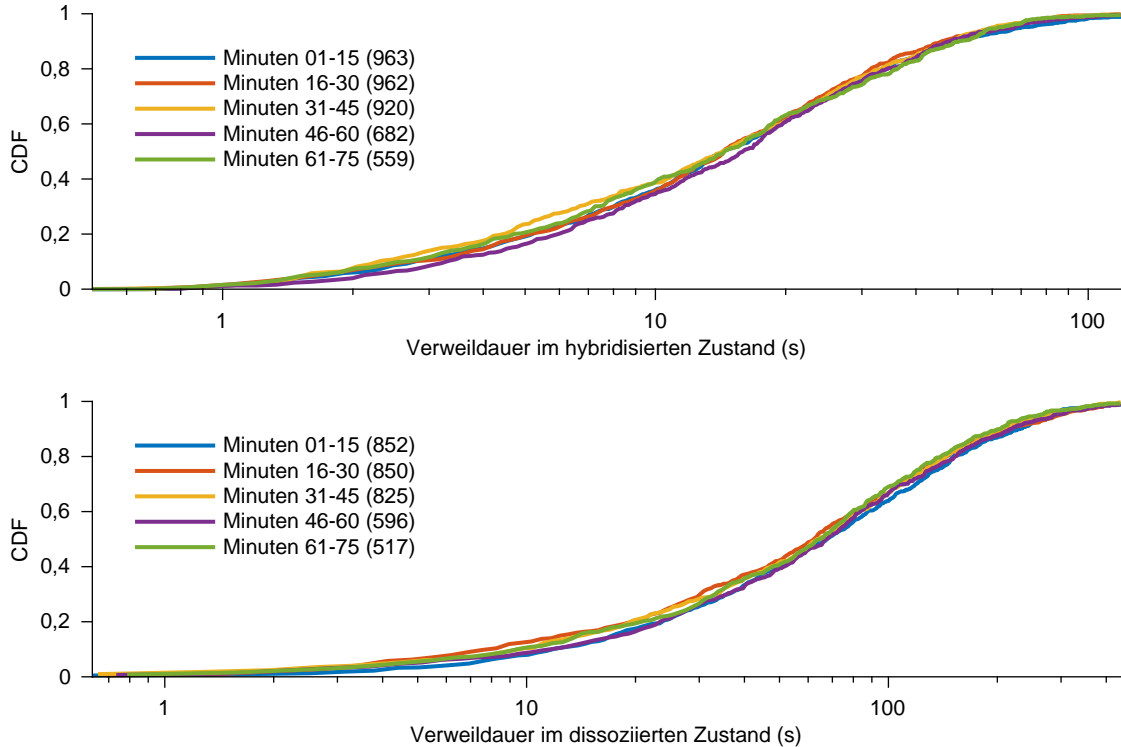
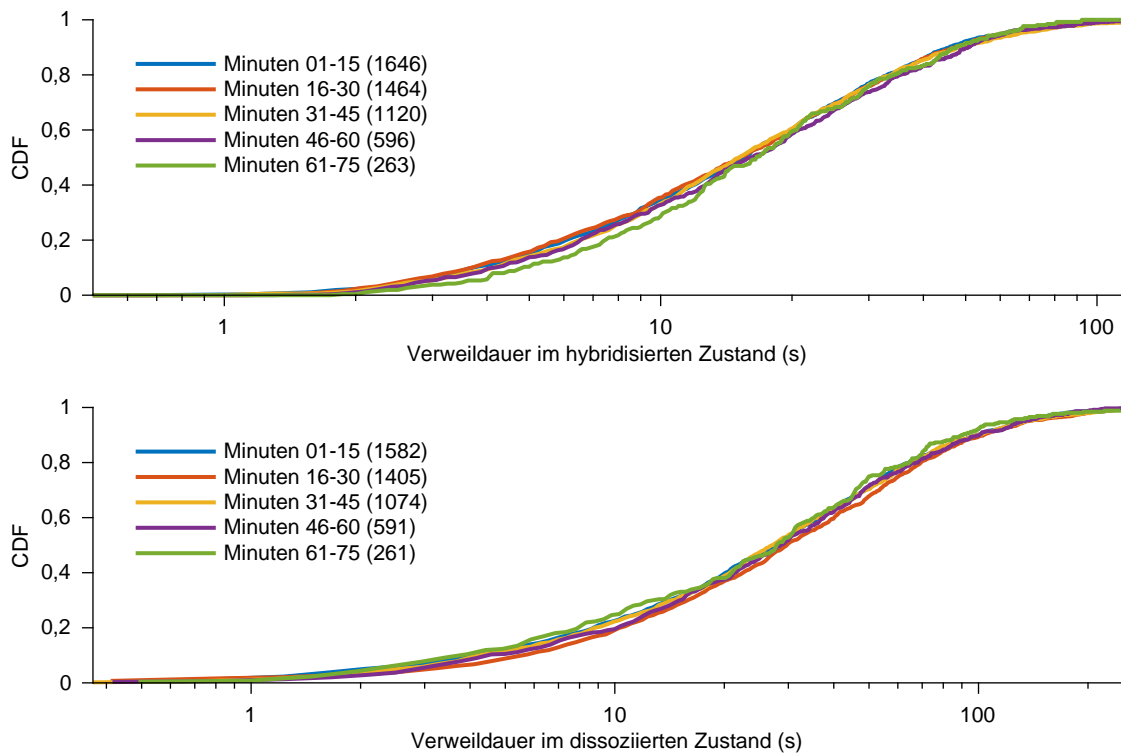
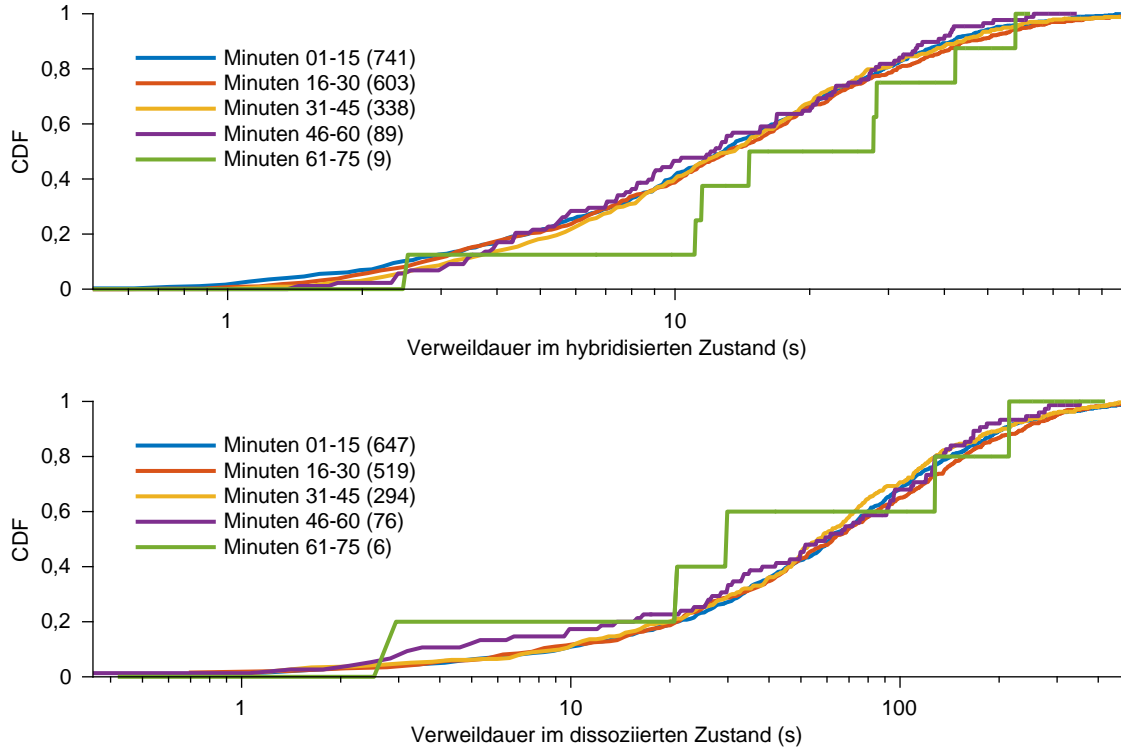


Abb. C.12.3: **CAGGAACG:CGTTCCTG, L = 2562 bp, distale Verankerungsgeometrie**

Abb. C.12.4: **CAGGAACG:CGTTCCTG, L = 546 bp, proximale Verankerungsgeometrie**Abb. C.12.5: **CAGGAACG:CGTTCCTG, L = 1050 bp, proximale Verankerungsgeometrie**

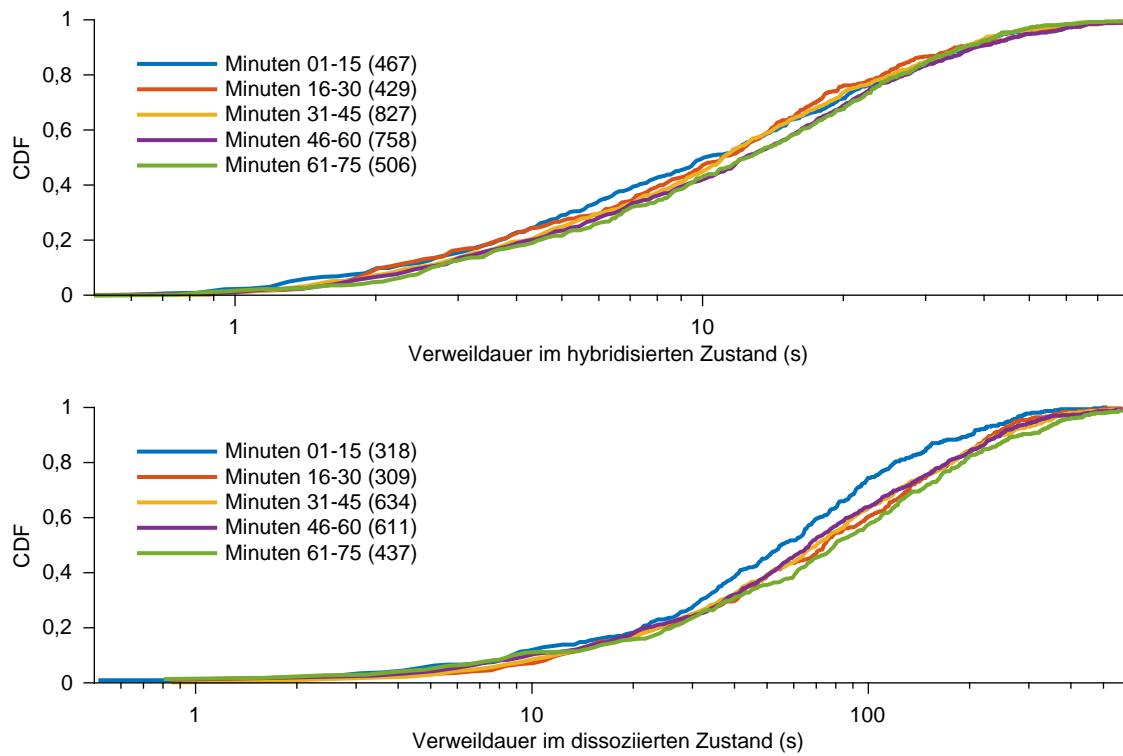


Abb. C.12.6: CAGGAACG:CGTTCCTG, L = 1554 bp, proximale Verankerungsgeometrie

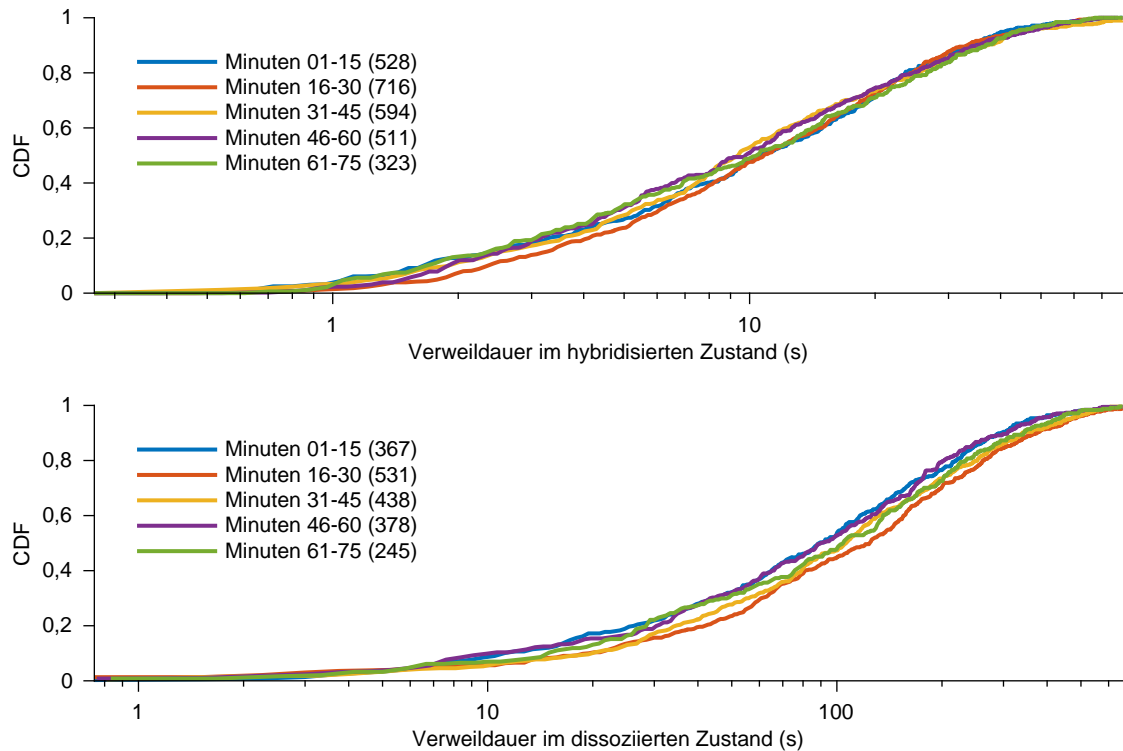


Abb. C.12.7: CAGGAACG:CGTTCCTG, L = 2058 bp, proximale Verankerungsgeometrie

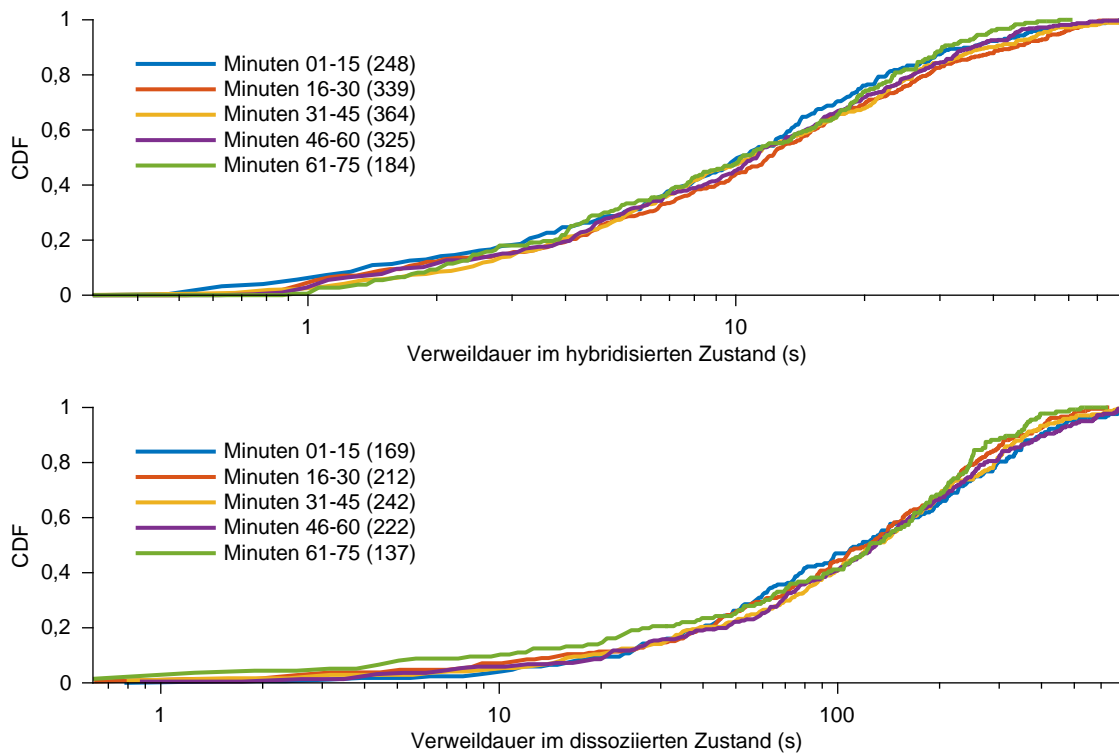


Abb. C.12.8: CAGGAACG:CGTTCCTG, L = 2562 bp, proximale Verankerungsgeometrie

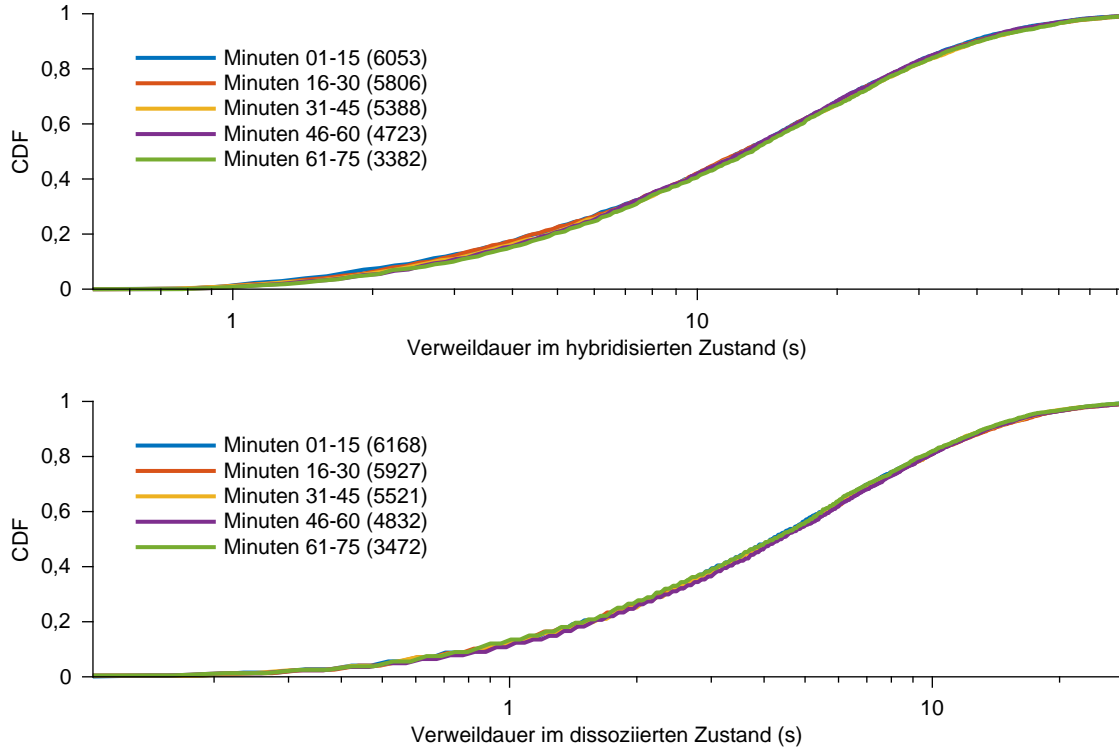
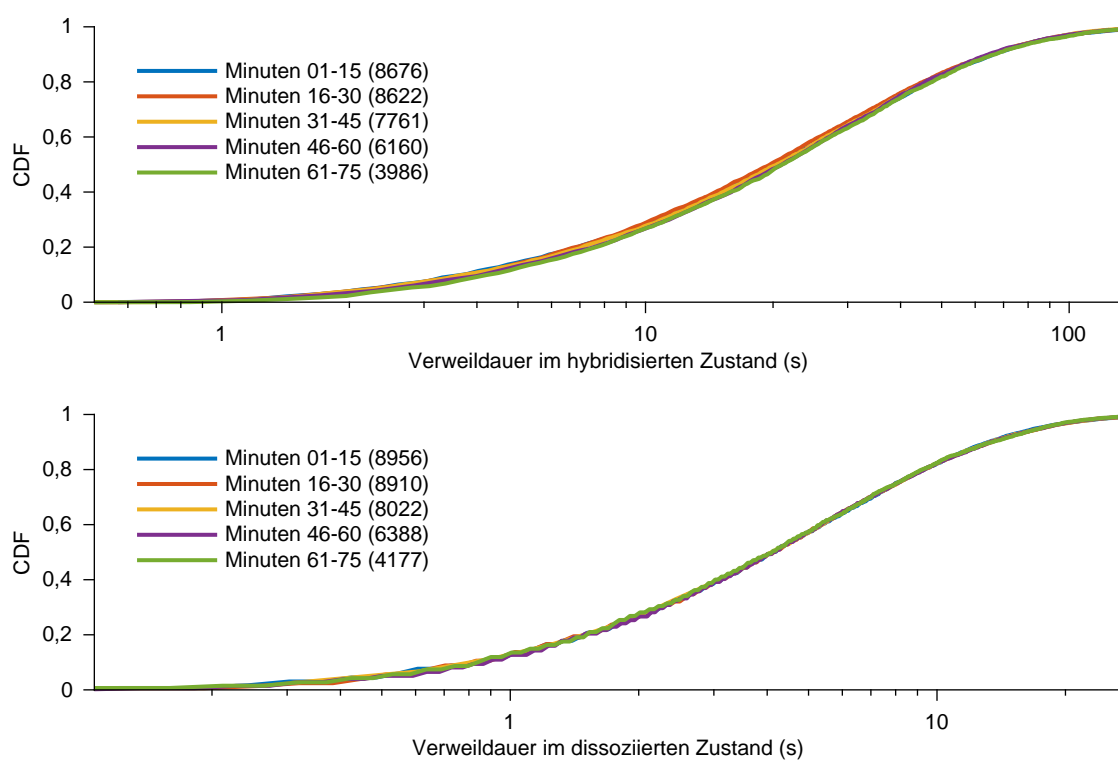
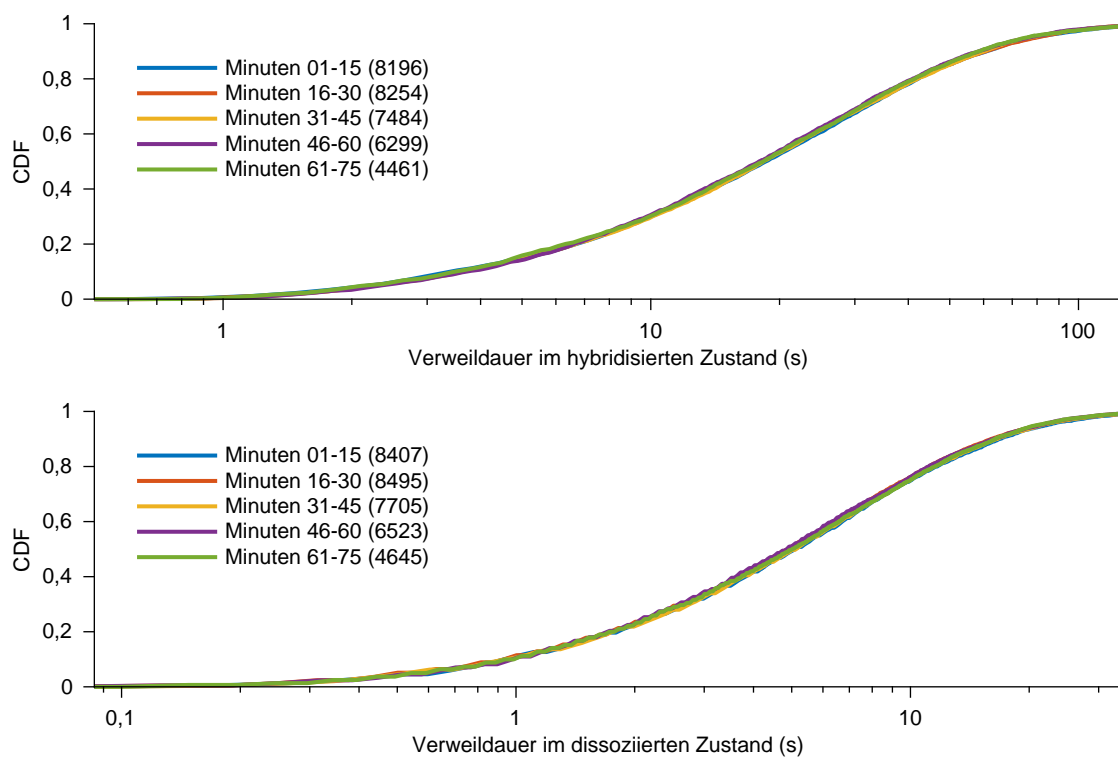


Abb. C.12.9: TACGCCAT:ATGGCGTA, L = 1554 bp, distale Verankerungsgeometrie

Abb. C.12.10: **ACGCATCA:TGATGCGT, L = 1554 bp, distale Verankerungsgeometrie**Abb. C.12.11: **CGCGATAT:ATATCGCG, L = 1554 bp, distale Verankerungsgeometrie**

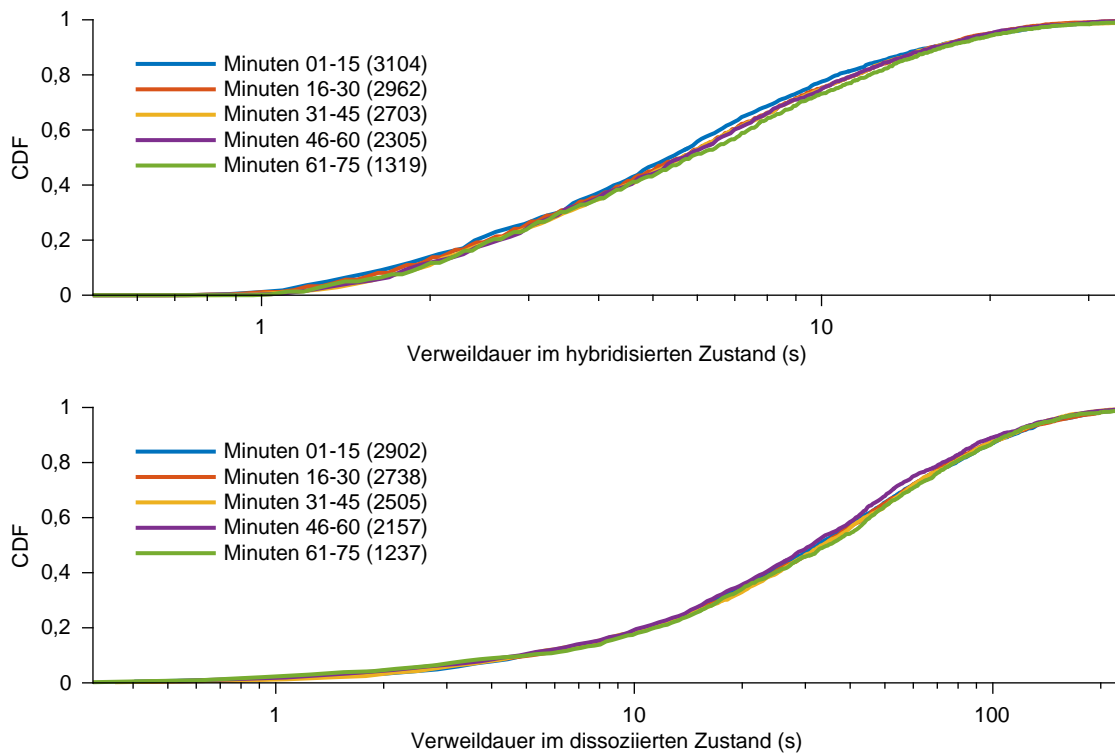


Abb. C.12.12: ACTCTCGT:ACGAGAGT, L = 1554 bp, distale Verankerungsgeometrie

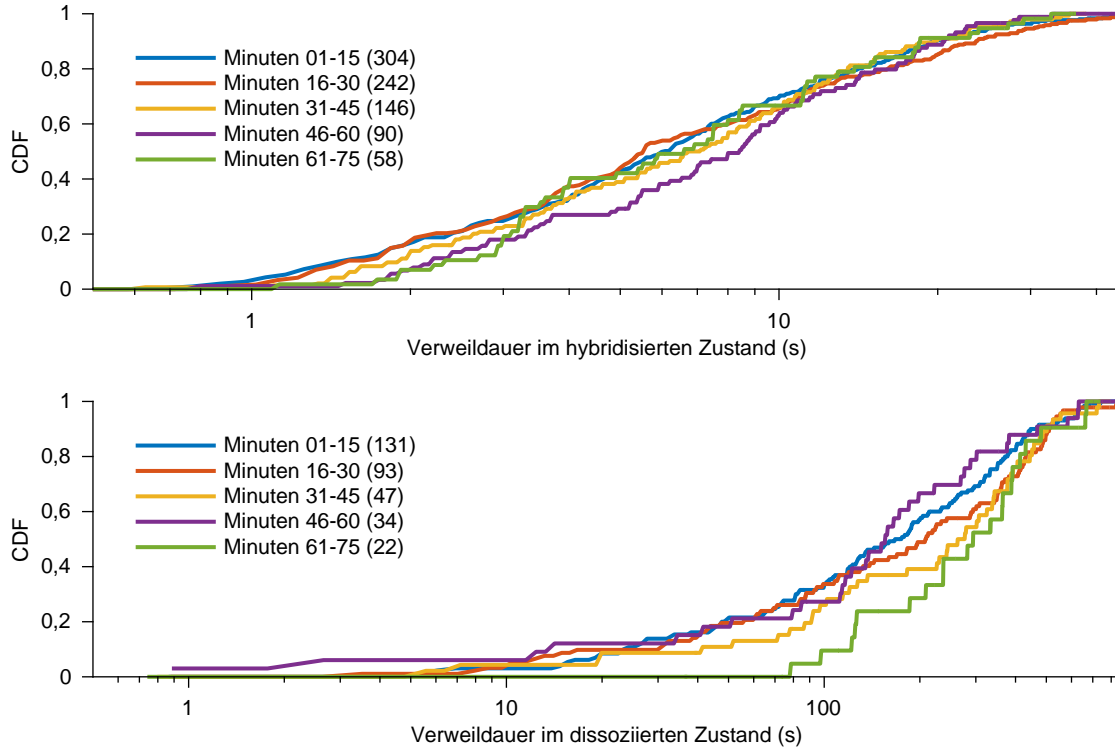


Abb. C.12.13: ACGCAGTA:TACTGCGT, L = 1554 bp, distale Verankerungsgeometrie

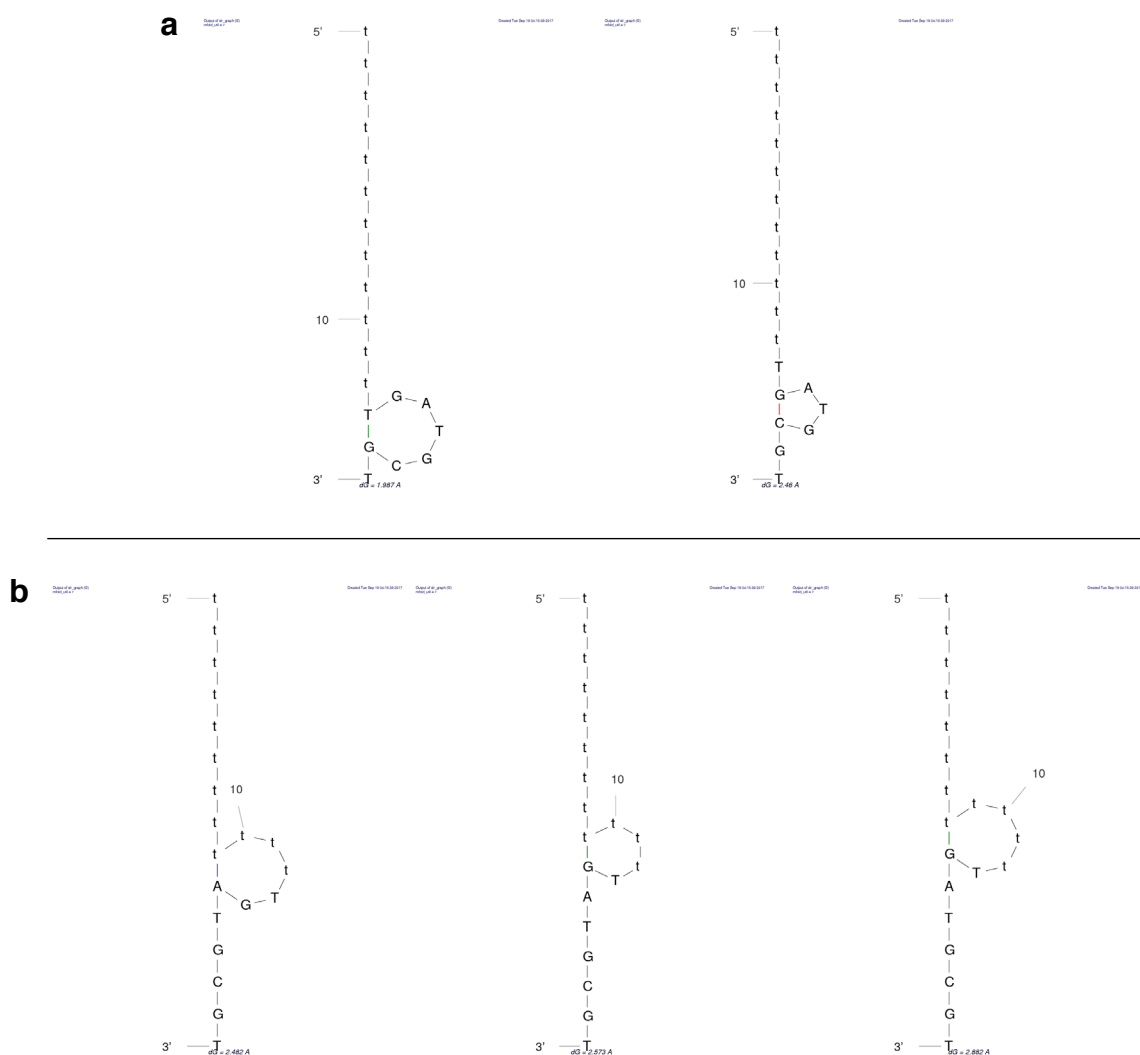


Abbildung C.13: **Berücksichtigte und vernachlässigte Hairpins in Strang B' der Sequenzvariante v3**
 Strukturformeln der Hairpins, erstellt mit DINAmelt [95]. In jedem Diagramm sind das 5'-Ende, das 3'-Ende und die zehnte Stelle in der Basensequenz gekennzeichnet. Der T_{12} -Verbinder ist in Kleinbuchstaben, die Interaktionssequenz in Großbuchstaben dargestellt. Die Linien zwischen den Basen repräsentieren das Rückgrat des Einzelstrangs (schwarz), A:T-Basenpaarungen (blau), G:C-Basenpaarungen (rot) und G:T-Wobble-Basenpaarungen (grün). Die Hairpins in **(a)** wurden bei der Berechnung von p_{Hairpin} nach Gl. (4.9) berücksichtigt, da jeweils zwei der acht Basen der Interaktionssequenz Teil des Hairpin-Stammes sind. Die Hairpins in **(b)** wurden bei der Berechnung von p_{Hairpin} nach Gl. (4.9) vernachlässigt, da jeweils nur eine der acht Basen der Interaktionssequenz Teil des Hairpin-Stammes ist.

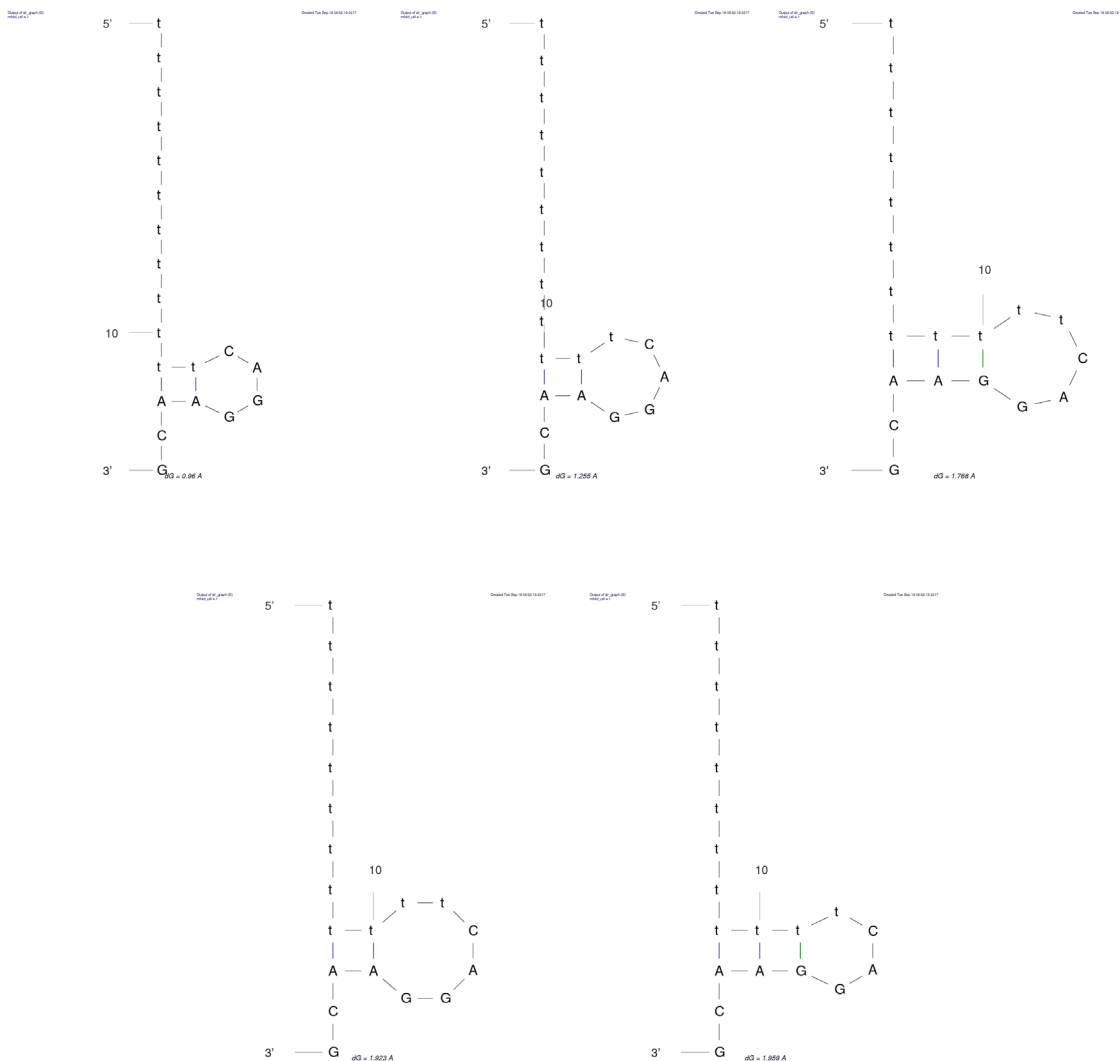


Abbildung C.14: Hairpins in Strang B der Sequenzvariante v5

Strukturformeln der Hairpins, erstellt mit DINAmelt [95]. In jedem Diagramm sind das 5'-Ende, das 3'-Ende und die zehnte Stelle in der Basensequenz gekennzeichnet. Der T₁₂-Verbinder ist in Kleinbuchstaben, die Interaktionssequenz in Großbuchstaben dargestellt. Die Linien zwischen den Basen repräsentieren das Rückgrat des Einzelstrangs (schwarz), A:T-Basenpaarungen (blau) und G:T-Wobble-Basenpaarungen (grün). Alle fünf Hairpins wurden bei der Berechnung von p_{Hairpin} nach Gl. (4.9) berücksichtigt, da jeweils mindestens zwei der acht Basen der Interaktionssequenz Teil des Hairpin-Stammes sind.

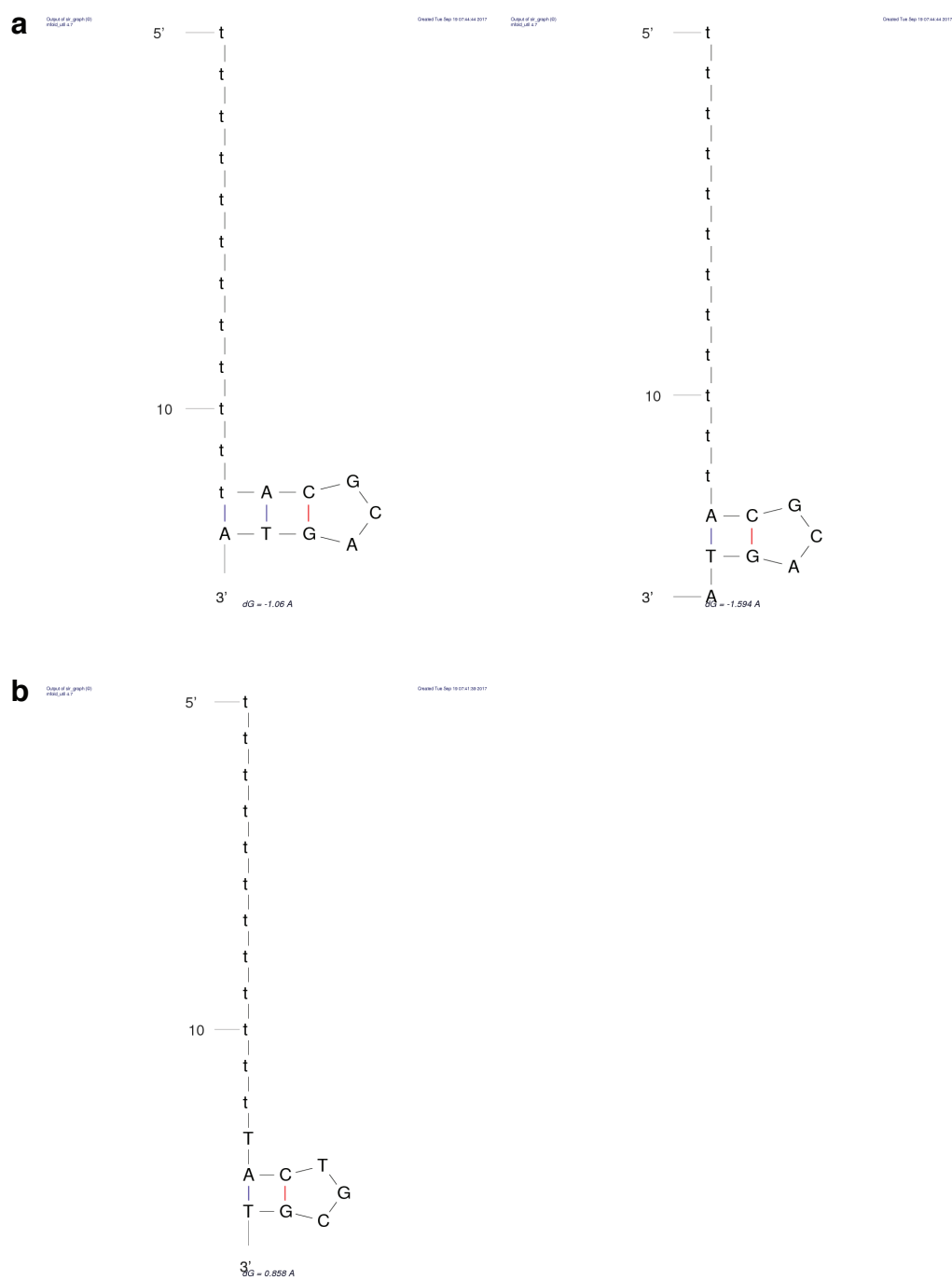


Abbildung C.15: **Hairpins der Sequenzvariante v7**

Strukturformeln der Hairpins, erstellt mit DINAmelt [95]. In jedem Diagramm sind das 5'-Ende, das 3'-Ende und die zehnte Stelle in der Basensequenz gekennzeichnet. Der T_{12} -Verbinder ist in Kleinbuchstaben, die Interaktionssequenz in Großbuchstaben dargestellt. Die Linien zwischen den Basen repräsentieren das Rückgrat des Einzelstrangs (schwarz), A:T-Basenpaarungen (blau) und G:C-Basenpaarungen (rot). **(a)** Mögliche Hairpins in Strang B, **(b)** Mögliches Hairpin in Strang B'. Alle Hairpins wurden bei der Berechnung von p_{Hairpin} nach Gl. (4.9) berücksichtigt, da jeweils mindestens zwei der acht Basen der Interaktionssequenz Teil des Hairpin-Stammes sind.

D Theoretische Ausführungen zu Kapitel 5

D.1 Herleitung von Modell 1 für C_{rel}

Man betrachte das Differentialgleichungssystem aus Abschnitt 5.3:

$$\frac{d}{dt}C_{rel} = k_{on,bi} \cdot J \cdot U_{rel} - k_{off} \cdot C_{rel} \quad (D.1a)$$

$$\frac{d}{dt}U_{rel} = -k_{on,bi} \cdot (J + X) \cdot U_{rel} + k_{off} \cdot (C_{rel} + D_{rel}) \quad (D.1b)$$

$$\frac{d}{dt}D_{rel} = k_{on,bi} \cdot X \cdot U_{rel} - k_{off} \cdot D_{rel} \quad (D.1c)$$

Darin bezeichnen C_{rel} , U_{rel} und D_{rel} die relativen Besetzungen der Zustände C, U und D. X bezeichnet die Konzentration des Kompetitors. $k_{on,bi}$ bezeichnet die bimolekulare Assoziationsrate, k_{off} die Dissoziationsrate des Duplex. Aufgrund der hohen Bindungsenergie des 12-bp-Duplex ($\Delta G_{NN} = -18,1$ kcal/mol) werden folgende zwei Annahmen getroffen:

$$U_{rel} \ll 1 \quad \Leftrightarrow \quad (C_{rel} + D_{rel}) \approx 1 \quad (D.2)$$

$$\frac{d}{dt}U_{rel} \approx 0 \quad (D.3)$$

Damit folgt aus Gl. (D.1b):

$$U_{rel} \approx \frac{k_{off}}{k_{on,bi} \cdot (J + X)} \quad (D.4)$$

und durch Einsetzen in Gl. (D.1a):

$$\frac{d}{dt}C_{rel} = k_{off} \cdot \frac{J}{J + X} - k_{off} \cdot C_{rel} = k_{off} \cdot \frac{1}{1 + \frac{X}{J}} - k_{off} \cdot C_{rel} \quad (D.5)$$

Ist außerdem X sehr viel größer als die Gesamtkonzentration der DNA-Nano-Objekte, so kann X über den Verlauf des Experiments als annähernd konstant betrachtet werden. Dann erhält man durch Differenzieren beider Seiten von (D.5):

$$\frac{d^2}{dt^2}C_{rel} = -k_{off} \cdot \frac{d}{dt}C_{rel} \quad (D.6)$$

Zur Lösung der inhomogenen Differentialgleichung (D.5) wird der folgende Ansatz gewählt:

$$C_{rel}(t) = c(t) \cdot e^{-k_{off} \cdot t} \quad (D.7)$$

Dieser Ansatz ergibt mit der Notation $c'(t) = \frac{d}{dt}c(t)$ und $c''(t) = \frac{d^2}{dt^2}c(t)$:

$$\frac{d}{dt}C_{\text{rel}}(t) = c'(t) \cdot e^{-k_{\text{off}} \cdot t} - k_{\text{off}} \cdot c(t) \cdot e^{-k_{\text{off}} \cdot t} \quad (\text{D.8})$$

$$\frac{d^2}{dt^2}C_{\text{rel}}(t) = c''(t) \cdot e^{-k_{\text{off}} \cdot t} - 2 \cdot k_{\text{off}} \cdot c'(t) \cdot e^{-k_{\text{off}} \cdot t} + k_{\text{off}}^2 \cdot c(t) \cdot e^{-k_{\text{off}} \cdot t} \quad (\text{D.9})$$

Einsetzen von (D.8) und (D.9) in Gl. (D.6) liefert

$$c''(t) = k_{\text{off}} \cdot c'(t) \quad (\text{D.10})$$

Eine Lösung von (D.10) ist $c'(t) = \alpha \cdot e^{k_{\text{off}} \cdot t}$, bzw.

$$c(t) = \frac{\alpha}{k_{\text{off}}} \cdot e^{k_{\text{off}} \cdot t} + \beta \quad (\text{D.11})$$

mit den Integrationskonstanten α und β .

Einsetzen von (D.11) in Gl. (D.7) liefert

$$C_{\text{rel}}(t) = \frac{\alpha}{k_{\text{off}}} + \beta \cdot e^{-k_{\text{off}} \cdot t} \quad (\text{D.12})$$

Einsetzen von (D.12) in Gl. (D.5) ergibt

$$\begin{aligned} -k_{\text{off}} \cdot \beta \cdot e^{-k_{\text{off}} \cdot t} &= k_{\text{off}} \cdot \frac{1}{1 + \frac{X}{J}} - k_{\text{off}} \cdot \frac{\alpha}{k_{\text{off}}} - k_{\text{off}} \cdot \beta \cdot e^{-k_{\text{off}} \cdot t} \\ \Leftrightarrow \alpha &= \frac{k_{\text{off}}}{1 + \frac{X}{J}} \end{aligned} \quad (\text{D.13})$$

Aus der Randbedingung $C_{\text{rel}}(0) = 1$ folgt

$$\beta = 1 - \frac{\alpha}{k_{\text{off}}} = \frac{1}{1 + \frac{J}{X}} \quad (\text{D.14})$$

Einsetzen der Terme für α und β in Gl. (D.12) liefert schlussendlich

$$C_{\text{rel}}(t) = \frac{1}{1 + \frac{X}{J}} + \frac{1}{1 + \frac{J}{X}} \cdot e^{-k_{\text{off}} \cdot t} \quad (\text{D.15})$$

D.2 Theorie zur Analyse möglicher Einzelmolekül-Kompetitions-Experimente

Bezeichnen C, U und D weiterhin die Zustände des Systems aus DNA-Nano-Objekt und Kompetitor wie in Kapitel 5, X die Konkurrentkonzentration und $k_{\text{on,bi}}$ die bimolekulare Assoziationsrate der Interaktionspartner B und B' (in $\text{M}^{-1}\text{s}^{-1}$). Die Assoziationsraten

aus dem Zustand U in die Zustände C bzw. D sind:

$$k_{\text{on,C}} = k_{\text{on,bi}} \cdot J \quad (\text{D.16})$$

$$k_{\text{on,D}} = k_{\text{on,bi}} \cdot X \quad (\text{D.17})$$

Der Zustand U zerfällt demnach mit einer Gesamtrate $k_{\text{on,bi}} \cdot (X + J)$.

Man betrachte ein DNA-Nano-Objekt, das gerade in den Zustand U übergegangen ist. Die Wahrscheinlichkeit, dass der nächste Zustandsübergang dieses Objekts in den Zustand D erfolgt, ist

$$f_D = \frac{k_{\text{on,bi}} \cdot X}{k_{\text{on,bi}} \cdot (X + J)} = \frac{1}{1 + \frac{J}{X}} \quad (\text{D.18})$$

Die Wahrscheinlichkeit, dass dieses Objekt n Mal vom Zustand U in den Zustand D und zurück wechselt und anschließend in den Zustand C übergeht, ist folglich $f_D^n \cdot (1 - f_D)$.

Im TMFM-Experiment kann nicht zwischen den Zuständen U und D unterschieden werden, da die mobile Reporter-Einheit in beiden Zuständen die gleiche Beweglichkeit besitzt. Die Zustände U und D können somit im Kontext der TMFM-Messung zu einem Zustand O vereinigt werden. Die mittlere Verweildauer τ_O des Systems im Zustand O kann mit den obigen Überlegungen durch die mittleren Verweildauern in den Zuständen U und D, τ_U und τ_D , ausgedrückt werden:

$$\begin{aligned} \tau_O &= \tau_U + \sum_{n=1}^{\infty} f_D^n \cdot (1 - f_D) \cdot n \cdot (\tau_U + \tau_D) \\ &= \tau_U + (\tau_U + \tau_D) \cdot (1 - f_D) \cdot \sum_{n=1}^{\infty} n \cdot f_D^n \\ &= \tau_U + (\tau_U + \tau_D) \cdot (1 - f_D) \cdot \frac{f_D}{(1 - f_D)^2} \\ &= \tau_U + (\tau_U + \tau_D) \cdot \frac{X}{J} \\ &= \tau_U \cdot \left(1 + \frac{X}{J}\right) + \tau_D \cdot \frac{X}{J} \\ &= \frac{1}{k_{\text{on,bi}} \cdot (X + J)} \cdot \left(\frac{X + J}{J}\right) + \frac{1}{k_{\text{off,D}}} \cdot \frac{X}{J} \\ &= \frac{1}{J} \cdot \left(\frac{1}{k_{\text{on,bi}}} + \frac{X}{k_{\text{off,D}}}\right) \end{aligned} \quad (\text{D.19})$$

Unter der Annahme $k_{\text{off,D}} = k_{\text{off,C}}$ lässt sich $k_{\text{off,D}}$ aus den inversen mittleren Verweildauern der DNA-Nano-Objekte im Zustand C ermitteln. Damit geht der J-Faktor aus der Steigung des Verlaufs von τ_O über der Konkretorkonzentration X hervor.

Literaturverzeichnis

- [1] N. C. Seeman, "Nucleic acid junctions and lattices," *Journal of Theoretical Biology*, vol. 99, no. 2, pp. 237–247, 1982.
- [2] N. C. Seeman, "Nanomaterials based on DNA," *Annual Review of Biochemistry*, vol. 79, pp. 65–87, 2010.
- [3] F. A. Aldaye, A. L. Palmer, and H. F. Sleiman, "Assembling materials with DNA as the guide," *Science*, vol. 321, no. 5897, pp. 1795–1799, 2008.
- [4] V. Linko and H. Dietz, "The enabled state of DNA nanotechnology," *Current Opinion in Biotechnology*, vol. 24, no. 4, pp. 555–561, 2013.
- [5] F. Zhang, J. Nangreave, Y. Liu, and H. Yan, "Structural DNA nanotechnology: State of the art and future perspective," *Journal of the American Chemical Society*, vol. 136, no. 32, pp. 11198–11211, 2014.
- [6] F. Hong, F. Zhang, Y. Liu, and H. Yan, "DNA Origami: Scaffolds for Creating Higher Order Structures," *Chemical Reviews*, vol. 117, no. 20, pp. 12584–12640, 2017.
- [7] H. Ramezani and H. Dietz, "Building machines with DNA molecules," *Nature Reviews Genetics*, vol. 21, no. 1, pp. 5–26, 2020.
- [8] B. Yurke, A. J. Turberfield, A. P. Mills, F. C. Simmel, and J. L. Neumann, "A DNA-fuelled molecular machine made of DNA," *Nature*, vol. 406, pp. 605–608, 2000.
- [9] H. Yan, X. Zhang, Z. Shen, and N. C. Seeman, "A robust DNA mechanical device controlled by hybridization topology," *Nature*, vol. 415, no. 6867, pp. 62–65, 2002.
- [10] F. C. Simmel and B. Yurke, "A DNA-based molecular device switchable between three distinct mechanical states," *Applied Physics Letters*, vol. 80, no. 5, pp. 883–885, 2002.
- [11] H. Gu, J. Chao, S. J. Xiao, and N. C. Seeman, "A proximity-based programmable DNA nanoscale assembly line," *Nature*, vol. 465, no. 7295, pp. 202–205, 2010.
- [12] T. G. Cha, J. Pan, H. Chen, J. Salgado, X. Li, C. Mao, and J. H. Choi, "A synthetic DNA motor that transports nanoparticles along carbon nanotubes," *Nature Nanotechnology*, vol. 9, no. 1, pp. 39–43, 2014.
- [13] J. Pan, F. Li, T. G. Cha, H. Chen, and J. H. Choi, "Recent progress on DNA based walkers," *Current Opinion in Biotechnology*, vol. 34, pp. 56–64, 2015.
- [14] A. J. Thubagere, W. Li, R. F. Johnson, Z. Chen, S. Doroudi, Y. L. Lee, G. Izatt, S. Wittman, N. Srinivas, D. Woods, E. Winfree, and L. Qian, "A cargo-sorting DNA robot," *Science*, vol. 357, no. 6356, 2017.
- [15] E. S. Andersen, M. Dong, M. M. Nielsen, K. Jahn, R. Subramani, W. Mamdouh, M. M. Golas, B. Sander, H. Stark, C. L. Oliveira, J. S. Pedersen, V. Birkedal, F. Besenbacher, K. V. Gothelf, and J. Kjems, "Self-assembly of a nanoscale DNA box with a controllable lid," *Nature*, vol. 459, no. 7243, pp. 73–76, 2009.
- [16] S. M. Douglas, I. Bachelet, and G. M. Church, "A logic-gated nanorobot for targeted transport of molecular payloads," *Science*, vol. 335, no. 6070, pp. 831–834, 2012.

- [17] E. Kopperger, J. List, S. Madhira, F. Rothfischer, D. C. Lamb, and F. C. Simmel, "A self-assembled nanoscale robotic arm controlled by electric fields," *Science*, vol. 359, no. 6373, pp. 296–301, 2018.
- [18] X. Chen and A. D. Ellington, "Shaping up nucleic acid computation," *Current Opinion in Biotechnology*, vol. 21, no. 4, pp. 392–400, 2010.
- [19] Y. Benenson, "Biomolecular computing systems: Principles, progress and potential," *Nature Reviews Genetics*, vol. 13, no. 7, pp. 455–468, 2012.
- [20] R. Jungmann, C. Steinhauer, M. Scheible, A. Kuzyk, P. Tinnefeld, and F. C. Simmel, "Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami," *Nano Letters*, vol. 10, no. 11, pp. 4756–4761, 2010.
- [21] J. Schnitzbauer, M. T. Strauss, T. Schlichthaerle, F. Schueder, and R. Jungmann, "Super-resolution microscopy with DNA-PAINT," *Nature Protocols*, vol. 12, no. 6, pp. 1198–1228, 2017.
- [22] Y. Yin and X. S. Zhao, "Kinetics and dynamics of DNA hybridization," *Accounts of Chemical Research*, vol. 44, no. 11, pp. 1172–1181, 2011.
- [23] K. D. Whitley, M. J. Comstock, and Y. R. Chemla, "Elasticity of the transition state for oligonucleotide hybridization," *Nucleic Acids Research*, vol. 45, no. 2, pp. 547–555, 2017.
- [24] J. X. Zhang, J. Z. Fang, W. Duan, L. R. Wu, A. W. Zhang, N. Dalchau, B. Yordanov, R. Petersen, A. Phillips, and D. Y. Zhang, "Predicting DNA hybridization kinetics from sequence," *Nature Chemistry*, vol. 10, no. 1, pp. 91–98, 2018.
- [25] S. Xiao, D. J. Sharpe, D. Chakraborty, and D. J. Wales, "Energy Landscapes and Hybridization Pathways for DNA Hexamer Duplexes," *Journal of Physical Chemistry Letters*, vol. 10, no. 21, pp. 6771–6779, 2019.
- [26] S. Howorka, L. Movileanu, O. Braha, and H. Bayley, "Kinetics of duplex formation for individual DNA strands within a single protein nanopore," *Proceedings of the National Academy of Sciences*, vol. 98, no. 23, pp. 12996–13001, 2001.
- [27] H. P. Lehr, M. Reimann, A. Brandenburg, G. Sulz, and H. Klapproth, "Real-time detection of nucleic acid interactions by total internal reflection fluorescence," *Analytical Chemistry*, vol. 75, no. 10, pp. 2414–2420, 2003.
- [28] M. M. Sekar, W. Bloch, and P. M. St John, "Comparative study of sequence-dependent hybridization kinetics in solution and on microspheres," *Nucleic Acids Research*, vol. 33, no. 1, pp. 366–375, 2005.
- [29] Y. Gao, L. K. Wolf, and R. M. Georgiadis, "Secondary structure effects on DNA hybridization kinetics: A solution versus surface comparison," *Nucleic Acids Research*, vol. 34, no. 11, pp. 3370–3377, 2006.
- [30] C. Chen, W. Wang, Z. Wang, F. Wei, and X. S. Zhao, "Influence of secondary structure on kinetics and reaction mechanism of DNA hybridization," *Nucleic Acids Research*, vol. 35, no. 9, pp. 2875–2884, 2007.
- [31] W. Michel, T. Mai, T. Naiser, and A. Ott, "Optical study of DNA surface hybridization reveals DNA surface density as a key parameter for microarray hybridization kinetics," *Biophysical Journal*, vol. 92, no. 3, pp. 999–1004, 2007.
- [32] X. Chen, Y. Zhou, P. Qu, and S. Z. Xin, "Base-by-base dynamics in DNA hybridization probed by fluorescence correlation spectroscopy," *Journal of the American Chemical Society*, vol. 130, no. 50, pp. 16947–16952, 2008.
- [33] D. Y. Zhang and E. Winfree, "Control of DNA strand displacement kinetics using toehold exchange," *Journal of the American Chemical Society*, vol. 131, no. 47, pp. 17303–17314, 2009.

- [34] I. I. Cisse, H. Kim, and T. Ha, "A rule of seven in Watson-Crick base-pairing of mismatched sequences," *Nature structural & molecular biology*, vol. 19, no. 6, pp. 623–627, 2012.
- [35] N. F. Dupuis, E. D. Holmstrom, and D. J. Nesbitt, "Single-molecule kinetics reveal cation-promoted DNA duplex formation through ordering of single-stranded helices," *Biophysical Journal*, vol. 105, no. 3, pp. 756–766, 2013.
- [36] T. E. Ouldrige, P. Šulc, F. Romano, J. P. Doye, and A. A. Louis, "DNA hybridization kinetics: Zippering, internal displacement and sequence dependence," *Nucleic Acids Research*, vol. 41, no. 19, pp. 8886–8895, 2013.
- [37] A. Johnson-Buck and N. G. Walter, "Discovering anomalous hybridization kinetics on DNA nanostructures using single-molecule fluorescence microscopy," *Methods*, vol. 67, no. 2, pp. 177–184, 2014.
- [38] J. S. Schreck, T. E. Ouldrige, F. Romano, P. Šulc, L. P. Shaw, A. A. Louis, and J. P. K. Doye, "DNA hairpins destabilize duplexes primarily by promoting melting rather than by inhibiting hybridization," *Nucleic Acids Research*, vol. 43, no. 13, pp. 6181–6190, 2015.
- [39] E. M. Peterson, M. W. Manhart, and J. M. Harris, "Single-Molecule Fluorescence Imaging of Interfacial DNA Hybridization Kinetics at Selective Capture Surfaces," *Analytical Chemistry*, vol. 88, no. 2, pp. 1345–1354, 2016.
- [40] C. Liu, J. M. Obliosca, Y. L. Liu, Y. A. Chen, N. Jiang, and H. C. Yeh, "3D single-molecule tracking enables direct hybridization kinetics measurement in solution," *Nanoscale*, vol. 9, no. 17, pp. 5664–5670, 2017.
- [41] J. C. Traeger and D. K. Schwartz, "Surface-Mediated DNA Hybridization: Effects of DNA Conformation, Surface Chemistry, and Electrostatics," *Langmuir*, vol. 33, no. 44, pp. 12651–12659, 2017.
- [42] H. Hata, T. Kitajima, and A. Suyama, "Influence of thermodynamically unfavorable secondary structures on DNA hybridization kinetics," *Nucleic Acids Research*, vol. 46, no. 2, pp. 1–10, 2017.
- [43] R. J. Menssen and A. Tokmakoff, "Length-Dependent Melting Kinetics of Short DNA Oligonucleotides Using Temperature-Jump IR Spectroscopy," *Journal of Physical Chemistry B*, vol. 123, no. 4, pp. 756–767, 2019.
- [44] Y. I. Chen, Y. J. Chang, T. D. Nguyen, C. Liu, S. Phillion, Y. A. Kuo, H. T. Vu, A. Liu, Y. L. Liu, S. Hong, P. Ren, T. E. Yankeelov, and H. C. Yeh, "Measuring DNA Hybridization Kinetics in Live Cells Using a Time-Resolved 3D Single-Molecule Tracking Method," *Journal of the American Chemical Society*, vol. 141, no. 40, pp. 15747–15750, 2019.
- [45] J. G. Wetmur and N. Davidson, "Kinetics of renaturation of DNA," *Journal of Molecular Biology*, vol. 31, no. 3, pp. 349–370, 1968.
- [46] M. E. Craig, D. M. Crothers, and P. Doty, "Relaxation Kinetics of Dimer Formation by Self Complementary Oligonucleotides," *Journal of Molecular Biology*, vol. 62, no. 2, pp. 383–401, 1971.
- [47] D. Pörschke, O. C. Uhlenbeck, and F. H. Martin, "Thermodynamics and kinetics of the helix-coil transition of oligomers containing GC base pairs," *Biopolymers*, vol. 12, no. 6, pp. 1313–1335, 1973.
- [48] Y. G. Chu and I. Tinoco, "Temperature-jump kinetics of the dC-G-T-G-A-A-T-T-C-G-C-G double helix containing a G · T base pair and the dC-G-C-A-G-A-A-T-T-C-G-C-G double helix containing an extra adenine," *Biopolymers*, vol. 22, no. 4, pp. 1235–46., 1983.

- [49] S. Ikuta, K. Takagi, R. Wallace, and K. Itakura, "Dissociation kinetics of 19 base paired oligonucleotide-DNA duplexes containing different single mismatched base pairs," *Nucleic Acids Research*, vol. 15, no. 2, pp. 797–811, 1987.
- [50] A. P. Williams, C. E. Longfellow, S. M. Freier, R. Kierzek, and D. H. Turner, "Laser Temperature-Jump, Spectroscopic, and Thermodynamic Study of Salt Effects on Duplex Formation by dGCATGC," *Biochemistry*, vol. 28, no. 10, pp. 4283–4291, 1989.
- [51] L. E. Morrison and L. M. Stols, "Sensitive Fluorescence-Based Thermodynamic and Kinetic Measurements of DNA Hybridization in Solution," *Biochemistry*, vol. 32, no. 12, pp. 3095–3104, 1993.
- [52] K. M. Parkhurst and L. J. Parkhurst, "Kinetic Studies by Fluorescence Resonance Energy Transfer Employing a Double-Labeled Oligonucleotide: Hybridization to the Oligonucleotide Complement and to Single-Stranded DNA," *Biochemistry*, vol. 34, no. 1, pp. 285–292, 1995.
- [53] M. R. Henry, P. Wilkins Stevens, J. Sun, and D. M. Kelso, "Real-time measurements of DNA hybridization on microparticles with fluorescence resonance energy transfer," *Analytical Biochemistry*, vol. 276, no. 2, pp. 204–214, 1999.
- [54] A. Johnson-Buck, X. Su, M. D. Giraldez, M. Zhao, M. Tewari, and N. G. Walter, "Kinetic fingerprinting to identify and count single nucleic acids," *Nature Biotechnology*, vol. 33, no. 7, pp. 730–732, 2015.
- [55] X. Su, L. Li, S. Wang, D. Hao, L. Wang, and C. Yu, "Single-Molecule Counting of Point Mutations by Transient DNA Binding," *Scientific Reports*, vol. 7, no. 1, pp. 1–9, 2017.
- [56] E. M. Peterson and J. M. Harris, "Identification of Individual Immobilized DNA Molecules by their Hybridization Kinetics Using Single-Molecule Fluorescence Imaging," *Analytical Chemistry*, vol. 90, no. 8, pp. 5007–5014, 2018.
- [57] D. A. Schafer, J. Gelles, M. P. Sheetz, and R. Landick, "Transcription by single molecules of RNA polymerase observed by light microscopy," *Nature*, vol. 352, pp. 444–448, 1991.
- [58] P. F. J. May, J. N. M. Pinkney, P. Zawadzki, G. W. Evans, D. J. Sherratt, and A. N. Kapanidis, "Tethered fluorophore motion: Studying large DNA conformational changes by single-fluorophore imaging," *Biophysical Journal*, vol. 107, no. 5, pp. 1205–1216, 2014.
- [59] H. Yin, R. Landick, and J. Gelles, "Tethered particle motion method for studying transcript elongation by a single RNA polymerase molecule," *Biophysical Journal*, vol. 67, no. 6, pp. 2468–2478, 1994.
- [60] H. Yin, I. Artsimovitch, R. Landick, and J. Gelles, "Nonequilibrium mechanism of transcription termination from observations of single RNA polymerase molecules," *Proceedings of the National Academy of Sciences*, vol. 96, no. 23, pp. 13124–13129, 1999.
- [61] S. F. Tolić-Nørrelykke, A. M. Engh, R. Landick, and J. Gelles, "Diversity in the Rates of Transcript Elongation by Single RNA Polymerase Molecules," *Journal of Biological Chemistry*, vol. 279, no. 5, pp. 3292–3299, 2004.
- [62] L. Finzi and J. Gelles, "Measurement of Lactose Repressor-Mediated Loop Formation and Breakdown," *Science*, vol. 267, no. 5196, pp. 378–380, 1995.
- [63] G. Lia, D. Bensimon, V. Croquette, J.-F. Allemand, D. Dunlap, D. E. A. Lewis, S. Adhya, and L. Finzi, "Supercoiling and denaturation in Gal repressor/heat unstable nucleoid protein (HU)-mediated DNA looping," *Proceedings of the National Academy of Sciences*, vol. 100, no. 20, pp. 11373–7, 2003.

- [64] N. Pouget, C. Turlan, N. Destainville, L. Salomé, and M. Chandler, "IS911 transpososome assembly as analysed by tethered particle motion," *Nucleic Acids Research*, vol. 34, no. 16, pp. 4313–4323, 2006.
- [65] F. Vanzi, C. Broggio, L. Sacconi, and F. S. Pavone, "Lac repressor hinge flexibility and DNA looping: Single molecule kinetics by tethered particle motion," *Nucleic Acids Research*, vol. 34, no. 12, pp. 3409–3420, 2006.
- [66] J. F. Beausang and P. C. Nelson, "Diffusive hidden Markov model characterization of DNA looping dynamics in tethered particle experiments," *Physical Biology*, vol. 4, no. 3, pp. 205–219, 2007.
- [67] D. Normanno, F. Vanzi, and F. S. Pavone, "Single-molecule manipulation reveals supercoiling-dependent modulation of lac repressor-mediated DNA looping," *Nucleic Acids Research*, vol. 36, no. 8, pp. 2505–2513, 2008.
- [68] O. K. Wang, M. Guthold, D. A. Erie, and J. Gelles, "Interconvertible lac repressor-DNA loops revealed by single-molecule experiments," *PLoS Biology*, vol. 6, no. 9, pp. 2028–2042, 2008.
- [69] L. Han, H. G. Garcia, S. Blumberg, K. B. Towles, J. F. Beausang, P. C. Nelson, and R. Phillips, "Concentration and length dependence of DNA looping in transcriptional regulation," *PLoS ONE*, vol. 4, no. 5, pp. 1–61, 2009.
- [70] D. Rutkauskas, H. Zhan, K. S. Matthews, F. S. Pavone, and F. Vanzi, "Tetramer opening in LacI-mediated DNA looping," *Proceedings of the National Academy of Sciences*, vol. 106, no. 39, pp. 16627–32, 2009.
- [71] N. Laurens, S. R. Bellamy, A. F. Harms, Y. S. Kovacheva, S. E. Halford, and G. J. Wuite, "Dissecting protein-induced DNA looping dynamics in real time," *Nucleic Acids Research*, vol. 37, no. 16, pp. 5454–5464, 2009.
- [72] C. Zurla, C. Manzo, D. Dunlap, D. E. A. Lewis, S. Adhya, and L. Finzi, "Direct demonstration and quantification of long-range DNA looping by the λ bacteriophage repressor," *Nucleic Acids Research*, vol. 37, no. 9, pp. 2789–2795, 2009.
- [73] D. A. Rusling, N. Laurens, C. Pernstich, G. J. L. Wuite, and S. E. Halford, "DNA looping by FokI: The impact of synapse geometry on loop topology at varied site orientations," *Nucleic Acids Research*, vol. 40, no. 11, pp. 4977–4987, 2012.
- [74] N. Laurens, D. A. Rusling, C. Pernstich, I. Brouwer, S. E. Halford, and G. J. L. Wuite, "DNA looping by FokI: The impact of twisting and bending rigidity on protein-induced looping dynamics," *Nucleic Acids Research*, vol. 40, no. 11, pp. 4988–4997, 2012.
- [75] C. Manzo, C. Zurla, D. D. Dunlap, and L. Finzi, "The effect of nonspecific binding of lambda repressor on DNA looping dynamics," *Biophysical Journal*, vol. 103, no. 8, pp. 1753–1761, 2012.
- [76] S. Johnson, M. Lindén, and R. Phillips, "Sequence dependence of transcription factor-mediated DNA looping," *Nucleic Acids Research*, vol. 40, no. 16, pp. 7728–7738, 2012.
- [77] S. Johnson, J. W. Van De Meent, R. Phillips, C. H. Wiggins, and M. Lindén, "Multiple LacI-mediated loops revealed by Bayesian statistics and tethered particle motion," *Nucleic Acids Research*, vol. 42, no. 16, pp. 10265–10277, 2014.
- [78] T. Plénat, C. Tardin, P. Rousseau, and L. Salomé, "High-throughput single-molecule analysis of DNA-protein interactions by tethered particle motion," *Nucleic Acids Research*, vol. 40, no. 12, pp. e89–e89, 2012.

- [79] P. C. Nelson, C. Zurla, D. Brogioli, J. F. Beausang, L. Finzi, and D. Dunlap, "Tethered particle motion as a diagnostic of DNA tether length," *Journal of Physical Chemistry B*, vol. 110, no. 34, pp. 17260–17267, 2006.
- [80] S. Brinkers, H. R. C. Dietrich, F. H. De Groot, I. T. Young, and B. Rieger, "The persistence length of double stranded DNA determined using dark field tethered particle motion," *Journal of Chemical Physics*, vol. 130, no. 21, p. 06B607, 2009.
- [81] S. Kumar, C. Manzo, C. Zurla, S. Ucuuncuoglu, L. Finzi, and D. Dunlap, "Enhanced tethered-particle motion analysis reveals viscous effects," *Biophysical Journal*, vol. 106, no. 2, pp. 399–409, 2014.
- [82] J. P. Mumm, A. Landy, and J. Gelles, "Viewing single lambda site-specific recombination events from start to finish," *EMBO Journal*, vol. 25, no. 19, pp. 4586–4595, 2006.
- [83] H. F. Fan, M. M. Cox, and H. W. Li, "Developing Single-Molecule TPM experiments for direct observation of successful RecA-Mediated strand exchange reaction," *PLoS ONE*, vol. 6, no. 7, p. e21359, 2011.
- [84] J. N. M. Pinkney, P. Zawadzki, J. Mazuryk, L. K. Arciszewska, D. J. Sherratt, and A. N. Kapanidis, "Capturing reaction paths and intermediates in Cre-loxP recombination using single-molecule fluorescence," *Proceedings of the National Academy of Sciences*, vol. 109, no. 51, pp. 20871–20876, 2012.
- [85] H. F. Fan, C. H. Ma, and M. Jayaram, "Real-time single-molecule tethered particle motion analysis reveals mechanistic similarities and contrasts of FliP site-specific recombinase with Cre and λ Int," *Nucleic Acids Research*, vol. 41, no. 14, pp. 7031–7047, 2013.
- [86] P. Zawadzki, P. F. J. May, R. A. Baker, J. N. M. Pinkney, A. N. Kapanidis, D. J. Sherratt, and L. K. Arciszewska, "Conformational transitions during FtsK translocase activation of individual XerCD-dif recombination complexes," *Proceedings of the National Academy of Sciences*, vol. 110, no. 43, pp. 17302–17307, 2013.
- [87] P. F. J. May, P. Zawadzki, D. J. Sherratt, A. N. Kapanidis, and L. K. Arciszewska, "Assembly, translocation, and activation of XerCD-dif recombination by FtsK translocase analyzed in real-time by FRET and two-color tethered fluorophore motion," *Proceedings of the National Academy of Sciences*, vol. 112, no. 37, pp. E5133–E5141, 2015.
- [88] P. W. Rothmund, "Folding DNA to create nanoscale shapes and patterns," *Nature*, vol. 440, no. 7082, pp. 297–302, 2006.
- [89] S. M. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf, and W. M. Shih, "Self-assembly of DNA into nanoscale three-dimensional shapes," *Nature*, vol. 459, no. 7245, pp. 414–418, 2009.
- [90] H. Dietz, S. M. Douglas, and W. M. Shih, "Folding DNA into twisted and curved nanoscale shapes," *Science*, vol. 325, no. 5941, pp. 725–730, 2009.
- [91] C. E. Castro, F. Kilchherr, D. N. Kim, E. L. Shiao, T. Wauer, P. Wortmann, M. Bathe, and H. Dietz, "A primer to scaffolded DNA origami," *Nature Methods*, vol. 8, no. 3, pp. 221–229, 2011.
- [92] B. Kick, F. Praetorius, H. Dietz, and D. Weuster-Botz, "Efficient Production of Single-Stranded Phage DNA as Scaffolds for DNA Origami," *Nano Letters*, vol. 15, no. 7, pp. 4672–4676, 2015.
- [93] S. M. Douglas, S. Teerapittayanon, A. Vazquez, W. M. Shih, A. H. Marblestone, and G. M. Church, "Rapid prototyping of 3D DNA-origami shapes with caDNAno," *Nucleic Acids Research*, vol. 37, no. 15, pp. 5001–5006, 2009.

- [94] D.-N. Kim, F. Kilchherr, H. Dietz, and M. Bathe, "Quantitative prediction of 3D solution shape and flexibility of nucleic acid nanostructures," *Nucleic acids research*, vol. 40, no. 7, pp. 2862–8, 2012.
- [95] N. R. Markham and M. Zuker, "DINAMelt web server for nucleic acid melting prediction," *Nucleic Acids Research*, vol. 33, no. SUPPL. 2, pp. W577–W581, 2005.
- [96] E. Yarmola, H. Sokoloff, and A. Chrambach, "The relative contributions of dispersion and diffusion to band spreading (resolution) in gel electrophoresis," *Electrophoresis*, vol. 17, no. 9, pp. 1416–1419, 1996.
- [97] E. Yarmola, P. P. Calabrese, A. Chrambach, and G. H. Weiss, "Interaction with the Matrix: The Dominant Factor in Macromolecular Band Spreading in Gel Electrophoresis," *The Journal of Physical Chemistry B*, vol. 101, no. 13, pp. 2381–2387, 2002.
- [98] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*. Springer Science+Business Media, LLC, third ed., 2006.
- [99] I. Rasnik, S. A. McKinney, and T. Ha, "Nonblinking and long-lasting single-molecule fluorescence imaging," *Nature Methods*, vol. 3, no. 11, pp. 891–893, 2006.
- [100] M. Swoboda, J. Henig, H. M. Cheng, D. Brugger, D. Haltrich, N. Plumeré, and M. Schlierf, "Enzymatic Oxygen Scavenging for Photostability without pH Drop in Single-Molecule Experiments," *ACS Nano*, vol. 6, no. 7, pp. 6364–6369, 2012.
- [101] L. Han, B. H. Lui, S. Blumberg, J. F. Beausang, P. C. Nelson, and R. Phillips, "Calibration of Tethered Particle Motion Experiments," in *Mathematics of DNA Structure, Function and Interactions*, pp. 123–138, Springer, New York, NY, 2009.
- [102] L. Firestone, K. Cook, K. Culp, N. Talsania, and K. Preston, "Comparison of autofocus methods for automated microscopy," *Cytometry*, vol. 12, no. 3, pp. 195–206, 1991.
- [103] A. J. Berglund, M. D. McMahon, J. J. McClelland, and J. A. Liddle, "Fast, bias-free algorithm for tracking single particles with variable size and shape," *Optics Express*, vol. 16, no. 18, pp. 14064–14075, 2008.
- [104] K. F. Wagenbauer, C. H. Wachauf, and H. Dietz, "Quantifying quality in DNA self-assembly," *Nature Communications*, vol. 5, no. 1, pp. 1–7, 2014.
- [105] L. R. Rabiner and B. H. Juang, "An introduction to hidden Markov models," *IEEE Signal Proc. Magazine*, vol. 3, no. 1, pp. 4–16, 1986.
- [106] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [107] J. D. Chodera, P. Elms, F. Noé, B. Keller, C. M. Kaiser, A. Ewall-Wice, S. Marqusee, C. Bustamante, and N. S. Hinrichs, "Bayesian hidden Markov model analysis of single-molecule force spectroscopy: characterizing kinetics under measurement uncertainty," *arXiv preprint arXiv:1108.1430*, 2011.
- [108] J. Stigler and M. Rief, "Hidden Markov analysis of trajectories in single-molecule experiments and the effects of missed events," *ChemPhysChem*, vol. 13, no. 4, pp. 1079–1086, 2012.
- [109] D. Pörschke and M. Eigen, "Co-operative non-enzymic base recognition. 3. Kinetics of the helix-coil transition of the oligoribouridylic–oligoriboadenylic acid system and of oligoriboadenylic acid alone at acidic pH.," *Journal of molecular biology*, vol. 62, no. 2, pp. 361–381, 1971.
- [110] B. Roux and R. Sauvé, "A general solution to the time interval omission problem applied to single channel analysis," *Biophysical Journal*, vol. 48, no. 1, pp. 149–158, 1985.

- [111] A. L. Blatz and K. L. Magleby, "Correcting single channel data for missed events," *Biophysical Journal*, vol. 49, no. 5, pp. 967–980, 1986.
- [112] S. C. Crouzy and F. J. Sigworth, "Yet another approach to the dwell-time omission problem of single-channel analysis," *Biophysical Journal*, vol. 58, no. 3, pp. 731–743, 1990.
- [113] R. Phillips, J. Kondev, and J. Theriot, *Physical Biology of the Cell*. New York: Garland Science, 2009.
- [114] H. Yamakawa and W. H. Stockmayer, "Statistical mechanics of wormlike chains. II. Excluded volume effects," *The Journal of Chemical Physics*, vol. 57, no. 7, pp. 2843–2854, 1972.
- [115] D. Shore and R. L. Baldwin, "Energetics of DNA twisting. I. Relation between twist and cyclization probability," *Journal of Molecular Biology*, vol. 170, no. 4, pp. 957–981, 1983.
- [116] J. Yan, R. Kawamura, and J. F. Marko, "Statistics of loop formation along double helix DNAs," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 71, no. 6, p. 061905, 2005.
- [117] D. E. Segall, P. C. Nelson, and R. Phillips, "Volume-exclusion effects in tethered-particle experiments: Bead size matters," *Physical Review Letters*, vol. 96, no. 8, p. 088306, 2006.
- [118] P. Furrer, J. Bednar, A. Z. Stasiak, V. Katritch, D. Michoud, A. Stasiak, and J. Dubochet, "Opposite effect of counterions on the persistence length of nicked and non-nicked DNA," *Journal of Molecular Biology*, vol. 266, no. 4, pp. 711–721, 1997.
- [119] P. Yakovchuk, E. Protozanova, and M. D. Frank-Kamenetskii, "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix," *Nucleic Acids Research*, vol. 34, no. 2, pp. 564–574, 2006.
- [120] J. SantaLucia, "A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics," *Proceedings of the National Academy of Sciences*, vol. 95, no. 4, pp. 1460–1465, 1998.
- [121] J. SantaLucia and D. Hicks, "The Thermodynamics of DNA Structural Motifs," *Annual Review of Biophysics and Biomolecular Structure*, vol. 33, no. 1, pp. 415–440, 2004.
- [122] W. H. Braunlin and V. A. Bloomfield, "¹H NMR Study of the Base-Pairing Reactions of d(GGAATTCC): Salt Effects on the Equilibria and Kinetics of Strand Association," *Biochemistry*, vol. 30, no. 3, pp. 754–758, 1991.
- [123] A. Savelyev, "Do monovalent mobile ions affect DNA's flexibility at high salt content?," *Physical Chemistry Chemical Physics*, vol. 14, no. 7, pp. 2250–2254, 2012.
- [124] T. Odijk, "Polyelectrolytes Near the Rod Limit.," *J Polym Sci Polym Phys Ed*, vol. 15, no. 3, pp. 477–483, 1977.
- [125] J. Skolnick and M. Fixman, "Electrostatic Persistence Length of Wormlike Polyelectrolyte," *Macromolecules*, vol. 10, no. 5, pp. 944–948, 1977.
- [126] D. Pörschke, "Persistence length and bending dynamics of DNA from electrooptical measurements at high salt concentrations," *Biophysical Chemistry*, vol. 40, no. 2, pp. 169–179, 1991.
- [127] W. H. Taylor and P. J. Hagerman, "Application of the method of phage T4 DNA ligase-catalyzed ring-closure to the study of DNA structure. II. NaCl-dependence of DNA flexibility and helical repeat," *Journal of Molecular Biology*, vol. 212, no. 2, pp. 363–376, 1990.
- [128] C. Baumann, S. Smith, V. Bloomfield, and C. Bustamante, "Ionic effects on the elasticity of single DNA molecules," *Proceedings of the National Academy of Sciences*, vol. 94, no. 12, pp. 6185–6190, 1997.

- [129] R. Podgornik, P. L. Hansen, and V. A. Parsegian, "Elastic moduli renormalization in self-interacting stretchable polyelectrolytes," *Journal of Chemical Physics*, vol. 113, no. 20, pp. 9343–9350, 2000.
- [130] N. Borochoy, H. Eisenberg, and Z. Kam, "Dependence of DNA Conformation on the Concentration of Salt," *Biopolymers*, vol. 20, no. 1, pp. 231–235, 1981.
- [131] G. S. Manning, "The persistence length of DNA is reached from the persistence length of its null isomer through an internal electrostatic stretching force," *Biophysical Journal*, vol. 91, no. 10, pp. 3607–3616, 2006.
- [132] A. Brunet, C. Tardin, L. Salomé, P. Rousseau, N. Destainville, and M. Manghi, "Dependence of DNA Persistence Length on Ionic Strength of Solutions with Monovalent and Divalent Salts: A Joint Theory-Experiment Study," *Macromolecules*, vol. 48, no. 11, pp. 3641–3652, 2015.
- [133] S. Guilbaud, L. Salomé, N. Destainville, M. Manghi, and C. Tardin, "Dependence of DNA Persistence Length on Ionic Strength and Ion Type," *Physical Review Letters*, vol. 122, no. 2, p. 028102, 2019.
- [134] S. P. Walton, G. N. Stephanopoulos, M. L. Yarmush, and C. M. Roth, "Thermodynamic and Kinetic Characterization of Antisense Oligodeoxynucleotide Binding to a Structured mRNA," *Biophysical Journal*, vol. 82, no. 1, pp. 366–377, 2002.
- [135] S. A. Kushon, J. P. Jordan, J. L. Seifert, H. Nielsen, P. E. Nielsen, and B. A. Armitage, "Effect of secondary structure on the thermodynamics and kinetics of PNA hybridization to DNA hairpins," *Journal of the American Chemical Society*, vol. 123, no. 44, pp. 10805–10813, 2001.
- [136] M. Nakano, E. M. Moody, J. Liang, and P. C. Bevilacqua, "Selection for Thermodynamically Stable DNA Tetraloops Using Temperature Gradient Gel Electrophoresis Reveals Four Motifs : d(cGNNAg), d(cGNABg), d(cCNNGg), and d(gCNNGc)," *Biochemistry*, vol. 41, no. 48, pp. 14281–14292, 2002.
- [137] S. Yoshizawa, G. Kawai, K. Watanabe, K.-i. Miura, and I. Hirao, "GNA Trinucleotide Loop Sequences Producing Extraordinarily Stable DNA," *Biochemistry*, vol. 36, no. 16, pp. 4761–4767, 1997.
- [138] M. Schickinger, M. Zacharias, and H. Dietz, "Tethered multifluorophore motion reveals equilibrium transition kinetics of single DNA double helices," *Proceedings of the National Academy of Sciences*, vol. 115, no. 32, pp. E7512–E7521, 2018.
- [139] J. B. Trads, T. Tørring, and K. V. Gothelf, "Site-Selective Conjugation of Native Proteins with DNA," *Accounts of Chemical Research*, vol. 50, no. 6, pp. 1367–1374, 2017.
- [140] K. Lang and J. W. Chin, "Cellular incorporation of unnatural amino acids and bioorthogonal Labeling of Proteins," *Chemical Reviews*, vol. 114, no. 9, pp. 4764–4806, 2014.
- [141] W. Shen, H. Zhong, D. Neff, and M. L. Norton, "NTA directed protein nanopatterning on DNA origami nanoconstructs," *Journal of the American Chemical Society*, vol. 131, no. 19, pp. 6660–6661, 2009.
- [142] B. Saccà, R. Meyer, M. Erkelenz, K. Kiko, A. Arndt, H. Schroeder, K. S. Rabe, and C. M. Niemeyer, "Orthogonal protein decoration of DNA origami," *Angewandte Chemie - International Edition*, vol. 49, no. 49, pp. 9378–9383, 2010.
- [143] N. Stephanopoulos, M. Liu, G. J. Tong, Z. Li, Y. Liu, H. Yan, and M. B. Francis, "Immobilization and one-dimensional arrangement of virus capsids with nanoscale precision using DNA origami," *Nano Letters*, vol. 10, no. 7, pp. 2714–2720, 2010.

- [144] J. Fu, M. Liu, Y. Liu, N. W. Woodbury, and H. Yan, "Interenzyme substrate diffusion for an enzyme cascade organized on spatially addressable DNA nanostructures," *Journal of the American Chemical Society*, vol. 134, no. 12, pp. 5516–5519, 2012.
- [145] Y. Fu, D. Zeng, J. Chao, Y. Jin, Z. Zhang, H. Liu, D. Li, H. Ma, Q. Huang, K. V. Gothelf, and C. Fan, "Single-step rapid assembly of DNA origami nanostructures for addressable nanoscale bioreactors," *Journal of the American Chemical Society*, vol. 135, no. 2, pp. 696–702, 2013.
- [146] A. Shaw, V. Lundin, E. Petrova, F. Fördos, E. Benson, A. Al-Amin, A. Herland, A. Blokzijl, B. Högberg, and A. I. Teixeira, "Spatial control of membrane receptor function using ligand nanocalipers," *Nature Methods*, vol. 11, no. 8, pp. 841–846, 2014.
- [147] A. Shaw, E. Benson, and B. Högberg, "Purification of functionalized DNA origami nanostructures," *ACS Nano*, vol. 9, no. 5, pp. 4968–4975, 2015.
- [148] G. Ke, M. Liu, S. Jiang, X. Qi, Y. R. Yang, S. Wootten, F. Zhang, Z. Zhu, Y. Liu, C. J. Yang, and H. Yan, "Directional Regulation of Enzyme Pathways through the Control of Substrate Channeling on a DNA Origami Scaffold," *Angewandte Chemie - International Edition*, vol. 55, no. 26, pp. 7483–7486, 2016.
- [149] R. E. Kohman, S. S. Cha, H. Y. Man, and X. Han, "Light-Triggered Release of Bioactive Molecules from DNA Nanostructures," *Nano Letters*, vol. 16, no. 4, pp. 2781–2785, 2016.
- [150] S. Sagredo, T. Pirzer, A. Aghebat Rafat, M. A. Goetzfried, G. Moncalian, F. C. Simmel, and F. De La Cruz, "Orthogonal Protein Assembly on DNA Nanostructures Using Relaxases," *Angewandte Chemie - International Edition*, vol. 55, no. 13, pp. 4348–4352, 2016.
- [151] D. Wang, Y. Zhang, M. Wang, Y. Dong, C. Zhou, M. A. Isbell, Z. Yang, H. Liu, and D. Liu, "A switchable DNA origami nanochannel for regulating molecular transport at the nanometer scale," *Nanoscale*, vol. 8, no. 7, pp. 3944–3948, 2016.
- [152] Z. Zhao, J. Fu, S. Dhakal, A. Johnson-Buck, M. Liu, T. Zhang, N. W. Woodbury, Y. Liu, N. G. Walter, and H. Yan, "Nanocaged enzymes with enhanced catalytic activity and increased stability against protease digestion," *Nature Communications*, vol. 7, no. 1, pp. 1–9, 2016.
- [153] P. Ketterer, A. N. Ananth, D. S. Laman Trip, A. Mishra, E. Bertosin, M. Ganji, J. Van Der Torre, P. Onck, H. Dietz, and C. Dekker, "DNA origami scaffold for studying intrinsically disordered proteins of the nuclear pore complex," *Nature Communications*, vol. 9, no. 1, pp. 1–8, 2018.
- [154] E. Pound, J. R. Ashton, H. A. Becerril, and A. T. Woolley, "Polymerase chain reaction based scaffold preparation for the production of thin, branched DNA origami nanostructures of arbitrary sizes," *Nano Letters*, vol. 9, no. 12, pp. 4302–4305, 2009.
- [155] M. Erkelenz, D. M. Bauer, R. Meyer, C. Gatsogiannis, S. Raunser, B. Saccà, and C. M. Niemeyer, "A facile method for preparation of tailored scaffolds for DNA-origami," *Small*, vol. 10, no. 1, pp. 73–77, 2014.
- [156] J. Kozyra, A. Ceccarelli, E. Torelli, A. Lopiccolo, J. Y. Gu, H. Fellermann, U. Stimming, and N. Krasnogor, "Designing Uniquely Addressable Bio-orthogonal Synthetic Scaffolds for DNA and RNA Origami," *ACS Synthetic Biology*, vol. 6, no. 7, pp. 1140–1149, 2017.
- [157] R. Veneziano, T. R. Shepherd, S. Ratanalert, L. Bellou, C. Tao, and M. Bathe, "In vitro synthesis of gene-length single-stranded DNA," *Scientific Reports*, vol. 8, no. 1, pp. 1–7, 2018.
- [158] E. Krieg and W. M. Shih, "Selective Nascent Polymer Catch-and-Release Enables Scalable Isolation of Multi-Kilobase Single-Stranded DNA," *Angewandte Chemie - International Edition*, vol. 57, no. 3, pp. 714–718, 2018.

- [159] P. M. Nafisi, T. Aksel, and S. M. Douglas, "Construction of a novel phagemid to produce custom DNA origami scaffolds," *Synthetic Biology*, vol. 3, no. 1, pp. 1–8, 2018.
- [160] T. R. Shepherd, R. R. Du, H. Huang, E. C. Wamhoff, and M. Bathe, "Bioproduction of pure, kilobase-scale single-stranded DNA," *Scientific Reports*, vol. 9, no. 1, pp. 1–9, 2019.
- [161] F. A. Engelhardt, F. Praetorius, C. H. Wachauf, G. Brüggenthies, F. Kohler, B. Kick, K. L. Kadletz, P. N. Pham, K. L. Behler, T. Gerling, and H. Dietz, "Custom-Size, Functional, and Durable DNA Origami with Design-Specific Scaffolds," *ACS Nano*, vol. 13, no. 5, pp. 5015–5027, 2019.
- [162] F. Praetorius, B. Kick, K. L. Behler, M. N. Honemann, D. Weuster-Botz, and H. Dietz, "Biotechnological mass production of DNA origami," *Nature*, vol. 552, no. 7683, pp. 84–87, 2017.
- [163] F. Schueder, J. Stein, F. Stehr, A. Auer, B. Sperl, M. T. Strauss, P. Schwille, and R. Jungmann, "An order of magnitude faster DNA-PAINT imaging by optimized sequence design and buffer conditions," *Nature Methods*, vol. 16, no. 11, pp. 1101–1104, 2019.
- [164] S. Strauss and R. Jungmann, "Up to 100-fold speed-up and multiplexing in optimized DNA-PAINT," *Nature Methods*, vol. 17, no. 8, pp. 789–791, 2020.
- [165] D. Han, S. Pal, Y. Yang, S. Jiang, J. Nangreave, Y. Liu, and H. Yan, "DNA gridiron nanostructures based on four-arm junctions," *Science*, vol. 339, no. 6126, pp. 1412–1415, 2013.
- [166] F. Zhang, S. Jiang, S. Wu, Y. Li, C. Mao, Y. Liu, and H. Yan, "Complex wireframe DNA origami nanostructures with multi-arm junction vertices," *Nature Nanotechnology*, vol. 10, no. 9, pp. 779–784, 2015.
- [167] E. Benson, A. Mohammed, J. Gardell, S. Masich, E. Czeizler, P. Orponen, and B. Högborg, "DNA rendering of polyhedral meshes at the nanoscale," *Nature*, vol. 523, no. 7561, pp. 441–444, 2015.
- [168] M. Matthies, N. P. Agarwal, and T. L. Schmidt, "Design and Synthesis of Triangulated DNA Origami Trusses," *Nano Letters*, vol. 16, no. 3, pp. 2108–2113, 2016.
- [169] R. Veneziano, S. Ratanalert, K. Zhang, F. Zhang, H. Yan, W. Chiu, and M. Bathe, "Designer nanoscale DNA assemblies programmed from the top down," *Science*, vol. 352, no. 6293, pp. 1534–1534, 2016.
- [170] H. Jun, T. R. Shepherd, K. Zhang, W. P. Bricker, S. Li, W. Chiu, and M. Bathe, "Automated sequence design of 3D polyhedral wireframe DNA origami with honeycomb edges," *ACS Nano*, vol. 13, no. 2, pp. 2083–2093, 2019.
- [171] E. C. Wamhoff, J. L. Banal, W. P. Bricker, T. R. Shepherd, M. F. Parsons, R. Veneziano, M. B. Stone, H. Jun, X. Wang, and M. Bathe, "Programming Structured DNA Assemblies to Probe Biophysical Processes," *Annual Review of Biophysics*, vol. 48, pp. 395–419, 2019.
- [172] J. Stein, F. Stehr, J. Bauer, C. Niederauer, R. Jungmann, K. Ganzinger, and P. Schwille, "Tracking Single Particles for Hours via Continuous DNA-mediated Fluorophore Exchange," *bioRxiv*, 2020.
- [173] K. K. H. Chung, Z. Zhang, P. Kidd, Y. Zhang, N. D. Williams, B. Rollins, Y. Yang, C. Lin, D. Baddeley, and J. Bewersdorf, "Fluorogenic probe for fast 3D whole-cell DNA-PAINT," *bioRxiv*, 2020.
- [174] P. Ketterer, E. M. Willner, and H. Dietz, "Nanoscale rotary apparatus formed from tight-fitting 3D DNA components," *Science Advances*, vol. 2, no. 2, p. e1501209, 2016.

Danksagung

Ich bedanke mich herzlich bei Prof. Dr. Hendrik Dietz für die Möglichkeit, dieses Forschungsthema bearbeiten zu dürfen, für die verständnisvolle Betreuung sowie für die angenehme und offene Arbeitsatmosphäre.

Den Mitgliedern der AG Dietz und AG Simmel – nicht zuletzt meinen verschiedenen Büro-Mitbewohner*innen – sei für die kollegiale und produktive Arbeitsatmosphäre gedankt, für die wissenschaftlichen Diskussionen und dafür, dass auch die zwischenmenschliche Komponente nicht zu kurz kam.

Prof. Dr. Martin Zacharias (Lehrstuhl für Biomolekulare Dynamik, Physik Dept., TUM) danke ich für die angenehme und erfolgreiche Kooperation.

Desweiteren möchte ich allen Student*innen danken, die im Rahmen von Abschlussarbeiten, Praktika oder HiWi-Tätigkeiten zu diesem Forschungsprojekt beigetragen haben.

Danke an Elke Fehsenfeld, Erika Bischofs, Helene Budjarek, Sonja Kraus, Susanne Kinzel, Daniela Härtwig und Dr. Petra Bele für ihre Hilfe in technischen und administrativen Angelegenheiten.

Aufmerksame Korrekturleser*innen sind goldwert. Herzlichen Dank an Flo, Jonas, Evi, Tom G, Diana, Pierre, Johanna, Elija, Christian und Jessica.

Besonders großer Dank gebührt meinen Eltern für ihre bedingungslose Unterstützung in allen Lebenslagen. Außerdem möchte ich mich bei meinen Geschwistern sowie bei allen Freunden und Freundinnen bedanken, die mir in den letzten paar Jahren Rückhalt gegeben und/oder für den notwendigen Ausgleich zur Arbeit gesorgt haben.

Liebe Diana,

ich bin unfassbar dankbar, dass Du mich auf diesem Weg durch alle Höhen und Tiefen begleitet hast. Dafür und für noch so viel mehr danke ich Dir von ganzem Herzen!