

# Molekulare Erkennung am Beispiel der DNA-Hybridisierung

Stabilität von DNA-Duplexen mit einzelsträngigen Schleifen-Strukturen,  
Spezifität der molekularen Erkennung in DNA-Konkurrenzsystemen

Dissertation  
zur Erlangung des Grades  
des Doktors der Naturwissenschaften  
der Naturwissenschaftlich-Technischen Fakultät II  
- Physik und Mechatronik -  
der Universität des Saarlandes

von

Christian Trapp

Saarbrücken

2013

Tag des Kolloquiums: 21.11.2013

Dekan: Univ.-Prof. Dr. Christian Wagner

Mitglieder des  
Prüfungsausschusses: Univ.-Prof. Dr. Albrecht Ott  
Univ.-Prof. Dr. Karin Jacobs  
Univ.-Prof. Dr. Joachim Rädler  
Dr. Matthias Brust

---

## Kurzzusammenfassung

Die spezifische molekulare Erkennung zwischen Bindungspartnern ist für alle biologische Systeme von größter Wichtigkeit. Jedoch ist die zugrundeliegende Physik nur teilweise verstanden. Ein Beispiel einer molekularen Erkennung ist die DNA-Hybridisierung, unter der man die Formation einer DNA-Doppelhelix aus zwei komplementären DNA-Einzelsträngen versteht. Die vorliegende Arbeit beschäftigt sich experimentell und theoretisch mit zwei Aspekten der DNA-Hybridisierung im thermodynamischen Gleichgewicht: i) Das Auftreten nicht-kanonischer Bindungskonformationen bei der Hybridisierung von teilweise komplementären DNA-Strängen macht den Prozess der molekularen Erkennung deutlich komplexer. Daher wird zum einen der Einfluss einzelsträngiger Schleifenstrukturen (*Loops*), als Beispiel einer nicht-kanonischen Bindungskonfiguration, untersucht, die bei Hybridisierung von DNA-Strängen ungleicher Länge entstehen können. ii) Zum anderen wird die Spezifität der molekularen Erkennung in Konkurrenz untersucht. Gerade in einer Konkurrenzsituation zwischen einer Vielzahl von ähnlichen Molekülen muss die molekulare Erkennung äußerst spezifisch sein, damit mit hoher Wahrscheinlichkeit nur Bindungen zwischen passenden Bindungspartnern auftreten. Es wurde gezeigt, dass die Untersuchung dieser Aspekte ein tiefgreifenderes Verständnis der DNA-Hybridisierung ermöglicht und sich dadurch ein umfassenderes Bild der molekularen Erkennung ergibt.



---

## Abstract

Molecular recognition between matching macromolecules is central to the biological organism. However, to date, the underlying physical mechanisms are poorly understood. An example of molecular recognition is the process of DNA hybridization, that describes the formation of a stable DNA double-helix from two complementary single strands. The present thesis focuses, experimentally and theoretically, on two aspects of DNA hybridization in thermal equilibrium. i) Non-canonical binding conformations, emerging upon hybridization of partly complementary DNA strands, may complicate the process of molecular recognition. As an example of a non-canonical binding conformation, the influence of bulged loops that can occur upon hybridization of DNA strands of unequal lengths is investigated. ii) The second part deals with the investigation of the specificity of molecular recognition in competition. In a competitive environment between many similar molecules the process of molecular recognition needs to be highly specific, so that only matching binding partners form stable bonds with a high probability. We have shown that the investigation of these two aspects helps to gain a deeper understanding of DNA hybridization and of molecular recognition in general.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Grundlagen</b>	<b>5</b>
2.1	DNA . . . . .	5
2.1.1	Aufbau und Struktur der DNA . . . . .	5
2.2	Biologische Funktion der DNA . . . . .	7
2.2.1	Das zentrale Dogma der Molekularbiologie . . . . .	7
2.2.2	Proteinbiosynthese . . . . .	8
2.3	DNA-Hybridisierung . . . . .	11
2.3.1	Kinetik der DNA-Hybridisierung . . . . .	11
2.3.2	Schmelztemperatur . . . . .	14
2.3.3	Das Nearest-Neighbor-Modell . . . . .	15
2.4	Polymerase-Kettenreaktion . . . . .	18
2.5	DNA-Microarrays . . . . .	20
2.5.1	Anwendungen . . . . .	22
<b>3</b>	<b>Experimentelle Methoden und Material</b>	<b>23</b>
3.1	DNA-Microarrays . . . . .	23
3.1.1	Lichtgesteuerte in situ-Synthese . . . . .	23
3.1.2	Aufbau des Synthesizers . . . . .	26
3.1.3	Datenauswertung . . . . .	30
3.2	Förster-Resonanz-Energietransfer . . . . .	31
3.3	Fluoreszenzkorrelationsspektroskopie . . . . .	34
<b>4</b>	<b>Stabilität von DNA-Duplexen mit einzelsträngigen Loop-Strukturen</b>	<b>39</b>
4.1	Loops in der Probe . . . . .	40
4.1.1	Experiment . . . . .	40
4.1.2	Ergebnisse und Diskussion . . . . .	42
4.2	Ergänzende Untersuchungen zum Einfluss der Loops in der Probe . . . . .	46
4.2.1	Einfluss der Oberfläche . . . . .	46
4.2.2	Loops anderer Sequenzen . . . . .	46

4.2.3	Vorversuch zur theoretischen Beschreibung in Abschnitt 6.1 . . .	48
4.3	Grenze der Hybridisierung - Loops im Target . . . . .	50
4.3.1	Experiment . . . . .	50
4.3.2	Ergebnisse und Diskussion . . . . .	51
4.4	Ergänzende Untersuchungen zur Grenze der Hybridisierung . . . . .	55
4.4.1	Einfluss der Synthesefehler . . . . .	55
4.4.2	Einfluss des Fluoreszenz-Markers . . . . .	58
4.4.3	Einfluss von Linkern . . . . .	61
4.5	Probes mit zwei Loops . . . . .	62
4.5.1	Vorversuch . . . . .	63
4.5.2	Einfluss des Abstandes der beiden Loops . . . . .	65
4.5.3	Theoretische Überlegungen . . . . .	69
4.5.4	FCS-Untersuchungen . . . . .	73
4.6	Zusammenfassung der Ergebnisse . . . . .	77
4.6.1	Loops in der Probe . . . . .	77
4.6.2	Loops im Target . . . . .	77
4.6.3	Probes mit zwei Loops . . . . .	78
<b>5</b>	<b>DNA-Hybridisierung in Konkurrenz</b>	<b>79</b>
5.1	Effekt der hochspezifischen DNA-Hybridisierung . . . . .	80
5.1.1	Experiment . . . . .	80
5.1.2	Ergebnisse und Diskussion . . . . .	80
5.2	Ausschluss von Artefakten . . . . .	84
5.2.1	Nicht-Gleichgewicht . . . . .	84
5.2.2	Messungen in Lösung . . . . .	87
5.2.3	Wahl des Sequenz-Motivs . . . . .	88
5.3	Charakterisierung des Effekts der hochspezifischen Hybridisierung . . .	90
5.3.1	Unterteilung in Systeme . . . . .	90
5.3.2	Schmelztemperaturen . . . . .	92
5.3.3	FCS-Untersuchungen . . . . .	95
5.3.4	Positionsabhängigkeit . . . . .	97
5.4	Erklärungsansatz . . . . .	101
5.5	Ordnungsverhalten . . . . .	103
5.5.1	Vorversuch: 4-Komponenten-System . . . . .	103
5.5.2	49-Komponenten-System . . . . .	105
5.6	Zusammenfassung der Ergebnisse . . . . .	113
5.6.1	Effekt der hochspezifischen Hybridisierung . . . . .	113
5.6.2	Ausschluss von Artefakten . . . . .	113

5.6.3	Charakterisierung des Effekts der hochspezifischen Hybridisierung	113
5.6.4	Erklärungsansatz . . . . .	114
5.6.5	Ordnungsverhalten . . . . .	114
<b>6</b>	<b>Theoretische Arbeiten</b>	<b>115</b>
6.1	Stabilität von DNA-Duplexen mit einzelsträngigen Loop-Strukturen . .	115
6.1.1	Die Zustandssumme des Zipper-Modells . . . . .	116
6.1.2	Die Zustandssumme des erweiterten Zipper-Modells . . . . .	120
6.2	Hochspezifische DNA-Hybridisierung in Konkurrenz . . . . .	131
6.2.1	Freie Energie einer <i>pre-melted</i> -Konfiguration . . . . .	133
6.2.2	Freie Energie von koexistierenden <i>pre-melted</i> - und Helix-Konfigurationen . . . . .	134
6.2.3	Einzelhybridisierung . . . . .	136
6.2.4	Konkurrenzhybridisierung ohne Wechselwirkung: standardspezifische Systeme . . . . .	137
6.2.5	Konkurrenzhybridisierung mit Wechselwirkung: hochspezifische Systeme . . . . .	140
6.2.6	Helmholtz Energie der Oberfläche . . . . .	142
<b>7</b>	<b>Zusammenfassung und Ausblick</b>	<b>149</b>
	<b>Literaturverzeichnis</b>	<b>155</b>
	<b>Anhang A</b>	<b>165</b>
A.1	Loops anderer Sequenzen . . . . .	167
A.2	Probes mit zwei Loops: FCS-Rohdaten . . . . .	168
A.2.1	Messungen bei 44 °C . . . . .	168
A.2.2	Messungen bei 55 °C . . . . .	168
A.3	DNA-Konkurrenzhybridisierung: FCS-Rohdaten und Fits . . . . .	169
A.3.1	Messungen bei Raumtemperatur . . . . .	169
A.3.2	Messungen bei 44 °C . . . . .	170
	<b>Anhang B</b>	<b>171</b>
B.1	Verwendete Target-Sequenzen . . . . .	173
B.2	Hybridisierungs-Puffer . . . . .	175
B.3	Protokoll für die Herstellung der Dendrimer-Oberflächen . . . . .	176
B.4	Optischer Aufbau des Synthesizers . . . . .	178



# 1 Einleitung

Die Steuerung eines Organismus erfolgt durch ein komplexes Zusammenspiel verschiedenster Zelltypen, Stoffwechselforgänge und Moleküle. Um die Funktionalität des Organismus zu gewährleisten und sicherzustellen, dass der Informationsfluss im System nicht gestört wird, ist es entscheidend, dass Bindungen auch bei Anwesenheit einer Vielzahl von Konkurrenten mit einer hohen Wahrscheinlichkeit nur zwischen passenden Bindungspartnern stattfinden. Die spezifische molekulare Erkennung zwischen Bindungspartnern ist daher für alle biologischen Systeme von größter Wichtigkeit und zentrales Thema der vorliegenden Arbeit.

Ein Beispiel einer Informationsübertragung anhand molekularer Erkennung in einem biologischen System ist die Proteinbiosynthese. Sie beschreibt die Herstellung eines funktionalen Proteins auf Basis einer genetischen Information. Die Synthese hängt dabei empfindlich von der korrekten molekularen Erkennung ab, da der Einbau von falschen Bausteinen in das Protein vermieden werden muss, um die gewünschte Funktion nicht zu beeinträchtigen. Andere Beispiele für molekulare Erkennung sind unter anderem die Wechselwirkung zwischen Antikörpern und Antigenen, die Bindung von Transkriptionsfaktoren an die DNA oder die DNA-Replikation bei Zellteilungen.

Im Rahmen dieser Arbeit wird als weiteres Beispiel für molekulare Erkennung die Hybridisierung von DNA im thermodynamischen Gleichgewicht untersucht. Unter Hybridisierung versteht man im Allgemeinen die Bindung zweier komplementärer Nukleinsäuren zu einer stabilen Doppelhelixstruktur. Die Stabilität der Doppelhelix wird dabei durch Wasserstoffbrückenbindungen zwischen komplementären Basen der beiden Nukleinsäuren (Watson-Crick-Paarung) und Stapelwechselwirkungen zwischen benachbarten Basen generiert. Die Hybridisierung ist äußerst spezifisch: Selbst eine einzelne Fehlstelle in der Basenabfolge hat Auswirkungen auf die Stabilität der DNA-Doppelhelix. *In vivo* spielt die Hybridisierung unter anderem eine wichtige Rolle in der bereits erwähnten DNA-Replikation bei Zellteilungen. Auch hier ist eine hohe Spezifität gefordert, da selbst eine einzelne falsche Base im Genom negative Auswirkungen auf die Proteinbiosynthese der entsprechenden Zelle haben kann.

Diese Arbeit beschäftigt sich experimentell und theoretisch mit zwei Aspekten der DNA-Hybridisierung. Zum einen wird die Stabilität von DNA-Duplexen mit einzel-

strängigen Schleifenstrukturen, sog. *Loops*, untersucht. Durch die Ausbildung von *Loops* können auch nur teilweise komplementäre DNA-Stränge stabile nicht-kanonische Bindungskonformationen eingehen. Der Prozess der molekularen Erkennung kann dadurch deutlich komplexer werden. Der Einfluss von *Loops* auf die Hybridisierung wird untersucht, indem *Loop*-Position und -Länge durch Einfügen zusätzlicher Basensequenzen in die beteiligten DNA-Stränge systematisch variiert werden. Dabei zeigt sich eine hohe Stabilität der DNA-Duplexe mit *Loop*-Strukturen. Das entwickelte Modell zur theoretischen Beschreibung des Einflusses von *Loop*-Strukturen kann die experimentellen Beobachtungen quantitativ wiedergeben.

Der zweite Teil dieser Arbeit untersucht die DNA-Hybridisierung in Konkurrenz zwischen verschiedenen DNA-Spezies um eine begrenzte Anzahl von Bindungsstellen. Die molekulare Erkennung muss gerade in einer Konkurrenzsituation, in der viele Moleküle mit ähnlichen Eigenschaften um Bindungsstellen konkurrieren, äußerst spezifisch sein. Die Spezifität der Bindung in Konkurrenz wird untersucht, indem die Hybridisierung einer DNA-Spezies mit konstanter Konzentration als Funktion der Konzentrationen der Konkurrenz-Spezies gemessen wird. Es zeigt sich, dass die Spezifität der Bindung unter geeigneten Bedingungen um mehrere Größenordnungen höher ist als erwartet und sich die Eigenschaften eines Konkurrenzsystems, anders als allgemein angenommen, nicht immer aus den Eigenschaften der jeweiligen Einzelsysteme ergeben. Auch diese Ergebnisse können mit einem erarbeiteten Modell zur DNA-Konkurrenzhybridisierung quantitativ reproduziert werden.

Zur Untersuchung der Eigenschaften der Hybridisierung werden in dieser Arbeit hauptsächlich DNA-Microarrays verwendet. Ein großer Vorteil der DNA-Microarray-Technologie besteht in der Möglichkeit, hochgradig parallel zu messen und so wesentlich effizienter und kostengünstiger arbeiten zu können. DNA-Microarrays bestehen aus einer Vielzahl regelmäßig angeordneter *Features*. Als *Features* bezeichnet man kleine abgegrenzte Bereiche auf der Microarray-Oberfläche, die ihrerseits immobilisierte DNA-Einzelstränge (*Probes*) einer bekannten Sequenz enthalten. Wird ein DNA-Microarray durch Zugabe einer zu analysierenden Mischung aus fluoreszenzmarkierten einzelsträngigen DNA-Spezies (*Targets*) hybridisiert, entstehen an den *Features* durch Akkumulation von komplementären *Targets* messbare Fluoreszenzsignale. Da Position und Sequenz der entsprechenden *Features* bekannt sind, eignen sich DNA-Microarrays bestens, um Rückschlüsse auf die Spezifität der molekularen Erkennung zwischen *Probes* und *Targets* im Falle der DNA-Hybridisierung zu ziehen.

Ein erfolgreiches Modell zur Beschreibung der DNA-Hybridisierung in Lösung ist das *Nearest-Neighbor*-Modell (NN-Modell), das die Stabilisierung der DNA-Doppelhelix aufgrund von Wasserstoffbrücken zwischen komplementären Basen und Stapelwechselwirkungen berücksichtigt. Die Hybridisierung auf DNA-Microarrays unterscheidet sich

---

jedoch aufgrund von oberflächen- und herstellungsbedingten Effekten erheblich von der in Lösung. Naiser *et al.* konnte allerdings zeigen, dass der Einfluss von Einzelbasenfehlern auf die Hybridisierung auf Grundlage des *double-ended zipper*-Modells (unter Verwendung der NN-Energien) auch auf DNA-Microarrays quantitativ wiedergegeben werden kann [1–3].

Neben dem generellen Interesse, die DNA-Hybridisierung aufgrund der oben genannten *in vivo*-Beispiele besser zu verstehen, ist ein besseres Verständnis der Hybridisierung auch im Hinblick auf biotechnologische Anwendungen, wie der Polymerasekettenreaktion oder DNA-Hochdurchsatzverfahren wie DNA-Microarrays, von großer Bedeutung. So kann in vielen Microarray-Experimenten die Längenverteilung der eingesetzten Targets nur schwer kontrolliert werden. Dadurch können sich unterschiedliche Längen der Probes auf der Oberfläche und der Targets in Lösung ergeben, wodurch es auch hier zu nicht-kanonischen Bindungskonformationen zwischen Probe und Target, wie beispielsweise Loops, kommen kann.

Im Falle der Genexpressionsanalyse mit Hilfe von DNA-Microarrays ist die Spezifität der Bindungen in Konkurrenz entscheidend, da hier viele Targets um Bindungsstellen auf der Oberfläche konkurrieren. Um die Genexpression zweier biologischer Proben miteinander zu vergleichen, wird die für eine bestimmte Genaktivität charakteristische mRNA (von engl. *messenger RNA*) der jeweiligen Probe isoliert und nach Fluoreszenzmarkierung auf ein Microarray gebracht. Durch Analyse der Fluoreszenzsignale der verschiedenen Target-Mischungen erhält man „Fingerabdrücke“ der biologischen Proben, anhand derer man eventuell auftretende Unterschiede in der genetischen Aktivität nachweisen und so Zusammenhänge zwischen verschiedenen Genen aufdecken kann. Die Qualität und Interpretation der Microarray-Daten wird maßgeblich durch die Spezifität der Bindung zwischen Probes und Targets bestimmt. Jedoch basiert die Datenanalyse im Falle von DNA-Microarrays immer noch weitestgehend auf empirischen, statistischen Methoden. Weiterhin fehlen allgemein anerkannte Standards zum Vergleich von Microarray-Daten von verschiedenen Experimenten oder Laboren [4–8]. Beide Aspekte zeigen, dass ein tiefergehendes Verständnis der DNA-Hybridisierung auch wünschenswert ist, um das volle Potential von biotechnologischen Anwendungen ausschöpfen zu können.

Die vorliegende Arbeit gliedert sich in insgesamt sieben Kapitel. In Kapitel 2 werden die biologischen und physikalischen Grundlagen wie z.B. der Aufbau der DNA oder die allgemeine Funktionsweise von DNA-Microarrays vorgestellt, die zum Verständnis dieser Arbeit notwendig sind. Die verwendeten experimentellen Aufbauten und Methoden werden in Kapitel 3 behandelt. Kapitel 4 und 5 widmen sich der experimentellen Untersuchung von DNA-Duplexen mit einzelsträngigen Loop-Strukturen bzw. der DNA-Hybridisierung in Konkurrenz. Im Anschluss an die experimentellen Arbeiten werden

in Kapitel 6 Gleichgewichtsmodelle vorgestellt, mit denen sich die experimentellen Ergebnisse der Kapitel 4 und 5 quantitativ reproduzieren lassen. Den Abschluss dieser Arbeit bildet Kapitel 7, das die erzielten Ergebnisse noch einmal zusammenfasst und Ausblick auf zukünftige Arbeiten auf diesem Gebiet gibt.

## 2 Grundlagen

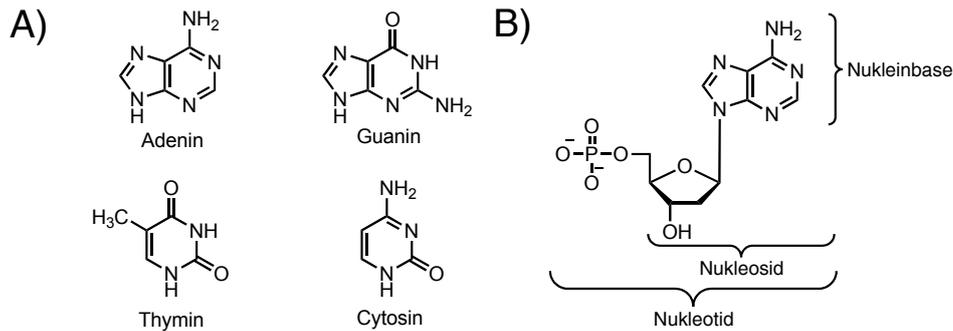
### 2.1 DNA

Die Desoxyribonukleinsäure, kurz DNA (von engl. *deoxyribonucleic acid*), ist Träger der genetischen Erbinformation und grundlegender Bestandteil aller Lebewesen und der meisten Viren. Zusammen mit RNA (von engl. *ribonucleic acid*) und Proteinen ist die DNA eines der Makromoleküle, die für alle bekannten Lebensformen essentiell ist. Die genetische Information ist anhand einer Abfolge von Nukleinbasen der DNA gespeichert und wird beispielsweise bei Zellteilungen durch den Prozess der DNA-Replikation an die Tochterzellen weitergegeben. Die DNA enthält als einziges Molekül Anweisungen für ihre eigene Replikation [9].

#### 2.1.1 Aufbau und Struktur der DNA

Die molekulare Zusammensetzung der DNA wurde 1919 von P.A. Levene aufgeklärt [10]. Demnach ist die DNA aus einzelnen Bausteinen, den Nukleotiden, aufgebaut. Nukleotide bestehen aus Phosphat, dem Zucker Desoxyribose und einer heterozyklischen Nukleinbase (vereinfacht auch Base genannt). Die Phosphat- und Zuckeruntereinheiten sind in jedem Nukleotid gleich und bilden das Rückgrat eines Strangs der doppelhelikalen DNA. Die Verbindung aus Zucker und Base nennt man Nukleosid (siehe Abb. 2.1 B)). Bei den in der DNA vorkommenden Basen handelt es sich um Adenin, Guanin, Thymin und Cytosin. Die beiden erstgenannten gehören zur Familie der Purine, die letzteren zur Familie der Pyrimidine (siehe Abb. 2.1 A)). Da sich die verschiedenen Nukleotide nur in ihren Basen unterscheiden, werden für sie auch die Abkürzungen A, G, T und C verwendet.

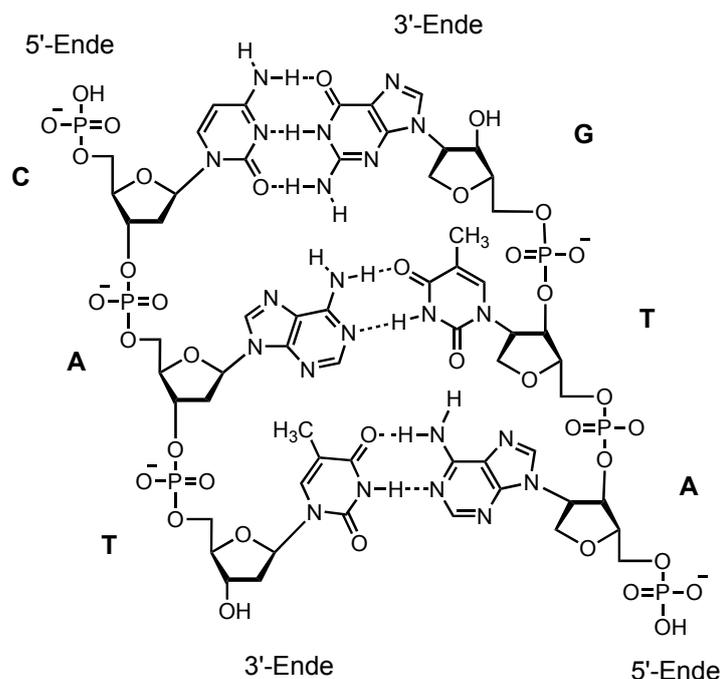
Die fünf Kohlenstoffatome der Desoxyribose sind von 1' bis 5' nummeriert. Die Basen sind am 1'-Ende gebunden, der Phosphatrest am 5'-Ende. Am 3'-Ende befindet sich eine OH-Gruppe, die mit der entsprechenden Phosphatgruppe am 5'-Ende einer weiteren Desoxyribose eine Phosphordiesterbindung eingehen kann. Dadurch besitzen DNA-Stränge immer ein 5'- und ein 3'-Ende. Da DNA-Polymerasen neue Nukleotide



**Abbildung 2.1:** Bausteine der DNA. **A)** Gezeigt sind die Strukturformeln der vier in der DNA vorkommenden Nukleinbasen Adenin, Guanin, Thymin und Cytosin. **B)** Aufbau eines Nukleotids. Die Verbindung aus Nukleobase (hier Adenin) und Zucker wird Nukleosid genannt. Bindet der Zucker zusätzlich noch einen Phosphatrest, so ergibt sich das entsprechende Nukleotid.

nur an die OH-Gruppe am 3'-Ende anfügen können, wächst ein DNA-Einzelstrang *in vivo* immer vom 5'- zum 3'-Ende.

Die Verbindung zweier DNA-Einzelstränge resultiert in der bekannten Doppelhelix-Struktur der DNA, die 1953 von Watson und Crick aufgeklärt wurde [11]. Während einzelsträngige DNA sehr flexibel ist und die Persistenzlänge je nach vorherrschenden Bedingungen zwischen 0.8 und 3 nm liegt [12–14], so ist doppelhelikale DNA mit einer Persistenzlänge von 40 bis 50 nm wesentlich steifer [15]. Die Anlagerung der beiden Einzelstränge geschieht in entgegengesetzter Richtung, d.h. ein Ende der Doppelhelix besteht aus dem 5'-Ende des einen Strangs und dem 3'-Ende des anderen. Die Nukleobasen der beiden komplementären Einzelstränge bilden bei Formation einer Doppelhelix Wasserstoffbrückenbindungen aus. Es verbindet sich immer die komplementären Basen Adenin mit Thymin (AT bzw. TA) und Cytosin mit Guanin (CG bzw. GC). Dies wird als Watson-Crick-Paarung bezeichnet. Da Adenin und Thymin nur über zwei Wasserstoffbrücken, Guanin und Cytosin aber über drei Wasserstoffbrücken miteinander verbunden sind, ist die Bindung zwischen Guanin und Cytosin stabiler. Zur weiteren Stabilisierung tragen neben den Wasserstoffbrückenbindungen noch Stapelwechselwirkungen ( $\pi$ -stacking) zwischen aufeinanderfolgenden Nukleobasen bei [16]. Diese sind am günstigsten für gestapelte GC-GC Basen, weniger günstig für AT-AT Basen [17]. Daher wird zur Kennzeichnung der thermodynamischen Stabilität eines DNA-Fragments oft der CG-Anteil in Prozent angegeben. Abb. 2.2 zeigt den Aufbau einer DNA-Doppelhelix mit den entsprechenden Bindungen. Für weiterführende Literatur sei an dieser Stelle beispielsweise auf [18] verwiesen.



**Abbildung 2.2:** Chemische Struktur eines DNA-DNA-Duplexes. Zur Bildung eines DNA-Duplexes lagern sich die beiden Einzelstränge in entgegengesetzter Richtung an (hier 5'-CAT-3' mit 3'-GTA-5'). Die Nukleinbasen der beiden komplementären Einzelstränge sind über Wasserstoffbrückenbindungen miteinander verbunden (gestrichelte Linien). Bei Adenin und Thymin sind es zwei, bei Guanin und Cytosin drei Wasserstoffbrücken. Zusätzlich üben die Nukleotide eines Strangs noch Stapelwechselwirkungen auf benachbarte Nukleotide aus, die zur Stabilisierung des Duplexes beitragen. Das Rückgrat des jeweiligen Einzelstrangs besteht aus Phosphat und Zucker.

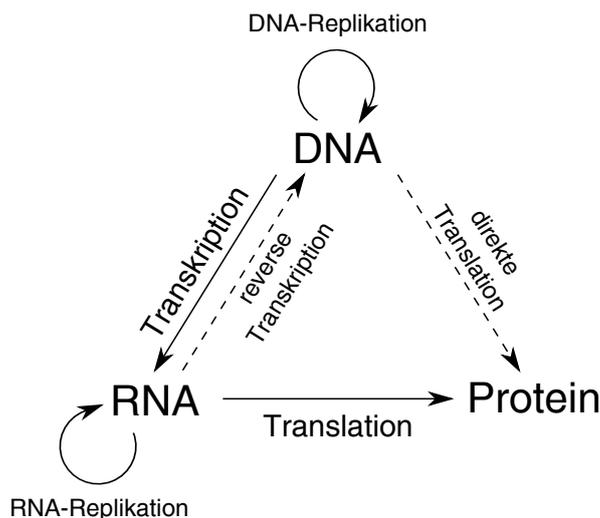
## 2.2 Biologische Funktion der DNA

### 2.2.1 Das zentrale Dogma der Molekularbiologie

Das zentrale Dogma der Molekularbiologie beschreibt den Fluss genetischer Information in einem biologischen System. Der Begriff wurde 1958 von F. Crick geprägt und 1970 erneut aufgegriffen [20, 21] (siehe Abb. 2.3):

„The central dogma of molecular biology deals with the detailed residue-by-residue transfer of sequential information. It states that such information cannot be transferred from protein to either protein or nucleic acid.“

Dies bedeutet, dass eine sequenzielle genetische Information sobald sie in ein Protein übersetzt wurde nicht an ein anderes Protein oder eine Nukleinsäure weitergegeben werden kann. Ein Rückfluss der Information ins Genom findet nicht statt. Das sog. „Spiegelman-Monster“, als Beispiel eines *in vitro*-Systems, zeigt allerdings, dass Informationen von RNA zu RNA weitergegeben werden kann [22]. Jedoch gilt das zentrale



**Abbildung 2.3:** Das zentrale Dogma der Molekularbiologie nach F. Crick. Es beschreibt den Informationsfluss innerhalb eines biologischen Systems. Die durchgezogenen Pfeile beschreiben allgemeine Mechanismen, die gestrichelten Pfeile spezielle Formen der Informationsübertragung. Beispielsweise wird bei der Proteinbiosynthese die in der DNA enthaltene Information zuerst in mRNA transkribiert und dann ausgehend von dieser mRNA in eine Polypeptidkette translatiert, die sich letztendlich zu einem Protein falten kann. Die DNA-Replikation spielt für die Zellteilung eine wichtige Rolle. Ein Beispiel für eine spezielle Informationsübertragung ist die reverse Transkription, die bei der Reproduktion von Retro-Viren auftritt (siehe z.B. [19]).

Dogma heute trotz kritischer Stimmen als einer der Grundpfeiler der Molekularbiologie [23].

## 2.2.2 Proteinbiosynthese

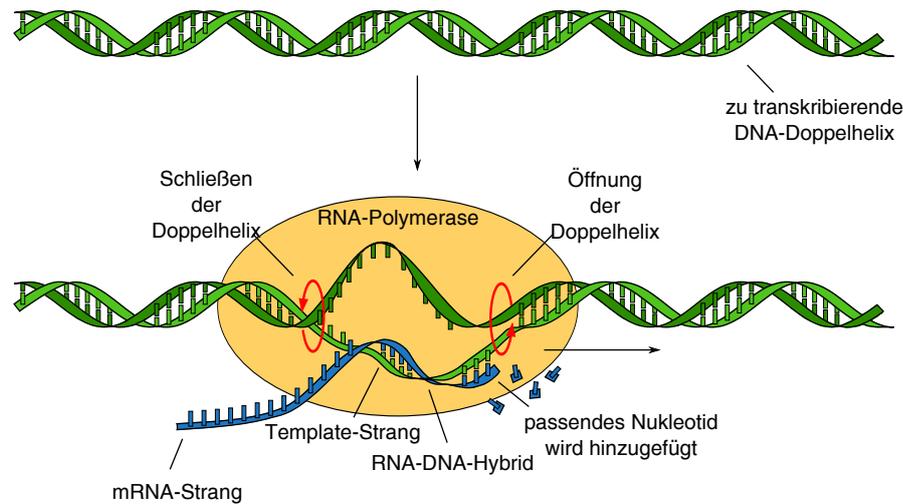
Unter Proteinbiosynthese versteht man die Synthese eines Proteins anhand der in der DNA gespeicherten Information an den Ribosomen lebender Zellen. Die kleinste Einheit dieser biologischen Erbinformation wird als Gen bezeichnet. Ein Gen ist eine Sequenz aus Nukleotiden der DNA, die den Bauplan des herzustellenden Proteins codiert. Proteine sind essentiell für die Steuerung und die Struktur einer Zelle bzw. des Organismus. Zum jetzigen Zeitpunkt konnten 80 % des Genoms biochemischen Funktionen im Organismus zugeordnet werden [24, 25].

Der Ablauf der Proteinbiosynthese lässt sich in die Prozesse Transkription und Translation unterteilen (siehe Abb. 2.4):

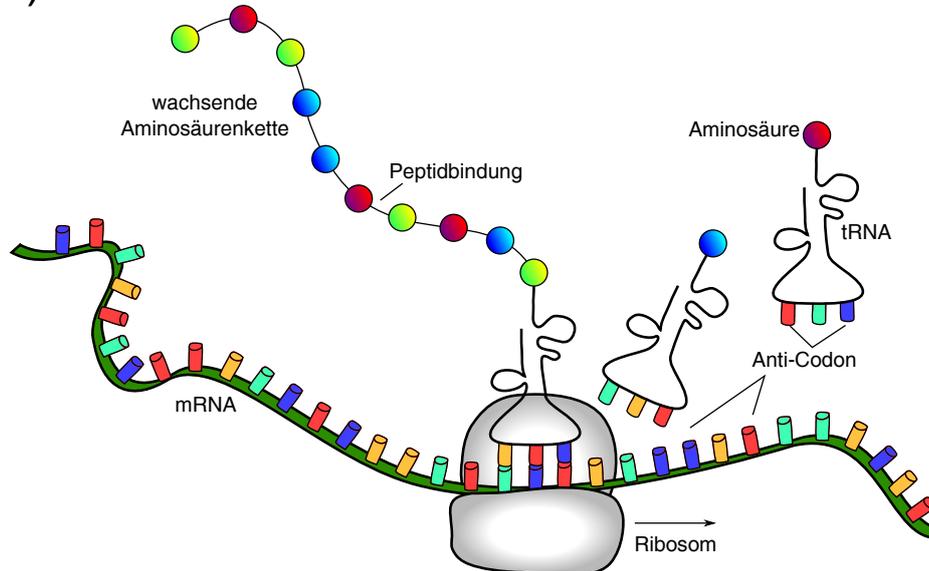
- **Transkription:**

Nach Bindung des Enzyms RNA-Polymerase an eine sog. Promoter-Region der DNA wird die DNA auf einer Länge von ca. 10-20 Basen entspiralisiert. Einer der beiden Einzelstränge der DNA dient als Vorlage für die Paarung komple-

## A) Transkription



## B) Translation



**Abbildung 2.4:** Prozesse der Proteinbiosynthese. **A)** Transkription: Die zu transkribierende Region der DNA-Doppelhelix wird durch das Enzym RNA-Polymerase auf einer Länge von ca. 10-20 Basen entspiralisiert. Die freigelegten Basen eines der beiden Stränge dient als Vorlage für die Synthese eines komplementären RNA-Einzelstrangs. Nach Initialisierung der Elongation der mRNA rückt die RNA-Polymerase schrittweise vor und baut sukzessive komplementäre Ribonukleotide in den wachsenden mRNA-Strang ein. Der entstandene mRNA-Strang wird nach Ende der Transkription freigegeben. **B)** Translation: Der im vorherigen Schritt hergestellte mRNA-Strang dient nun als Vorlage zur Herstellung eines Proteins an den Ribosomen. Eine Aminosäure eines Proteins wird durch drei aufeinanderfolgende Ribonukleotide (Codon) der mRNA codiert. Durch die mit dem passenden Anti-Codon versehene tRNA wird die entsprechende Aminosäure zu den Ribosomen transportiert. Die Aminosäuren zweier tRNAs werden durch eine Peptidbindung miteinander verknüpft. So entsteht durch das schrittweise Vorrücken des Ribosoms auf der mRNA eine wachsende Kette aus Aminosäuren, die sich nach Ende der Translation in ein Protein falten kann.

mentärer Ribonukleotide. Ribonukleotide bestehen im Gegensatz zur DNA aus dem Zucker Ribose. Anstelle der Nukleinbase Thymin wird Uracil verwendet. Zur Initialisierung des Transkriptionsprozesses verbindet die RNA-Polymerase zwei Ribonukleotide. Dabei entsteht eine Esterbindung zwischen Phosphat und Ribose, wobei Pyrophosphat eliminiert wird. Zur Elongation der mRNA rückt die RNA-Polymerase schrittweise vorwärts und legt den codierenden DNA-Strang zur Paarung weiterer Ribonukleotide frei. Die Ableserichtung der DNA verläuft dabei vom 3'-Ende zum 5'-Ende, sodass die Synthese der komplementären mRNA (von engl. *messenger* RNA) entsprechend von 5'- zum 3' verläuft. Der Elongationsprozess läuft solange weiter, bis die Polymerase auf eine zweite spezielle Sequenz des zu transkribierenden DNA-Strangs trifft, der Terminator-Sequenz. In Prokaryoten wird die produzierte mRNA direkt von den Ribosomen zur Translation verwendet. In Eukaryoten verlässt die produzierte mRNA erst nach weiteren Prozessen den Zellkern durch eine Kernpore und gelangt zu den Ribosomen [18].

- **Translation:**

Die im Schritt der Transkription gebildete mRNA dient als Vorlage für die Herstellung eines Proteins. Drei aufeinanderfolgende Nukleotide der mRNA, die Basentriplets bzw. Codons, codieren eine bestimmte Aminosäure. Proteine ergeben sich dann als Sequenz von aufeinanderfolgenden Aminosäuren. Da es  $4^3 = 64$  verschiedene Kombinationen aus drei Nukleotiden, aber nur zwanzig proteinogene Aminosäuren gibt, spricht man von einer Degeneration des genetischen Codes: Mehrere Codes werden für die gleiche Aminosäure verwendet. Durch die Degeneration sinkt die Fehlerrate des Translationsprozesses, da trotz Fehlern beim Ablesen der Basen die richtige Aminosäure in das Protein eingebaut werden kann. Gelangt die mRNA zu den Ribosomen, so wird sie in 5'-3'-Richtung zur Herstellung des Proteins bearbeitet (dies entspricht der gleichen Richtung, in der die mRNA im Transkriptionsprozess synthetisiert wurde): Ein Codon der mRNA wird abgelesen und die mit der entsprechenden Aminosäure beladenen tRNA (von engl. *transfer* RNA) dockt mit Hilfe des Ribosoms an. Die tRNA enthält dazu am unteren Ende ein zum Codon der mRNA komplementäres Gegenstück, das Anticodon. Nachdem sich eine zweite tRNA mit der passenden Aminosäure neben die erste tRNA auf mRNA gesetzt hat, werden die beiden Aminosäuren mit einer Peptidbindung miteinander verknüpft und die erste tRNA verlässt das Ribosom. Durch Wiederholung dieses Prozesses entsteht so nacheinander eine immer länger werdende Kette von Aminosäuren. Erreicht das Ribosom ein Stopp-Codon, so wird der Prozess abgebrochen. Das gebildete Protein löst sich vom Ribosom und kann nach etwaiger Faltung in eine räumliche Struktur (Sekundärstruktur und Tertiärstruktur) Aufgaben im Organismus übernehmen.

Die Menge und Häufigkeit der Proteinproduktion bezeichnet man als Genexpression. Wie in Abschnitt 2.5.1 erläutert, können mit Hilfe von DNA-Microarrays die Genexpressionsprofile verschiedenster Zelltypen bei variierenden äußeren Bedingungen analysiert und Rückschlüsse auf die beteiligten Proteine gezogen werden. Auf die Regulierung der Genexpression wird an dieser Stelle nicht eingegangen, sondern auf Standardliteratur zum Thema verwiesen.

## 2.3 DNA-Hybridisierung

Als Hybridisierung bezeichnet man im Allgemeinen die Bindung zweier (teilweise) komplementärer Nucleinsäuren zu einer stabilen Doppelhelix. Die Stabilität der Doppelhelix beruht auf Wasserstoffbrückenbindungen zwischen den Basen der beiden Nucleinsäuren und Stapelwechselwirkungen zwischen benachbarten Basen. Die Hybridisierung der komplementären Basen der Nucleinsäuren ist äußerst spezifisch. Bereits eine einzelne Fehlstelle kann die Stabilität der Doppelhelix merklich beeinflussen. Diese Spezifität der molekularen Erkennung wird sowohl in der Natur als auch in vielen molekularbiologischen Anwendungen wie beispielsweise bei DNA-Hochdurchsatzverfahren ausgenutzt. Im Folgenden werden die für diese Arbeit relevanten Aspekte der DNA-Hybridisierung wie Kinetik, Schmelztemperatur und Stabilität der DNA-Doppelhelix näher vorgestellt.

### 2.3.1 Kinetik der DNA-Hybridisierung

In diesem Abschnitt soll die Kinetik der Hybridisierung zweier DNA-Einzelstränge  $S_1$  und  $S_2$  zu einem Duplex  $D$  mathematisch beschrieben werden. Nimmt man an, dass die Einzelstränge  $S_1$  und  $S_2$  nicht selbst-komplementär sind und keine stabilen Sekundärstrukturen ausbilden, so kann die Hybridisierung im einfachsten Fall mit einem Zweizustandsmodell beschrieben werden, in dem der Duplex  $D$  entweder vollständig geschlossen oder denaturiert ist. Für die Hybridisierungsreaktion ergibt sich dann:



$k_+$  bzw.  $k_-$  ist die Assoziations- bzw. die Dissoziationsrate der Reaktion. Dieses Zweizustandsmodell kann zur Beschreibung der Hybridisierung nur dann herangezogen werden, wenn keine signifikanten Zwischenzustände zwischen vollständig hybridisierten und denaturiertem Duplex vorliegen, da die Bindung von individuellen Basenpaaren nicht berücksichtigt wird. Das *Zipper*-Modell, das individuelle Basenpaarung und -dissoziation berücksichtigt, wird in Abschnitt 6.1.1 vorgestellt.

Im Falle von DNA-Microarrays (siehe Abschnitt 2.5) ist eine der beiden Spezies  $S_1$  bzw.  $S_2$  auf einer Oberfläche fixiert und wird als *Probe*  $P$  bezeichnet. Die Spezies in Lösung wird als *Target*  $T$  bezeichnet. Diese Nomenklatur wird im Hinblick auf Abschnitt 2.5 für die weiteren Ausführungen übernommen.

Ist die Target-Konzentration, wie in den meisten in dieser Arbeit durchgeführten Experimenten, wesentlich größer als die Probe-Konzentration auf der Oberfläche ( $[T] \gg [P]$ ), so ist die Reduzierung der Target-Moleküle in Lösung aufgrund des Hybridisierungsvorgangs vernachlässigbar und kann als konstant angenommen werden ( $[T] = [T_0]$ ). Die Konzentration  $[P(t)]$  der unhybridisierten Probes auf der Oberfläche zum Zeitpunkt  $t$  entspricht gerade der Differenz aus Anfangskonzentration  $[P_0]$  und bereits hybridisierten Probes:  $[P(t)] = [P_0] - [D(t)]$ . Da die Bildung eines Duplexes eine Reaktion 2. Ordnung [26] und die Denaturierung eine Reaktion 1. Ordnung ist, ergibt sich für die Duplexkonzentration  $[D(t)]$  zum Zeitpunkt  $t$  folgende Differentialgleichung:

$$\frac{d[D(t)]}{dt} = k_+[T_0]([P_0] - [D(t)]) - k_-[D(t)] \quad (2.2)$$

Durch die Einführung von  $\theta = [D]/[P_0]$  als den Anteil von hybridisierten Probe-Molekülen und  $K = k_+/k_-$  als Bindungskonstante der Target-Moleküle ergibt die stationäre Lösung ( $d[D(t)]/dt = 0$ ) dieser Differentialgleichung die klassische Langmuir-Isotherme [27]:

$$\theta = \frac{K \cdot [T_0]}{1 + K \cdot [T_0]} \quad (2.3)$$

Der Anteil hybridisierter Probes ist experimentell direkt zugänglich. In Microarray-Experimenten ist  $\theta$  proportional zur Intensität des Hybridisierungssignals und wird daher in dieser Arbeit als Hybridisierungssignal bezeichnet.

Gl. 2.2 beschreibt die Hybridisierung einer Target-Spezies. Mit den gleichen Überlegungen können auch Differentialgleichungen für zwei nicht wechselwirkende Target-Spezies  $A$  und  $B$  aufgestellt werden, die um die selben Bindungsplätze konkurrieren. Es ergibt sich für Spezies  $A$

$$\frac{d[D_A(t)]}{dt} = k_{A,+}[A_0]([P_0] - [D_A(t)] - [D_B(t)]) - k_{A,-}[D_A(t)] \quad (2.4)$$

und entsprechend für Spezies  $B$

$$\frac{d[D_B(t)]}{dt} = k_{B,+}[B_0]([P_0] - [D_B(t)] - [D_A(t)]) - k_{B,-}[D_B(t)] \quad (2.5)$$

Die stationäre Lösung von Gl. 2.4 (und entsprechend für Gl. 2.5) lautet:

$$[D_A] = \frac{K_A \cdot [P_0] \cdot [A_0]}{1 + K_A \cdot [A_0] + K_B \cdot [B_0]} \quad (2.6)$$

Bildet man das Verhältnis  $[D_A]/[D_B]$  der beiden Duplexkonzentrationen im Gleichgewicht, so erhält man:

$$\frac{[D_A]}{[D_B]} = \frac{[A_0] \cdot K_A}{[B_0] \cdot K_B} = \frac{[A_0]}{[B_0]} \cdot \exp\left(\frac{\Delta\Delta G}{RT}\right) \quad (2.7)$$

Gl. 2.7 besagt, dass das Verhältnis der Duplexkonzentrationen nur vom Verhältnis der Konzentrationen und der Differenz der Bindungsenergien  $\Delta\Delta G = \Delta G_A - \Delta G_B$  der beiden Spezies  $A$  und  $B$  aus der Einzelhybridisierung abhängt. In Kapitel 5 wird gezeigt, dass Gl. 2.7 unter bestimmten Voraussetzungen versagt: Selbst bei ähnlichen Bindungskonstanten der beiden Spezies in der Einzelhybridisierung (kleines  $\Delta\Delta G$ ) und einer wesentlich höheren Konzentration von Spezies  $B$  im Vergleich zur Spezies  $A$ , kann die Duplexkonzentration  $[D_A]$  der Spezies  $A$  unbeeinflusst von der Anwesenheit von Spezies  $B$  bleiben.

In Gl. 2.7 wurde ausgenutzt, dass die temperaturabhängige Bindungskonstante  $K$  einer Spezies über

$$K(T) \propto \exp\left(\frac{\Delta G_D}{RT}\right) \quad (2.8)$$

mit der Änderung der freien Energie  $\Delta G_D$  bei Duplexformation in Verbindung gebracht werden kann (mit freier Energie ist - wenn nicht anders angegeben - in dieser Arbeit immer die Gibbsche freie Energie  $G$  und nicht die Helmholtzsche freie Energie  $F$  bezeichnet).  $T$  bezeichnet die Temperatur, bei der die Hybridisierung stattfindet, und  $R = 1.987 \cdot 10^{-3}$  kcal/(mol·K) die universelle Gaskonstante. Die freie Energie  $\Delta G_D$  kann über die Van't-Hoff-Gleichung

$$\frac{1}{T_m} = \frac{R}{\Delta H_D} \ln\left(\frac{C_T}{4}\right) + \frac{\Delta S_D}{\Delta H_D} \quad (2.9)$$

bestimmt werden.  $C_T$  bezeichnet die Gesamtkonzentration der Moleküle und  $T_m$  die Schmelztemperatur des Duplexes.  $T_m$  entspricht per Definition der Temperatur, bei der 50% der Duplexe denaturiert sind (siehe Abschnitt 2.3.2). Ist die Enthalpieänderung  $\Delta H_D$  und die Entropieänderung  $\Delta S_D$  durch Messung der Schmelzkurve bekannt, kann die Änderung der freien Energie über

$$\Delta G_D = \Delta H_D - T\Delta S_D \quad (2.10)$$

berechnet werden. Darüber hinaus kann  $\Delta G_D$  auch anhand des *Nearest-Neighbor*-Modell (NN-Modell, siehe Abschnitt 2.3.3) theoretisch vorhergesagt werden.

**Anmerkung:**

Bei vergleichbaren Konzentrationen von Probe und Target ( $[P] \approx [T]$ ) muss die Reduzierung der Target-Moleküle aufgrund der Hybridisierungsreaktion in Betracht gezogen werden. Gl. 2.2 wird dann abgeändert zu [28]:

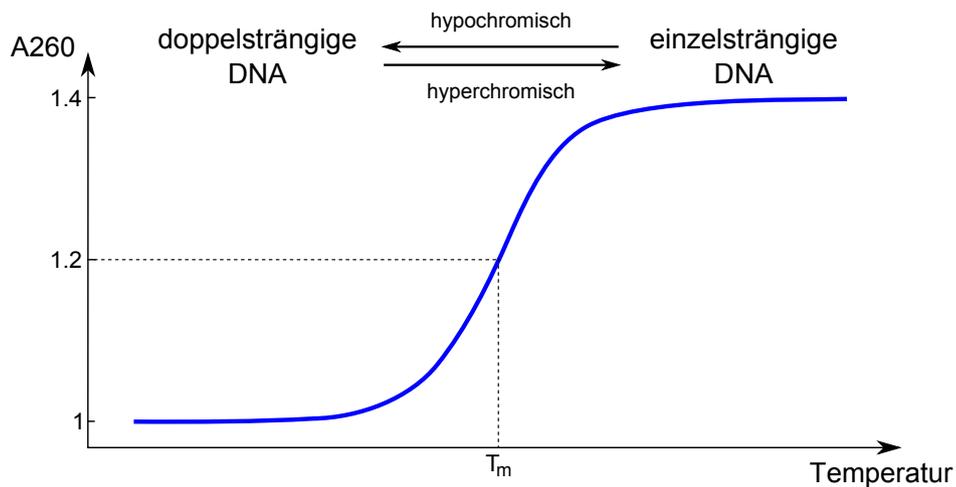
$$\frac{d[D(t)]}{dt} = k_+([T_0] - [D(t)])([P_0] - [D(t)]) - k_-[D(t)] \quad (2.11)$$

Entsprechendes gilt für Gl. 2.4 und 2.5 in der Konkurrenzhybridisierung. Die stationären Lösungen dieser Gleichungen sind jedoch von der selben Form, wobei  $[A_0]$  bzw.  $[B_0]$  durch  $[A_{eq}] = [A_0] - [D_A(t)]$  bzw.  $[B_{eq}] = [B_0] - [D_B(t)]$  ersetzt werden müssen.

### 2.3.2 Schmelztemperatur

In den fünfziger Jahren des letzten Jahrhunderts wurde experimentell festgestellt, dass eine Temperaturerhöhung in der kooperativen Denaturierung von doppelhelikaler DNA in zwei Einzelstränge resultiert [29, 30]. Die thermische Denaturierung von DNA ist seitdem Gegenstand aktueller Forschung, da so wichtige Informationen über die Wechselwirkung von einzelnen Basenpaaren, die Stabilität der Doppelhelix oder die Einflüsse von Lösungsmittel und Salzkonzentration erhalten werden können [31–33]. Zur Untersuchung der thermischen Denaturierung von DNA stehen mehrere experimentelle Methoden zur Verfügung, wie z.B. UV-Absorption, dynamische Differenzkalorimetrie, Circular dichroismus, Kernspinresonanz (NMR von engl. *nuclear magnetic resonance*), Fluoreszenzemission und Temperaturgradient-Gelelektrophorese [34–37].

Eine wichtige Größe, die aus diesen Untersuchungen gewonnen werden kann, ist die Schmelztemperatur des betreffenden DNA-Duplexes. Die Schmelztemperatur eines DNA-Duplexes entspricht per Definition der Temperatur, bei der 50 % der Duplexe denaturiert sind. Für praktische Anwendungen, wie das Design von *Primern* in der Polymerase-Kettenreaktion (PCR von engl. *polymerase chain reaction*, siehe Abschnitt 2.4) und der Probes in Microarray-Experimenten, ist die Schmelztemperatur ein sehr wichtiger Parameter. Eine häufig eingesetzte Methode zur Untersuchung des Schmelzübergangs von DNA ist die bereits erwähnte Messung der UV-Absorption [34]. Die UV-Absorption von doppelsträngiger DNA ist bei einer Wellenlänge von  $\lambda = 260$  nm aufgrund erhöhter Stapelwechselwirkung zwischen Basen entlang des Duplexes (siehe Abschnitt 2.1) niedriger als bei einzelsträngiger DNA (Hyperchromizität). Wird die Temperatur langsam erhöht, so zeigt sich der Schmelzübergang von doppelsträngiger DNA in einer Zunahme der UV-Absorption  $A_{260}$  um 20-40 % (siehe Abb. 2.5). Dabei muss bedacht werden, dass die UV-Absorption nur den Anteil von Basen wiedergibt,



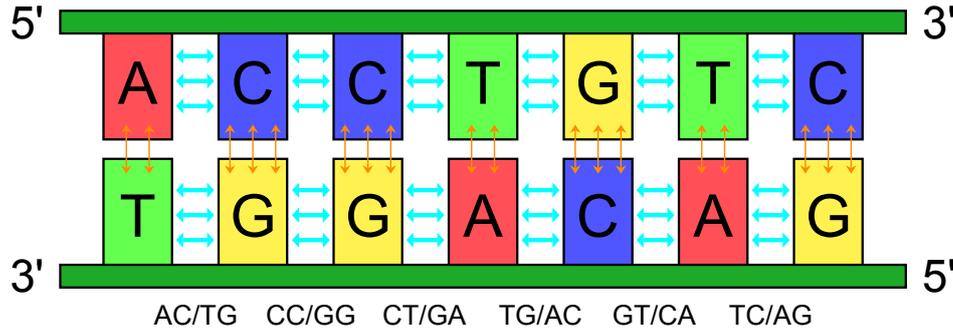
**Abbildung 2.5:** DNA-Schmelzübergang. Eine typische Methode zur Bestimmung thermodynamischer Größen wie beispielsweise der Schmelztemperatur ist die Messung der UV-Absorption von DNA als Funktion der Temperatur. Doppelsträngige DNA hat aufgrund höherer Stapelwechselwirkungen eine niedrigere UV-Absorption als einzelsträngige DNA. Schmilzt die DNA bei Temperaturerhöhung, so zeigt sich dies in einer Erhöhung der UV-Absorption um 20-40 %. Die Schmelztemperatur liegt per Definition die Temperatur in der Mitte des Schmelzübergangs.

bei denen keine Stapelwechselwirkungen untereinander vorliegen. Dieser Anteil ist nicht unbedingt gleichzusetzen mit dem Anteil dissoziierter Duplexen. Nimmt man jedoch an, dass der Schmelzübergang ein Zwei-Zustands-Prozess ist und somit keine signifikanten teilweise denaturierten Zwischenzustände vorliegen (dies ist insbesondere bei kurzen DNA-Strängen der Fall), so ist der Anteil von geschmolzenen Basen gleich dem Anteil von geschmolzenen Duplexen [38].

Es gibt verschiedene Ansätze die Schmelztemperatur eines DNA-Duplexes empirisch zu bestimmen. Beispielsweise lässt sich anhand der Wallace-Regel die Schmelztemperatur für kurze Duplexe zwischen 5 und 20 Basenpaaren berechnen [39]. Für größere Duplexlängen (Duplexlänge  $> 50$  Basenpaare) kann der von Wetmur *et al.* entwickelte Ansatz verwendet werden [40]. Darüber hinaus existieren auch thermodynamische Zugänge wie z.B. der DINAMelt-Server [41], die basierend auf den sog. *Nearest-Neighbor*-Parametern (siehe Abschnitt 2.3.3) sehr erfolgreich die Schmelztemperatur vorhersagen.

### 2.3.3 Das Nearest-Neighbor-Modell

Wie in Abschnitt 2.1 erläutert, wird die Stabilität der DNA-Doppelhelix durch Wasserstoffbrücken zwischen gegenüberliegenden Basen der beiden Einzelstränge und Stapelwechselwirkungen zwischen benachbarten Basen entlang eines Strangs generiert. Der Beitrag der Stapelwechselwirkungen ist jedoch wesentlich größer als der Beitrag der



**Abbildung 2.6:** *Nearest-Neighbor-Modell* (NN-Modell, Abb. nach [1]). Das Modell beinhaltet Wasserstoffbrückenbindungen zwischen den Basen der beiden Einzelstränge (orange Pfeile) und Stapelwechselwirkungen zwischen benachbarten Basen (hellblaue Pfeile). AT-Paare werden durch zwei, CG-Paare durch drei Wasserstoffbrücken verknüpft. Die Stärke der Stapelwechselwirkungen hängt von den entsprechenden Nachbarbasen ab. Im hier verwendeten Formalismus werden die beiden Beiträge zur Stabilität eines Duplets von Basenpaaren, z.B. CT/GA (5'-CT-3' mit 3'-GA-5'), zu einem NN-Parameter zusammengefasst. Die Stabilität eines DNA-Duplexes kann dann als Summe der einzelnen NN-Parameter betrachtet werden.

Wasserstoffbrückenbindungen [42]. Im Formalismus von Gotoh *et al.*, der im Rahmen dieser Arbeit verwendet wird, sind die beiden Beiträge zur Stabilität zu einem *Nearest-Neighbor*-Parameter (NN-Parameter)  $\Delta G_{NN}$  kombiniert [43]. In Abb. 2.6 ist das NN-Modell veranschaulicht: Die Wasserstoffbrückenbindungen zwischen gegenüberliegenden Basen des Duplexes sind durch vertikale Pfeile, die Stapelwechselwirkungen entlang eines Strangs durch horizontale Pfeile gekennzeichnet. Die entsprechenden NN-Paare sind unter dem Duplex angegeben. Mit Hilfe des NN-Modells lässt sich die freie Energie des Duplexes  $\Delta G_{D,T}$  für eine bestimmte Temperatur  $T$  berechnen zu:

$$\Delta G_{D,T} = \Delta G_{\text{init},T} + \sum \Delta G_{\text{NN},T} + \Delta G_{\text{sym},T} + \Delta G_{\text{AT-Term},T} \quad (2.12)$$

$\Delta G_{\text{init},T}$  ist die zur Duplexinitialisierung notwendige Energie und  $\Delta G_{\text{NN},T}$  die Energie der jeweiligen NN-Paare. Befindet sich ein A/T Basenpaar am Ende des Duplexes, so muss  $\Delta G_{D,T}$  durch  $\Delta G_{\text{AT-Term},T}$  korrigiert werden: Da A/T Basenpaare nur durch zwei Wasserstoffbrücken aneinander gebunden sind, ergibt sich bei A/T-Paaren am Ende des Duplexes eine erhöhte Öffnungswahrscheinlichkeit und dadurch eine Destabilisierung des Duplexes. Ist der Duplex selbst-komplementär, so muss noch  $\Delta G_{\text{sym},T}$  berücksichtigt werden.

Auf analoge Weise können die Änderungen der Enthalpie  $\Delta H_D$  und der Entropie  $\Delta S_D$  bei Duplexformation anhand der NN-Parameter berechnet werden. Damit können die NN-Parameter  $\Delta G_{T,D}$  für verschiedenen Temperaturen berechnet werden:

$$\Delta G_{D,T} = \Delta H_D - T \cdot \Delta S_D \quad (2.13)$$

**Tabelle 2.1:** *Nearest-Neighbor-Parameter* für Watson-Crick-Basenpaarung bei 1M NaCl nach [44]. Ein NN-Parameter gibt die Stabilität eines Duplets von Basenpaaren an, hier beispielsweise CT/GA. Das bedeutet, dass sich eine Sequenz bestehend aus zwei Basen 5'-CT-3' mit einer komplementären Sequenz 3'-GA-5' verbindet. Aus Symmetriegründen ergeben sich so 10 verschiedene NN-Parameter. Zusätzlich muss noch ein Term für die Duplexinitialisierung, ein Term zur Korrektur von Duplexen mit AT-Endterm und ein Symmetrieterm zur Korrektur von selbst-komplementären Sequenzen berücksichtigt werden. Die rechte Spalte enthält die aus Spalte zwei und drei berechneten NN-Parameter für eine Temperatur von 44 °C. Die Energieskala ist hier so gewählt, dass Werte von  $\Delta g > 0$  den Duplex stabilisieren.

NN-Paar	$\Delta H_D$ [kcal/mol]	$\Delta S_D$ [cal/(K·mol)]	$\Delta G_{NN,44^\circ\text{C}}$ [kcal/mol]
AA/TT	7.6	21.3	0.85
AT/TA	7.2	20.4	0.73
TA/AT	7.2	21.3	0.45
CA/GT	8.5	22.7	1.30
GT/CA	8.4	22.4	1.30
CT/GA	7.8	21.0	1.14
GA/CT	8.2	22.2	1.16
CG/GC	10.6	27.2	1.98
GC/CG	9.8	24.4	2.07
GG/CC	8.0	19.9	1.69
Initialisierung	-0.2	5.7	-2.00
AT-Endterm	-2.2	-6.9	-0.01
Symmetrie	0.0	1.4	-0.44

Es wird angenommen, dass  $\Delta H_D$  und  $\Delta S_D$  temperaturunabhängig sind, was nach [45, 46] eine sehr gute Näherung für Nukleinsäuren darstellt.

Die NN-Parameter wurden von verschiedenen Arbeitsgruppen experimentell ermittelt [43, 47–52] und von SantaLucia *et al.* zu einem Parametersatz vereinigt [53]. Dieser Parametersatz wird zusammen mit den Parametern für Einzelfehlstellen (MMs von engl. *mismatches*), die von Allawi *et al.* etabliert wurden [54], für die theoretischen Berechnungen in Kapitel 6 verwendet. Mit den NN-Parametern aus Tab. 2.1 kann über die Van't-Hoff-Gleichung (Gl. 2.9) die Schmelztemperatur  $T_m$  in °C eines Duplexes berechnet werden [44]:

$$T_m = \Delta H_D^{ges} \times 1000 / (\Delta S_D^{ges} + R \cdot \ln(C_T/4)) - 273.15 \quad (2.14)$$

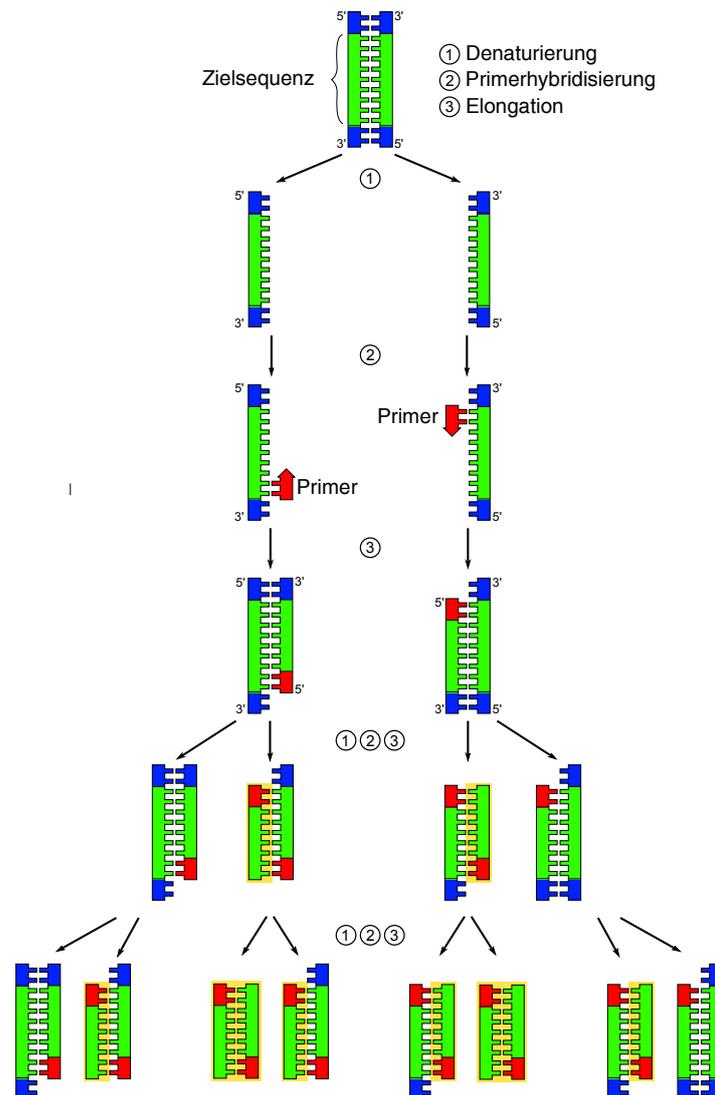
$\Delta H_D^{ges}$  in kcal/mol bzw.  $\Delta S_D^{ges}$  in cal/(K·mol) ist die gesamte Änderung der Enthalpie bzw. Entropie bei Duplexdenaturierung,  $R = 1.9872$  cal/(K·mol) die ideale Gas-konstante und  $C_T$  die Gesamtkonzentration der DNA. Gl. 2.14 gilt nur für Natrium-Konzentrationen  $[\text{Na}^+]$  von 1 mol/l und im Falle gleicher Konzentrationen der beiden

komplementären DNA-Stränge. Wird die  $\text{Na}^+$ -Konzentration verringert, wird die Abschirmung der negativ geladenen Phosphatgruppen der jeweiligen Einzelstränge durch den Puffer schwächer. Die sich daraus ergebende erhöhte Abstoßung der Einzelstränge führt zu einer erniedrigten Duplexstabilität. Höhere  $\text{Na}^+$ -Konzentration als 1 mol/l resultieren nicht in einer signifikanten Erhöhung der Duplexstabilität, da die elektrostatische Abschirmung nicht weiter verbessert werden kann. Darüber hinaus existieren eine Vielzahl von empirischen Korrekturen der Schmelztemperatur von der Salzkonzentration [55].

## 2.4 Polymerase-Kettenreaktion

Die Polymerase-Kettenreaktion (PCR von engl. *polymerase chain reaction*) ist eine *in vitro*-Methode zur selektiven und schnellen, millionenfachen Vervielfältigung einer bestimmten Nukleotidsequenz der DNA, die 1971 von Kleppe *et al.* erfunden und von Mullis *et al.* weiterentwickelt wurde [56, 57]. Voraussetzung ist, dass mindestens eine Teilsequenz der zu vervielfältigenden DNA bekannt ist. Anhand der bekannten Teilsequenz werden zwei sog. *Primer* konstruiert, die jeweils komplementär zu einem Strang der Doppelhelix sind und sich an gegenüberliegenden Enden der zu replizierenden DNA-Sequenz anlagern. Die PCR-Technik umfasst einen Drei-Schritte-Zyklus, der in Abb. 2.4 dargestellt ist.

Durch Aufbrechen der Wasserstoffbrückenbindungen mittels Temperaturerhöhung auf ca. 95 °C wird im ersten Schritt die DNA-Doppelhelix in zwei Einzelstränge denaturiert. Anschließendes Abkühlen der Reaktionstemperatur auf ca. 50-60 °C ermöglicht eine Anlagerung der Primer an die einzelsträngigen DNA-Templates. Die Temperatur wird normalerweise 3-5 °C unterhalb der Schmelztemperatur der Primer eingestellt, um eine spezifische Anlagerung der Primer zu gewährleisten. Bei einer zu niedrigen Temperatur kann es zur Anlagerung an nicht 100 % komplementäre Sequenzen kommen, die zu unspezifischen Produkten führt. Eine zu hohe Temperatur kann im schlechtesten Fall die Anlagerung der Primer verhindern. Daher ist die genaue Kenntnis der Schmelztemperaturen der beiden Primer (die von der Länge und der Sequenz der Primer abhängt) von großer Wichtigkeit (siehe auch Abschnitte 2.3.2 und 2.3.3). Nach Anlagerung der Primer wird die Temperatur auf die Arbeitstemperatur der verwendeten DNA-Polymerase erhöht (die Taq-Polymerase arbeitet beispielsweise bei 75-80 °C [58]). Die DNA-Polymerase fügt den beiden freien Strängen ausgehend vom 3'-Ende freie Nukleotide hinzu und verbindet das Rückgrat (Ligation). Der Primer wird nicht wieder abgelöst, er bildet den Anfang des neuen Einzelstrangs. Durch mehrmalige Wiederholung dieses Zyklus wird die gewünschte DNA-Sequenz exponentiell repliziert.

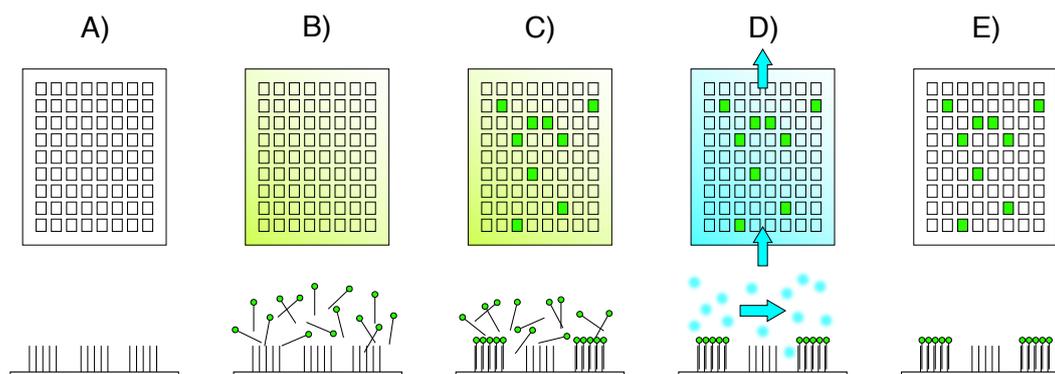


**Abbildung 2.7:** PCR-Zyklus. Ziel der PCR ist es, den in grün markierten Duplexbereich exponentiell zu vervielfältigen. Dazu wird im ersten Schritt durch eine Erhöhung der Temperatur auf ca. 95 °C der DNA-Duplex denaturiert (1). Die Hybridisierung der speziell für die zu vervielfältigende Sequenz ausgewählten Primer geschieht im nächsten Schritt bei einer Temperatur von ca. 50-60 °C (2). Nach Erhöhung der Temperatur auf die Arbeitstemperatur der gewählten DNA-Polymerase verlängert diese die beiden Primer (3). Nach Ende der Primer-Elongation liegen nun zwei DNA-Duplexe vor, die jedoch noch nicht der gewünschten Sequenz entsprechen. Eine Zykluswiederholung resultiert in vier DNA-Duplexen. Die gelb hinterlegten Teile der beiden mittleren Duplexe entsprechen nun erstmals der gewünschten Sequenz. Diese dienen nun als Vorlage für die exponentielle Vervielfältigung des gewünschten DNA-Duplexes: Im nächsten Schritt entstehen acht DNA-Duplexe, wovon zwei der gewünschten Sequenz entsprechen. Ab diesem Punkt wird durch mehrmaliges Wiederholen der Zyklen eine exponentielle Zunahme des gewünschten DNA-Duplexes erzielt.

Die PCR-Technik wird für eine Vielzahl von Experimenten und Analysen eingesetzt. Aufgrund ihrer extrem hohen Empfindlichkeit kann sie beispielsweise in der Forensik zur Erstellung von genetischen Fingerabdrücken auch bei sehr kleinen Probenmengen eingesetzt werden. Darüber kann sie zur Erkennung von Erbkrankheiten bzw. Virusinfektionen im Frühstadium oder zur Klonierung von Genen verwendet werden. Auch bei DNA-Microarrays können die durch die PCR-Technik gewonnenen Stränge als *Targets* (siehe Abschnitt 2.5) eingesetzt werden. Den für die Microarray-Experimente notwendigen Fluoreszenzmarker kann man entweder während der PCR durch Verwendung von fluoreszenzmarkierten Nukleotidbausteinen anbringen oder in einem separaten Schritt nach Ende der PCR.

## 2.5 DNA-Microarrays

DNA-Microarrays erlauben es, komplexe Gemische aus einer Vielzahl von einzelsträngigen Oligonukleotiden, den *Targets*, parallel zu analysieren. Dazu werden bekannte Oligonukleotid-Sequenzen, die sog. *Probes*, in einer regelmäßigen Anordnung an definierten Positionen, den *Features*, mit einem Ende auf einer Oberfläche immobilisiert. Ein Feature besteht dabei aus einer Vielzahl von Probes der gleichen Sequenz. Um Informationen über die zu untersuchenden Target-Mischungen zu erhalten (wie beispielsweise die Konzentration oder die Sequenz einer bestimmten Target-Spezies), werden diese mit einem Fluoreszenzmarker versehen und in Kontakt mit der Microarray-Oberfläche gebracht. Durch Hybridisierung an komplementären Probes werden die Targets an den Features akkumuliert und man erhält messbare Fluoreszenzsignale. Nachdem alle Features mit Targets hybridisiert sind, was aufgrund des langsamen Diffusionsprozesses mehrere Stunden dauern kann, werden ungebundene, unspezifische Targets abgewaschen, während die hybridisierten Targets auf der Microarray-Oberfläche verbleiben (siehe Abb. 2.8). Da die Position und die Sequenz jedes Features bekannt ist, kann man so Rückschlüsse auf das Vorhandensein einer Targetsequenz in der Lösung ziehen. Prinzipiell kann man außerdem anhand der Fluoreszenzintensität auf die entsprechende Target-Konzentration schließen. Jedoch sind Microarrays aufgrund zahlreicher Faktoren, wie der schwierigen Vorhersagbarkeit von Probe-Target-Affinitäten, Konkurrenzhybridisierungen zwischen verschiedenen Target-Spezies oder variierenden Synthesequalitäten nur semi-quantitativ. Dies betrifft allerdings nur den Vergleich zwischen verschiedenen Probe-Target-Paaren. Werden Signalunterschiede bei der Hybridisierung von verschiedenen Target-Mischungen an eine bestimmte Probe-Sequenz festgestellt (siehe Genexpressionsanalyse, Abschnitt 2.5.1), so spricht dies für unterschiedliche Target-Konzentrationen der entsprechenden Spezies in den beiden Mischungen.



**Abbildung 2.8:** Hybridisierung eines DNA-Microarrays (Abb. nach [1]). Die obere Zeile zeigt die Microarray-Oberfläche mit den einzelnen Features. In der unteren Zeile sind die auf der Oberfläche immobilisierten Probes auf mikroskopischer Ebene dargestellt. **A)** Das Array wurde noch nicht hybridisiert und die Probes haben noch keine Targets gebunden. **B)** Hybridisierung des Arrays mit der zu untersuchenden Target-Lösung. **C)** Nach Erreichen des Gleichgewichts (was aufgrund des langsamen Diffusionsprozesses mehrere Stunden dauern kann) sollten alle Probes, für die komplementäre Targets in Lösung existieren, hybridisiert sein und sich für die entsprechenden Features messbare Fluoreszenzsignale ergeben. Im hier dargestellten Beispiel befinden sich zu den mittig befindlichen Probes keine komplementären Targets in der Lösung. **D)** Die in der Lösung verbleibenden Targets werden vom Microarray gewaschen. **E)** Die hybridisierten Targets verbleiben auf der Oberfläche. Da Position und Sequenz der Features bekannt ist, können so Rückschlüsse auf die Zusammensetzung der Target-Lösung gezogen werden.

Die Länge der Probes spielt bei DNA-Microarrays eine wichtige Rolle. Durch die Verwendung von kurzen Probes wird eine hohe Selektivität des Microarrays erzielt, da so die verschiedenen Target-Spezies fast ausschließlich an den zu ihnen komplementären Probes hybridisieren können. Eine möglichst hohe Sensitivität, d.h. ein messbares Fluoreszenzsignal bei niedrigerer Target-Konzentration, wird durch lange Probes realisiert. Anhand der verwendeten Probes der DNA-Microarrays lassen sich zwei Typen von Arrays unterscheiden. Im Falle von cDNA-Microarrays (cDNA von engl. *complementary DNA*) werden cDNA-Probes oder PCR-Produkte immobilisiert. Da die Probes hier bis zu mehreren hundert Basen lang sein können und somit wie oben beschrieben keine Unterscheidung zwischen ähnlichen Target-Sequenzen möglich ist, werden cDNA-Microarrays hauptsächlich für Genexpressionsanalysen verwendet (siehe Abschnitt 2.5.1). Die Probes auf Oligonukleotid-Microarrays werden durch Immobilisierung vorsynthetisierter Probes oder durch *in situ*-Synthese hergestellt und haben eine Länge zwischen 15 und 100 Basen. Dieser Art von Microarrays findet aufgrund der hohen Sensitivität Anwendung im Bereich *Genotyping* und Resequenzierung. Ein Überblick über die möglichen DNA-Microarray-Systeme und ihre Anwendungen ist in [59, 60] gegeben. Für weitere Literatur zu DNA-Microarrays sei auf [1] verwiesen.

### 2.5.1 Anwendungen

- **Messung von Genexpressionsprofilen**

Bei der Messung von Genexpressionsprofilen soll bestimmt werden, welche Gene einer biologischen Probe in verschiedenen Situationen transkribiert werden. So kann beispielsweise die Genexpression einer gesunden Zelle mit der einer Krebszelle verglichen oder die Änderung der Transkription einer Zelle bei mechanischer Belastung im Bezug auf eine Referenzprobe untersucht werden. Dadurch können Rückschlüsse auf die Gene gezogen werden, die für die Veränderung der Genexpression beispielsweise bei Krankheit oder Stress verantwortlich sind. Um cDNA aus einer biologischen Probe zu gewinnen, wird im ersten Schritt mRNA aus der Probe isoliert. Durch reverse Transkription erhält man daraus die notwendige cDNA, die zusätzlich noch mit Fluoreszenzfarbstoffen markiert wird. Die cDNA-Mischung wird dann auf ein Microarray gebracht, das im optimalen Fall alle Gene des zu untersuchenden Organismus enthält. Anhand der Fluoreszenzsignale der einzelnen Features kann man auf die Expression und die Beteiligung der einzelnen Gene schließen (siehe z.B. [9, 61–63]).

- **Detektion von SNPs, *Genotyping***

Einzelnukleotid-Polymorphismen (SNPs von engl. *single nucleotide polymorphisms*) bezeichnen die Variation einzelner Nukleotide in einem DNA-Strang zwischen Mitgliedern der selben biologischen Spezies oder zwischen Allelen des selben Chromosoms in diploiden Zellen. Es wird davon ausgegangen, dass SNPs die meisten Unterschiede im menschlichen Genom ausmachen. SNPs haben großen Einfluss auf den Phänotyp. Sie können aber auch Ursache für die Ausbildung einer bestimmten Krankheit oder die Reaktion eines Individuums auf Medikamente, Impfungen, Chemikalien oder ähnliches sein. *Genotyping*-Microarrays sind für die Genomforschung und pharmazeutische Forschung von großer Bedeutung [64, 65].

- **Detektion/Identifikation von Pathogenen**

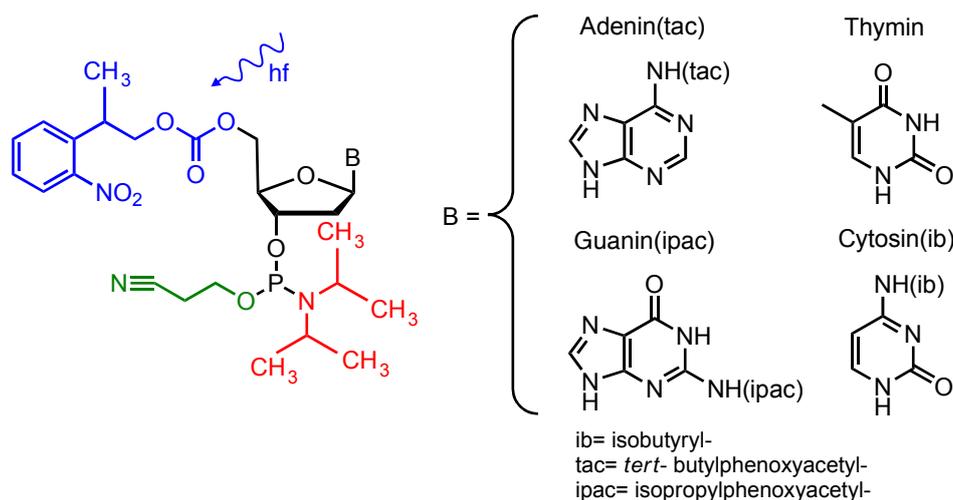
Enthält ein Microarray pathogenspezifische Probe-Sequenzen, so können Viren und Bakterien in einer biologischen Probe (z.B. Blutkulturen, Essen) schnell nachgewiesen werden [66, 67].

# 3 Experimentelle Methoden und Material

## 3.1 DNA-Microarrays

### 3.1.1 Lichtgesteuerte in situ-Synthese

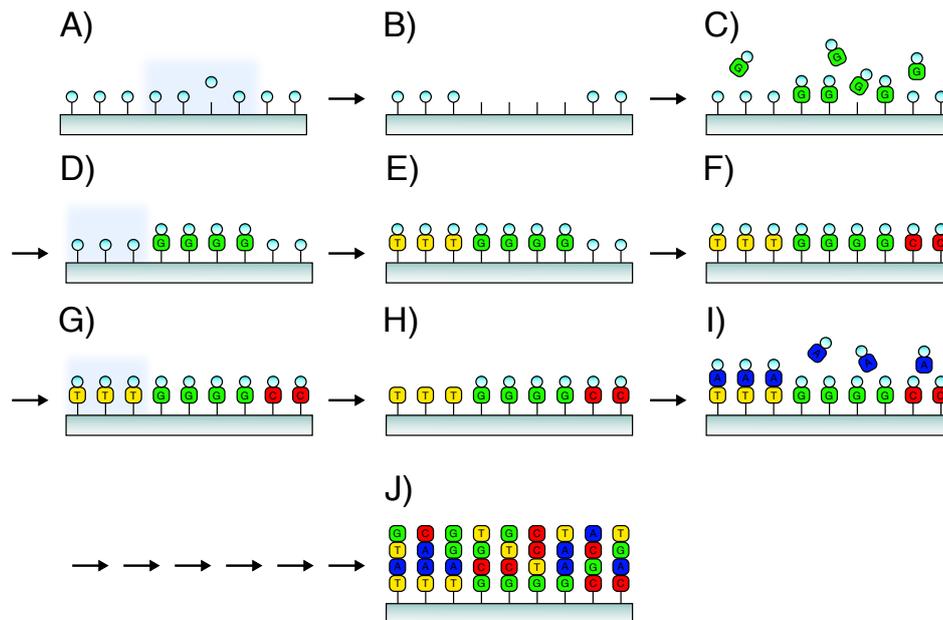
Die Probes auf der Microarray-Oberfläche sind aus einzelnen Bausteinen, den Phosphoramiditen [68–70], aufgebaut. Phosphoramidite (siehe Abb. 3.1) bestehen aus Nukleinbasen, Desoxyribose und einer Phosphorgruppe am 3'-Ende des Zuckers, die bei chemischer Aktivierung mit der Hydroxyl-Gruppe am 5'-Ende des 2-Desoxyribosering eines weiteren Phosphoramidits eine Bindung eingehen kann. Zur Vermeidung ungewollter Nebenreaktionen ist die Phosphorgruppe durch eine Diisopropylamino- und eine 2-Cyanoethyl-Gruppe geschützt. Um den kontrollierten Aufbau der Probes zu gewährleisten, verfügen die hier verwendeten Phosphoramidite über eine photolabile NPPOC (2-(2-Nitrophenyl)-propoxycarbonyl)-Schutzgruppe am 5'-Ende. Durch Bestrahlung mit UV-Licht ( $\lambda = 365 \text{ nm}$ ) kann diese Schutzgruppe abgespalten und dort ein weiterer Phosphoramidit-Baustein angekoppelt werden. Dadurch ergibt sich eine Synthese in 3'-5'-Richtung. Am 1'-Ende befindet sich eine der vier möglichen Basen Adenin, Thymin, Cytosin und Guanin. Um die Basen während der Synthese vor unerwünschten Nebenreaktionen zu schützen, sind weitere Schutzgruppen (*ib*, *tac*, *ipac*) notwendig. Alle Schutzgruppen werden am Ende der Synthese mithilfe einer 1:1 Entschützungslösung aus Ethanol und Ethylendiamin entfernt. Die verwendeten RayDite™ Phosphoramidite NPPOC-dA(*tac*), NPPOC-dC(*ib*), NPPOC-dG(*ipac*) und NPPOC-dT werden kommerziell bei Sigma-Aldrich erworben. Da Phosphoramidite extrem sensibel auf Wasser reagieren (autokatalytische Degradation, [71]), werden sie erst unmittelbar vor der Synthese in wasserfreiem Acetonitril (Reinheit 99.8 %, Sigma-Aldrich) gelöst. Das verwendete Acetonitril wird zusätzlich noch mit Hilfe eines Molekularsiebs (Trap-Paks™, Applied Biosystems) getrocknet. Der Transport der Chemikalien während der Synthese wird durch hochreines Argon 6.0 vorgenommen. Die Kopplung der



**Abbildung 3.1:** Strukturformel eines Phosphoramidits. Ein Phosphoramidit besteht aus einer Nukleinbase, dem Zucker Desoxyribose und einer Phosphatgruppe am 3'-Ende des Zuckers, die durch eine Diisopropylamino- (rot) und eine 2-Cyanoethyl-Gruppe (grün) vor ungewollten Reaktionen geschützt ist. Eine photolabile NPPOC (2-(2-Nitrophenyl)-propoxycarbonyl)-Schutzgruppe (blau) am 5'-Ende kann mittels UV-Licht abgespalten werden. So kann ein DNA-Strang durch kontrolliertes Koppeln mit einem weiteren Phosphoramidit sukzessive aufgebaut werden. Die funktionelle Gruppe B am 1'-Ende steht für eine der vier Basen Adenin, Thymin, Cytosin und Guanin, die mit weiteren Schutzgruppen (*ib, tac, ipac*) versehen sind.

Amidite erfolgt in einer Lösung aus Acetonitril und einem Aktivator (Activator42<sup>TM</sup>, 0.25 M, Proligo®Reagents). Die Abspaltung der NPPOC-Schutzgruppe, findet in einem leicht basischen Milieu bestehend aus einer 25 mM Lösung aus Piperidin (Reinheit 99 %, Sigma-Aldrich) und Acetonitril statt. Nach jedem fünften Kopplungsschritt wird der Synthesekreislauf mit einer auf Iod basierenden Oxidatorlösung (Oxidizer, Proligo®Reagents) gespült, um die während der Deprotektion entstandenen instabilen Phosphitriester-Bindungen zu oxidieren und in stabile Phosphortriester umzuwandeln. Dadurch wird sichergestellt, dass die Probes während der Synthese stabil bleiben und nicht brechen. Der Synthesekreislauf zum Aufbau der Probes besteht im Wesentlichen aus drei Schritten: Entfernung der NPPOC-Schutzgruppen mittels UV-Licht, Ankopplung von neuen Amidit-Bausteinen und Oxidation der Probes. Dies ist in Abb. 3.2 ohne den Oxidationsschritt dargestellt.

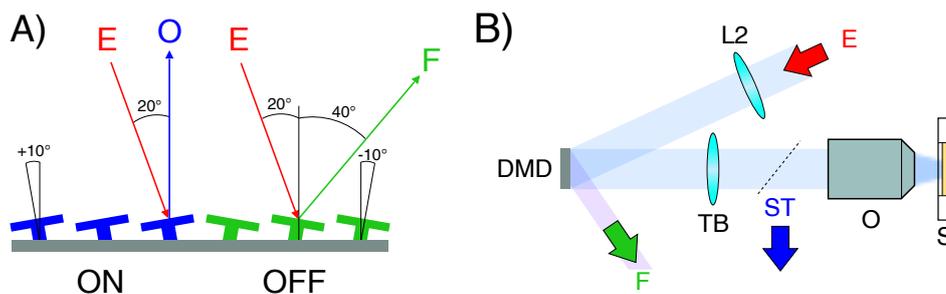
Als Substratoberfläche für die Synthese der Microarrays werden mit APTES silanisierte und mit Dendrimeren beschichtete Glasoberflächen (Menzel Deckgläser, Ø 20 mm, 0.1 mm dick) verwendet [72]. Die Dendrimere dienen dabei als Linker zwischen Oberfläche und Probes. Vor Beginn der Synthese der Probes wird zur Abstandserzeugung zwischen Dendrimeren und eigentlicher Probe-Sequenz die Base Thymin an die gesamte Microarray-Oberfläche gekoppelt. Das Protokoll zur Herstellung der Oberflächen findet sich in Anhang B.3. Alle weiteren Details finden sich in [1, 70].



**Abbildung 3.2:** Syntheszyklus (Abb. nach [1]). **A)+B)** Die NPPOC-Schutzgruppen (blaue Kugeln) werden durch Bestrahlung mit UV-Licht an den gewünschten Positionen des DNA-Microarrays entfernt. **C)** An den frei gewordenen Positionen können sich die gewünschten Amiditbausteine mit Schutzgruppe (hier Guanin) anlagern. **D)+E)** Entfernung der Schutzgruppen und Anlagerung von Thymin. **F)-J)** Durch erneute Wiederholungen der Prozedur zum Ankoppeln von neuen Amiditbausteinen können so an beliebigen Stellen des Microarray beliebige Probe-Sequenzen generiert werden.

### Fehlerquellen während der Synthese

Die größte Fehlerquelle während der Synthese stellt Streulicht dar. Streulicht kann an ungewollten Positionen auf der Oberfläche des Arrays die Schutzgruppen der Phosphoramidite entfernen. Werden an diesen Positionen nicht vorgesehene Bausteine eingebracht (*insertions*), entstehen fehlerhafte Probe-Sequenzen. Auch der umgekehrte Fall ist möglich, bei dem trotz Bestrahlung die Schutzgruppe nicht entfernt wird und somit eine Base in der Probe-Sequenz fehlt (*deletions*). Da die Kopplungseffizienz der Phosphoramidite nicht bei 100% liegt ([73, 74]), kann es zu weiteren Fehlern im Probe-Aufbau kommen: Ist ein Baustein entschützt und die vorgesehene Base koppelt nicht an, sondern wird erst im nächsten Schritt durch eine ungewollte Base ersetzt, ergeben sich Probes mit Fehlstellen (MMs von engl. *mismatches*). Darüber hinaus existiert ein weiterer Reaktionsmechanismus bei Bestrahlung der NPPOC-Schutzgruppe mit UV-Licht. Dabei entstehen Nitroso-Verbindungen, die permanent gebunden bleiben und so die betreffende Probe von der weiteren Synthese ausschließen [75]. Durch eine fehlerhafte Fokussierung der Masken auf die Substratoberfläche und durch verunreinigte Chemikalien kann sich die Qualität der Synthese weiter verschlechtern.

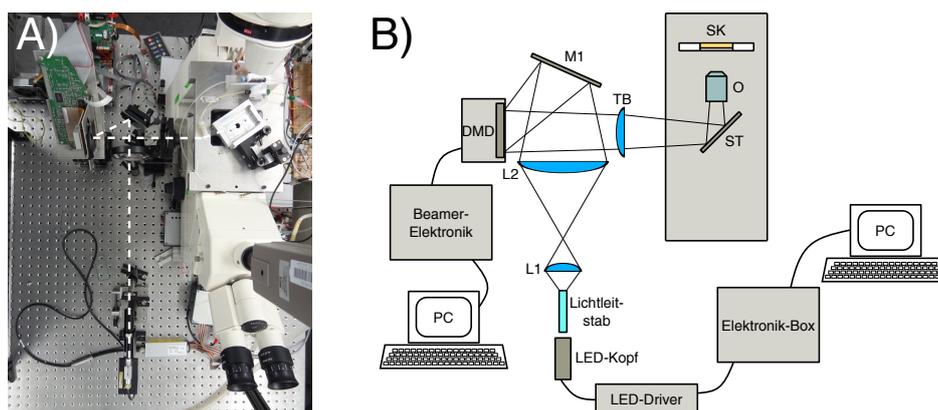


**Abbildung 3.3:** Funktionsweise des DMD (von engl. *digital micromirror device*, Abb. nach [1]). **A)** Die Spiegel des DMD können um  $\pm 10^\circ$  relativ zur Achse des DMD geneigt werden. Bei einer Neigung der Spiegel von  $+10^\circ$  wird einfallendes Licht E in die Mikroskopoptik O reflektiert (AN-Position). Sind die Spiegel  $-10^\circ$  geneigt, wird das einfallende Licht in eine Lichtfalle F gelenkt (AUS-Position). **B)** Hier ist der DMD inklusive der Mikroskopoptik dargestellt. Einfallendes Licht E wird mittels Linse L2 und DMD (bei Spiegel in AN-Position) durch die Tubuslinse TB gelenkt, trifft danach auf einen Strahlteiler ST und wird mit dem Mikroskopobjektiv O auf die Chipoberfläche fokussiert. Durch den Strahlteiler kann die Abbildung auf der Chipoberfläche zusätzlich mit einer CCD-Kamera betrachtet werden. Befinden sich Spiegel in der AUS-Position, wird das Licht in eine Lichtfalle F reflektiert.

### 3.1.2 Aufbau des Synthesizers

Der laboreigene DNA-Synthesizer wurde im Rahmen von [1] entwickelt und aufgebaut. Der Aufbau besteht aus einem Optik- und einem Fluidikteil, die beide über einen PC angesteuert werden. Den Fluidikteil bildet ein käuflich erworbener Oligonukleotidsynthesizer (ABI 381A), der für den Transport der benötigten Chemikalien während der Synthese zuständig ist.

Kernstück des optischen Teils des Synthesizers ist ein *digital micromirror device* (DMD<sup>TM</sup>, Texas Instruments), mit dessen Hilfe virtuelle Masken zur orts aufgelösten Belichtung der Chip-Oberfläche generiert werden: Der DMD besteht aus  $1024 \times 768 = 786432$  Einzelspiegeln, die um einen Winkel von  $\pm 10^\circ$  relativ zur Achse des DMD geneigt werden können, was in einer AN- und AUS-Position resultiert. Spiegel in der AN-Position reflektieren das einfallende Licht senkrecht zur Ebene des DMD in eine Mikroskopoptik, während Spiegel in der AUS-Position das Licht in eine Lichtfalle lenken (siehe Abb. 3.3 A)). Mit der Mikroskopoptik werden die virtuellen Masken auf die Substratoberfläche projiziert und fokussiert (siehe Abb. 3.3 B)). Die Fokussierung auf die Substratoberfläche wird mit einer CCD-Kamera, die am Mikroskop angeschlossen ist, vorgenommen. Im aktuellen Aufbau wird ein Feature auf dem Chip durch  $6 \times 6$  Spiegel des DMD erzeugt. Diese Einstellung gewährleistet in der Praxis einen guten Kompromiss zwischen Bildauflösung und -kontrast. Werden weniger Spiegel zur Erzeugung eines Features benutzt, verringert sich die Bildqualität und damit die Qualität der Synthese aufgrund von Streulicht und Beugungseffekten erheblich.



**Abbildung 3.4:** Optischer Aufbau des Synthesizers. **A)** Foto des Optikteils des Synthesizers mit eingezeichnetem Lichtweg (weiß gestrichelte Linien). **B)** Schemazeichnung des Optikteils inklusive aller Komponenten: Lichtquelle (Hochleistungs-LED inklusive Ansteuerungselektronik), Lichtleitstab, Linsen L1, L2 und Tubuslinse TB, Metallspiegel M1, DMD (inklusive Ansteuerungselektronik), Strahlteiler ST, Mikroskopobjektiv O und Synthesekammer SK. Weitere technische Details sind in Anhang B.4 gegeben.

Durch die Verwendung von virtuellen Masken zur Belichtung der Substratoberfläche ist die Flexibilität des Syntheseprozesses im Vergleich zur Verwendung von Chrommasken, wie sie bei manchen kommerziellen Microarray-Produkten eingesetzt werden, wesentlich größer, da die Masken je nach Anforderung an das Microarray auf unkomplizierte Art modifiziert werden können. Der Herstellungsprozess des Microarrays an sich ist aber deutlich langsamer [70].

Der optische Teil des Synthesizers wurde im Vergleich zu [1] im Rahmen dieser Arbeit modifiziert, um eine höhere Belichtungshomogenität der Microarray-Oberfläche zu erzielen. Der neue optische Aufbau ist dem eines handelsüblichen Beamers nachempfunden. Er besteht aus einer Hochleistungs-LED (UV-LED module LC-L2, Hamamatsu) der Wellenlänge  $\lambda = 365$  nm, einem Lichtleitstab (8 mm Fused Silica Hexagonal Light Pipe, Edmunds), mehreren Linsen, einem Metallspiegel, einem invertierten Mikroskop (Zeiss Axiovert 135 TV) zur Fokussierung des Lichts auf die Substratoberfläche und dem bereits erwähnten DMD. Dieser Aufbau hat im Vergleich zu dem in [1] vorgestellten mehrere Vorteile: Durch den Austausch der 250 W UHP Quecksilberlampe gegen die Hochleistungs-LED der korrekten Wellenlänge zur Deprotektion der Phosphoramidite müssen zum einen keine Filter verwendet werden, zum anderen verringert sich die thermische Belastung der Komponenten erheblich. Im alten Aufbau führte die Verwendung der UHP-Lampe zu einer inhomogenen Ausleuchtung des DMD und damit der Chip-Oberfläche, da sich die Lichtbogenstrecke der UHP-Lampe im Fokuspunkt eines Parabolspiegels befand und parallel zur optischen Achse ausgerichtet war. Dadurch ergab sich ein sehr inhomogenes Strahlprofil. Es wurde versucht das inhomogene Strahlprofil durch eine *intensity leveling mask*, ein invertiertes Bild des Lichtquellenprofils,

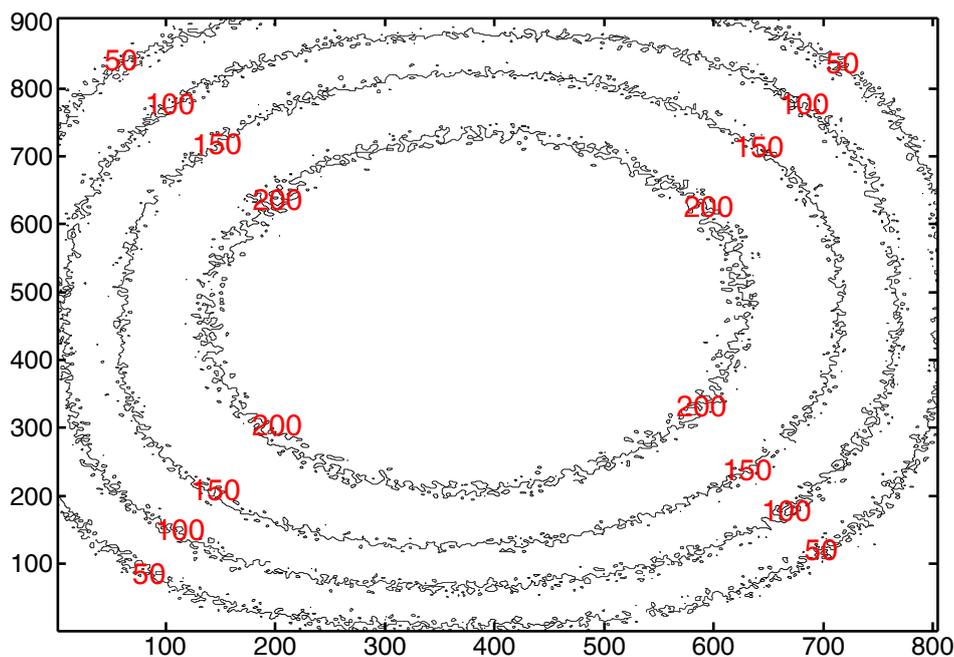
zu korrigieren. Im modifizierten Aufbau ist dies nicht nötig, da man durch die vielfache Totalreflexion des Lichts im Innern des Lichtleiters eine sehr homogene Leuchtfläche am Lichtleiterausgang erhält, die mit Linse L1 ( $f=30$  mm,  $\varnothing=25.4$  mm, UV Fused Silica, Thorlabs) ins Unendliche abgebildet wird. Mit Hilfe der Feldlinse L2 ( $f=150$  mm,  $\varnothing=75$  mm, UV Fused Silica, Thorlabs) und dem Metallspiegel M1 kann der DMD homogen ausgeleuchtet werden. Die Feldlinse L2 sorgt dafür, dass die Austrittspupille der vorhergehenden Stufe (Linse L1) auf die Eintrittspupille der nachfolgenden Stufe (Tubuslinse TB) abgebildet wird. Zur Projektion des DMD-Bildes auf die Chipoberfläche befindet sich der DMD in der Bildebene des invertierten Mikroskops, die außerhalb des Mikroskopkörpers liegt. Da das verwendete Mikroskopobjektiv (Zeiss Fluor  $5\times$ , 0.25 NA) unendlich korrigiert ist, muss zur Projektion des DMD-Bildes ins Unendliche eine Tubuslinse TB ( $f=164.5$  mm, Carl Zeiss) verwendet werden (siehe Abb. 3.4).

#### Homogenität der Belichtung

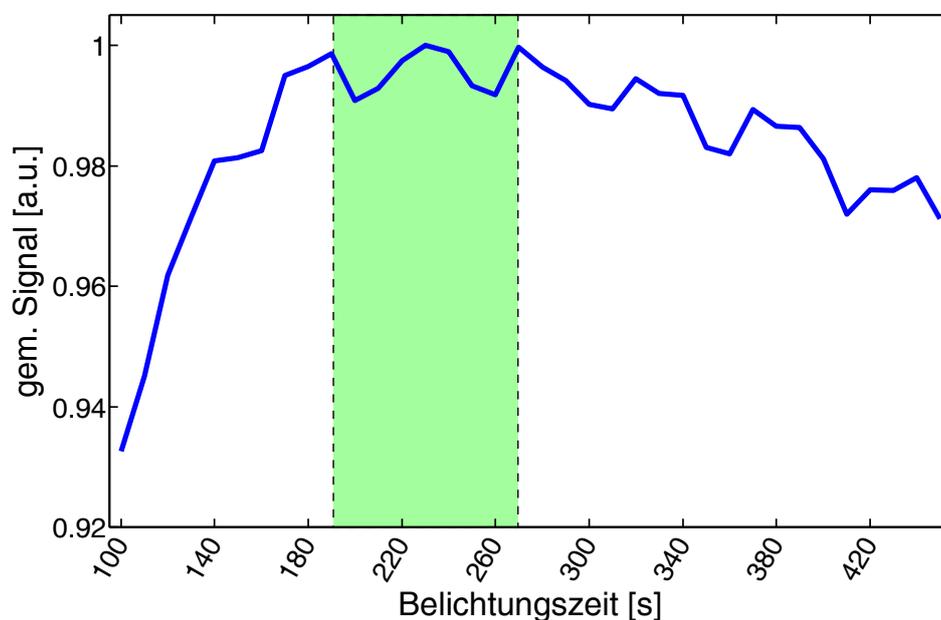
Um zu überprüfen, ob der modifizierte optische Aufbau des Synthesizers tatsächlich in einer homogenen Ausleuchtung der Chip-Oberfläche resultiert, wird das Bild des DMD (alle Spiegel in der AN-Position) mit Mikroskop-Objektiv im Strahlengang auf einen Schirm projiziert. Die Projektion wird mit einer Nikon Coolpix 4500 Digitalkamera fotografiert und am PC mittels MatLab ausgewertet. Das Ergebnis ist in Abb. 3.5 als Konturdiagramm dargestellt: Jede Linie kennzeichnet einen Bereich gleicher Helligkeit. Die Ausleuchtung der Chip-Oberfläche ist über einen großen Bereich in der Mitte des Bildes homogen und fällt zu den Rändern aufgrund Vignettierung ab: Reflektiertes Licht von den Ecken des DMD, die sich nah am Rand der Eintrittspupille befinden, wird durch die Apertur der Tubuslinse (und das Mikroskop-Objektiv) blockiert. Die Vignettierung durch das Objektiv ist im Vergleich zur Vignettierung durch die Tubuslinse vernachlässigbar. Aufgrund dieser unvermeidbaren Vignettierung wurde bereits im alten Aufbau nie die Gesamtfläche des DMD zur Erzeugung der Features ausgenutzt. Der nun homogen ausgeleuchtete mittige Bereich des DMD ist für die Chip-Synthese absolut ausreichend, sodass die Verwendung einer *intensity leveling mask* nicht mehr notwendig ist.

#### Ermittlung der Belichtungsdauer

Der modifizierte optische Aufbau macht eine erneute Ermittlung der optimalen Belichtungsdauer für eine erfolgreiche Abspaltung der Schutzgruppen der Phosphoramidite (Kopplungsschritt) notwendig. Die Intensität in der Fokusebene des Mikroskops



**Abbildung 3.5:** Helligkeitsverteilung auf der Chipoberfläche. Die Helligkeitsverteilung wird mit Objektiv im Strahlengang aufgezeichnet. Jede Linie des Konturdiagramms kennzeichnet einen Bereich gleicher Helligkeit. Gut zu erkennen ist, dass die Helligkeitsverteilung über einen großen Bereich in der Mitte der Abbildung homogen ist. Dies ist gerade der Bereich, der normalerweise für die Synthese des Chips genutzt wird. Endnahe Bereiche werden aufgrund von unvermeidbarer Vignettierung weniger stark ausgeleuchtet. Eine *intensity leveling mask* ist bei Verwendung des modifizierten Aufbaus nicht mehr notwendig.



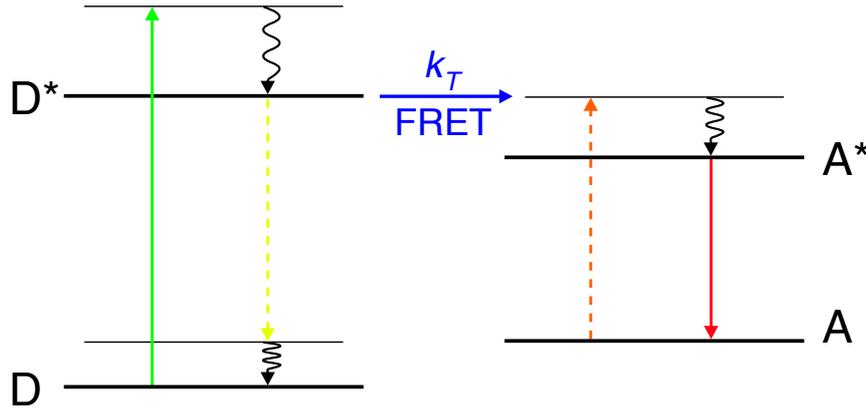
**Abbildung 3.6:** Fluoreszenzintensität der Features als Funktion der Belichtungszeit. Nach Anstieg der Signalintensität auf ein Maximum bei ca. 190 s bleibt das Signal bis ca. 270 s weitgehend konstant (grüner Bereich). Danach fällt die Signalstärke mit zunehmender Belichtungsdauer ab. Anhand des Graphen wird die Belichtungszeit für die Abspaltung der Schutzgruppen der Amidite auf 200 s festgelegt.

wird mit einem Laser Power Sensor (PS10Q, Coherent Inc.) gemessen. Wird ein komplett weißes Bild mittels des DMD in die Bildebene projiziert (alle Spiegel in der AN-Position), so ergibt sich die mittlere Intensität der Bildebene zu  $44 \text{ mW/cm}^2$  (im alten Aufbau betrug die Intensität  $77 \text{ mW/cm}^2$ ).

Die Auswirkung auf die Belichtungszeit pro Kopplung wird untersucht, indem Features der Sequenz Map2k18817 (siehe Anhang B.1) während der Synthese unterschiedlich lang belichtet werden. Die Belichtungszeit wird von 100 s auf 450 s in 10 s Schritten erhöht. Die Ergebnisse sind in Abb. 3.6 dargestellt. Dort sind die Signale der Features gegen ihre Belichtungszeit aufgetragen. Im Belichtungszeitfenster von 190 s bis 270 s (grüner Bereich) ist die Intensität am größten und nahezu identisch. Wird die Belichtungsdauer weiter erhöht, nimmt die Intensität ab und die Features werden unscharf. Die Belichtungsdauer wird für die Synthesen mit dem neuen optischen Aufbau auf 200 s festgesetzt (im alten Aufbau betrug die Belichtungszeit 160 s). Die alte Belichtungsdauer würde auch im neuen Aufbau zur erfolgreichen Deprotektion der Phosphoramidite ausreichen und in einem nur um ca. 2 % reduzierten Signal resultieren. Die geringere Intensität in der Bildebene ist daher für die Chip-Synthese von keinem großen Nachteil und wird durch die homogenere Belichtung (siehe oben) mehr als ausgeglichen.

#### 3.1.3 Datenauswertung

Nach Entfernung der restlichen Schutzgruppen der Phosphoramidite mit einer 1:1 Mischung aus Ethanol und Ethylendiamin nach Ende der Synthese wird das Array zur besseren Handhabung mit UV-Kleber (Norland Optical Adhesive 61, Edmund optics) auf einem Metallplättchen fixiert. Danach kann der Chip mit einer Hybridisierungskammer verschraubt und die Kammer mit der zu untersuchenden Target-Lösung befüllt werden. Die Hybridisierungskammer besteht aus einem Ein- und Auslasskanal für die Target-Lösung, einer Heizfolie und einem PT-100 Widerstand (siehe [1]). Sowohl Heizfolie als auch Widerstand sind mit einem PC verbunden, der über ein mit *ProfiLab Expert* geschriebenes Programm die Temperaturregelung übernimmt. Die Target-Lösung besteht aus fluoreszenzmarkierter einzelsträngiger DNA, die in meisten Hybridisierungsexperimenten in 5xSSPE-Puffer (siehe Anhang B.2) gelöst ist. Da sich die Target-DNA durch Hybridisierung mit den Probes an den Features akkumuliert, entsteht ein messbares Fluoreszenzsignal. Nach Erreichen des thermodynamischen Gleichgewichts werden Bilder der Microarray-Oberfläche mittels eines Fluoreszenzmikroskops (IX81, Zeiss) und einer CCD-Kamera (EM-CCD C9100-02, Hamamatsu) aufgenommen. Zur exakten Positionierung der Hybridisierungskammer ist eine mikrometeregenaue Positionierungsplattform der Firma Märzhäuser installiert. Die Steuerung dieser Komponenten erfolgt über die Software *Simple PCI*.



**Abbildung 3.7:** Jablonski-Diagramm der Energieniveaus von Donator- und Akzeptor-Fluorophor. Wird der Donator angeregt (grüner Pfeil), relaxiert anschließend durch innere Konversion zum angeregten Zustand D\* und von dort in seinen Grundzustand D (grünelber Pfeil), so kann die dabei frei werdende Energie strahlungsfrei einem in der Nähe befindlichen Akzeptor-Fluorophor mit der Ratenkonstante  $k_T$  transferiert werden. Dadurch wird der Akzeptor seinerseits angeregt (oranger Pfeil). Dieser Energieübertag wird mit FRET (Förster-Resonanz-Energietransfer) bezeichnet. FRET kann aufgrund der Energieerhaltung nur zwischen resonanten Übergängen gleicher Energie stattfinden. Relaxiert der Akzeptor in seinen Grundzustand (roter Pfeil), so äußert sich das in der Fluoreszenz des Akzeptors.

Anhand der gewonnenen TIFF-Bilder können die Fluoreszenzsignale mit dem Java-Programm *ScanArray* in Intensitäten umgewandelt und schließlich mit MatLab analysiert werden. Alle Details finden sich in [1, 76].

## 3.2 Förster-Resonanz-Energietransfer

Im Folgenden sollen die für diese Arbeit wichtigsten theoretischen Aspekte des Förster-Resonanz-Energietransfer (FRET) kurz vorgestellt werden. Für weiterführende Literatur sei auf [77–82] verwiesen.

Relaxiert ein angeregter elektronischer Zustand D\* eines Donator-Moleküls in den Grundzustand und transferiert seine Energie dabei strahlungsfrei zu einem in der Nähe befindlichen Akzeptor-Molekül A, so spricht man von FRET [77]. Um zu betonen, dass der Akzeptor ein Fluorophor ist, wird FRET häufig auch als Fluoreszenz-Resonanz-Energietransfer bezeichnet. Der Energietransfer zwischen resonanten Übergängen erfolgt mittels einer Dipol-Dipol-Wechselwirkung mit der Ratenkonstanten  $k_T$  (siehe Abb. 3.7). Befinden sich Donator und Akzeptor im Abstand  $r$ , so ergibt sich im einfachsten Fall die Ratenkonstante  $k_T$  eines Resonanz-Energietransfers zu [80]:

$$k_T(r) = \frac{\kappa^2 \Phi_D^0}{\tau_D^0 r^6} \left( \frac{9000 \ln(10)}{128 \pi^5 N_A n^4} \right) \int_0^\infty I_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda \quad (3.1)$$

Dabei ist  $\kappa^2$  der Orientierungsfaktor zwischen den Übergangsdipolen von Donator und Akzeptor,  $\Phi_D^0$  die Fluoreszenzquantenausbeute des Donators in Abwesenheit des Akzeptors,  $\tau_D^0$  die Lebensdauer des Donators in Abwesenheit des Akzeptors,  $N_A$  die Avogadro-Konstante,  $n$  der Brechungsindex des Mediums in dem Wellenlängenbereich, in dem der Überlapp von Emissionsspektrum des Donators und Absorptionsspektrum des Akzeptors signifikant ist,  $I_D(\lambda)$  das normalisierte Fluoreszenzspektrum des Donators ( $\int_0^\infty I_D(\lambda) d\lambda = 1$ ) in Abwesenheit des Akzeptors,  $\epsilon_A$  der molare Absorptionskoeffizient des Akzeptors und  $\lambda$  die Wellenlänge.

### Förster-Radius

Gl. 3.1 kann umgeschrieben werden zu:

$$k_T(r) = \frac{1}{\tau_D^0} \left( \frac{R_0}{r} \right)^6 \quad (3.2)$$

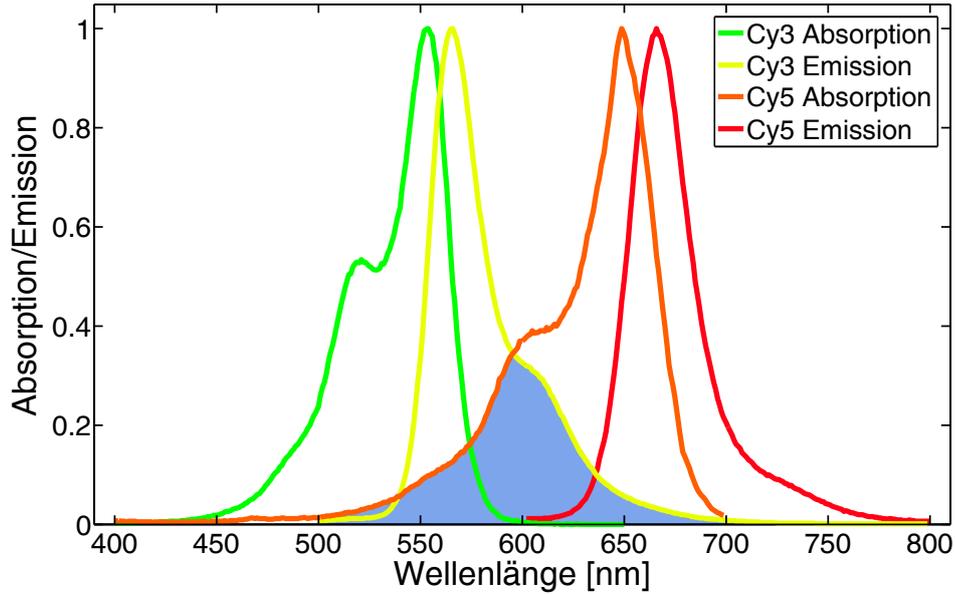
$R_0$  hat die Dimension einer Länge und wird als Förster-Radius bezeichnet. Es gilt:

$$R_0 = \kappa^2 \Phi_D^0 \left( \frac{9000 \ln(10)}{128\pi^5 N_A n^4} \right) \int_0^\infty I_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda \quad (3.3)$$

Ist  $r = R_0$ , so ergibt sich für die Ratenkonstante  $k_T = \frac{1}{\tau_D^0}$ . Der Förster-Radius entspricht also gerade dem kritischen Abstand, bei dem der Energietransfer und der spontane Zerfall der angeregten Donator-Moleküle durch die üblichen Relaxationsmechanismen gleich wahrscheinlich sind. Der Förster-Radius liegt typischerweise zwischen 2 nm und 9 nm [79].

### Spektraler Überlapp

Damit ein Energieübertrag stattfinden kann, müssen die beiden Fluorophore so gewählt werden, dass die vom Donator emittierte Energie vom Akzeptor absorbiert werden kann. Dazu müssen sich Teile des Emissionsspektrums des Donators mit Teilen des Absorptionsspektrums des Akzeptors überlappen. Dies ist am Beispiel des Donator-Fluorophors Cy3 und des Akzeptor-Fluorophors Cy5 in Abb. 3.8 gezeigt. Der Überlapp des Emissions- und Absorptionsspektrums ist in blau markiert. Je größer dieser Überlapp ist, desto besser funktioniert der Energietransfer. Dieser spektrale Überlapp wird durch das Überlappungsintegral  $J(\lambda)$  beschrieben.



**Abbildung 3.8:** Absorptions- und Emissionsspektrum der Fluorophore Cy3 und Cy5. Überlappen sich Teile des Emissionsspektrums des Donators mit Teilen des Absorptionsspektrums des Akzeptors (hier blau markiert), kann es zu einem Energieübertrag kommen. Der Energieübertrag ist umso effektiver, je größer der spektrale Überlapp der beiden Spektren ist.

Es gilt:

$$J(\lambda) = \int_0^{\infty} I_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda \quad (3.4)$$

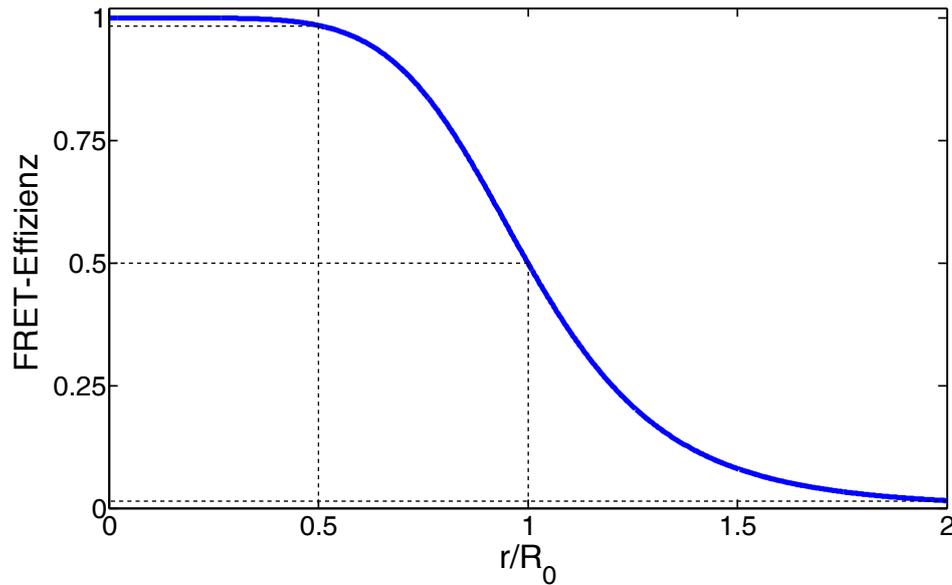
Es wird angemerkt, dass das Überlappungsintegral nicht einfach nur die Fläche unter den beiden sich überschneidenden Spektren darstellt. Durch den Faktor  $\lambda^4$  werden auch die Wellenlängen berücksichtigt, bei denen sich die Spektren überlappen, wodurch größere Wellenlängen wesentlich stärker gewichtet werden.

### FRET-Effizienz

Die FRET-Effizienz  $E(r)$  ist definiert als das Verhältnis zwischen angeregten Donator-Molekülen, die ihre Energie zu einem Akzeptor transferieren  $\#_{D^* \rightarrow A}$ , und der Anzahl aller angeregten Donator-Molekülen  $\#_{D^*}$ .

$$E(r) = \frac{\#_{D^* \rightarrow A}}{\#_{D^*}} = \frac{k_T}{\frac{1}{\tau_D} + k_T} \quad (3.5)$$

Im zweiten Schritt wurden die entsprechenden Raten eingesetzt: Die FRET-Effizienz ist gerade das Verhältnis aus Übergangsrates und gesamter Zerfallsrate des Donators. Ist die Übergangsrates  $k_T$  schneller als die Relaxationsrate  $\frac{1}{\tau_D}$  des Donators in den Grundzustand bei Abwesenheit des Akzeptors, so ist der Energietransfer effizient. Umgekehrt



**Abbildung 3.9:** FRET-Effizienz  $E(r)$  als Funktion des Verhältnisses zwischen Abstand  $r$  und Förster-Radius  $R_0$ . Beträgt die FRET-Effizienz bei einem Verhältnis von  $r/R_0 = 0.5$  noch nahezu 1, so hat sie bereits bei  $r/R_0 = 1$  auf 0.5 abgenommen. Bei  $r/R_0 = 2$  ist ein Energieübertrag kaum noch möglich. Diese starke Abhängigkeit der FRET-Effizienz vom Abstand  $r$  macht Messungen mit FRET hochempfindlich.

ist der Prozess ineffizient. Mit Gl. 3.2 ergibt sich

$$E(r) = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6} \quad (3.6)$$

Die FRET-Effizienz  $E(r)$  hängt also empfindlich vom Abstand zwischen Donator und Akzeptor ab. Dies ist in Abb. 3.9 noch einmal veranschaulicht: Während bei  $r = 0.5 \cdot R_0$  die FRET-Effizienz noch ca. 98 % beträgt, ist sie bei  $r = 2 \cdot R_0$  schon auf 1.5 % abgefallen. Diese starke Abstandsabhängigkeit macht Untersuchungen mittels FRET zu einer sehr empfindlichen Messmethode. Eine mögliche praktische Anwendung wird im folgenden Abschnitt vorgestellt.

### 3.3 Fluoreszenzkorrelationsspektroskopie

Mit Hilfe der Fluoreszenzkorrelationsspektroskopie (FCS von engl. *fluorescence correlation spectroscopy*) können Gleichgewichtsfluktuationen von dynamischen molekularen Ereignissen, wie z.B. der Diffusion oder der Konformationsänderung von Biomolekülen, durch statistische Analyse auf nicht invasive Art und Weise untersucht werden. Koppelt man ein Fluoreszenzsignal an die verschiedenen Zustände des zu untersuchenden Systems, so erzeugen spontane Fluktuationen der Systemzustände Variationen im

Fluoreszenzsignal. Die Autokorrelationsfunktion dieser Fluoreszenzfluktuationen enthält unter anderem Informationen über die charakteristischen Zeitskalen des Systems. Die FCS-Technik wurde bereits 1972 von Madge *et al.* zur Untersuchung der Diffusion und Bindung von Ethidiumbromid an doppelsträngige DNA erfunden [83] und 1993 von Rigler *et al.* durch die Einführung von konfokaler Belichtung weiterentwickelt [84]. Der dort beschriebene experimentelle Aufbau wird so ähnlich auch in dieser Arbeit verwendet (siehe Abb. 3.10).

Wird die FCS-Technik mit FRET (siehe Abschnitt 3.2) kombiniert, kann die Dynamik intramolekularer Prozesse studiert werden. So können beispielsweise Gleichgewichtsfluktuationen von Basenpaaren intern markierter doppelsträngiger DNA aufgrund der sehr empfindlichen Abhängigkeit der FRET-Effizienz vom Abstand von Donator- und Akzeptor-Fluorophor (Gl. 3.6), die sich in Fluktuationen des Fluoreszenzsignals manifestieren, mit Hilfe der FCS analysiert werden. In dieser Arbeit werden zur Messung der Fluktuationsdynamik doppelsträngiger DNA in Lösung die Probenmoleküle mit dem Fluorophor Cy3 und die Targets entsprechend mit dem Quencher Cy5 intern markiert. Für die Autokorrelationsfunktion  $G_{it}(t)$  der Fluoreszenzintensität  $I(t)$  der DNA-Duplexe gilt:

$$G_{it}(t) = \frac{\langle I(0)I(t) \rangle_t - \langle I(0) \rangle^2}{\langle I(0) \rangle^2} \quad (3.7)$$

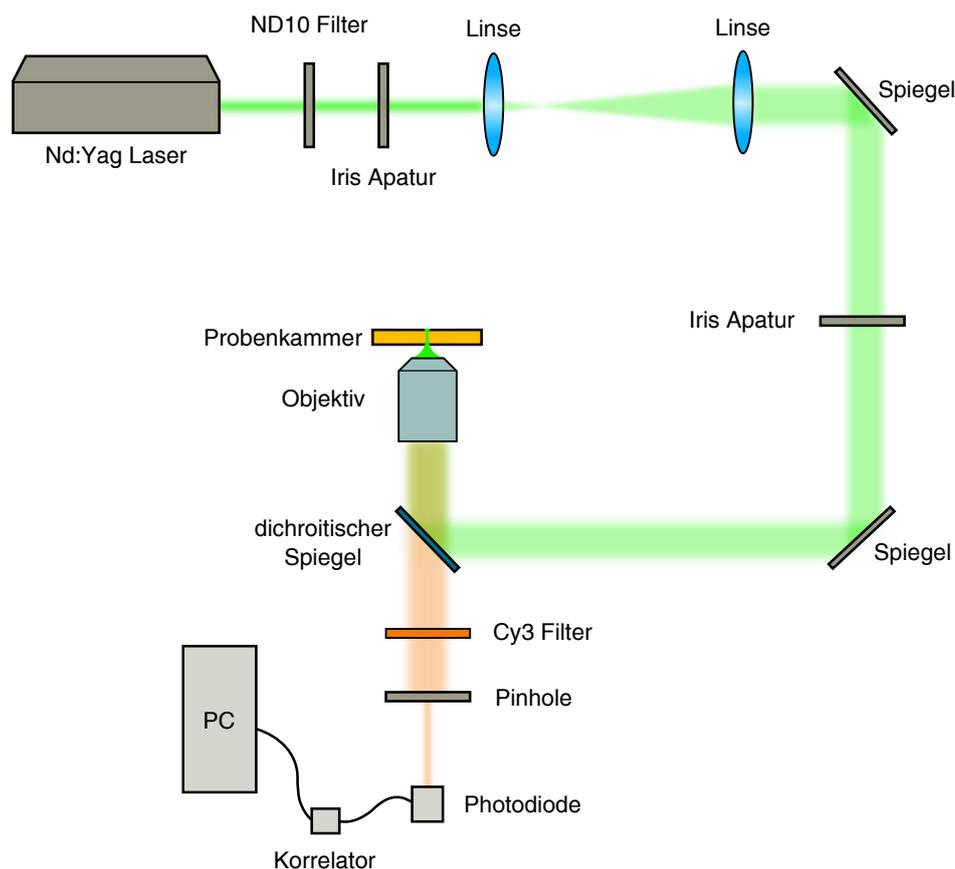
Um den Beitrag der Diffusion der DNA-Duplexe zu eliminieren, werden wie in [85] beschrieben neben den intern mit Cy5 markierten Target-Molekülen auch Targets ohne Markierung verwendet. So wird nur der Beitrag der Diffusion  $G_{Diffusion}(t)$  gemessen. Für die diffusionsbefreite Autokorrelationsfunktion  $G(t)$  gilt dann:

$$G(t) = \frac{G_{it}(t)}{G_{Diffusion}(t)} \quad (3.8)$$

Ein guter Überblick über die Geschichte der FCS-Technik, die theoretischen Grundlagen und Anwendungen ist in [86] gegeben. Für weiterführende Literatur zur Photonenkorrelation sei auf [87] verwiesen.

### Experimenteller Aufbau

Die Autokorrelationsfunktionen  $G(t)$  der Basenpaardynamik werden mit dem experimentellen Aufbau in Abb. 3.10 untersucht. Als Anregungslicht für die Cy3-Fluorophore wird ein frequenzverdoppelter Nd:Yag Laser ( $\lambda = 532nm$ ) verwendet. Der Laserstrahl wird zuerst mittels zweier Linsen aufgeweitet und durch den hinteren Mikroskopport in eine Mikroskopoptik (Axiovert 135, Zeiss, Deutschland) eingekoppelt. Der Laser wird



**Abbildung 3.10:** Experimenteller Aufbau des FCS-Experiments. Der Laserstrahl eines Nd:YAG Lasers ( $\lambda = 532 \text{ nm}$ ) wird durch die beiden Linsen aufgeweitet und mit dem Mikroskopobjektiv in die Probenkammer fokussiert. Die markierten DNA-Duplexe in der Probenkammer werden so zur Fluoreszenz angeregt, das emittierte Licht durch das selbe Objektiv gesammelt und auf ein  $50 \text{ }\mu\text{m}$  Pinhole fokussiert. Ein dichroitischer Spiegel und ein Cy3-Filter lassen nur emittiertes Licht passieren. Die Emission wird von einem Photonenzähler detektiert und von einem Korrelator analysiert. Das Korrelatorsignal wird von einem PC aufgenommen. Photo-Bleaching und Triplet-Formation der Fluorophore wird durch die Begrenzung der Anregungsleistung des Lasers durch einen ND10-Filter vermieden.

mit einem Wasserimmersionsobjektiv (C-Apochromat 40X, NA 1.2, Zeiss, Deutschland) auf die Probenkammer fokussiert. Die Kammer besteht aus einer PDMS-Dichtung, die oben und unten von jeweils einem Deckglässchen bedeckt ist. Der Aufbau der Probenkammer entspricht im Wesentlichen dem aus Abschnitt 3.1.3. Die Temperatur der Kammer wird über einen Pt100-Widerstand ausgelesen und mit einer Heizfolie kontrolliert. Das von der Probe emittierte Licht wird mit dem selben Objektiv gesammelt und auf ein Pinhole ( $\text{\O} 50 \text{ }\mu\text{m}$ ) fokussiert. Ein dichroitischer Spiegel und ein Cy3-Filter sorgen dafür, dass nur das von den Duplexen emittierte Licht auf das Pinhole fokussiert wird. Das Licht wird von einem Photonenzähler (H8259-01, Hamamatsu) detektiert und das Signal mittels eines Hardware-Korrelators (Flex 99r480, Correlator.com) analysiert. Die Autokorrelationsfunktion  $G(t)$  und der Photonenstrom werden als Funktion

der Zeit von einem PC ausgenommen. Um Photo-Bleaching und Triplet-Formation der Fluoreszenzfarbstoffe zu vermeiden, ist die Anregungsleistung des Laser durch einen ND10-Filter begrenzt. Da der Beitrag des *afterpulsing* des Photonenzählers (ca. 3 $\mu$ s) mindestens eine Größenordnung schneller ist als die eigentlich zu messenden Basenpaarfluktuationen, muss es nicht in Betracht gezogen werden.

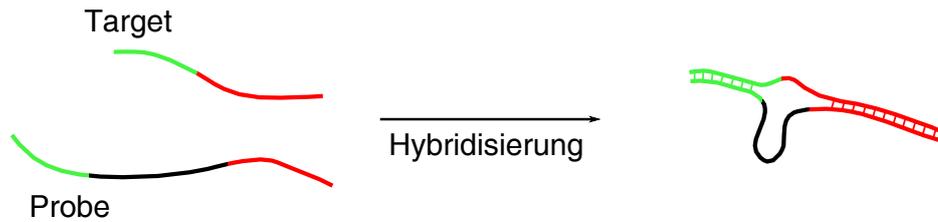
Die DNA-Moleküle werden vor den Messungen in Standardreaktionsgefäßen vorbereitet. Um zu gewährleisten, dass alle Proben mit Targets hybridisiert sind, beträgt das Probe-Target-Verhältnis 1:2. Die Gefäße werden in kochendem Wasser inkubiert und relaxieren über Nacht auf Raumtemperatur. Dies stellt sicher, dass die Hybridisierung den Gleichgewichtszustand erreicht. Kurz vor der Messungen werden die Target-Lösungen in die Probenkammer gegeben.



## 4 Stabilität von DNA-Duplexen mit einzelsträngigen Loop-Strukturen

In früheren Arbeiten [1–3, 88] wurde der Einfluss von Einzelfehlstellen (MMs, von engl. *mismatches*) auf das Hybridisierungsverhalten kurzer DNA-Stränge untersucht, bei denen Probe und Target jeweils die gleiche Länge hatten. In realen Microarray-Experimenten aber ist die Längenverteilung der Target-Sequenzen oft nur schwer kontrollierbar, da die verwendeten Targets häufig durch PCR-Methoden synthetisiert werden. Dadurch kann es zur Hybridisierung von Sequenzen ungleicher Länge kommen, wodurch der Hybridisierungsprozess aufgrund nicht-kanonischer Bindungskonformationen wesentlich komplexer wird. Um das Verständnis der DNA-Hybridisierung weiter zu vertiefen, ist es daher notwendig die Bedingung der gleichen Längen von Probe und Target aufzugeben und den Einfluss der Längendifferenz zwischen Probe und Target auf die Hybridisierung zu untersuchen.

Um eine Längendifferenz zwischen den hybridisierenden Strängen zu generieren, werden im Weiteren gezielt zusätzliche Basen in Probe- oder Target-Sequenz eingebracht. Da die Stränge abgesehen von den zusätzlichen Basen perfekt komplementär zueinander sind, können so bei geeigneter Einbringung der Basen einzelsträngige Schleifenstrukturen, sog. *Loops*, während der Hybridisierung entstehen. Die Ausbildung eines Loops in der Probe ist in Abb. 4.1 dargestellt. Die zusätzlichen Basen (schwarz) sind dort zwischen einem roten und grünen Teilbereich, die jeweils komplementär zu den entsprechenden Teilsequenzen im Target sind, in die Ausgangssequenz der Probe eingebettet. Probe und Target können in diesem Fall nur dann hybridisieren, wenn der schwarze Teilbereich als Loop ausgebildet wird. Variiert man die Anzahl der zusätzlichen Basen und die Position, an der sie in die Probe eingefügt werden, können Loops verschiedener Länge und -Position in kontrollierter Weise generiert werden. Die Einbringung eines Loops in die Probe hat den Vorteil, dass mit Hilfe von DNA-Microarrays und einer einzigen Target-Spezies eine große Anzahl verschiedener Kombinationen von Loop-Länge und -Position parallel in einem Experiment untersucht werden kann, ohne dass es zu einer Konkurrenzsituation verschiedener Targets in Lösung kommt. Bei Einbau des Loops in das Target besteht der Vorteil, dass die synthetisierten Probes



**Abbildung 4.1:** Loop-Formation bei Hybridisierung von Probe und Target. Der grüne und rote Bereich der Probe ist komplementär zu dem entsprechenden Bereich des Targets. Hybridisieren die beiden DNA-Stränge, so wird der schwarze Bereich, der die zusätzlich eingefügten Basen enthält, als Loop ausgebildet.

auf der Microarray-Oberfläche die gleiche Länge haben. Jedoch muss bei Variation von Loop-Länge oder -Position immer eine neue Target-Spezies verwendet werden. Will man auch hier eine große Anzahl verschiedener Kombinationen von Loop-Länge und -Position parallel untersuchen, müssen viele verschiedene Target-Spezies gleichzeitig zur Hybridisierung der Probes verwendet werden. Dadurch kann es zu einer ungewollten Konkurrenz um Bindungsstellen kommen, die den eigentlich zu untersuchenden Einfluss von Loops auf das Hybridisierungsverhalten verfälschen können.

Im Folgenden soll der Einfluss von einzelsträngigen Loop-Strukturen auf die Hybridisierung untersucht werden.

## 4.1 Loops in der Probe

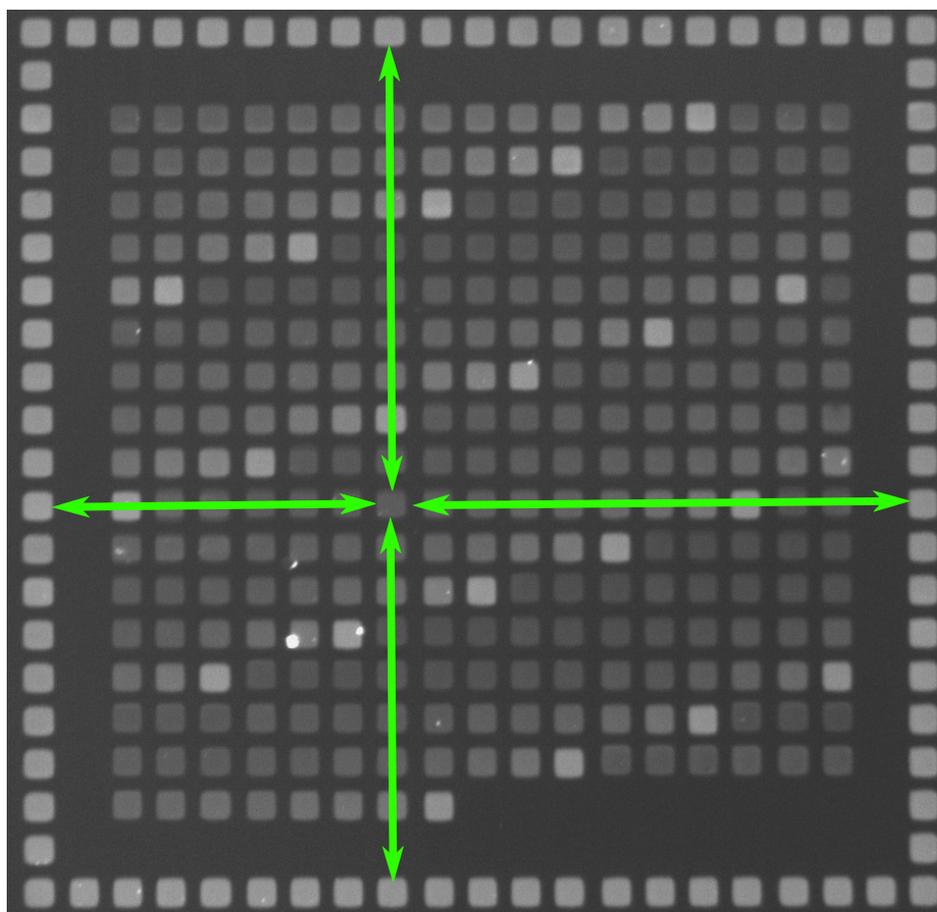
### 4.1.1 Experiment

Um einen Loop in der Probe zu erzeugen, wird eine zusätzliche poly-T-Sequenz, die zum Target nicht komplementär ist, in das Sequenzmotiv des PM eingefügt. Die Ausgangssequenz des PM ist die Sequenz Map2k18833c-reverse (Länge 33 Basen, Anhang B.1).

Es werden poly-T-Sequenzen bis zu einer Länge  $L$  von 13 Basen realisiert, die an 20 verschiedenen Positionen  $P$  entlang des Strangs in die PM-Sequenz eingebracht werden (in Abschnitt 4.2 werden Loops anderer Sequenz untersucht). Die so entstehenden 260 verschiedenen Probe-Sequenzen ergeben zusammen mit 20 weiteren PM-Features (zur Kontrolle der Synthesqualität) einen Featureblock. Tab. 4.1 gibt einen Überblick über die verwendeten Probe-Sequenzen. Der Featureblock ist zusätzlich noch von PM-Features umgeben, die der Kontrolle der Belichtungshomogenität während der Synthese dienen. Jedes Feature innerhalb des Featureblocks kann so anhand des Mittelwertes der vier Signale der entsprechenden PM-Features skaliert und eventuell auftretende Belichtungsgradienten - zumindest linear - beseitigt werden. Dies wird durch die grünen Pfeile

**Tabelle 4.1:** Synthetisierte Probe-Sequenzen. Die aufgeführten Probe-Sequenzen bilden die Features auf der Oberfläche des DNA-Microarrays. Die rot markierten Buchstaben kennzeichnen die zusätzlich eingefügten Basen. Insgesamt entstehen so 261 verschiedene Features.

Loop-Position	Loop-Länge	Probe-Sequenz (3'-5')	Länge (Basen)
PM Probe		TTCAATACTACTCATAATTACCAACAACATTAC	33
7	1	3'-TTCAATA <sup>T</sup> CTACTCATAATTACCAACAACATTAC-5'	34
7	2	3'-TTCAATA <sup>TT</sup> CTACTCATAATTACCAACAACATTAC-5'	35
⋮	⋮	3'-TTCAATA <sup>TT...</sup> CTACTCATAATTACCAACAACATTAC-5'	⋮
7	13	3'-TTCAATA <sup>TTTTTTT</sup> CTACTCATAATTACCAACAACATTAC-5'	46
8	1	3'-TTCAATACT <sup>T</sup> TACTCATAATTACCAACAACATTAC-5'	34
8	2	3'-TTCAATACT <sup>TT</sup> TACTCATAATTACCAACAACATTAC-5'	35
⋮	⋮	⋮	⋮
26	13	3'-TTCAATACTACTCATAATTACCAACA <sup>TTTTTTTTTTT</sup> TACATTAC-5'	46



**Abbildung 4.2:** Skalierung der Hybridisierungssignale anhand der PM-Umrandung. Die Fluoreszenzintensität jedes Features kann anhand des Mittelwertes der vier umliegenden PM-Features (grüne Pfeile) skaliert werden. Synthesebedingte Belichtungsgradienten werden so linear korrigiert.

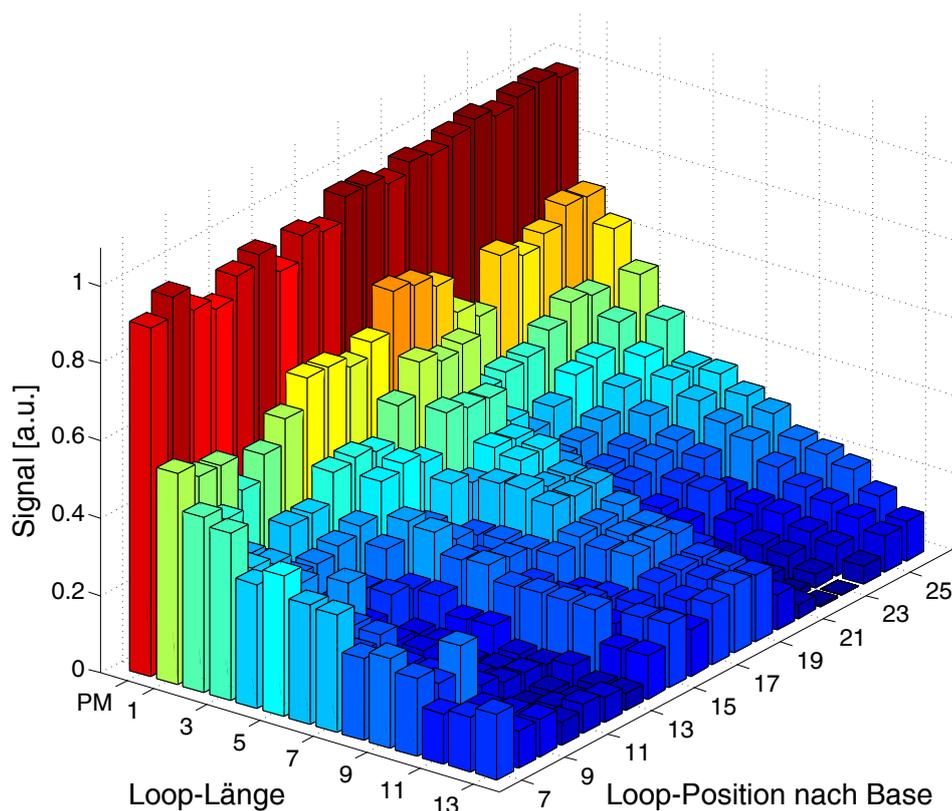
in Abb. 4.2 angedeutet. Zur Erhöhung der statistischen Signifikanz der Daten und der Minimierung des experimentellen Fehlers wird der Featureblock insgesamt an vier verschiedenen Positionen der Microarray-Oberfläche synthetisiert. Experimentelle Fehler können beispielsweise durch Belichtungsgradienten während der Synthese, Verschmutzungen oder Inhomogenitäten der Oberfläche auftreten. Um die endgültigen Daten zu erhalten, wird jeder Featureblock anhand der PM-Umrandung skaliert und dann der Mittelwert der vier Featureblöcke gebildet.

### 4.1.2 Ergebnisse und Diskussion

Die Hybridisierungssignale des endgültigen Datensatzes sind als Funktion der Loop-Länge und -Position in Abb. 4.3 gezeigt. Stärkstes und schwächstes Hybridisierungssignal sind jeweils auf 1 bzw. 0 normiert. Loop-Position  $P$  bedeutet dabei, dass die zusätzlichen poly-T-Basen nach Base  $P$  der Ausgangssequenz eingefügt werden (Zählrichtung ist vom 3'- zum 5'-Ende). Die Hybridisierungstemperatur beträgt 44 °C.

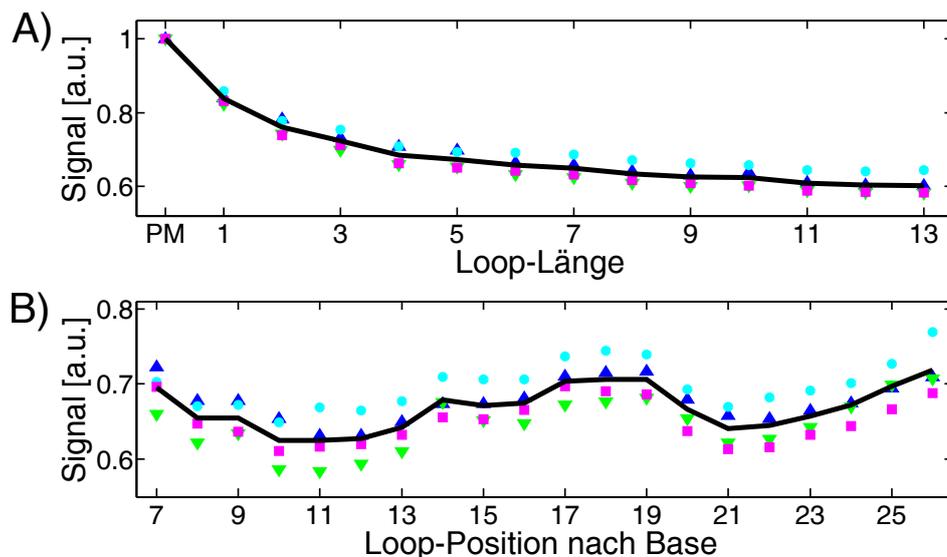
Um die Abhängigkeit der Hybridisierungssignale von der Loop-Länge besser sichtbar zu machen, sind die Daten in Abb. 4.4 A) über alle Loop-Positionen gemittelt. Das Hybridisierungssignal des PM ist wieder auf 1 normiert. Man erkennt einen monotonen Abfall der Signalintensität mit zunehmender Loop-Länge. Das Einfügen einer einzelnen zusätzlichen Base reicht bereits aus um das Signal auf 85% des PM-Signals zu reduzieren. Ein Loop der Länge 13 (größte Anzahl zusätzlicher Basen im Experiment) minimiert das Signal auf 60%. Der große Einfluss der zusätzlichen Basen auf das Hybridisierungssignal lässt sich durch die kurze Sequenzlänge erklären: Bei kurzen Duplexen ist die bei der Hybridisierung generierte freie Energie entsprechend klein, sodass Fehlstellen große Auswirkungen auf die Stabilität des Duplexes haben. Die Abhängigkeit des Hybridisierungssignals von der Loop-Länge kann auch durch entropische Betrachtungen erklärt werden. Die Anzahl der Konformationen der Probe wächst mit zunehmender Länge der zusätzlich eingefügten Basen an. Dies hat zur Folge, dass beim Übergang vom flexiblen denaturierten Zustand in die steife Doppelhelix eine immer größere Entropieänderung notwendig ist. Diese große Entropieänderung kann durch die bei der Bindung frei werdende Enthalpie nicht mehr kompensiert werden. Die freie Energie des Duplexes nimmt entsprechend ab. Erschwerend hinzu kommt, dass nicht alle Probe-Konformationen eine Duplex-Initialisierung zulassen und somit auch die Duplexinitialisierungsenergie von der Längendifferenz zwischen Probe und Target abhängig ist. Dies wird detaillierter in Abschnitt 6.1.2 besprochen.

Abb. 4.4 B) zeigt die Abhängigkeit des Hybridisierungssignals als Funktion der Loop-Position nach Mittelung über alle Loop-Längen. Auch hier wird das Signal des PM auf 1 normiert. Das entstehende „Defektprofil“ ist symmetrisch bezüglich der Mitte des



**Abbildung 4.3:** Hybridisierungssignale als Funktion der Loop-Länge und der Loop-Position. Die Loop-Länge  $L$  wird zwischen 0 (PM) und 13 zusätzlichen Thymin Basen variiert, die Loop-Position  $P$  zwischen 7 und 26.  $P$  bedeutet dabei, dass die zusätzlichen Basen nach Base  $P$  der Ausgangssequenz der Probe eingefügt werden (Zählrichtung ist vom 3'- zum 5'-Ende). Das stärkste Signal ist auf 1, das schwächste auf 0 normiert. Die Hybridisierungstemperatur beträgt 44 °C.

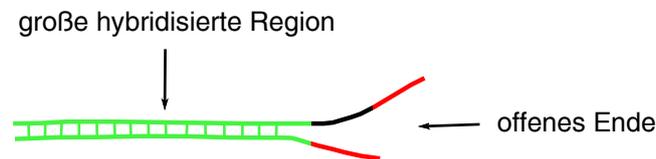
Duplexes. Das Signal ist am stärksten für mittige und endnahe Loop-Positionen. Zwischen diesen beiden Maxima, ungefähr drei bis vier Basen von der Mitte des Duplexes entfernt, sind die gemessenen Signale am schwächsten. Der Unterschied zwischen dem stärksten und schwächsten Signal beträgt nur ungefähr 10%, was im Vergleich zur Signalabhängigkeit von der Loop-Länge relativ gering ist. Vergleicht man das hier gemessene Defektprofil mit der Positionsabhängigkeit des Signals im Falle von MMs, siehe z.B. [3], so ergibt sich ein gravierender Unterschied. Im Falle von MMs erhält man zwar auch ein symmetrisches Defektprofil, jedoch ist der Einfluss der MMs bei mittigen Positionen am größten, sodass hier das schwächste Signal auftritt und das Signal bei MM-Positionen am Ende des Duplexes wieder ansteigt. Dies deutet darauf hin, dass im Falle der Loops der Signalabhängigkeit von der Loop-Position andere Mechanismen zugrunde liegen als im Falle der MMs. Die Loop-Positionsabhängigkeit kann qualitativ verstanden werden (siehe Abb. 4.5). Loops in der Nähe eines Endes des Duplexes haben wenige potentielle Bindungsstellen in Richtung des Endes. Das führt dazu, dass diese Loops meistens nicht geschlossen sind und der Duplex ein offenes Ende hat, ein



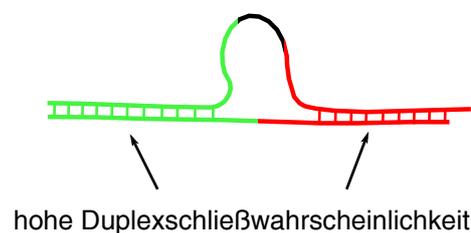
**Abbildung 4.4:** Experimentell bestimmte Hybridisierungssignale von DNA-Duplexen mit Loop. Symbole: Featureblock 1 - blaue Dreiecke; Featureblock 2 - cyanblaue Kreise; Featureblock 3 - grüne Dreiecke; Featureblock 4 - magentarote Vierecke; experimenteller Mittelwert - schwarze Linie. **A)** Hybridisierungssignale als Funktion der Loop-Länge (von Loop-Länge 0 (PM) bis Loop-Länge 13) in Basen nach Mittelung über alle Loop-Positionen und Normierung auf das PM-Signal. **B)** Hybridisierungssignale als Funktion der Loop-Position nach Mittelung über alle Loop-Längen. Loop-Position 7 bedeutet beispielsweise, dass der Loop nach Base 7 der Ausgangssequenz eingefügt wird. Die Zählrichtung ist von der Oberfläche weg (3'-5'-Richtung).

sog. *dangling end*. Die andere Seite des Duplexes kann jedoch eine große hybridisierte Region mit entsprechend großer freier Energie ausbilden, sodass der Duplex insgesamt eine hohe Stabilität und damit ein gut messbares Hybridisierungssignal aufweist (Abb. 4.5 A)). Befindet sich der Loop in einer mittigen Position, so stehen rechts und links des Loops viele Bindungsstellen zur Verfügung. Dies resultiert in einem geschlossenen Loop und einer ebenfalls hohen Duplexstabilität (Abb. 4.5 B)). Wie experimentell beobachtet nimmt das Hybridisierungssignal zwischen den oben genannten Extremen ein Minimum an. In einem solchen Fall stehen dem Loop weniger Bindungspartner zur Verfügung als in einer mittigen Position und die maximale Länge des hybridisierten Duplexbereiches ist ebenfalls kleiner als bei Loops in der Nähe des Duplexendes (Abb. 4.5 C)). Die theoretische Beschreibung des Einflusses von einzelsträngigen Loop-Strukturen auf das Hybridisierungsverhalten wird in Abschnitt 6.1 vorgestellt (siehe auch [89, 90]).

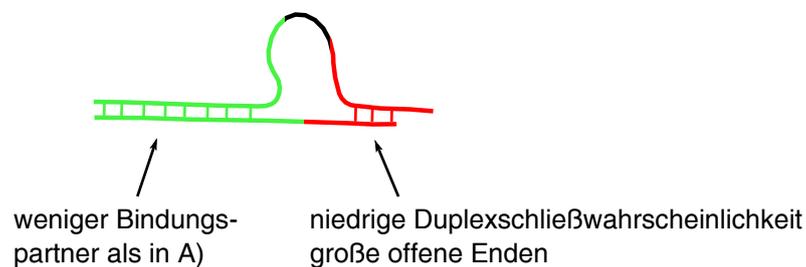
## A) Endposition



## B) mittige Position



## C) Position zwischen Mitte und Ende



**Abbildung 4.5:** Abhängigkeit des Hybridisierungssignals von der Loop-Position. **A)** Die Einbringung eines Loops nahe am Duplexende bewirkt, dass sich der Duplex an diesem Ende nicht schließen kann und somit ein offenes Ende vorliegt. Dies wird jedoch durch die große grüne Region kompensiert, die eine hohe Duplexstabilität und ein gut messbares Hybridisierungssignal liefert. **B)** Befindet sich der Loop in der Mitte des Duplexes, so ist der Duplex mit hoher Wahrscheinlichkeit aufgrund der Vielzahl an potentiellen Bindungspartnern auf beiden Loop-Seiten geschlossen. Dies generiert eine hohe Duplexstabilität. **C)** Befindet sich der Loop in einer Position zwischen Mitte und Ende des Duplexes, so ist der Duplex weniger stabil als in A) und B), da zum einen weniger Bindungspartner auf einer Seite des Loops zur Verfügung stehen und zum anderen die Wahrscheinlichkeit von großen offenen Enden sehr hoch ist.

## 4.2 Ergänzende Untersuchungen zum Einfluss der Loops in der Probe

### 4.2.1 Einfluss der Oberfläche

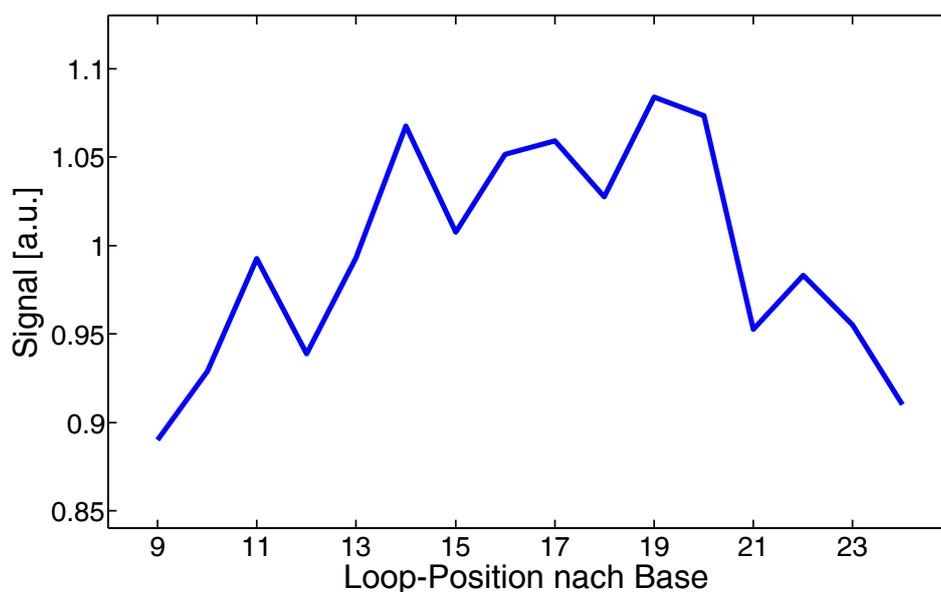
Laut zahlreicher Literaturangaben (siehe z.B. [91]) besteht eine große Diskrepanz zwischen der Hybridisierung in Lösung und der auf Oberflächen. So unterscheidet sich beispielsweise die Hybridisierungskinetik aufgrund elektrostatischer Effekte und sterischer Hinderung auf Oberflächen deutlich. Auch ist die Übergangsregion der Langmuir-Absorptionsisothermen aufgrund unvermeidbarer Synthesefehler während der *in situ*-Synthese auf Microarrays verbreitert. Obwohl wie in [3] gezeigt die Hybridisierung auf den hier verwendeten Microarray-Oberflächen trotzdem mit Hilfe der in Lösungsexperimenten gewonnenen NN-Energien [53] beschrieben werden kann, soll im Folgenden der Einfluss der Microarray-Oberfläche auf das Hybridisierungsverhalten von Probes mit Loop untersucht werden.

Dazu wird das Experiment aus Abschnitt 4.1 mit der umgekehrten Probe-Sequenz wiederholt: Das 5'-Ende der Sequenz Map2k18833c-reverse aus Abschnitt 4.1 entspricht gerade dem 3'-Ende der hier verwendeten Sequenz Map2k18833c und umgekehrt (siehe Anhang B.1). Aufgrund von Platzmangel auf dem Chip werden die zwischen 1 und 13 Basen langen poly-T-Sequenzen aber nur an 16 verschiedenen Positionen in die PM-Sequenz eingefügt, so dass sich 208 verschiedene Probe-Sequenzen ergeben.

Abb. 4.6 zeigt die gemessenen Hybridisierungssignale als Funktion der Loop-Position nach Mittelung über alle Loop-Längen bei einer Temperatur von 20 °C. Das Hybridisierungssignal ist für mittige Loop-Positionen am größten und fällt für endnahe Positionen ab. Vergleicht man das erhaltene Loop-Positionsprofil mit dem Profil aus Abschnitt 4.1, so stellt man fest, dass sich bei Umkehrung der Probe-Sequenz keine wesentlichen Veränderungen im Signalverlauf festmachen lassen. Der Einfluss der Oberfläche auf das Hybridisierungsverhalten der Duplexe mit Loop ist daher vernachlässigbar. Dies ist auch ein wichtiger Aspekt für die theoretische Beschreibung des Einflusses von Loops in Abschnitt 6.1. Dort werden die beiden Enden des Duplexes gleichberechtigt behandelt und keine Unterscheidung in oberflächennahes und -fernes Ende getroffen.

### 4.2.2 Loops anderer Sequenzen

Um Loops in der Probe zu generieren, wurden bislang immer zusätzliche poly-T-Sequenzen in die Ausgangssequenz des PM eingefügt. Um die Untersuchung von Loop-Strukturen in der Probe auszuweiten, werden im Folgenden Loops mit anderer Sequenz

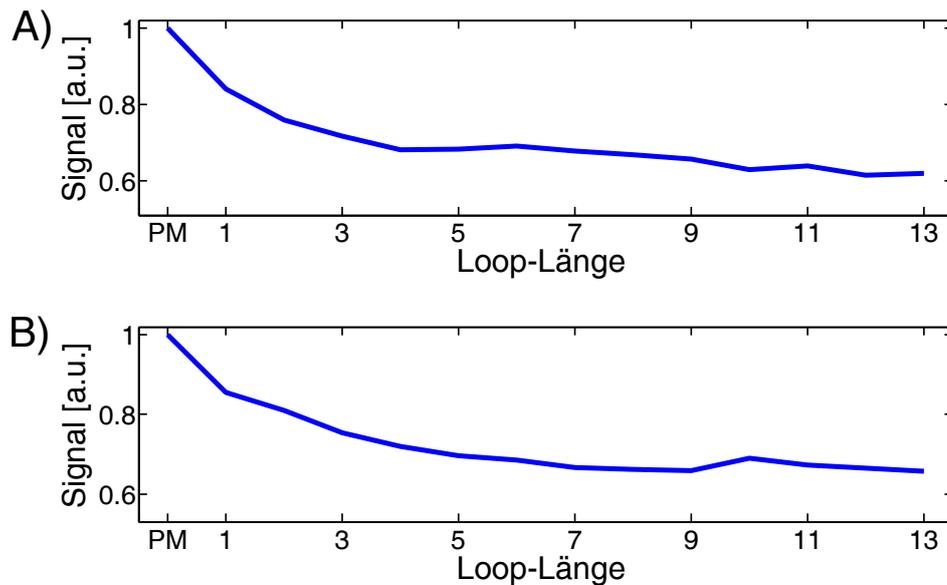


**Abbildung 4.6:** Einfluss der Microarray-Oberfläche. Zur Untersuchung des Einflusses der Microarray-Oberfläche wird ein Chip mit der zu Map2118833c-reverse umgekehrten Probe-Sequenz Map2k178833c synthetisiert. Hier ist das Hybridisierungssignal als Funktion der Loop-Position nach Mittelung über alle Loop-Längen bei einer Temperatur von 20 °C gezeigt. Die Loop-Position gibt an, nach welcher Base der Ausgangssequenz der Loop eingefügt wird. Es kann kein signifikanter Unterschied im Vergleich zum Signalverlauf bei Hybridisierung mit der Sequenz Map2118833c-reverse festgestellt werden.

untersucht. Dazu werden analog zu den vorherigen Experimenten poly-C bzw. beliebige Sequenzen (ausgehend vom Sequenzmotiv „CAGT“) zwischen 1 und 13 Basen Länge in die Ausgangssequenz Map2k18833c-reverse des PM eingefügt.

Poly-A-Sequenzen können zur Erzeugung der Loops nicht gewählt werden: Um Abstand zwischen der eigentlichen Probe-Sequenz und der Dendrimer-Oberfläche der Microarrays zu erzielen, wird vor jeder Synthese auf der gesamten zur Verfügung stehenden Fläche die Base Thymin an die Oberfläche gekoppelt (siehe Abschnitt 3.1). Bei Verwendung von poly-A-Sequenzen wäre daher eine Bindung der Probes an die Oberfläche sehr wahrscheinlich. Auch poly-G-Sequenzen eignen sich zur Generierung von Loops nur bedingt, da es aufgrund der molekularen Eigenschaften von Guanin schwierig ist eine große Anzahl von Guanin-Basen hintereinander zu synthetisieren. Guanin hat zum einen mit 151.13 g/mol die größte molare Masse aller Basen und nimmt durch seinen zusätzlichen Substituenten wesentlich mehr Platz in Anspruch als die zweite Purinbase Adenin.

Aufgrund von Platzmangel auf dem Chip werden nur 3 verschiedene Positionen für das Einbringen der Loops gewählt: nach Base 9, 16 bzw. 23 der PM-Sequenz (Zählrichtung ist wieder vom 3'- zum 5'-Ende). Ein Überblick über die verwendeten Probe-Sequenzen ist in Anhang A.1 gegeben. Abb. 4.7 A) zeigt das Hybridisierungssignal als Funktion

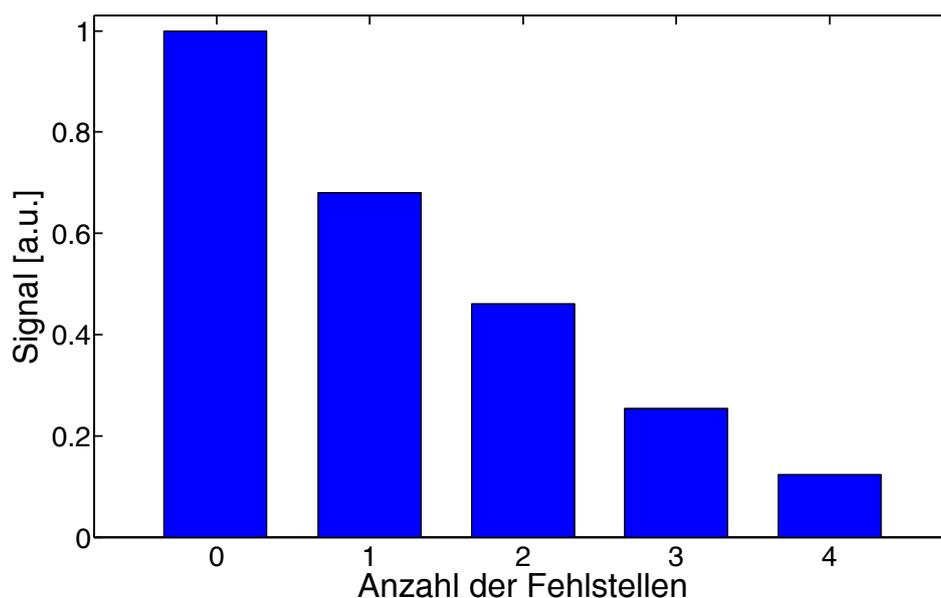


**Abbildung 4.7:** Loops anderer Sequenzen. Gezeigt sind die Hybridisierungssignale als Funktion der Loop-Länge nach Mittelung über alle Loop-Position und Normierung auf das PM-Signal. **A)** poly-C-Loops **B)** beliebige Sequenzen

der Loop-Länge nach Mittelung über alle Loop-Positionen für poly-C-Loops, Abb. 4.7 B) für Loop-Sequenzen ausgehend vom Sequenz-Motiv „CAGT“ bei einer Temperatur von 44 °C. Es ergibt sich für das Hybridisierungssignal als Funktion der Loop-Länge die gleiche Abhängigkeit wie in Abschnitt 4.1. Das Signal nimmt mit zunehmender Loop-Länge monoton ab. Die bereits in Abschnitt 4.1 beobachtete Abhängigkeit des Hybridisierungssignals von der Loop-Länge ist also nicht nur für poly-T-Sequenzen gültig. Der zugrundeliegende Mechanismus, der für die Signalabnahme bei zunehmender Loop-Länge verantwortlich ist, ist also universeller Natur.

### 4.2.3 Vorversuch zur theoretischen Beschreibung in Abschnitt 6.1

In Abschnitt 6.1 wird der Einfluss von Loops auf das Hybridisierungssignal mit Hilfe eines Gleichgewichtsmodells theoretisch untersucht. Zur korrekten Beschreibung müssen Synthesefehler in den Probes, die während der *in situ*-Synthese unvermeidbar auftreten, im Modell berücksichtigt werden. Dabei wird angenommen, dass die Synthesefehler einer Binominalverteilung folgen. Um den Rechenaufwand des Modells zu minimieren, soll im Folgenden die maximale Anzahl der Synthesefehler pro Probe, die im Modell berücksichtigt werden muss, experimentell bestimmt werden. Dazu werden in die Ausgangssequenz Map2k18833c-reverse an verschiedenen Positionen absichtlich Fehlstellen eingebaut, indem die sich dort befindenden Basen durch Adenin ersetzt werden (befindet sich in der Ausgangssequenz bereits Adenin an der betroffenen Stelle, so wird



**Abbildung 4.8:** Einfluss von Fehlstellen auf das Hybridisierungssignal. Gezeigt sind die Mittelwerte der Feature-Signale als Funktion der Anzahl der Fehlstellen nach Subtraktion des Hintergrundes und Normierung auf das PM-Signal (0 Fehlstellen). Die Hybridisierungstemperatur beträgt 44 °C. Die Signale nehmen mit wachsender Anzahl von Fehlstellen monoton ab. Das Signal einer Probe mit 4 Fehlstellen beträgt beispielsweise nur noch 12% des PM-Signals.

Adenin durch Thymin ersetzt). Dadurch würden sich bei einer 33 Basen langen Sequenz theoretisch  $2^{33} - 1$  verschiedene Probe-Sequenzen ergeben, die sich von der gewünschten Probe-Sequenz unterscheiden. Aufgrund der enormen Anzahl an möglichen Probe-Sequenzen und der Tatsache, dass Probes mit einer großen Anzahl von Synthesefehlern extrem unwahrscheinlich sind, werden 8 Positionen in der Probe-Sequenz ausgewählt, an denen Einzelfehlstellen eingebaut werden. Es werden an maximal 4 der 8 Positionen gleichzeitig Fehlstellen eingefügt. Dadurch ergeben sich 162 verschiedene Probe-Sequenzen, die sich entweder durch die Position oder die Anzahl der Fehlstellen unterscheiden.

Abb. 4.8 zeigt die nach der Anzahl der Fehlstellen geordneten gemittelten Signale der Features bei Hybridisierung mit der Target-Sequenz Map2k18833c-reverse bei einer Temperatur von 44 °C. Der Hintergrund wird subtrahiert und alle Signale auf das PM-Signal normiert. Null Fehlstellen kennzeichnen dabei das PM-Feature. Wie zu erwarten nimmt das Signal mit zunehmender Fehlstellenanzahl monoton ab. Der Signalmittelwert der Features mit 4 Fehlstellen beträgt nur noch 12% des PM-Signals. Die Wahrscheinlichkeit für das Auftreten von Probes mit mehr als 3 Synthesefehlern beträgt bei einer Synthesefehlerwahrscheinlichkeit von  $p = 0.084$  (siehe Abschnitt 6.1) weniger als 30%. Dadurch ergibt sich bei der Berücksichtigung von maximal 3 Synthesefehlern ein numerischer Fehler von weniger als 4% (siehe Gl. (6.3) und Abschnitt 6.1.2). Da der experimentelle Fehler größer ist, können somit Probe-Sequenzen mit 4

oder mehr Fehlern in der theoretischen Beschreibung vernachlässigt werden.

## 4.3 Grenze der Hybridisierung - Loops im Target

Die monotone Abnahme des Hybridisierungssignals mit zunehmender Loop-Länge in Abschnitt 4.1 wirft die Frage auf, bis zu welcher Längendifferenz zwischen Probe und Target noch eine erfolgreiche Hybridisierung auf dem Microarray stattfinden kann. Dies ist unter anderem auch wichtig für die maximale Anzahl an Sequenzen, die prinzipiell auf einem Microarray unterschieden werden können. Die Bestimmung der Informationsentropie kollektiver DNA-Bindung auf einem Microarray und damit des physikalischen Limits der molekularen DNA-Erkennung ist ein zentraler Aspekt der Forschung im Arbeitskreis.

Um die Grenze der Hybridisierung als Funktion der Längendifferenz untersuchen zu können, müssen immer längere Probes synthetisiert werden. Jedoch ist die maximale Länge der Probes, die auf Oberflächen mit hoher Qualität noch synthetisiert werden können, aufgrund der *in situ*-Synthese nicht beliebig. Bei kommerziellen Anbietern wie Affymetrix beträgt die maximale Länge beispielsweise 25 Basen [92]. Bei der bisher maximalen Längendifferenz von 13 Basen in Abschnitt 4.1 mussten bei der Ausgangssequenz Map2k18833c-reverse bereits Probes von 46 Basen Länge synthetisiert werden. Da es nicht sinnvoll ist die Länge der Probes auf der Microarray-Oberfläche weiter zu erhöhen, werden die zusätzlichen Basen zum Erzeugen noch größerer Längendifferenzen im Folgenden im Target eingebettet (da die zusätzlichen Basen bei Hybridisierung hier immer einen Loop ausbilden, ist die Untersuchung der maximalen Längendifferenz mit der Untersuchung der maximalen Loop-Länge gleichzusetzen).

### 4.3.1 Experiment

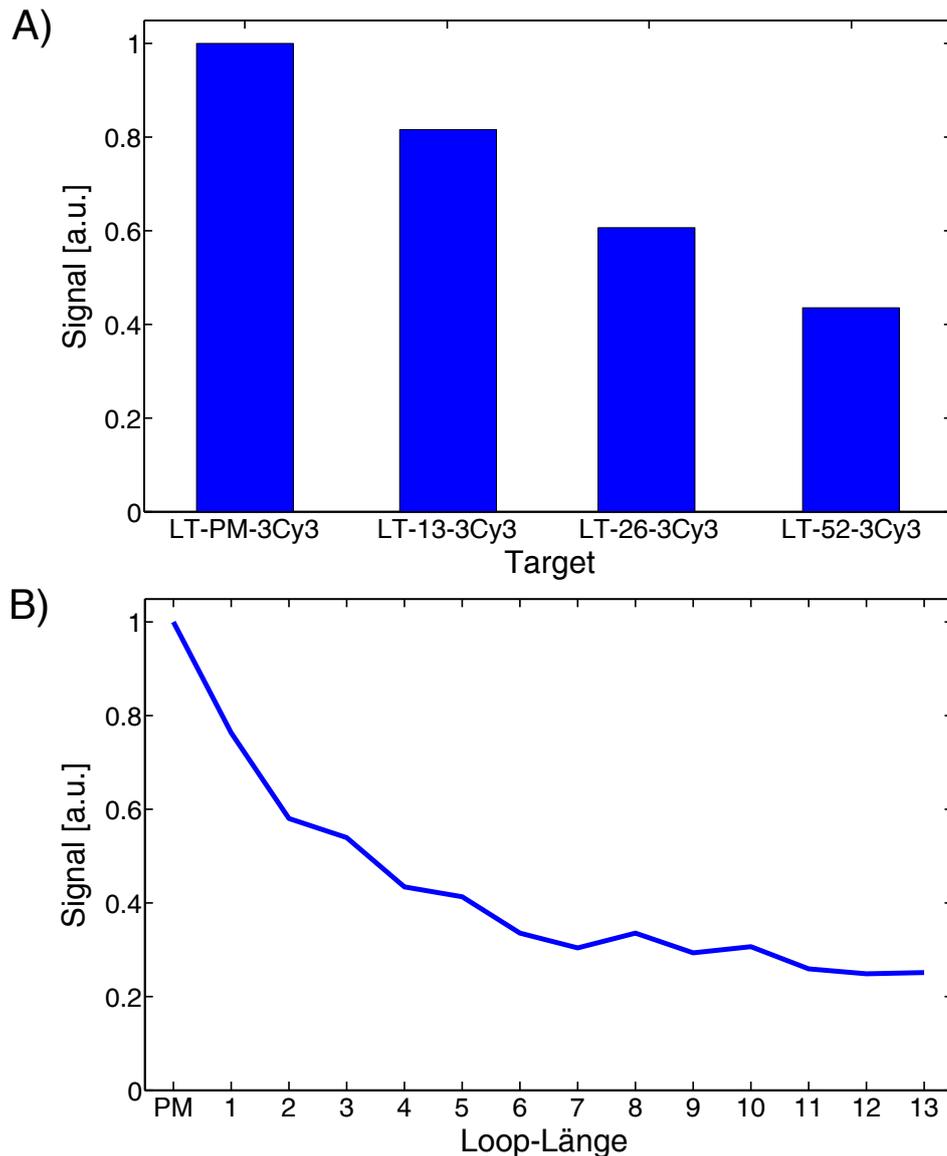
Um die gewonnenen Daten bei Einbringung eines Loops in das Target mit denen aus Abschnitt 4.1 „Loops in der Probe“ vergleichbar zu machen, muss darauf geachtet werden, dass die Sequenzumgebung des Loops unverändert bleibt. So wird sichergestellt, dass die NN-Paare der Duplexe in beiden Fällen identisch sind. Dies hat zur Folge, dass die Ausgangssequenz des Targets für die Einbettung eines Loops im vorliegenden Experiment der Ausgangssequenz der Probes im Abschnitt 4.1 entsprechen muss: Vertauscht man die Enden der Ausgangssequenz der Probes Map2k18833c-reverse, so erhält man die hier verwendete Ausgangssequenz der Targets LT-PM-3Cy3. Die Probe-Sequenz wird so zur Target-Sequenz (siehe Anhang B.1). Da bei jeder Änderung von

Loop-Länge oder -Position eine weitere Target-Spezies kommerziell erworben werden muss, werden nur drei weitere Target-Spezies mit den Loop-Längen 13, 26 bzw. 52 entworfen. Die entsprechenden Target-Sequenzen sind: LT-13-3Cy3, LT-26-3Cy3 und LT-52-3Cy3. Die Loop-Position nach Base 16 der Ausgangssequenz der Targets (Zählrichtung vom 3'- zum 5'-Ende) stimmt bei allen drei Target-Sequenzen überein. Es wird eine mittige Loop-Position gewählt, da die vorherigen Experimente gezeigt haben, dass Probes mit mittigem Loop am stabilsten sind und damit das stärkste Signal aufweisen. Die Loop-Sequenz besteht wie in Abschnitt 4.1 aus poly-T-Basen. Die vier Target-Spezies werden einzeln nacheinander auf den Chip „Loops im Target“ gebracht. Die Sequenz der Features auf der Oberfläche entspricht immer der zu LT-PM-3Cy3 komplementären Sequenz.

Das experimentelle Vorgehen ist wie folgt: Die erste Target-Lösung wird in die Kammer gegeben und die Temperatur auf 80 °C erhöht. Nach 2 min wird die Temperatur auf 44 °C abgesenkt. Dadurch sind die Targets homogen in der Kammer verteilt und es kommt nicht zu Inhomogenitäten durch den Einfüllprozess. Nach Erreichen der 44 °C werden die Probes 105 min lang hybridisiert und im Anschluss daran die Daten aufgenommen. Danach wird die Temperatur wieder auf 80 °C erhöht, die alte Target-Lösung ausgespült und die Kammer insgesamt 4 min mit Pufferlösung gereinigt. So wird sichergestellt, dass kein Restsignal auf dem Chip verbleibt und die nachfolgenden Hybridisierungen nicht verfälscht werden. Nachdem die mit Puffer gefüllte Kammer über Nacht bei 44 °C inkubiert wurde, erfolgt die Kontrolle des Restsignals. Danach kann die Prozedur mit einer neuen Target-Lösung wiederholt werden. Der Chip wird in der Reihenfolge LT-PM-3Cy3, LT-13-3Cy3, LT-26-3Cy3 und LT-52-3Cy3 hybridisiert. Zur weiteren Vergleichbarkeit der Ergebnisse wird der Chip „Loops in der Probe“ aus Abschnitt 4.1, der die Untersuchung des Einflusses von Loops in der Probe zur Aufgabe hatte, noch einmal synthetisiert. Der Loop wird hier jedoch nur nach Base 16 der Ausgangssequenz Map2k18833c-reverse (Zählrichtung wie immer vom 3'- zum 5'-Ende) in die Probes eingebracht. Das experimentelle Vorgehen ist hier analog dem oben beschriebenen.

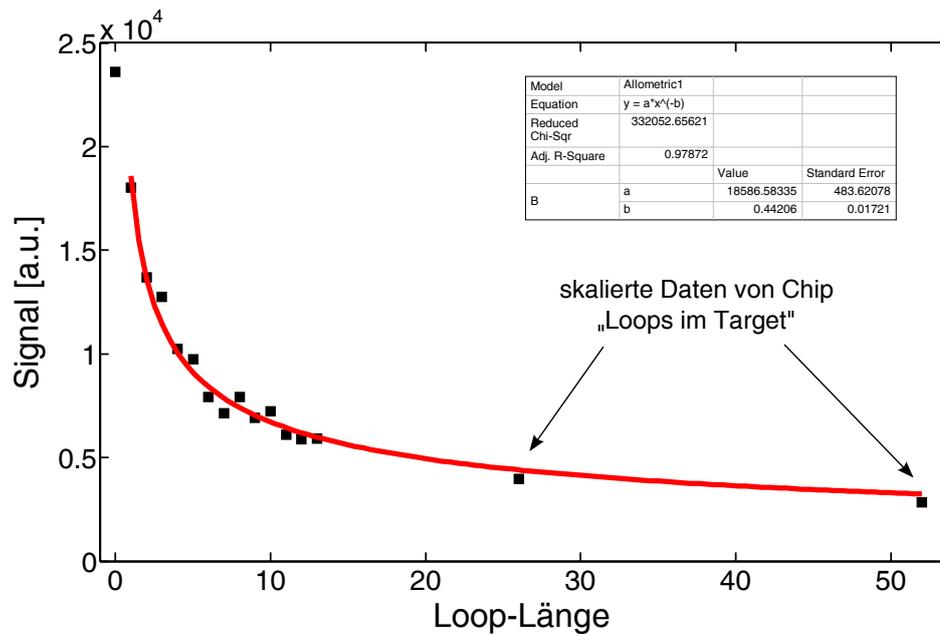
### 4.3.2 Ergebnisse und Diskussion

Abb. 4.9 A) zeigt die Signale der Features des Chips „Loops im Target“ nach Hybridisierung mit den entsprechenden Target-Sequenzen nach Subtraktion des Hintergrundes bei einer Temperatur von 44 °C. Die Signale sind zusätzlich auf das PM-Signal (Hybridisierung mit LT-PM-3Cy3) normiert. Wie zu erwarten, nimmt das Signal mit zunehmender Loop-Länge kontinuierlich ab. So erreicht die Hybridisierung mit LT-52-3Cy3 nur noch ca. 44% des PM-Signals. Abb. 4.9 B) zeigt die Signale des Chips „Loops in



**Abbildung 4.9:** Vergleich der Hybridisierungssignale bei Einbringung der Loops in Probe und Target. **A)** Signale bei Hybridisierung mit den verschiedenen Target-Sequenzen nach Subtraktion des Hintergrundes und Normierung auf das PM-Signal (das PM-Signal entspricht der Hybridisierung mit LT-PM-3Cy3). Zur Generierung der Loops werden zusätzliche poly-T-Sequenzen verschiedener Länge in die Targets eingefügt. So enthält beispielsweise LT-52-3Cy3 52 zusätzliche Thymin Basen. LT-PM-3Cy3 ist perfekt komplementär zu den Probes auf der Oberfläche. Das Signal nimmt mit zunehmender Loop-Länge monoton ab und erreicht bei der Hybridisierung mit LT-52-3Cy3 ein Minimum von ca. 44% des PM-Signals. **B)** Hybridisierungssignale als Funktion der Loop-Länge nach Subtraktion des Hintergrundes und Normierung auf PM-Signal. Hier werden die zusätzlichen Basen in die Probes auf der Oberfläche eingebracht und der Chip mit der Target-Spezies Map2k18833c-reverse hybridisiert. Es ergibt sich eine monotone Signalabnahme mit zunehmender Loop-Länge, wobei das Signal bei einer Loop-Länge von 13 Basen nur noch 25% des PM-Signals entspricht. Vergleicht man dieses Signal mit dem entsprechenden Signal, das bei der Hybridisierung mit LT-13-3Cy3 in Teil A) erzielt wurde, so ergibt sich ein deutlicher Unterschied, obwohl identische Duplexe bei der Hybridisierung entstehen.

der Probe“ als Funktion der Loop-Länge nach Hybridisierung mit der Target-Sequenz Map2k18833c-reverse bei einer Temperatur von 44 °C. Auch hier wird der Hintergrund subtrahiert und auf das PM-Signal normiert. Vergleicht man das Signal der Hybridisierung mit LT-13-3Cy3 aus Abb. 4.9 A) mit dem sich ergebenden Signal bei einer Loop-Länge von 13 Basen aus Abb. 4.9 B), so stellt man fest, dass die beiden Signale sich mit 82% bzw. 25% des PM-Signals deutlich unterscheiden, obwohl die bei der Hybridisierung entstehenden Duplexe eigentlich identisch sein sollten. Da die Sequenzen, die auf der Oberfläche synthetisiert werden, bei beiden Chips nicht identisch sind, können sich die auftretenden Synthesefehler unterschiedlich auf die Stabilität der verschiedenen Probes auswirken und die Signalstärke beider Chips asymmetrisch beeinflussen. Der Einfluss der Synthesefehler wird in Abschnitt 4.4.1 detailliert besprochen. Ein weiterer Grund für die unterschiedlichen Signale kann in der sterischen Hinderung der Probes untereinander begründet liegen. Da die Probes auf dem Chip „Loops in der Probe“ mit maximal 46 Basen Länge bereits sehr lang sind, kann es hier zu einer Wechselwirkung benachbarter Probes kommen, die die Hybridisierung der Targets beeinflussen kann. Nimmt man die unterschiedlichen Signale der Features bei Einbringung der Loops in Target bzw. Probe als Tatsache an, so kann ein Skalierungsfaktor bestimmt werden, der beide Experimente miteinander verbindet. Der Skalierungsfaktor wird anhand des Signals der Features mit Loop-Länge 13 vom Chip „Loops in der Probe“ und des Signals nach Hybridisierung mit LT-13-3Cy3 des Chip „Loops im Target“ bestimmt. So ergibt sich ein Skalierungsfaktor von  $0.25/0.82 \approx 0.3$ , der auch für das Umrechnen der Signale bei Hybridisierung mit LT-26-3Cy3 und LT-52-3Cy3 als konstant angenommen wird. Abb. 4.10 zeigt die absoluten Daten von Chip „Loops in der Probe“ mit den entsprechend skalierten Daten von Chip „Loops im Target“ nach Subtraktion des Hintergrundes (Hintergrundsignal  $\theta_{background} = 2.693 \cdot 10^4$  a.u.). Um auf noch größere Loop-Längen bzw. Längendifferenzen zwischen Probe und Target extrapolieren zu können, sind die Daten zusätzlich noch mit einem Power-Law der Form  $a \cdot x^{-b}$  gefittet (rote Linie, die Längendifferenz null ist hier ausgenommen). Der Fit mit  $b = 0.44$  stimmt sehr gut mit den Daten überein. Der Wert von  $b$  liegt in der Nähe des Wertes für eine Gaußsche Kette  $b_{Gauss} = 0.5$ . Dies ist konsistent, da sich die DNA-Stränge bei geeigneten experimentellen Bedingungen (z.B. hohe Salzkonzentration), insbesondere bei großen Längen wie LT-26-3Cy3 bzw. LT-52-3Cy3, wie ideale Ketten verhalten. Nimmt man an, dass es bei Signalen  $\leq 1.1 \cdot \theta_{background}$  zu keiner nennenswerten Hybridisierung kommt, ergibt sich mit Hilfe des Power-Laws eine maximale Längendifferenz  $\Delta L_{max}$  von 79 Basen. Dies gilt natürlich nur für die hier verwendete Sequenz Map2k18833c-reverse und die gemessenen Absolutwerte der Hybridisierungssignale (der Parameter  $a$  des Power-Laws hängt von den Absolutwerten ab). Weitere Untersuchungen müssten klären, inwieweit die maximale Längendifferenz  $\Delta L_{max}$  beispielsweise vom



**Abbildung 4.10:** Hybridisierungssignale als Funktion der Loop-Länge. Symbole: schwarze Punkte - experimentelle Daten; rote Kurve - Fit. Dargestellt sind die Hybridisierungssignale des Chips „Loops in der Probe“ als Funktion der Loop-Länge bis zu einer Loop-Länge von 13 Basen nach Subtraktion des Hintergrundes. Zusätzlich zu diesen Daten sind die skalierten Daten von Chip „Loops im Target“ gezeigt (Skalierungsfaktor 0.3). Die gesamten Daten können mit einem Power-Law der Form  $a \cdot x^{-b}$  angefitet werden (Loop-Länge 0 (PM-Duplex) ist hier ausgenommen). Extrapoliert man die Signale mit Hilfe des Power-Laws zu noch größeren Loop-Längen, so ergibt sich eine maximale Loop-Länge von 79 Basen, ab der keine nennenswerte Hybridisierung mehr stattfindet.

CG-Gehalt der Ausgangssequenz, ihrer Länge oder der Position der eingebrachten Basen abhängt. Bringt man die zusätzlichen Basen nicht, wie hier geschehen, zusammenhängend in die Ausgangssequenz ein, sondern verteilt sie über den gesamten Strang, so können sich ebenfalls andere Abhängigkeiten ergeben: Im vorliegenden Experiment stehen immer noch große zusammenhängende Teile zur Hybridisierung zur Verfügung, die nicht durch zusätzliche nicht-komplementäre Basen unterbrochen werden. Bei beliebiger Verteilung der Basen über den gesamten Strang sollte sich daher eine kleinere maximale Längendifferenz ergeben. Zusammen mit der Tatsache, dass die hier gewählte Ausgangssequenz bereits sehr lang ist (große freie Energie) und eine mittige Position für das Einbringen der zusätzlichen Basen gewählt wurde, bei der die Signale erfahrungsgemäß am stärksten sind, stellt die hier gefundene maximale Längendifferenz eine obere Schranke für die Grenze der Hybridisierung dar.

Es soll darauf hingewiesen werden, dass bei den Daten aus Abschnitt 4.1 der Hintergrund nicht subtrahiert wurde. Wird der Hintergrund dort subtrahiert, so ergibt sich bei einer Loop-Länge von 13 Basen ein Signal von 29.5% des PM-Signals (hier 25%). Die Daten der beiden Chips sind also innerhalb des experimentellen Fehlers konsistent.

## 4.4 Ergänzende Untersuchungen zur Grenze der Hybridisierung

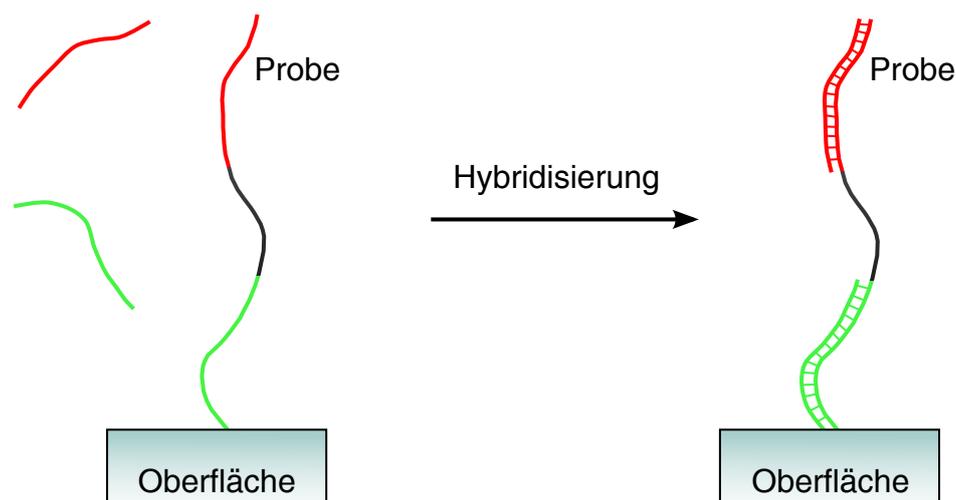
### 4.4.1 Einfluss der Synthesefehler

#### Motivation

Im vorherigen Abschnitt wurden unterschiedlichen Hybridisierungssignale bei Einbettung der Loops in das Target bzw. bei Einbettung in die Probe gemessen, obwohl in beiden Fällen identische Duplexe bei Hybridisierung entstehen sollten. Da bei Einbettung der Loops in die Probes je nach Loop-Länge sehr lange Probes synthetisiert werden müssen, wird hier untersucht, ob der Einbau von Synthesefehlern mit zunehmender Probe-Länge begünstigt wird und somit ihre Anzahl überproportional zunimmt. Dies wäre eine Erklärung für die stärkeren Signale bei Einbettung der Loops in die Targets, da hier die Probes auf der Oberfläche immer die gleiche Länge haben. Die Synthesequalität kann in diesem Fall keine Funktion der Probe-Länge sein und die Anzahl der Synthesefehler ist im Mittel für alle Probes gleich.

#### Experiment

Um die Abhängigkeit der Synthesefehler von der Probe-Länge zu untersuchen, wird der Chip „Loops in der Probe“ aus Abschnitt 4.3 mit zwei weiteren Target-Sequenzen - Map2k18833c-rev-low und Map2k18833c-rev-up - hybridisiert (siehe Anhang B.1). Die Länge der Probes wird auf diesem Chip durch 1-13 zusätzliche Basen nach Position 16 der Ausgangssequenz Map2k18833c-reverse variiert. Die beiden neuen Targets werden so entworfen, dass Map2k18833c-rev-low nur unterhalb der zusätzlich eingefügten Basen hybridisieren kann, Map2k18833c-rev-up nur oberhalb. Die Sequenz Map2k18833c-reverse wird also in zwei Target-Sequenzen aufgeteilt. So können die Regionen unterhalb und oberhalb der zusätzlich eingefügten Basen unabhängig voneinander auf die Häufigkeit der auftretenden Synthesefehler untersucht werden. Wird die Synthesequalität der Duplexregion oberhalb der eingefügten Basen mit zunehmender Anzahl der zusätzlichen Basen schlechter und weist somit mehr Synthesefehler auf als die Region unterhalb, würden sich bei den Hybridisierungen mit Map2k18833c-rev-low und Map2k18833c-rev-up verschiedene Abhängigkeiten der Signale von der Anzahl der eingefügten Basen ergeben: Die Synthese der unteren Region wird nicht durch die zusätzlich eingefügten Basen beeinflusst. In Abb. 4.11 ist der Grundgedanke des Experiments noch einmal veranschaulicht. Target Map2k18833c-rev-low ist dort in grün,



**Abbildung 4.11:** Untersuchung der Synthesefehler in Abhängigkeit der Probe-Länge. Die Target-Sequenz Map2k18833c-reverse wird in zwei Sequenzen aufgeteilt, von der die eine (grün) nur unterhalb der eingefügten Basen (schwarz) hybridisieren kann, die andere (rot) nur oberhalb. So kann der Einfluss der Synthesefehler der beiden Teilbereiche der Probe unabhängig voneinander untersucht werden. Die Hybridisierungen werden, anders als hier gezeigt, im Experiment nacheinander durchgeführt.

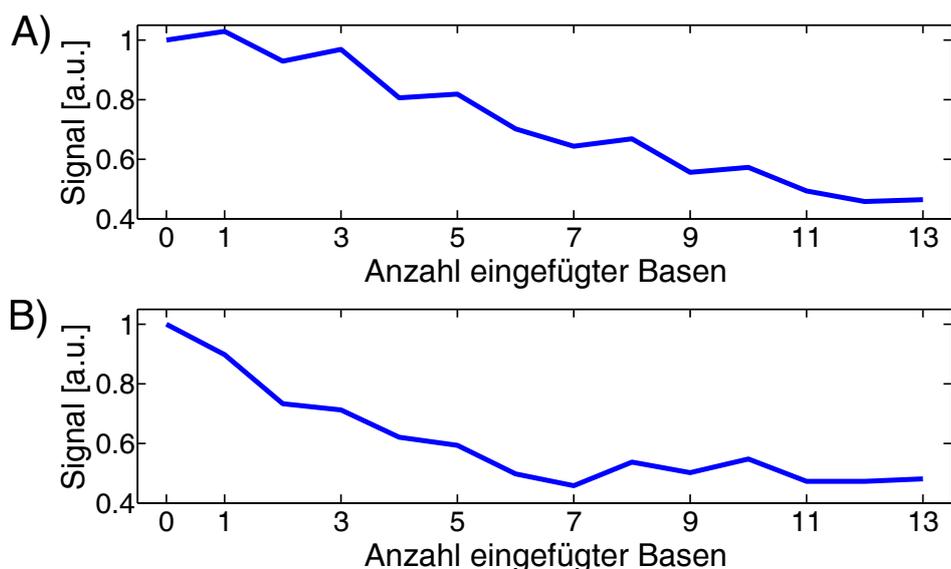
Target Map2k18833c-rev-up in rot gezeichnet. Die entsprechenden komplementären Regionen in der Probe sind mit den gleichen Farben gekennzeichnet. Die zusätzlich eingefügten Basen (schwarz) befinden sich in der Mitte des Probe.

Der Chip wird nacheinander einzeln mit den beiden Targets hybridisiert. Das experimentelle Vorgehen ist analog Abschnitt 4.3:

- Hybridisierung mit Map2k18833c-rev-low
- Temperaturerhöhung auf 80 °C für 2 min
- Abkühlen auf 20 °C
- Hybridisierung bei 20 °C für 60 min
- Bilderaufnahme
- Temperaturerhöhung auf 80 °C, Ausspülen der alten Target-Lösung und Spülen der Kammer mit Puffer für 4 min
- Wiederholung der Prozedur für Map2k18833c-rev-up

## Ergebnisse und Diskussion

Abb. 4.12 zeigt die Hybridisierungssignale als Funktion der Anzahl der eingefügten Basen bei Hybridisierung mit Map2k18833c-reverse-low (A)) bzw. mit Map2k18833c-



**Abbildung 4.12:** Abhängigkeit der Synthesefehler von der Anzahl der eingefügten Basen. Gezeigt sind die Signale bei Hybridisierung mit den Target-Sequenzen Map2k18833c-reverse-low bzw. Map2k18833c-reverse-up, die entweder unterhalb bzw. oberhalb der eingefügten Basen hybridisieren können. So kann die Abhängigkeit der Synthesefehler von der Probe-Länge für jeden Teilbereich der Probe unabhängig untersucht werden. Die Länge der Probes wird durch das Einbringen von zusätzlichen Sequenzen verschiedener Länge variiert. Die Signale erreichen bei beiden Hybridisierungen bei 13 zusätzlich eingefügten Basen ein Signal von ca. 50% des Signals bei 0 zusätzlichen Basen. **A)** Hybridisierung mit Map2k18833c-reverse-low **B)** Hybridisierung mit Map2k18833c-reverse-up

reverse-up (B)) nach Subtraktion des Hintergrundes und Normierung auf das Signal bei 0 zusätzlich eingefügten Basen bei einer Temperatur von 20 °C. Der schnellere Abfall des Signals in B) kann in der größeren Flexibilität der oberen Duplexregion begründet liegen. Die Duplexregion unterhalb der eingefügten Basen ist an einer Seite auf der Oberfläche fixiert, sodass weniger Konformationen zur Verfügung stehen und so die Entropieänderung bei Hybridisierung des entsprechenden Targets nicht so groß ist. Das untere Ende der oberen Duplexregion ist aufgrund des größeren Abstandes zur Oberfläche wesentlich flexibler. Dies erhöht die Anzahl der Konformationen der oberen Region und macht die Hybridisierung des entsprechenden Targets aufgrund der Einschränkung der entropischen Freiheitsgrade schwieriger. Bei weiterer Erhöhung der Anzahl der eingefügten Basen wird jedoch in beiden Fällen jeweils der gleiche Endwert von ca. 50% des PM-Signals bei 13 zusätzlichen Basen erreicht. Dies deutet darauf hin, dass die zunehmende Anzahl der eingefügten Basen keine Auswirkungen auf die Häufigkeit der Synthesefehler hat: Die absolute Anzahl der Synthesefehler ist in der Region unterhalb und oberhalb der eingefügten Basen ungefähr gleich.

## 4.4.2 Einfluss des Fluoreszenz-Markers

### Motivation

In den Experimenten in Abschnitt 4.3 zur maximalen Längendifferenz zwischen Probe und Target sind die verwendeten Target-Moleküle immer am 3'-Ende mit dem Fluorophor Cy3 markiert. Dies ist das Standardvorgehen bei den Hybridisierungsexperimenten der hier verwendeten DNA-Microarrays, da die Probes in 3'-5'-Richtung von der Oberfläche weg synthetisiert werden. Die bei der Hybridisierung entstehenden Duplexe sind so an dem oberflächenfernen Ende markiert, was eine Wechselwirkung zwischen dem Fluorophor und der Oberfläche verhindern soll. Im Folgenden Experiment soll untersucht werden, ob das Hybridisierungsverhalten von Targets, die am 5'-Ende (oberflächennah) modifiziert sind, von den bisherigen experimentellen Beobachtungen abweicht.

### Experiment

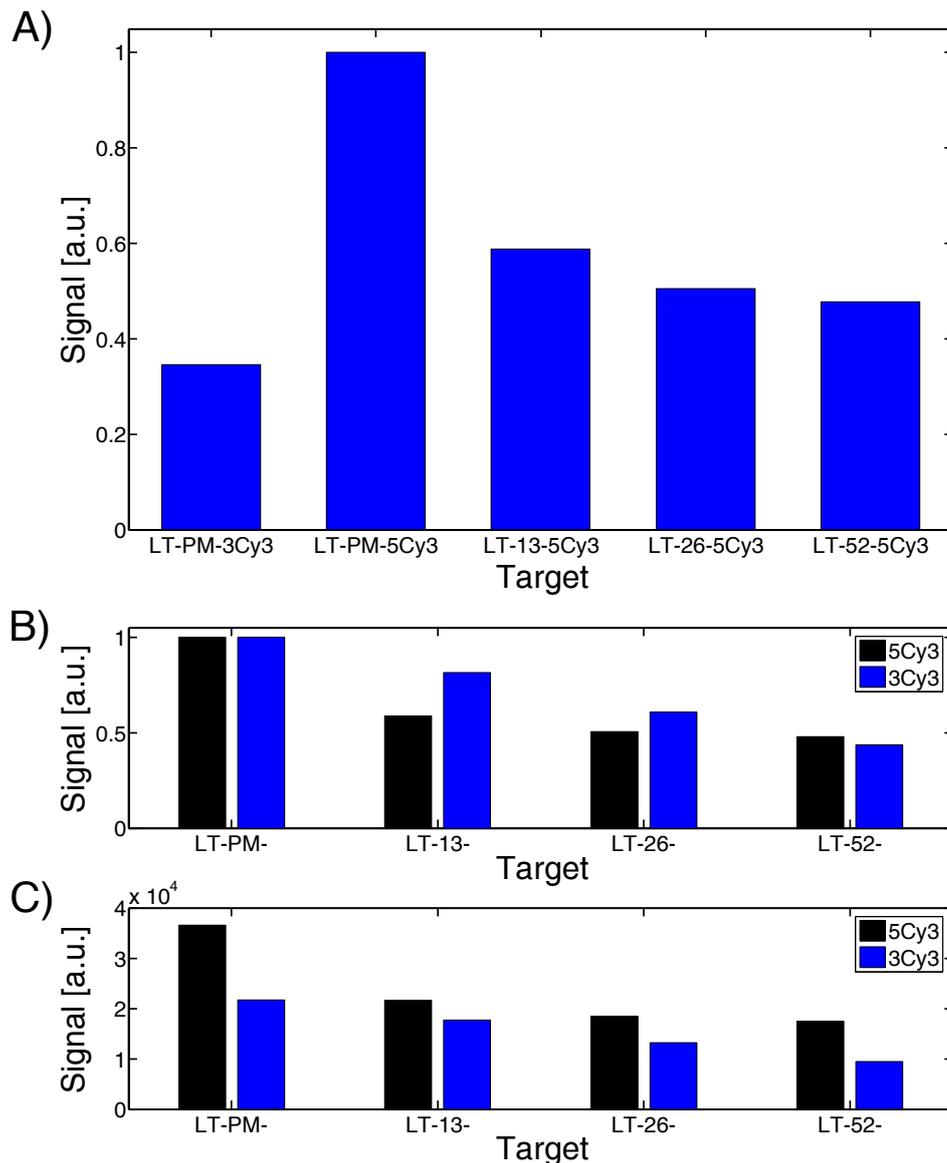
Zur Untersuchung des Einflusses des Fluoreszenzmarkers wird der Chip aus Abschnitt 4.3 noch einmal synthetisiert. Die Probes auf der Oberfläche sind also immer zum Target LT-PM-3Cy3 komplementär. Es werden Hybridisierungen mit den Targets LT-PM-3Cy3, LT-PM-5Cy3, LT-13-5Cy3, LT-26-5Cy3 und LT-52-5Cy3 nacheinander durchgeführt (siehe Anhang B.1). Die Sequenzen der hier verwendeten Targets sind mit den Sequenzen der entsprechenden Targets aus Abschnitt 4.3 identisch. Das Fluorophor Cy3 befindet sich jetzt lediglich am 5'-Ende. Anhand der Hybridisierung mit LT-PM-3Cy3 und LT-PM-5Cy3 kann die Abhängigkeit der Signalstärke von der Position des Fluoreszenzfarbstoff anhand identischer Target-Sequenzen untersucht werden. Die vier am 5'-Ende modifizierten Targets dienen analog den Experimenten aus Abschnitt 4.3 zur generellen Untersuchung der Abhängigkeit des Hybridisierungssignals von der Loop-Länge. Das experimentelle Vorgehen ist hier:

- Hybridisierung mit LT-PM-3Cy3
- Temperaturerhöhung auf 80 °C für 2 min
- Abkühlen auf 44 °C
- Hybridisierung bei 44°C für 90 min
- Bilderaufnahme

- Temperaturerhöhung auf 80 °C, Ausspülen der alten Target-Lösung und Spülen der Kammer mit Puffer für 4 min
- Wiederholung der Prozedur für weitere Hybridisierungen

## Ergebnisse und Diskussion

Abb. 4.13 A) zeigt die Signale bei Hybridisierung mit den entsprechenden Target-Sequenzen nach Subtraktion des Hintergrundes und Normierung auf das stärkste Signal (Hybridisierung mit LT-PM-5Cy3). Das Signal der am 5'-Ende markierten Targets nimmt auch hier mit zunehmender Loop-Länge kontinuierlich ab. Obwohl die Sequenzen LT-PM-3Cy3 und LT-PM-5Cy3 bis auf die Position der Modifikation identisch sind, beträgt das Signal von LT-PM-3Cy3 nur ca. 35% des Signals von LT-PM-5Cy3. In B) sind die Signale der am 5'-Ende markierten Targets (schwarz) mit den Ergebnissen aus Abschnitt 4.3 (blau) verglichen, wobei sich qualitativ keine großen Abweichungen feststellen lassen. Lediglich der Rückgang des Signals von LT-13-5Cy3 auf 59% des Signals von LT-PM-5Cy3 ist im Vergleich zum Rückgang von LT-13-3Cy3 auf ca. 82% des Signals von LT-PM-3Cy3 deutlich stärker. Die Ergebnisse bei Hybridisierung mit den längsten Targets LT-52-5Cy3 bzw. LT-52-3Cy3 sind mit 48% bzw. 44% des entsprechenden PM-Signals sehr ähnlich. Es soll an dieser Stelle angemerkt werden, dass die Absolutsignale der Targets mit 5'-Modifikation wesentlich stärker sind als die entsprechenden Signale aus Abschnitt 4.3, was die experimentelle Beobachtung aus A) bestätigt (Abb. 4.13 C)). Eine Modifikation der Targets mit Cy3 bewirkt generell eine Duplexstabilisierung und eine Erhöhung der Schmelztemperatur. Nach [93] ist die Stabilisierung außerdem abhängig von der Position der Modifikation, sodass ein am 5'-Ende angebrachtes Fluorophor sich stärker auf die Stabilität des Duplexes auswirken kann. Ein weiterer Grund für die höheren Signale der am 5'-Ende markierten Targets kann in der Synthesequalität der kommerziell erworbenen Targets begründet liegen: Da die Targets in 3'-5'-Richtung synthetisiert werden, wird bei einer Modifikation des 5'-Endes (im Gegensatz zur Modifikation des 3'-Endes) sichergestellt, dass nur Targets, die die gewünschte Länge erreicht haben, mit einem Fluorophor markiert werden. Des Weiteren sind die meisten modifizierten Targets hydrophiler als unmodifizierte, was dazu führt, dass die Trennung zwischen modifizierten Targets gewünschter Länge und unmodifizierten verkürzten Targets mittel HPLC bei 5'-Modifikation wesentlich besser funktioniert. Zwar soll sichergestellt werden, dass alle Targets markiert sind und die gewünschte Länge haben, jedoch können sich oben beschriebene Aspekte trotzdem auf die Synthesequalität auswirken. Daher ist die Modifikation des 5'-Endes des Targets eigentlich vorzuziehen, was aber die Wahrscheinlichkeit einer Wechselwirkung zwischen



**Abbildung 4.13:** Einfluss der Position des Fluoreszenzmarkers auf die Hybridisierung. **A)** Gezeigt sind die Hybridisierungssignale bei Verwendung der angegebenen Target-Sequenzen nach Subtraktion des Hintergrundes und Normierung auf das stärkste Signal (Hybridisierung mit LT-PM-5Cy3) bei einer Temperatur von 44 °C. Die Abnahme des Hybridisierungssignales mit zunehmender Loop-Länge wird auch hier deutlich. So beträgt das Signal bei der Hybridisierung mit LT-52-5Cy3 nur noch 48% des PM-Signals. Die Signale von LT-PM-3Cy3 und LT-PM-5Cy3 unterscheiden sich deutlich, obwohl sich die entstehenden Duplexe nur in der Position der Fluoreszenzmarkierung unterscheiden. **B)** Vergleicht man die normierten Signale bei Hybridisierung der am 5'-Ende markierten Targets mit den Signalen aus Abschnitt 4.3 (Markierung am 3'-Ende), so ergeben sich bis auf die Hybridisierungen mit den Targets LT-13-3Cy3 bzw. LT-13-5Cy3 in etwa die gleichen Abhängigkeiten von der Loop-Länge. **C)** Vergleich der absoluten Signale aus Abschnitt 4.4.2 mit denen aus Abschnitt 4.3. Alle am 5'-Ende markierten Targets liefern ein stärkeres Signal als ihre am 3'-Ende markierten Pendants.

Fluorophor und Oberfläche vergrößert. Ob es durch die Markierung der 5'-Enden der Targets zu einer verstärkten Wechselwirkung mit der Oberfläche kommt, die zu unspezifischen, starken Signalen führt, soll im folgenden Abschnitt untersucht werden.

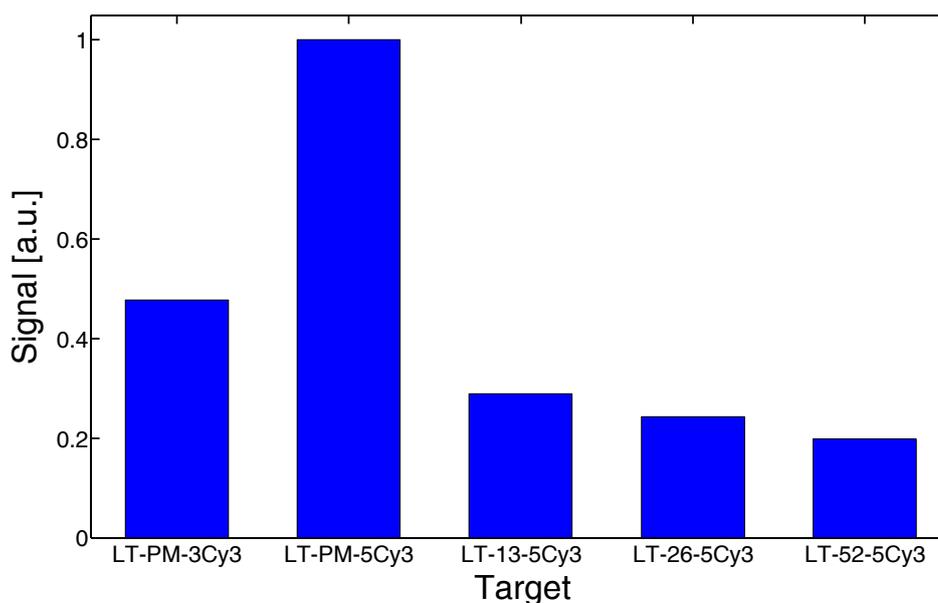
### 4.4.3 Einfluss von Linkern

#### Experiment

Damit die Wahrscheinlichkeit einer Wechselwirkung zwischen Oberfläche und Targets mit Fluorophor am 5'-Ende bei Hybridisierung der Probes minimiert wird, werden Abstandshalter zur Oberfläche, sog. *Linker*, vor die eigentliche Sequenz der Probes eingebaut. Die Linker bestehen aus poly-T-Sequenzen von 12 Basen Länge. Die eigentliche Sequenz der Probes, die zur Target-Sequenz LT-PM-3Cy3 komplementär ist, wird so um 12 Thymin Basen am 3'-Ende verlängert. Das experimentelle Vorgehen ist mit dem aus Abschnitt 4.4.2 identisch.

#### Ergebnisse und Diskussion

In Abb. 4.14 sind die Signale bei Hybridisierung mit den entsprechenden Target-Sequenzen nach Subtraktion des Hintergrundes und Normierung auf das stärkste Signal (Hybridisierung mit LT-PM-5Cy3) gezeigt. Auch in diesem Experiment stellt man einen deutlichen Signalunterschied bei Markierung der Targets am 3'- bzw. 5'-Ende fest. So beträgt das Signal des am 3'-Endes markierten Targets LT-PM-3Cy3 nur 48% des Signals von LT-PM-5Cy3. Die im vorherigen Abschnitt gefundenen Ergebnisse bestätigen sich hier also. In diesem Experiment ergibt sich allerdings eine andere Abhängigkeit des Signals von der Loop-Länge. Zwar nehmen auch hier die Signale mit zunehmender Loop-Länge ab, jedoch ist bereits das Signal von LT-13-5Cy3 mit ca. 29% des Signals von LT-PM-5Cy3 deutlich schwächer als das entsprechende Signal in Abschnitt 4.4.2 (dort betrug das Signal von LT-13-5Cy3 ca. 59% des Signals von LT-PM-5Cy3). Die Hybridisierung mit LT-52-5Cy3 liefert nur noch ein Signal von 20%. Durch das Hinzufügen eines Linkers soll die Probe für Hybridisierungen mit Targets besser zugänglich gemacht und eine Wechselwirkung des Fluorophors mit der Oberfläche minimiert werden. Allerdings ist die gesamte Probe-Länge inklusive Linkern von 12 Basen mit 45 Basen schon recht groß, sodass Wechselwirkungen der einzelnen Probes untereinander nicht ausgeschlossen werden können. Dies wiederum erschwert eine Hybridisierung von langen Target-Sequenzen. In [73] wurde gezeigt, dass die Signalstärke abhängig von der Linker-Länge ist. Für eine Sequenzlänge von 18 Basen wurde eine optimale Spacer-



**Abbildung 4.14:** Einfluss von Linkern auf die Hybridisierung. Gezeigt sind die Hybridisierungssignale bei Verwendung der angegebenen Target-Sequenzen nach Subtraktion des Hintergrundes und Normierung auf das stärkste Signal (Hybridisierung mit LT-PM-5Cy3) bei einer Temperatur von 44 °C. Um den Abstand der eigentlichen Probe-Sequenz Map2k18833c-reverse zur Oberfläche zu vergrößern, werden Linker bestehend aus 12 Thymin Basen eingebaut. Die Hybridisierungssignale nehmen mit zunehmender Loop-Länge ab und betragen bei der größten untersuchten Loop-Länge nur noch 20% des PM-Signals. Die PM-Duplex mit Markierung am 3'-Ende hybridisiert deutlich schwächer als der PM-Duplex mit 5'-Markierung.

Länge von 9 Basen ermittelt. Daher ist es auch möglich, dass die in diesem Experiment gewählte Linker-Länge von 12 Basen nicht die optimale Länge darstellt und sich bei einer anderen Länge die Signalunterschiede zwischen LT-PM-5Cy3 und den Targets mit Loops geringer sind.

Mit diesem Experiment konnte eine Wechselwirkung des Fluorophors mit der Oberfläche als Ursache der stärkeren Signale der am 5'-Ende markierten Targets ausgeschlossen werden, da sich die Signalunterschiede bei Hybridisierung mit LT-PM-3Cy3 bzw. LT-PM-5Cy3 bestätigt haben. Andere mögliche Ursachen der verschiedenen starken Signale bei 3'- und 5'-Markierung wurden in Abschnitt 4.4.2 diskutiert.

## 4.5 Probes mit zwei Loops

Eine Verallgemeinerung der in den vorherigen Abschnitten präsentierten Experimente besteht darin, das Einfügen von zusätzlichen Basen in die Ausgangssequenz der Probe an mehreren Stellen gleichzeitig zuzulassen. Dies kann vor allem bei der Hybridisierung von sehr langen Strängen, die in mehreren Teilbereichen komplementär

zueinander sind, ein wichtiger Faktor sein. Daher werden im Folgenden an zwei Stellen der Ausgangssequenz zusätzliche Basensequenzen eingefügt, die bei Hybridisierung jeweils einen Loop ausbilden sollen. Wichtige Fragen sind, ob man die Loops unabhängig voneinander behandeln kann und falls ja, welche Voraussetzungen dabei gelten müssen. Dazu sollen verschiedene Kombinationen von Loop-Längen und -Positionen der beiden Loops getestet und insbesondere der Einfluss des Abstandes der beiden Loops auf das Hybridisierungssignal untersucht werden.

### 4.5.1 Vorversuch

#### Experiment

Die zusätzlichen poly-T-Sequenzen werden an zwei verschiedenen Positionen in die Ausgangssequenz Map2k18833c-reverse der Probe eingefügt, sodass bei Hybridisierung zwei Loops entstehen. Um die Gesamtlänge der Probes nicht zu groß werden zu lassen, werden die Loop-Längen zwischen 3 und 6 Basen variiert. Da die Sequenz Map2k18833c-reverse 33 Basen lang ist, ist so die maximal erreichte Probe-Länge von 45 Basen um eine Base kürzer als die in Abschnitt 4.1. Aufgrund der Größe des Chips können nicht alle Kombinationen von Loop-Längen und -Positionen beider Loops realisiert werden. Bei Variation der Position von Loop B in Richtung von Loop A (zur Oberfläche hin) wird die Ausgangsposition von Loop A festgehalten und die Längen beider Loops variiert. Entsprechendes gilt für die Variation der Position von Loop A in Richtung von Loop B. Die Ausgangspositionen von Loop A bzw. Loop B befindet sich nach Base 10 bzw. 23 der Ausgangssequenz der Probe (Zählrichtung vom 3'- zum 5'-Ende). Bei Loop-Längen zwischen 3 und 6 Basen ergeben sich so insgesamt 448 verschiedene Probe-Sequenzen. Die realisierten Kombinationen von Längen und Positionen der beiden Loops bei Variation der Position von Loop B sind in Tab. 4.2 angegeben (Variation der Position von Loop A verläuft analog). Zur Signalkontrolle sind darüber hinaus noch 8 PM-Features über den Chip verteilt. Der Chip wird mit dem Target Map2k18833c-reverse bei 44 °C bis Erreichen des Gleichgewichts hybridisiert.

#### Ergebnisse und Diskussion

In Abb. 4.15 sind die Daten bei Variation der Position von Loop A in Richtung der Ausgangsposition von Loop B (nach Base 23 der Ausgangssequenz Map2k18833c-reverse, Variation von Oberfläche weg) nach Subtraktion des Hintergrundes bei einer Temperatur von 44 °C gezeigt. Die Teile A) - D) zeigen die Daten jeweils für eine feste Länge

**Tabelle 4.2:** Kombinationen von Länge und Position der beiden Loops bei Variation der Position von Loop B. In die Ausgangssequenz Map2k18833c-reverse werden an zwei unterschiedlichen Positionen Loops variabler Länge eingefügt. Die Positionsangaben beziehen sich dabei auf die Base der Ausgangssequenz, nach der die Loops eingefügt werden (Zählrichtung ist von 3' zu 5'). Die Position von Loop B wird in Richtung der Position von Loop A variiert, die Position von Loop A bleibt konstant. So ergeben sich 224 Kombinationen von Länge und Position der beiden Loops. Die Variation der Position von Loop A verläuft analog.

Loop A		Loop B	
Position	Länge	Position	Länge
10	3	23	3
10	3	23	4
10	3	23	5
10	3	23	6
10	3	22	3
10	3	22	4
⋮	⋮	⋮	⋮
10	4	23	3
10	4	23	4
⋮	⋮	⋮	⋮
10	6	10	3
10	6	10	4
10	6	10	5
10	6	10	6

von Loop B als Funktion des Abstandes der beiden Loops, wobei die Länge von Loop A zwischen 3 und 6 Basen variiert wird. Bei einem Abstand von 0 Basen liegt nur noch ein einzelner Loop nach Position 23 der Ausgangssequenz vor, dessen Länge sich als Summe der Längen von Loop A und Loop B ergibt. In allen vier Fällen steigt das Signal bei Annäherung der beiden Loops an. Es soll angemerkt werden, dass bei einer festen Länge von Loop B und einem festen Abstand der beiden Loops normalerweise vier Signale für die vier verschiedenen Längen von Loop A angegeben sind. Ist dies nicht der Fall, so konnte das entsprechende Signal aufgrund von Verunreinigungen auf dem Chip nicht ausgewertet werden. Um die Abhängigkeit des Signals vom Abstand der beiden Loops deutlicher hervorzuheben, zeigt Teil E) die Signale als Funktion des Abstandes der beiden Loops nach Mittelung über alle Längen von Loop A für die vier Längen von Loop B. Die blaue Linie entspricht z.B. den gemittelten Signalen aus Teil A). Die anderen Kurven werden entsprechend gewonnen. Das Signal bleibt bei Annä-

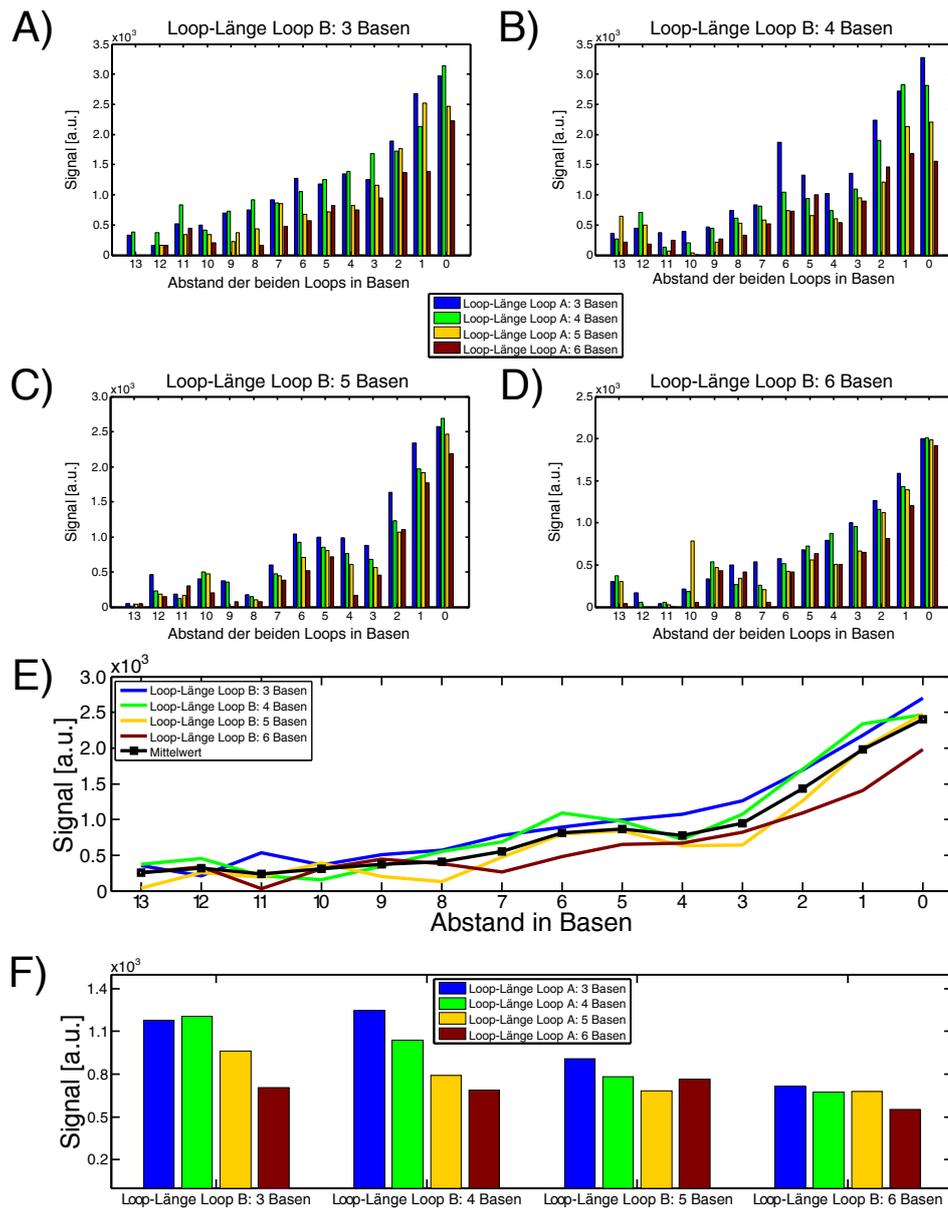
herung der beiden Loops bis zu einem gewissen Abstand konstant und steigt dann nach erst leichtem Anstieg ab einem Abstand von ca. 3 Basen relativ stark an. Dabei liefert - wie nach den Ergebnissen aus den vorherigen Abschnitten zu erwarten - die Länge von 6 Basen von Loop B das schwächste Signal. Teil F) zeigt die Signale bei Mittelung über alle Positionen von Loop A als Funktion der Längen von Loop A und Loop B. Wird die Länge von Loop B konstant gehalten und die Länge von Loop A erhöht, so nimmt das Signal im Allgemeinen ab. Daraus folgt auch, dass die Probes mit der größten Gesamtlänge der beiden Loops das schwächste Signal liefern: Vergleicht man die Signale der 4 Blöcke, so hat der Block bei einer Länge von 6 Basen von Loop B die geringsten Signale. Abb. 4.16 zeigt die entsprechenden Daten bei Variation der Position von Loop B in Richtung von Loop A (nach Base 10 der Ausgangssequenz, Variation Richtung Oberfläche). Auch hier ergeben sich qualitativ die gleichen Abhängigkeiten: Das Signal steigt bei Annäherung der beiden Loops an (siehe z.B. Teil E)); Probes mit der größten Anzahl zusätzlich eingefügter Basen haben das schwächste Signal (siehe z.B. Teil F), rechter Block). Ob die Variation der Loop-Position in Richtung der Oberfläche oder von der Oberfläche weg geschieht, scheint keinen Einfluss auf die Abhängigkeit der Signale vom Abstand der beiden Loops zu haben.

Die Tatsache, dass Probes mit zunehmender Anzahl zusätzlich eingefügter Basen ein immer schwächeres Signal liefern, ist mit den Ergebnissen aus den vorherigen Abschnitten konsistent. Der Signalanstieg bei Annäherung der beiden Loops ist jedoch unerwartet und soll in den folgenden Abschnitten genauer untersucht werden.

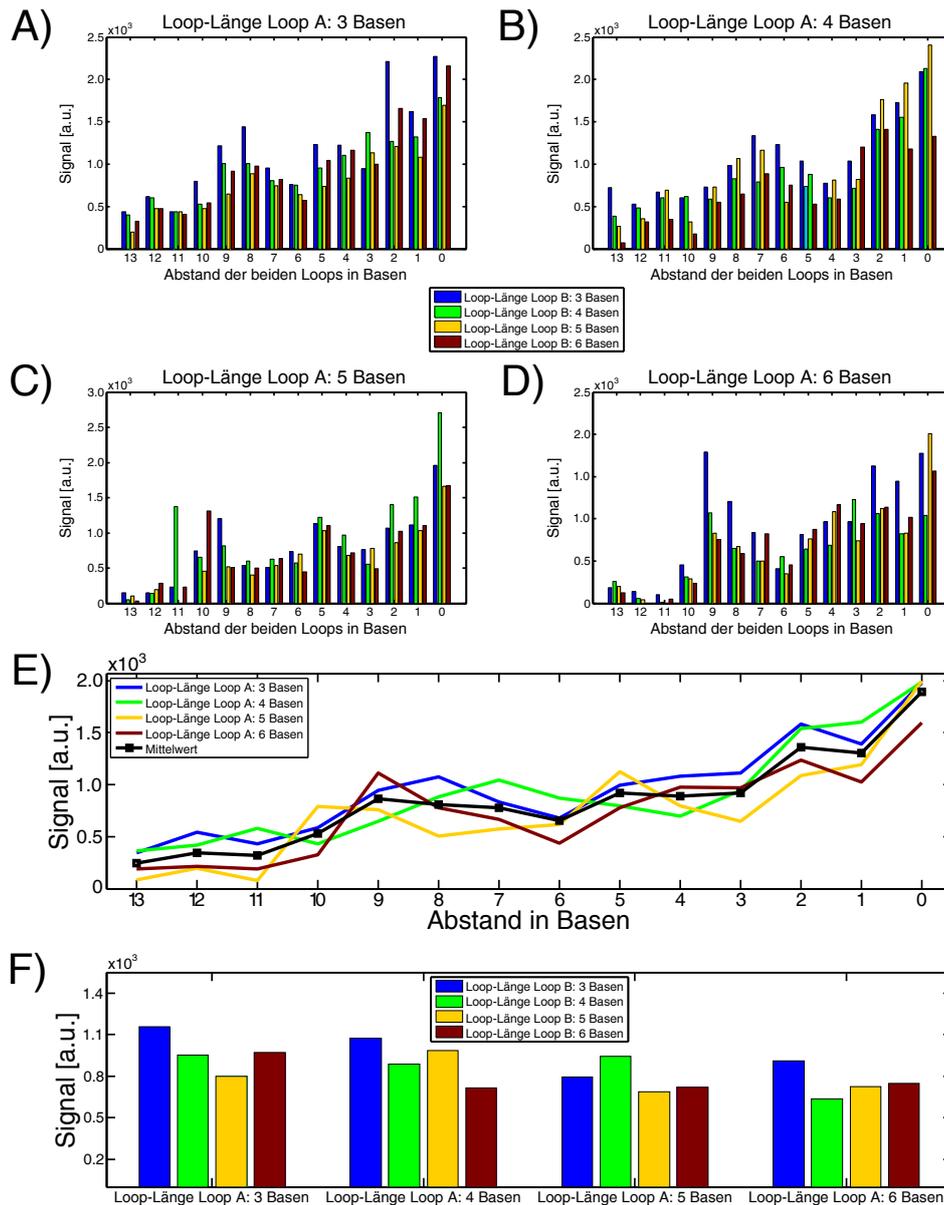
## 4.5.2 Einfluss des Abstandes der beiden Loops

### Experiment

In den Experimenten aus Abschnitt 4.5.1 konnte kein qualitativer Signalunterschied bei Variation der Loop-Position in Richtung der Oberfläche (Variation der Position von Loop B bei fixer Position von Loop A) oder von der Oberfläche weg (Variation der Position von Loop A bei fixer Position von Loop B) festgestellt werden. Daher wird in diesem Versuch nur die Position des oberflächennahen Loops A in Richtung des oberflächenfernen Loops B verändert. Dazu werden zwei verschiedene feste Positionen für Loop B gewählt: Position 1 befindet sich nach Base 16 der Ausgangssequenz der Probe, Position nach Base 26. Die Ausgangsposition von Loop A befindet sich immer nach Base 3 der Ausgangssequenz. Zählrichtung ist auch hier vom 3'- zum 5'-Ende des Probesequenz. Da in Abschnitt 4.5.1 die verschiedenen Kombinationen der Längen der beiden Loops lediglich eine Auswirkung auf die Signalstärke nicht aber auf den Signalverlauf bei Annäherung der Loops hatten, ist die Länge der beiden Loops in



**Abbildung 4.15:** Hybridisierungssignale bei Variation der Position von Loop A. **A) - D)** Jeder Teilgraph zeigt die Hybridisierungssignale als Funktion des Abstandes der beiden Loops jeweils für eine feste Länge und Position von Loop B nach Subtraktion des Hintergrundes bei einer Temperatur von 44 °C. Die Position von Loop B befindet sich immer nach Base 23 der Ausgangssequenz. Variiert wird dabei die Länge und die Position von Loop A (Ausgangsposition nach Base 10) in Richtung von Loop B (von der Oberfläche weg). Die Signale steigen mit sinkendem Abstand der beiden Loops an und erreichen ihr Maximum, wenn beide Loops zu einem Loop vereinigt sind (Abstand 0). Sind weniger als 4 Signale für jeden Abstand der beiden Loops angegeben, so konnte das entsprechende Signal aufgrund von Verunreinigungen auf dem Chip nicht ausgewertet werden. **E)** Gezeigt sind die Signale bei Mittelung über alle Längen von Loop A für jede der vier Teilabbildungen A)- D). Hier lässt sich gut erkennen, dass die Signale bei Annäherungen der Loops über einen weiten Abstandsbereich konstant bleiben und dann ansteigen. **F)** Mittelung der Signale aus den Teilabbildungen A) - D) über alle Positionen von Loop A. Die Proben mit der größten Gesamtlänge der beiden Loops liefern das schwächste Signal (Block ganz rechts).



**Abbildung 4.16:** Hybridisierungssignale bei Variation der Position von Loop B. **A) - D)** Jeder Teilgraph zeigt die Hybridisierungssignale als Funktion des Abstandes der beiden Loops jeweils für eine feste Länge und Position von Loop A nach Subtraktion des Hintergrundes bei einer Temperatur von 44 °C. Die Position von Loop A befindet sich immer nach Base 10 der Ausgangssequenz. Variiert wird dabei die Länge und die Position von Loop B (Ausgangsposition nach Base 23) in Richtung von Loop A (zur Oberfläche hin). Die Signale steigen mit sinkendem Abstand der beiden Loops an und erreichen ihr Maximum, wenn beide Loops zu einem Loop vereinigt sind (Abstand 0). Sind weniger als 4 Signale für jeden Abstand der beiden Loops angegeben, so konnte das entsprechende Signal aufgrund von Verunreinigungen auf dem Chip nicht ausgewertet werden. **E)** Gezeigt sind die Signale bei Mittelung über alle Längen von Loop B für jede der vier Teilabbildungen A) - D). Hier lässt sich gut erkennen, dass die Signale bei Annäherungen der Loops über einen weiten Abstandsbereich konstant bleiben und dann ansteigen. **F)** Mittelung der Signale aus den Teilabbildungen A) - D) über alle Positionen von Loop B.

diesem Experiment immer identisch. Es werden zwei verschiedene Loop-Längen von 3 bzw. 5 Basen realisiert. So ergeben sich insgesamt 76 verschiedene Probe-Sequenzen. Außerdem werden zusätzlich zu den Probes mit zwei Loops noch Probes mit nur einem Loop synthetisiert. Der Loop hat dabei eine Länge von 3 bzw. 5 Basen und wird an allen Position zwischen Base 3 und 26 der Ausgangssequenz eingefügt (analog den Experimenten aus Abschnitt 4.1). Diese Probes dienen der Untersuchung wie sich das Signal einer Probe mit zwei Loops aus den Signalen der Probes mit jeweils einem Loop der entsprechenden Länge und Position zusammensetzen lässt (siehe Abschnitt 4.5.3). Zur Signalkontrolle werden außerdem mehrere PM-Features synthetisiert. Der Chip wird mit dem Target Map2k18833c-reverse bei 44 °C bis Erreichen des Gleichgewichts hybridisiert.

### Ergebnisse und Diskussion

Die Teile A) - D) von Abb. 4.17 zeigen die Signale bei Variation der Position von Loop A in Richtung von Loop B für zwei verschiedene Längen der beiden Loops (die Länge von Loop A ist mit der von Loop B immer identisch) und zwei Positionen von Loop B nach Subtraktion des Hintergrundes. Position 1 von Loop B befindet sich nach Base 16 der Ausgangssequenz, Position 2 nach Base 26. Die farbigen Linien kennzeichnen dabei die Signale der einzelnen Featureblocks, die schwarze Kurve zeigt den experimentellen Mittelwert. Im Unterschied zu den Ergebnissen aus Abschnitt 4.5.1 fällt hier das Signal bei Annäherung der Loops zuerst leicht ab bevor es zu einem Anstieg kommt. Bei Position 1 von Loop B erfolgt der Signalanstieg für beide Loop-Längen ungefähr ab einem Abstand von 6 Basen, bei Position 2 von Loop B ab ungefähr 13 Basen. Bei einem Abstand von 0 Basen erreicht das Signal ein Maximum. Beim diesem Abstand entsteht aus den beiden Loops ein einzelner Loop dessen Länge gerade die Summe der Längen der beiden Loops ist. Teil E) und F) zeigen den Vergleich der Mittelwerte der Signale in Abhängigkeit der Längen der beiden Loops für eine feste Position von Loop B. Wie zu erwarten liefern die Probes mit den längeren Loops (blaue Linien) das schwächere Signal. Es zeigt sich hier deutlich, dass die Länge der einzelnen Loops keine Auswirkungen auf den qualitativen Signalverlauf bei Annäherung der Loops hat.

Der zuerst auftretende Signalabfall, der so in Abschnitt 4.5.1 nicht beobachtet wurde, kann mit der Ausgangsposition von Loop A begründet werden. Die Ausgangsposition von Loop A befindet sich näher am Ende des Strangs als im vorherigen Experiment. In den Untersuchungen zum Einfluss von einzelnen Loops in der Probe auf die Stabilität des Duplexes in Abschnitt 4.1 wurde beobachtet, dass die Einbringung von endnahen Loops die Stabilität des Duplexes weniger stark negativ beeinflusst und daher ein starkes Signal gemessen werden kann. Das bedeutet, dass der leichte Abfall des Signals

auch in den vorherigen Experimenten beobachten werden sollte, wenn die Ausgangsposition von Loop A entsprechend angepasst wird.

Der Signalanstieg bei Annäherung der Loops deutet darauf hin, dass es für die Stabilität des Duplexes günstiger ist, wenn die Defekte stärker lokalisiert sind und sich nicht über die gesamte Länge des Duplexes erstrecken (dies wurde in den Experimenten zur Grenze der Hybridisierung in Abschnitt 4.3 diskutiert). Dadurch stehen genügend große defektfreie Bereiche der Probe zur Hybridisierung zur Verfügung, sodass insgesamt die Stabilität weniger stark beeinflusst wird. Verteilen sich die Defekte hingegen über den gesamten Duplex, wird der Duplex in viele kleine defektfreie Bereiche unterteilt, die jeweils wenig zur Duplexstabilität beitragen können.

Es stellt sich jetzt die Frage, ob sich das Signal einer Probe mit zwei Loops aus den Signalen zweier Probes mit jeweils einem Loop geeigneter Position und Länge berechnen lässt. Darauf soll im folgenden Abschnitt eingegangen werden.

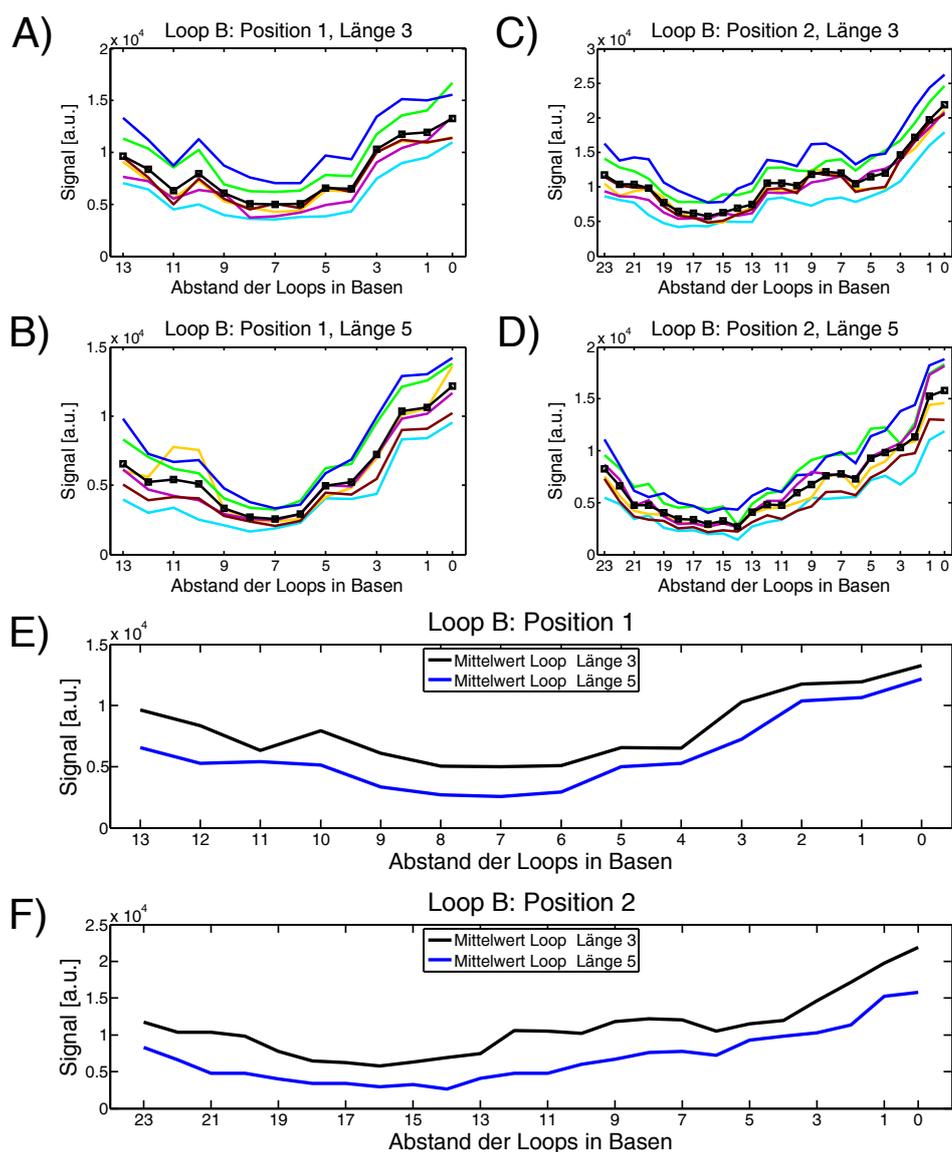
### 4.5.3 Theoretische Überlegungen

In diesem Abschnitt soll mit einfachen theoretischen Überlegungen eine grobe Abschätzung für den Wechselwirkungsabstand der beiden Loops getroffen werden. Haben die beiden Loops einen großen Abstand, so ist die Wahrscheinlichkeit einer Wechselwirkung gering und die beiden Loops können unabhängig voneinander betrachtet werden. Das Signal einer Probe mit zwei Loops sollte sich dann aus den entsprechenden Signalen zweier Probes mit jeweils einem Loop geeigneter Position und Länge berechnen lassen. Beispielweise würde sich das Signal einer Probe mit einem Loop der Länge  $L_A$  an Position A und einem Loop der Länge  $L_B$  an Position B aus dem Signal einer Probe mit Loop der Länge  $L_A$  an Position A und dem Signal einer Probe mit Loop der Länge  $L_B$  an Position B zusammensetzen. Nähern sich die Loops einander an, so treten sie ab einem gewissen Abstand (der wahrscheinlich eine Funktion der Länge und Position der einzelnen Loops ist, siehe Abschnitt 4.5.2) in Wechselwirkung. Das Signal einer Probe mit zwei Loops lässt sich dann nicht mehr aus den Signalen der Einzel-Loops berechnen. Zu den theoretischen Berechnungen in diesem Abschnitt werden neben den gemessenen Signalen der Probes mit zwei Loops auch die Signale der zusätzlich synthetisierten Probes mit nur einem Loop herangezogen. Alle Signale werden auf das PM-Signal normiert.

Das Hybridisierungssignal einer Target-Spezies ist gegeben durch (siehe Abschnitt 2.3, Gl. (2.3)):

$$\theta \propto \frac{C \cdot K \cdot [T_0]}{1 + C \cdot K \cdot [T_0]} \quad (4.1)$$

Dabei ist  $K$  die Bindungskonstante der Target-Spezies und  $[T_0]$  die Konzentration der



**Abbildung 4.17:** Hybridisierungssignale bei Variation der Position von Loop A. **A) - D)** Jeder Teilgraph zeigt die Hybridisierungssignale als Funktion des Abstandes der beiden Loops jeweils für eine feste Länge und Position von Loop B nach Subtraktion des Hintergrundes bei einer Temperatur von 44 °C. Variiert wird die Position von Loop A (Ausgangsposition nach Base 10) in Richtung von Loop B (von der Oberfläche weg). Position 1 von Loop B befindet sich nach Base 16 der Ausgangssequenz, Position 2 nach Base 26. Die beiden Loops haben immer die gleiche Länge. Die Signale fallen zunächst bei Verringerung des Abstandes leicht ab und steigen dann bei weiterer Verringerung an bis sie beim Abstand 0 das Maximum erreichen. Die farbigen Linien geben die Signale der einzelnen Featureblocks an, die schwarze Linie kennzeichnet den experimentellen Mittelwert. **E)** Vergleich der Signalmittelwerte als Funktion des Loop-Abstandes für die beiden untersuchten Loop-Längen und Position 1 von Loop B. Die Proben mit den größten Loop-Längen liefern das schwächere Signal. **F)** Entsprechende Mittelungen für Position 2 von Loop B.

Targets.  $C$  ist ein notwendiger Skalierungsfaktor, der die gemessenen Fluoreszenzsignale mit den theoretisch berechneten Bindungskonstanten in Verbindung bringt.  $C$  kann nicht vermieden werden, da die absolute Größe von Bindungskonstanten nur schwer vorhergesagt werden kann [94, 95]. Die absolute Größe der Bindungskonstanten des PM ist jedoch für die weiteren Überlegungen nicht entscheidend. Nimmt man an, dass die Bindungskonstante  $K_{Loop}$  einer Probe mit Loop um einen Faktor  $x = \exp(-\Delta G_{Defekt}/RT)$  kleiner ist als die Bindungskonstante des PM  $K_{PM}$ , so gilt:

$$K_{Loop} = x \cdot K_{PM} \quad (4.2)$$

Damit ergibt sich das Signalverhältnis  $y$  zwischen einer Probe mit einem Loop und der PM-Probe nach Gl. 4.1 zu:

$$y = \frac{\theta_{Loop}}{\theta_{PM}} = \frac{C \cdot x \cdot K_{PM} \cdot [T_0]}{1 + C \cdot x \cdot K_{PM} \cdot [T_0]} \cdot \frac{1 + C \cdot K_{PM} \cdot [T_0]}{C \cdot K_{PM} \cdot [T_0]} \quad (4.3)$$

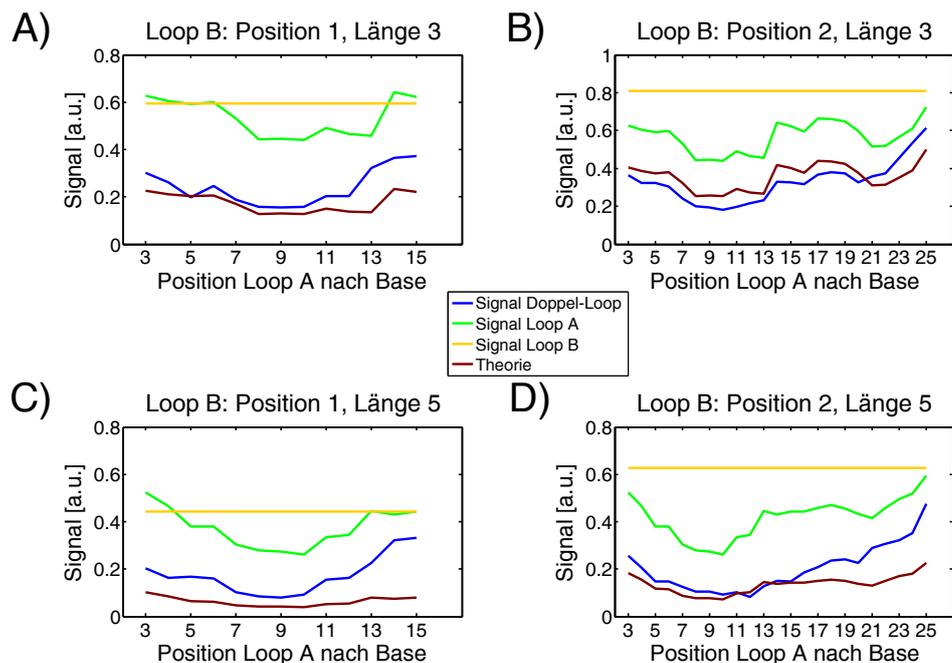
Löst man nach  $x$  auf, so erhält man:

$$x = \frac{y}{1 + C \cdot K_{PM} \cdot [T_0] (1 - y)} \quad (4.4)$$

Der Abschwächungsfaktor  $x$  gibt an, um wieviel sich die Bindungskonstante des PM ändern muss, um das gemessene Verhältnis  $y$  der Signale von PM und Probe mit Loop zu erhalten. Dabei muss natürlich  $0 < y < 1$  gelten. Sind die Abschwächungsfaktoren  $x$  für die Probes mit Einzel-Loop bestimmt, können die Signale der Probes mit zwei Loops im Bezug zum PM-Signal bestimmt werden. Unter der Annahme, dass die Einflüsse der beiden Loops A und B unabhängig voneinander sind, gilt für das Signalverhältnis:

$$\frac{\theta_{Doppel-Loop}}{\theta_{PM}} = \frac{C \cdot x_A \cdot x_B \cdot K_{PM} \cdot [T_0]}{1 + C \cdot x_A \cdot x_B \cdot K_{PM} \cdot [T_0]} \cdot \frac{1 + C \cdot K_{PM} \cdot [T_0]}{C \cdot K_{PM} \cdot [T_0]} \quad (4.5)$$

Abb. 4.18 zeigt den Vergleich zwischen den mit Gl. 4.5 berechneten Signalen der Probes mit zwei Loops (dunkelrote Kurve) und den gemessenen Signalen (blaue Kurve) als Funktion der Position von Loop A für die möglichen Kombinationen von Länge und Position von Loop B. Zusätzlich sind noch die Signale der Probes mit Einzel-Loop gezeigt (grüne und gelbe Kurve). Da die Position und Länge von Loop B in jeder der vier Teilabbildungen gleich ist, ergibt sich jeweils ein konstantes Signal (Position 1 von Loop B befindet sich nach Base 16 der Ausgangssequenz, Position 2 nach Base 26). Mit Hilfe des Skalierungsfaktors  $C$  können die dunkelrote und blaue Kurve teilweise zur Deckung gebracht werden. Jedoch ist es nicht möglich, die beiden Kurven gleichzeitig für große (niedrige Positionen von Loop A) und kleine Abstände (ähnliche Positionen von Loop A und Loop B) der beiden Loops zur Deckung zu bringen. Da die



**Abbildung 4.18:** Theoretische Überlegungen zum Einfluss der Doppel-Loops. Jede der vier Teilabbildungen zeigt die Hybridisierungssignale der Probes mit Doppel-Loops als Funktion der Position von Loop A für die möglichen Kombinationen von Länge und Position von Loop B (blaue Kurve). Position 1 von Loop B befindet sich nach Base 16 der Ausgangssequenz, Position 2 nach Base 26. Die Länge von Loop A und B ist identisch. Niedrige Positionen von Loop A entsprechen großen Abständen der beiden Loops, hohe Positionen kleinen Abständen. Zusätzlich sind noch die Signale der Probes mit Einzel-Loop gezeigt (grüne und gelbe Kurve). Die dunkelrote Kurve gibt die berechneten Signale nach Gl. 4.5 an.

Unabhängigkeit der Einflüsse der beiden Loops eher für große Abstände, d.h. niedrige Positionen von Loop A, gegeben ist, werden die Kurven im Bereich großer Abstände zur Deckung gebracht. Je kleiner der Abstand der beiden Loops wird, umso mehr unterscheiden sich die dunkelrote und blaue Kurve voneinander (siehe insbesondere Teil A) und D)). Somit kann man aus den Kurven einen ungefähren Abstand ablesen, ab dem die Unabhängigkeit der beiden Einflüsse  $x_A$  und  $x_B$  der beiden Loops nicht mehr gegeben ist. Dieser Abstand liegt bei ungefähr 5-8 Basen. Im nächsten Abschnitt wird der Wechselwirkungsabstand der beiden Loops mit einer sensitiveren Methode, der Fluoreszenzkorrelationsspektroskopie (FCS), untersucht.

## 4.5.4 FCS-Untersuchungen

### Experiment

Mit der Fluoreszenzkorrelationsspektroskopie (FCS von engl. *fluorescence correlation spectroscopy*) können unter anderem thermische Fluktuationen von doppelsträngiger DNA im Gleichgewicht untersucht werden. Dazu werden die DNA-Duplexe mit zwei Fluorophoren intern markiert, die als FRET-Paar dienen. Das FRET-Paar besteht in den hier durchgeführten Experimenten aus einer Probe, die mit Cy3 markiert ist, und einem mit Cy5 markiertem Target. Das Cy3-Fluorophor fungiert dabei als Donator des Cy5-Moleküls. Befinden sich die Fluorophore bei Anregung von Cy3 in engem räumlichen Abstand, so wird die Fluoreszenzemission von Cy3 durch Cy5 gequencht. Vergrößert sich die Distanz, kann wieder ein Fluoreszenzsignal von Cy3 gemessen werden. Die thermischen Fluktuationen einer DNA-Doppelhelix, die intern mit Donator und Akzeptor markiert ist, können so in Fluktuationen im Fluoreszenzsignal umgewandelt werden. Dadurch kann beispielsweise die Dynamik des Öffnungs- und Schließprozesses doppelsträngiger DNA (sog. *breathing*) untersucht werden [85].

Die im Folgenden vorgestellten Hybridisierungsexperimente werden in Lösung durchgeführt. Dazu werden vier verschiedene Probe-Sequenzen kommerziell erworben. Die Ausgangssequenz jeder Probe ist die Sequenz Map2k18833c-reverse, die zur verwendeten Target-Spezies perfekt komplementär ist (siehe Anhang B.1). Drei der vier Probe-Sequenzen enthalten im Vergleich zur Ausgangssequenz an zwei Positionen jeweils 3 zusätzliche Basen. Werden die Probes mit der Target-Spezies hybridisiert, bilden die zusätzlich eingefügten Basen in den Probes jeweils einen Loop aus. Da der Abstand der beiden gewählten Positionen je nach Probe-Sequenz unterschiedlich ist, ergeben sich dadurch drei unterschiedliche Abstände der Loops. In Tab. 4.3 sind die verwendeten Probe-Sequenzen und der entsprechende Abstand der beiden Loops angegeben. Die Positionsangaben der Loops beziehen sich dabei auf die Position, an der die zusätzlichen Basen in die Ausgangssequenz Map2k18833c-reverse der Probe eingefügt werden (Zählrichtung ist immer vom 3'- zum 5'-Ende). Die Probe-Sequenzen sind intern mit Cy3 markiert: PM-Probe an Position 18, die Probes mit zusätzlichen Basen an Position 21. Die Probe-Sequenzen bilden zusammen mit dem mit Cy5 intern markierten Target MAP2k18833c-reverse das FRET-Paar, welches sich immer zwischen den beiden Loops befindet. Die Markierung der Targets befindet sich an Basenposition 17. Nähert man die Loops einander und somit auch dem FRET-Paar an, so sollte dies Auswirkungen auf die Fluktuationen des Duplexes haben, die sich in einer veränderten Korrelationsfunktion niederschlagen. Um den Beitrag der Diffusion zur Korrelationsfunktion zu eliminieren, muss neben dem mit Cy5 markierten Target noch eine

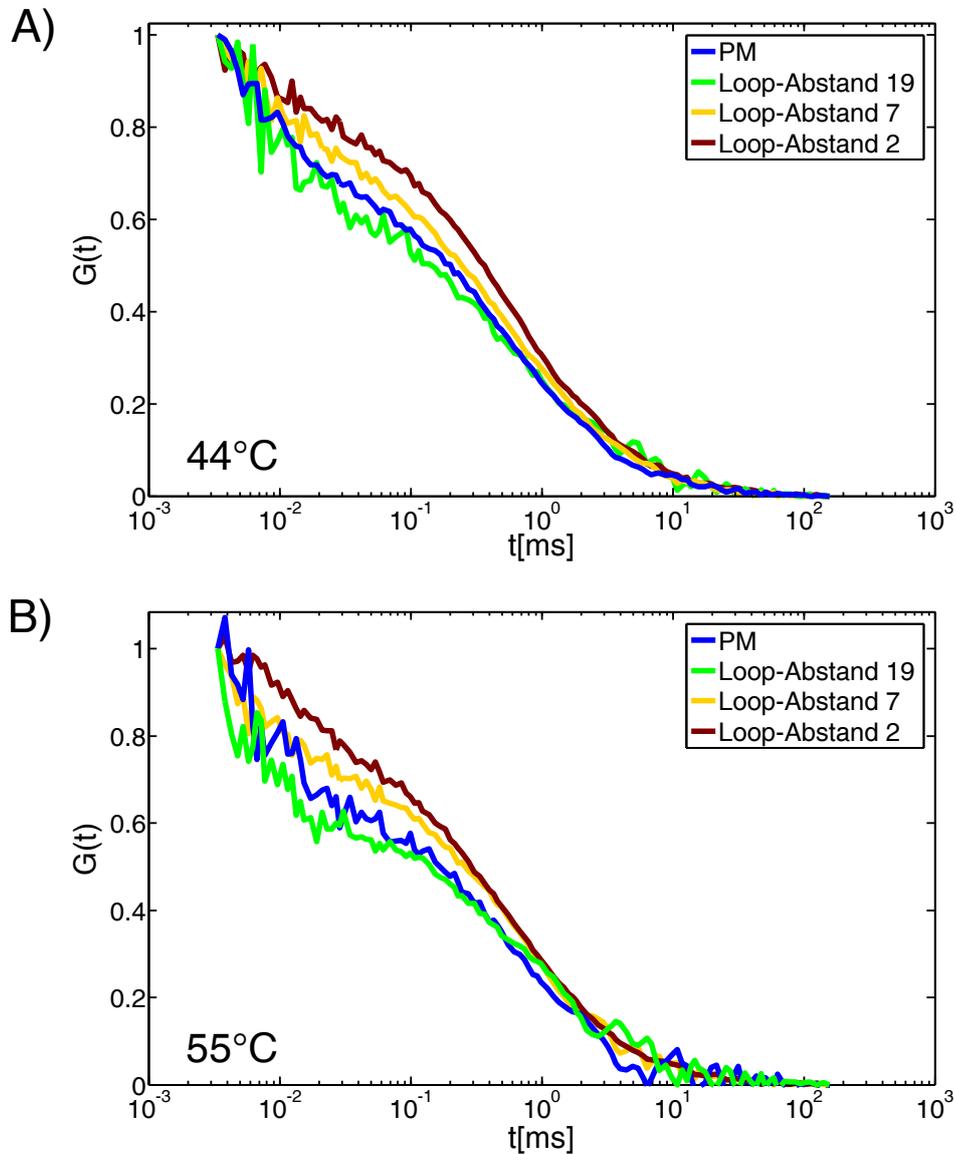
**Tabelle 4.3:** Verwendete Probe-Sequenzen für die FCS-Untersuchungen der Probes mit zwei Loops. Die Positionsangaben beziehen sich auf die Position, an der die Loops in die Ausgangssequenz des PM Map2k18833c-reverse eingebracht werden.

Name der Probe	Position Loop A	Position Loop B	Abstand der Loops
PM	/	/	/
Doppel-Loop-19	7	26	19
Doppel-Loop-7	13	20	7
Doppel-Loop-2	16	18	2

Target-Sequenz ohne interne Markierung verwendet werden. Bei Hybridisierung der Probes mit dem Target ohne interne Markierung wird die Korrelationsfunktion der Diffusion der DNA-Duplexe bestimmt. Dividiert man die Korrelationsfunktion, die bei Messung mit FRET-Paar erhalten wird, durch die Korrelationsfunktion bei Messung ohne FRET-Paar, so wird der Beitrag der Diffusion eliminiert (siehe Abschnitt 3.3 und [85]). Zur Messung der Korrelationsfunktionen der in Tab. 4.3 angegebenen Probes werden Lösungen aus 10 nM Probes und 20 nM Targets in 5xSSPE-Puffer hergestellt. Die Lösungen werden auf 90 °C erhitzt und relaxieren über Nacht auf Raumtemperatur. Durch die lange Relaxationszeit und die doppelte Target-Konzentration im Vergleich zur Probe-Konzentration wird sichergestellt, dass die meisten Probes mit Targets hybridisiert sind. Um die Statistik der Daten zu verbessern, werden für die Messungen der Diffusions- und FRET-Korrelationsfunktionen für jedes der vier Paare aus Probe und Target jeweils fünf Lösungen verwendet und daraus die Mittelwerte der Signale bestimmt (siehe Anhang A.2).

## Ergebnisse und Diskussion

Abb. 4.19 zeigt die Mittelwerte der diffusionsbereinigten Korrelationsfunktionen  $G(t)$  nach Subtraktion des Hintergrundes und Normierung der Amplitude auf 1 für die vier möglichen Probes bei zwei Temperaturen. Je nach Abstand der beiden Loops ergeben sich unterschiedliche Korrelationsfunktionen. Wie man anhand der Daten erkennen kann, bleibt der Unterschied in den Korrelationsfunktionen über einen großen Bereich von Loop-Abständen konstant: Die Korrelationsfunktionen für die Loop-Abstände von 7 bzw. 19 Basen unterscheiden sich nur wenig untereinander und von der Korrelationsfunktion des PM-Duplexes. Erst eine weitere Verringerung des Abstandes von 7 auf 2 Basen ändert die Korrelationsfunktion deutlich. Dies ist ein Hinweis darauf, dass sich die Öffnungs- und Schließdynamik der Duplexe erst bei Unterschreitung eines gewissen Abstandes der beiden Loops wesentlich von der Dynamik der PM-Probe unterscheidet. Daraus folgt, dass dieser Abstand im hier vorliegenden Fall bei der Einbringung



**Abbildung 4.19:** FCS-Untersuchungen der Probes mit zwei Loops. Die Abbildung zeigt die diffusionsbereinigten Korrelationsfunktionen  $G(t)$  für vier verschiedene Probe-Target-Duplexe bei Temperaturen von 44 °C (A)) bzw. 55 °C (B)). Erst bei Verringerungen des Loop-Abstandes auf zwei Basen ändert sich die entsprechende Korrelationsfunktion merklich von der des PM. Der Einfluss der Temperatur ist im untersuchten Temperaturbereich von 44 °C bis 55 °C marginal.

zweier Loops von je 3 Basen Länge in die Probe mit ungefähr 7 Basen beziffert werden kann. Unterschreiten die beiden Loops diesen Abstand, so treten sie miteinander in Wechselwirkung und können in theoretischen Beschreibungen nicht mehr unabhängig voneinander betrachtet werden. In Abschnitt 4.5.3 wurde der Wechselwirkungsabstand der beiden Loops anhand theoretischer Überlegungen mit 5-8 Basen angegeben. Die experimentellen Beobachtungen in diesem Abschnitt decken sich damit. Der erwartete Einfluss der Temperatur auf die Korrelationsfunktionen ist hier nicht zu erkennen. Der Grund dafür kann die aufgrund der Länge der Duplexe hohen Schmelztemperaturen sein (auch bei Einbringung zweiter Loops). Liegen die beiden realisierten Temperaturen unterhalb der Schmelztemperaturen aller verwendeter Duplexe, so sind die Duplexe relativ stabil und fluktuieren ähnlich. Wählt man zwei Temperaturen, von der die eine unterhalb und die andere oberhalb der Schmelztemperaturen der vier Duplexe liegt, sollten sich die Korrelationsfunktionen voneinander unterscheiden. Dies kann auch ausgenutzt werden, um die Unterschiede der Korrelationsfunktion für die verschiedenen Loop-Abstände deutlicher zu machen. Stellt man die Temperatur so ein, dass sie gerade mit dem niedrigsten Schmelzpunkt der vier Probe-Target-Duplexe zusammenfällt, so müssten sich die Korrelationsfunktionen deutlicher voneinander unterscheiden, da der DNA-Duplex am Schmelzpunkt stark fluktuiert. Die Berechnung und die experimentelle Bestimmung der Schmelztemperaturen der Duplexe ist aufgrund der Loops schwierig. Außerdem können wesentlich höhere Temperaturen im verwendeten experimentellen Aufbau nicht erreicht werden.

Im Falle der DNA-Microarrays waren die Signalunterschiede bei Annäherung der Loops bereits bei niedrigeren Temperaturen deutlich sichtbar. Die Ursache dafür sind die aufgrund der auftretenden Synthesefehler niedrigeren Schmelztemperaturen der Probe-Target-Duplexe. Auch hier sollte es eine ideale Temperatur geben, bei der die Signalunterschiede als Funktion des Abstandes der Loops am deutlichsten ist. Jedoch ist die Bestimmung der Schmelztemperatur mit den hier verwendeten Microarrays experimentell nicht möglich, da die Denaturierungskurve extrem verbreitert ist [1]. Es soll angemerkt werden, dass die Form Korrelationsfunktionen vom schnelleren der beiden Prozesse - Öffnung bzw. Schließung des Duplexes - bestimmt wird, sodass in der Nähe des Schmelzpunktes die Öffnung des Duplexes der dominierende Beitrag ist.

## 4.6 Zusammenfassung der Ergebnisse

### 4.6.1 Loops in der Probe

- Mit zunehmender Loop-Länge nimmt das Hybridisierungssignal ab. Dies bestätigte sich für drei unterschiedliche Loop-Sequenzen: poly-T-Loops, poly-C-Loops und beliebige Loop-Sequenzen. Die Grenze, ab der keine nennenswerte Hybridisierung mehr auftritt, wurde mit Hilfe eines Power-Laws der Form  $a \cdot x^{-b}$  zu 79 Basen bestimmt.
- Bei Variation der Loop-Position entlang des Duplexes entsteht ein bezüglich der Mitte des Duplexes symmetrisches „Loop-Defektprofil“. Probes mit Loops an mittleren oder endnahen Positionen liefern das stärkste Signal. Befindet sich der Loop zwischen diesen beiden Positionen, nimmt das Hybridisierungssignal ein Minimum an.
- Ein Einfluss der Microarray-Oberfläche auf das Hybridisierungssignal konnte nicht festgestellt werden.

### 4.6.2 Loops im Target

- Die Einbringung der Loops in die Target-Sequenz resultiert in stärkeren Signalen als bei Einbringung der Loops in die Probe. Mit Hilfe eines Skalierungsfaktors können die Daten ineinander umgerechnet werden. Die während der Synthese auftretenden Synthesefehler sind nicht für die unterschiedlichen Signale verantwortlich.
- Targets, die am 5'-Ende markiert sind, liefern bei Hybridisierung ein wesentlich stärkeres Fluoreszenzsignal als Targets mit Markierung am 3'-Ende. Als Grund konnte eine unspezifische Bindung der Fluorophore an die Oberfläche ausgeschlossen werden, da die Signale auch bei Verwendung von Abstandshaltern deutlich stärker sind. Die Markierung der Targets mit Fluorophoren hat generell Auswirkungen auf die Stabilität der Duplexe. Daher ist es möglich, dass Markierungen am 5'-Ende mehr zur Duplexstabilität beiträgt als eine entsprechende Markierung am 3'-Ende.

### 4.6.3 Probes mit zwei Loops

- Bei Verringerung des Abstandes der Loops steigt das Hybridisierungssignal an. Je länger die Einzel-Loops sind, desto schwächer ist das Hybridisierungssignal.
- Die Loops treten bei einem Abstand von ungefähr 5-8 Basen in Wechselwirkung (bei den untersuchten Einzel-Loop-Längen von 3 bzw. 5 Basen) und können in theoretischen Beschreibungen nicht mehr unabhängig voneinander betrachtet werden. Die Abhängigkeit des Wechselwirkungsabstandes von der Länge der Einzel-Loops muss in weiterführenden Experimenten untersucht werden.

## 5 DNA-Hybridisierung in Konkurrenz

Biologische Systeme basieren auf spezifischer molekularer Bindung und Erkennung. So müssen beispielsweise komplexe Moleküle mit einer Vielzahl von sehr ähnlichen Molekülen in einer Zelle um Bindungspartner konkurrieren. Um die Funktionalität der Zelle und des Organismus zu gewährleisten, ist es dabei von größter Wichtigkeit, dass Bindungen mit einer hohen Wahrscheinlichkeit nur zwischen passenden Molekülen stattfinden. Auch für biotechnologische Anwendungen, wie der Polymerase-Kettenreaktion (PCR) oder DNA-Microarrays, sowie für pharmazeutische Forschung, spielt das Prinzip der molekularen Erkennung eine zentrale Rolle. Die in der PCR-Technik eingesetzten Primer müssen an die richtigen Stellen des DNA-Templates binden, damit die gewünschte Sequenz vervielfältigt werden kann. Mit Hilfe der DNA-Microarray-Technologie können unter anderem komplexe Mischungen aus verschiedenen unbekannt DNA-Sequenzen untersucht und die Konzentration und Sequenz der einzelnen Spezies bestimmt werden. Die Qualität und die Genauigkeit der Analyse hängt auch davon ab, wie gut die verschiedenen Spezies nur an den für sie vorgesehenen komplementären Probes hybridisieren. Eine unspezifische Hybridisierung der Spezies ist nicht erwünscht, da dies die Messergebnisse verfälschen würde.

In den in Kapitel 4 vorgestellten Experimenten wurden die Microarrays nur mit Target-Lösungen bestehend aus einer einzelnen Target-Spezies hybridisiert, um eine Konkurrenzhybridisierung zu vermeiden. Im Folgenden wird diese Vereinfachung aufgehoben und die DNA-Konkurrenzhybridisierung zwischen zwei Target-Spezies ähnlicher Sequenz unter kontrollierten Bedingungen untersucht. Die in den Abschnitten 5.2 bis 5.3.2 präsentierten Experimente wurden im Rahmen von [96] durchgeführt und gemeinsam interpretiert (siehe auch [97]). Die Vorstellung der Ergebnisse ist zum Verständnis des Kapitels, insbesondere für Abschnitt 5.4, und für das erarbeitete Modell in Abschnitt 6.2 essentiell.

## 5.1 Effekt der hochspezifischen DNA-Hybridisierung

### 5.1.1 Experiment

In diesem Experiment soll die Konkurrenz-Hybridisierung zwischen der Target-Spezies PM und einer weiteren Target-Spezies MM1 bzw. MM2 untersucht werden. Alle verwendeten Targets sind jeweils 16 Basen lang. Die Target-Spezies PM ist am 3'-Ende mit Cy3 markiert und perfekt komplementär zu den Probes auf der Oberfläche. MM1 bzw. MM2 sind unmarkiert und unterscheiden sich von der PM-Sequenz nur an einer Position: Die Fehlstelle befindet sich bei der Target-Spezies MM1 in der Mitte des Strangs, bei Target-Spezies MM2 an einer endnahen Position (siehe Tab. 5.1 und Anhang B.1). Die in Tab. 5.1 angegebenen Bindungskonstanten  $K$  wurden in [96] auf einer Oberfläche mit der *total internal reflection fluorescence*-Methode (TIRF) bestimmt. Zur Untersuchung werden zwei Chips synthetisiert, die entweder mit verschiedenen Target-Lösungen bestehend aus PM und MM1 (System „PM vs. MM1“) oder vier Target-Lösungen bestehend aus PM und MM2 (System „PM vs. MM2“) hybridisiert werden. Die Target-Lösungen unterscheiden sich dabei in der Konzentration der Target-Spezies MM1 bzw. MM2: Während die Konzentration der PM-Spezies konstant auf 1 nM gehalten wird, wird die Konzentration der Targets mit Fehlstelle - MM1 und MM2 - schrittweise von 0 nM auf 1000 nM erhöht. Die Chips werden zunächst nur mit einer 1 nM Lösung aus PM hybridisiert (Einzelhybridisierung, Konzentration der Konkurrenz-Spezies beträgt 0 nM). Nach Erreichen des Gleichgewichts bei einer Temperatur von 44 °C werden die Hybridisierungssignale der PM-Spezies aufgenommen. Nach Ausspülen der alten Target-Lösung und mehrminütiger Reinigung der Hybridisierungskammer mit 5xSSPE-Puffer bei 80 °C werden die Chips mit einer neuen Target-Lösung hybridisiert und die Prozedur wiederholt.

### 5.1.2 Ergebnisse und Diskussion

Abb. 5.1 zeigt die Hybridisierungssignale des PM als Funktion der Konzentration der Konkurrenzspezies MM1 (A)) bzw. MM2 (B)) nach Subtraktion des Hintergrunds und Normierung der Signale auf das Hybridisierungssignal des PM bei Einzelhybridisierung (Konzentration der Konkurrenz-Spezies beträgt 0 nM) und auf den Hintergrund. Die Konzentration des PM beträgt immer 1 nM. Im Konkurrenzsystem „PM vs. MM2“ (B)) nimmt das Signal des PM bei Konzentrationserhöhung von MM2 kontinuierlich ab und erreicht bei einer MM2-Konzentration von 1000 nM nur noch 8% des Signals der Einzelhybridisierung. Im Fall A) hat die Konzentration von MM1 keinen Einfluss

**Tabelle 5.1:** Sequenzen und Bindungskonstanten aller verwendeter Targets. Das Target PM ist perfekt komplementär zu den Probes auf der Oberfläche. Die Targets MM1 und MM2 besitzen eine Fehlstelle, Targets MM3 und MM4 besitzen 2 (die entsprechenden Basen sind rot markiert). Die angegebenen Bindungskonstanten wurden in [96, 97] mittels TIRF auf der Oberfläche bestimmt.

Name des Targets	Sequenz (3'-5')	$K$ [1/M] bei 44 °C
PM	AAT-GCT-AGA-CTA-GGA-A	$(9 \pm 5) \cdot 10^8$
MM1	AAT-GCT-ACA-CTA-GGA-A	$(6 \pm 1) \cdot 10^7$
MM2	AAC-GCT-AGA-CTA-GGA-A	$(4 \pm 2) \cdot 10^8$
MM3	AAT-GCT-ACT-CTA-GGA-A	$(1 \pm 0.1) \cdot 10^7$
MM4	AAA-GCT-AGA-CTA-GAA-A	$(1 \pm 0.3) \cdot 10^8$

auf das Hybridisierungssignal des PM. Selbst bei 1000-fach erhöhter Konzentration der MM1-Targets im Vergleich zur PM-Konzentration bleibt das Signal des PM mit 97% des Signals der Einzelhybridisierung nahezu unverändert.

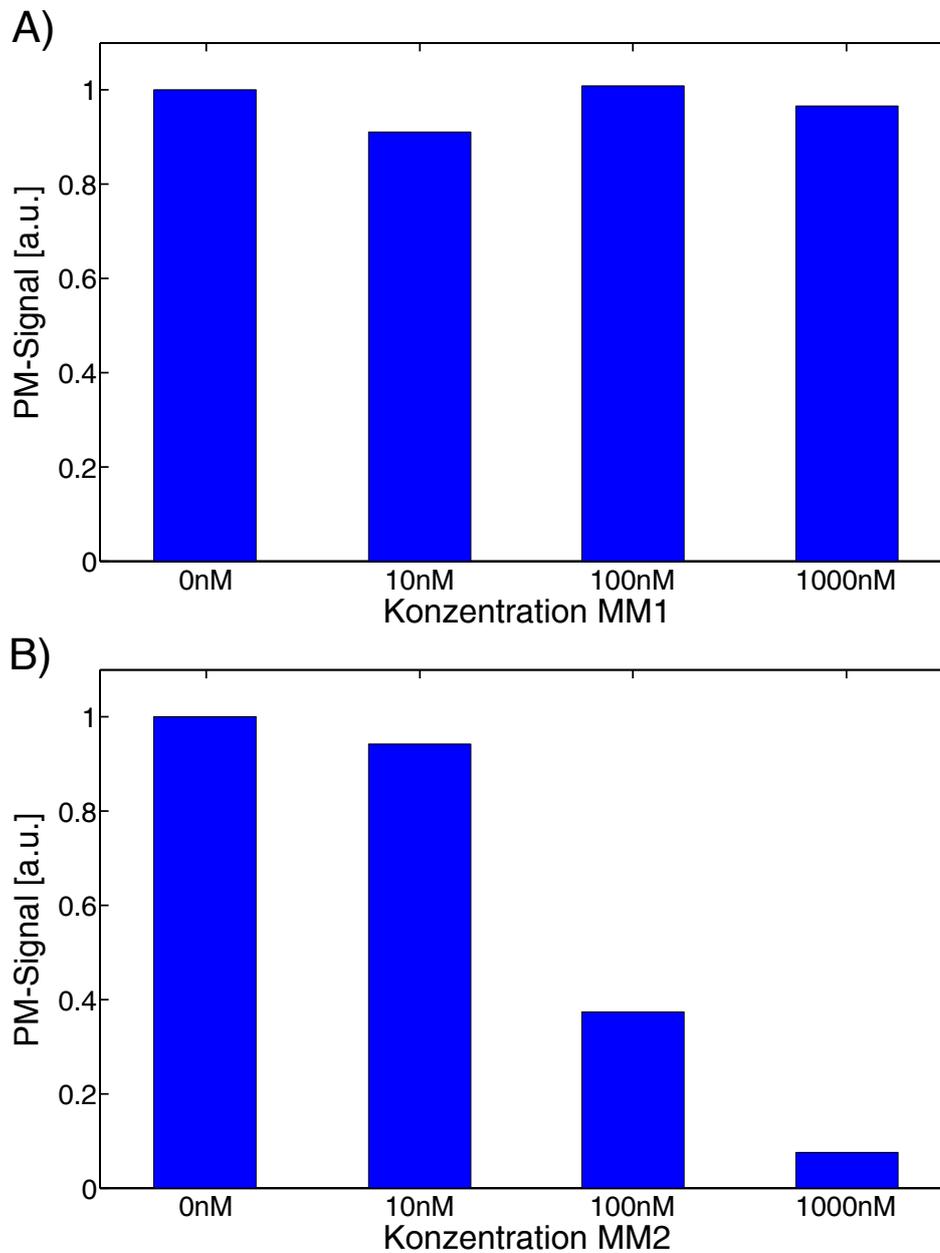
Im thermodynamischen Gleichgewicht berechnet sich die Oberflächenkonzentration  $[D_A]$  einer Target-Spezies A in Konkurrenz mit einer Target-Spezies B mit (siehe Gl. (2.6) aus Abschnitt 2.3):

$$[D_A] = \frac{K_A \cdot [P_0] \cdot [A_0]}{1 + K_A \cdot [A_0] + K_B \cdot [B_0]} \quad (5.1)$$

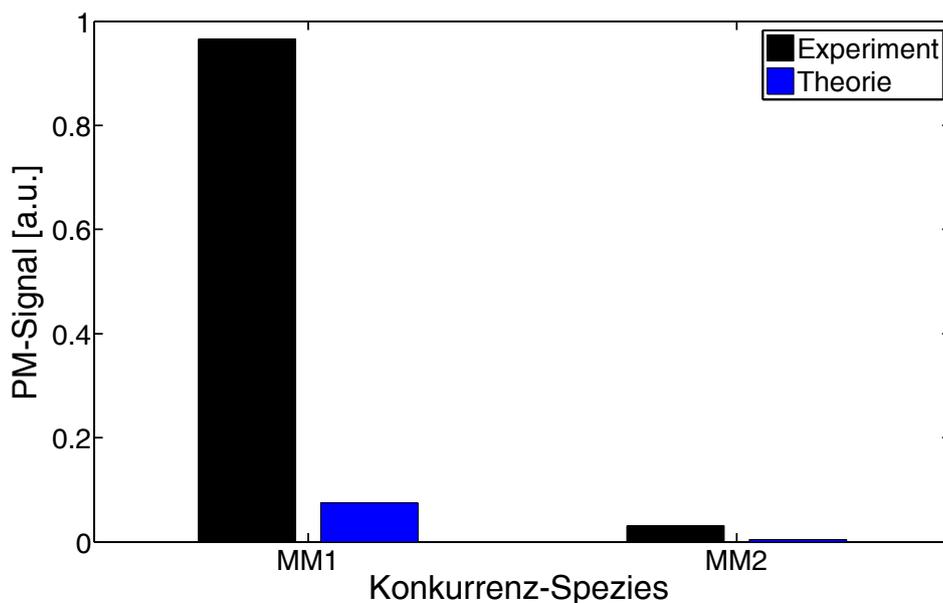
Dabei bezeichnen  $[A_0]$  bzw.  $[B_0]$  die Konzentrationen der beiden Spezies in Lösung,  $K_A$  bzw.  $K_B$  die individuellen Bindungskonstanten der Einzelhybridisierung und  $[P_0]$  die Probe-Konzentration. Für das Verhältnis der Duplexkonzentrationen  $[D_A]/[D_B]$  der beiden konkurrierenden Spezies A und B ergibt sich dann:

$$\frac{[D_A]}{[D_B]} = \frac{[A_0] \cdot K_A}{[B_0] \cdot K_B} = \frac{[A_0]}{[B_0]} \cdot e^{\beta \Delta \Delta G} \quad (5.2)$$

$\Delta \Delta G = \Delta G_A - \Delta G_B$  ist der Unterschied der effektiven freien Bindungsenergie der beiden Target-Spezies,  $\beta = 1/RT$  mit der Gaskonstanten  $R$  und der Temperatur  $T$ . Nach Gl. 5.2 hängt das Verhältnis der Oberflächenbedeckungen also nur vom Verhältnis der Target-Konzentrationen und dem Unterschied in der freien Bindungsenergie ab. Da sich die Bindungskonstanten von PM und MM1 nur um eine Größenordnung unterscheiden, die Konzentration von MM1 aber bis zu 1000-fach erhöht ist, ist die Hybridisierung des PM in Abb. 5.1 A) unerwartet spezifisch: Bei hohen Konzentrationen von MM1 sollte wie im Falle „PM vs. MM2“ nur noch ein stark vermindertes PM-Signal messbar sein. Die hochspezifische Hybridisierung des PM im Konkurrenzsystem „PM vs. MM1“ ist in



**Abbildung 5.1:** Hybridisierungssignale des PM bei verschiedenen Konzentration der Konkurrenz-Targets für die beiden Systeme „PM vs. MM1“ bzw. „PM vs. MM2“ bei einer Temperatur von 44 °C. Die Signale sind jeweils auf das Signal der Einzelhybridisierung ( $[PM]=1$  nM,  $[MM1]$  bzw.  $[MM2]=0$  nM) normiert. Die Konzentration des PM beträgt immer 1 nM. **A)** System „PM vs. MM1“: Auch eine 1000 fache Konzentrationserhöhung von MM1 beeinflusst das Hybridisierungssignal des PM nicht. **B)** „PM vs. MM2“: Mit zunehmender MM2-Konzentration nimmt das PM-Signal entsprechend Gl. 5.1 ab.



**Abbildung 5.2:** Vergleich der experimentellen Daten mit der theoretischen Vorhersage für die beiden Konkurrenzsysteme „PM vs. MM1“ und „PM vs. MM2“ bei einer Temperatur von 44 °C. Die Konzentration des PM beträgt 1 nM, die Konzentration der Konkurrenten MM1 bzw. MM2 1000 nM. Alle Daten sind auf den jeweiligen Wert der Einzelhybridisierung ( $[PM]=1$  nM,  $[MM1]$  bzw.  $[MM2]=0$  nM) normiert. Die experimentell bestimmten Signale sind in schwarz dargestellt. Die theoretischen Vorhersagen (blau) werden mit Hilfe von Gl. 5.1 und den Bindungskonstanten aus Tab. 5.1 bestimmt. Im System „PM vs. MM1“ auf der linken Seite kann das PM-Signal in Anwesenheit der 1000-fach höher konzentrierten MM1-Spezies mit Gl. 5.1 nicht vorhergesagt werden. System „PM vs. MM2“ lässt sich im Rahmen der Messgenauigkeit anhand dieser Gleichung beschreiben.

Abb. 5.2 verdeutlicht. Dort sind die experimentell bestimmten (schwarz) und die nach Gl. 5.1 berechneten (blau) normierten Signale des PM bei einer Konzentration von 1000 nM der Konkurrenzspezies MM1 bzw. MM2 gezeigt (zur Berechnung werden die Bindungskonstanten aus Tab. 5.1 benutzt.). Im Fall „PM vs. MM2“ können die Signale mit Gl. 5.1 im Rahmen der Messgenauigkeit korrekt wiedergegeben werden. Im Konkurrenzsystem „PM vs. MM1“ ist die Diskrepanz zwischen der Vorhersage der Theorie und den experimentell ermittelten Daten deutlich zu erkennen. Der Effekt der hochspezifischen Hybridisierung soll in den folgenden Abschnitten detailliert charakterisiert und die Bedingungen für das Auftreten des Effekts untersucht werden. Die Messungen auf der Oberfläche mit der *total internal reflection fluorescence*-Methode (TIRF) und die Messungen mit *Plate Reader* in Lösung mittels FRET wurden im Rahmen von [96] durchgeführt und gemeinsam interpretiert (siehe auch [97]). Die Präsentation der Ergebnisse ist zum Verständnis der hochspezifischen Hybridisierung notwendig.

## 5.2 Ausschluss von Artefakten

In diesem Abschnitt soll untersucht werden, ob der Effekt der hochspezifischen Hybridisierung durch Artefakte in den Messungen oder den verwendeten Messmethoden verursacht wird.

### 5.2.1 Nicht-Gleichgewicht

Das Auftreten der hochspezifischen Hybridisierung des PM in Abschnitt 5.1 ist nur im thermodynamischen Gleichgewicht unerwartet, da dann das Verhältnis der Oberflächenbedeckungen der beiden konkurrierenden Spezies lediglich von den jeweiligen Konzentrationen und den individuellen Bindungskonstanten aus den Einzelhybridisierungen abhängt (siehe Gl. 5.2). Da sich hier die Bindungskonstanten der Spezies nur um maximal eine Größenordnung unterscheiden, die Konzentration aber bis zu 1000-fach höher ist, sollte in beiden Systemen „PM vs. MM1“ und „PM vs. MM2“ die PM-Spezies von der Oberfläche verdrängt werden. Jedoch wird dies nur im System „PM vs. MM2“ beobachtet und nicht im System „PM vs. MM1“. In einer Nicht-Gleichgewichtssituation könnte der Effekt beispielsweise durch sehr lange Relaxationsraten erklärt werden: Das unerwartet starke PM-Signal wäre nur eine Momentaufnahme und kein stabiler Zustand des Systems. Daher soll im Folgenden untersucht werden, ob sich die Konkurrenzsysteme wirklich im thermodynamischen Gleichgewicht befinden.

### Experiment

Zur Überprüfung des Gleichgewichts werden drei Experimente durchgeführt (siehe [96, 97]):

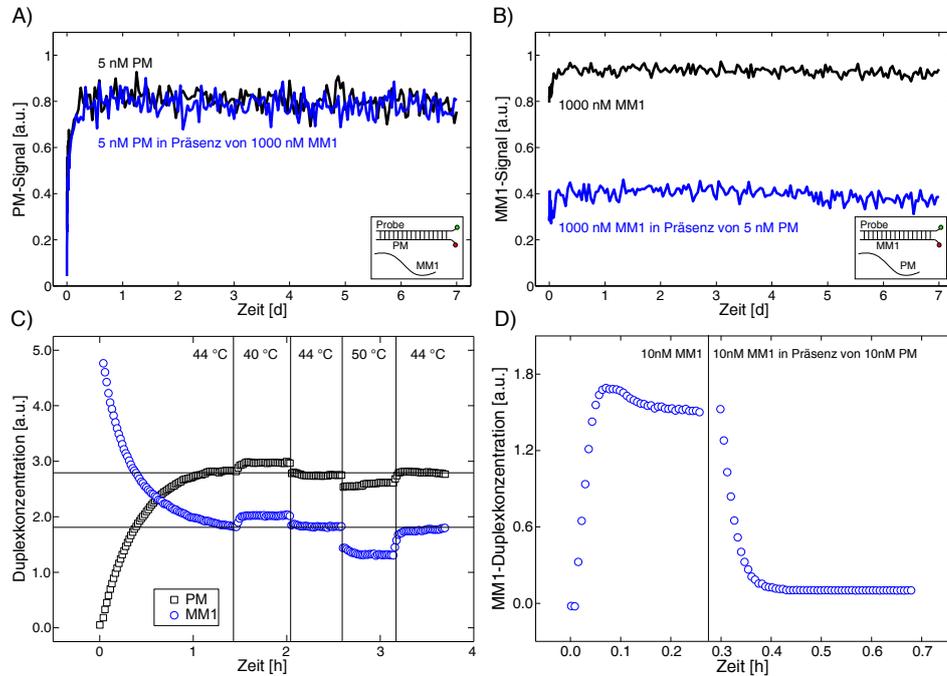
- Um lange Relaxationsraten des Systems „PM vs. MM1“ auszuschließen, werden Langzeit-Messungen in Lösung mittels FRET und *Plate Reader* durchgeführt. Die Messung in Lösung bietet sich an, da sich die an der Oberfläche immobilisierten Probes auf sehr langen Zeitskalen lösen könnten, was das Signal verfälschen würde. Das FRET-Paar besteht entweder aus PM und Probe oder MM1 und Probe. Probe und zu untersuchende Target-Spezies sind mit Cy3 bzw. mit Cy5 markiert, während die entsprechende Konkurrenzspezies unmarkiert bleibt. Bei Hybridisierung der markierten Target-Spezies mit der Probe und Anregung des Cy3-Fluorophors kann aufgrund von FRET ein Cy5-Signal gemessen werden. Hybridisiert die unmarkierte Target-Spezies, so findet kein Energieübertrag auf das Cy5-Fluorophor statt und es kann kein Cy5-Signal gemessen werden. Dadurch

kann unterschieden werden, welche Target-Spezies mit der Probe einen Duplex bildet. Die Konzentration der Probes beträgt 10 nM, die der PM-Targets 5 nM und die der MM1-Targets 1000 nM.

- Befindet sich das System im Gleichgewicht, so stellen sich bei einer bestimmten Temperatur immer die selben Gleichgewichtswerte der Duplexkonzentrationen der PM-Probe- und MM1-Probe-Duplexe ein. Ergibt sich bei Änderung der Temperatur und Rückkehr zur Ausgangstemperatur eine Hysterese, so ist dies ein starkes Indiz für eine Nicht-Gleichgewichtssituation. Daher wird in einer Oberflächenhybridisierung eine Lösung aus 5 nM PM (Cy5-Markierung) und 1000 nM MM1 (Cy3-Markierung) in die Hybridisierungskammer gegeben und die Duplexkonzentrationen der beiden Spezies mittels TIRF-Methode gemessen. Nach Erreichen des vermeintlichen Gleichgewichts bei 44 °C wird die Temperatur alternierend erhöht oder abgesenkt und überprüft, ob die Ausgangswerte der Duplexkonzentrationen bei 44 °C wieder erreicht werden.
- Ob die Reihenfolge der Hybridisierungen einen Einfluss auf das Gleichgewichtssignal hat und die Bindung einer Target-Spezies an die Oberfläche reversibel ist, wird überprüft, indem die Oberfläche zuerst mit der MM1-Spezies bis Erreichen des Gleichgewichts hybridisiert ( $[MM1]=10$  nM) und dann eine Lösung aus PM und MM1 ( $[PM]:[MM1]=10:10$  nM) in die Hybridisierungskammer gegeben wird. In den vorherigen Experimenten wurden die Probes zuerst mit der PM-Spezies hybridisiert. Die Fluoreszenzsignale werden mit der TIRF-Methode gemessen.

## Ergebnisse und Diskussion

Abb. 5.3 zeigt die Experimente, die zum Ausschluss von Artefakten durchgeführt werden. In A) und B) sind die normierten Signale der Langzeitmessungen in Lösung mittels FRET bei 44 °C gezeigt. Das FRET-Paar besteht in A) aus PM und Probe, in B) aus MM1 und Probe. Die schwarzen Kurven zeigen die Signale der Einzelhybridisierung von PM bzw. MM1. Die blauen Kurven entsprechen den Signalen der Konkurrenzhybridisierungen. Die Konzentration des PM ist immer 5 nM, die des MM1 1000 nM. Lange Relaxationszeiten der Systeme können ausgeschlossen werden, da die Fluoreszenzsignale über einen Zeitraum von mindestens einer Woche konstant sind. In C) sind die absoluten Duplexkonzentrationen beider Spezies PM und MM1 als Funktion der Zeit und der Temperatur gezeigt. Die Ausgangstemperatur von 44 °C wird alternierend erhöht oder abgesenkt. Nachdem das System bei den gewählten Temperaturen im Gleichgewicht ist (horizontale Linien), wird wieder die Ausgangstemperatur von 44 °C eingestellt. Die Rückkehr zur Ausgangstemperatur resultiert immer in den identischen



**Abbildung 5.3:** Experimente zum Ausschluss von Artefakten (durchgeführt von Marc Schenkelberger [96], siehe auch [97]). **A)+B)** Gezeigt sind die normierten Signale der Langzeitmessung in Lösung mittels FRET bei 44 °C. Das FRET-Paar besteht in (A)) aus PM und Probe, in (B)) aus MM1 und Probe. Die entsprechende Konkurrenzspezies ist unmarkiert. Die schwarzen Kurven entsprechen den Signalen der Einzelhybridisierung, die blauen geben die Signale in Konkurrenz an. Alle Signale sind über einen Zeitraum von mindestens einer Woche konstant, sodass lange Relaxationsraten ausgeschlossen werden können. **C)** Hier sind die absoluten Duplexkonzentrationen im System „PM vs. MM1“ bei einem Konzentrationsverhältnis von  $[PM]:[MM1]=5:1000$  nM als Funktion der Zeit gezeigt. Wird die Ausgangstemperatur von 44 °C alternierend erhöht oder abgesenkt (vertikale Linien), stellen sich die ursprünglichen Duplexkonzentrationen bei der Rückkehr zu 44 °C wieder ein (horizontale Linien). Es ist keine Hysterese messbar. **D)** Der Effekt der hochspezifischen Hybridisierung tritt auch dann auf, wenn die Oberfläche zuerst mit einer 10 nM MM1-Lösung gesättigt wird (linke Seite) und dann die PM-Spezies mit einer Konzentration von 10 nM zusätzlich in die Kammer gegeben wird (rechte Seite). Die Bindung der Targets an die Oberfläche ist reversibel. Anhand dieser Experimente können Artefakte in den Messungen ausgeschlossen werden.

Duplexkonzentrationen, sodass keine Hysterese messbar ist. Es soll angemerkt werden, dass sich die Variation des Signals bei Änderung der Temperatur aus zwei Beiträgen zusammensetzt: der Menge der hybridisierten Targets und der Emissionscharakteristik des Fluorophors. In D) ist gezeigt, dass der Effekt der hochspezifischen Hybridisierung nicht von der Reihenfolge der Hybridisierungen abhängt. Wird die Oberfläche zuerst mit MM1 hybridisiert und dann durch eine Target-Lösung bestehend aus PM und MM1 ersetzt, so verdrängt der PM die MM1-Targets trotzdem von der Oberfläche. Die Bindung einer Target-Spezies an die Oberfläche ist also reversibel. Zusammenfassend lässt sich sagen, dass keine Hinweise auf ein Nicht-Gleichgewicht der Systeme gefunden werden konnte. Der Effekt der hochspezifischen DNA-Hybridisierung in Konkurrenz tritt im thermodynamischen Gleichgewicht auf.

## 5.2.2 Messungen in Lösung

### Experiment

Um einen störenden Einfluss der Oberfläche und eine daraus möglicherweise resultierende Verfälschung der Daten zu vermeiden, wird die DNA-Konkurrenzhybridisierung der Systeme „PM vs. MM1“ und „PM vs. MM2“ in Lösung mittels FRET und *Plate Reader* untersucht (siehe [96, 97]). Das FRET-Paar besteht aus der mit Cy3 markierten Probe und dem mit Cy5 markierten Target. Das entsprechende Konkurrenz-Target ist unmarkiert. So ergeben sich pro System zwei mögliche Kombinationen von markierten und unmarkierten Targets. Die Konzentration der mit Cy5 markierten PM-Targets beträgt immer 5 nM, während die Konzentration der unmarkierten Spezies von 0 nM schrittweise auf 1000 nM erhöht wird. Ist umgekehrt MM1 bzw. MM2 mit Cy5 markiert, so wird die Konzentration dieser Spezies konstant auf 100 nM gehalten und die Konzentration der unmarkierten PM-Targets von 0 nM schrittweise auf 20 nM erhöht. Alle Messungen werden bei 44 °C durchgeführt. Die Bindungskonstanten der verwendeten Targets in Lösung wurden in [96, 97] bestimmt und sind in Tab. 5.2 angegeben. Anzumerken bleibt, dass die in Lösung und auf der Oberfläche bestimmten Bindungskonstanten sehr gut übereinstimmen.

### Ergebnisse und Diskussion

Abb. 5.4 zeigt die normierten Signale der beiden Konkurrenzsysteme „PM vs. MM1“ und „PM vs. MM2“ als Funktion der Zeit und Konzentration der unmarkierten Target-Spezies bei Messung in Lösung für die vier möglichen Kombinationen von Markierungen

**Tabelle 5.2:** In Lösung mittels FRET und *Plate Reader* bestimmte Bindungskonstanten der verwendeten Targets (durchgeführt von Marc Schenkelberger [96], siehe auch [97]).

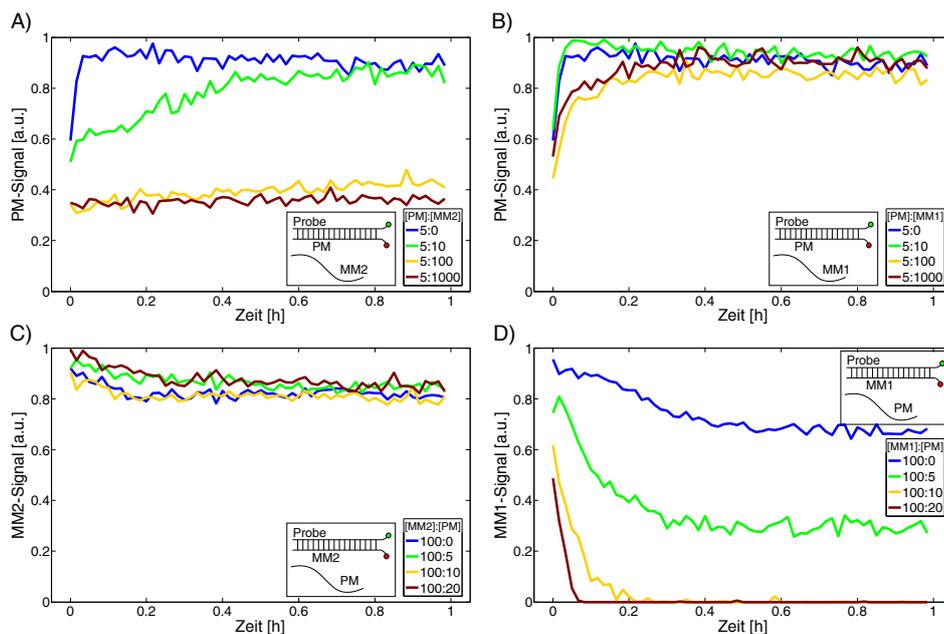
Name des Targets	$K$ [1/M] bei 44 °C
PM	$(2 \pm 0.5) \cdot 10^8$
MM1	$(1.3 \pm 0.5) \cdot 10^7$
MM2	$(2 \pm 0.4) \cdot 10^8$

der beiden Konkurrenz-Targets. Das verwendete FRET-Paar (Cy3 markierte Probe und Cy5 markiertes Target) und das unmarkierte Target sind in der jeweiligen Teilabbildung angegeben. Bei Erhöhung der MM2-Konzentration und konstanter PM-Konzentration in A) nimmt das PM-Signal ab. Dies ist konsistent mit dem Verhalten des Systems „PM vs. MM2“ in Abschnitt 5.1 (Abb. 5.1 B)) bei der Messung auf dem Microarray. Im Kontrollexperiment C), bei dem die Konzentration des PM erhöht und die Konzentration von MM2 konstant auf 100 nM gehalten wird, bleibt das MM2-Signal zeitlich konstant. Die geringe Konzentration des PM hat keinen Einfluss auf das MM2-Signal. Im Konkurrenzsystem „PM vs. MM1“ hat die Konzentrationserhöhung von MM1 in B) keinen Einfluss auf das PM-Signal. Dies entspricht ebenfalls dem in Abschnitt 5.1 (Abb. 5.1 A)) beobachteten Verhalten. Da sich die Bindungskonstanten von PM und MM1 auch in Lösung nur geringfügig unterscheiden, würde Gl. 5.1 hier einen Rückgang des PM-Signals bei 200-facher Überkonzentration des MM1 vorhersagen. Das Kontrollexperiment D) bestätigt, dass bereits eine leichte PM-Konzentrationserhöhung eine große Auswirkung auf das MM1-Signal hat. Bei einer PM-Konzentration von 10 nM liegt das MM1-Signal unter der Auflösungsgrenze des *Plate Readers*. Zusammenfassend lässt sich sagen, dass der beobachtete Effekt der hochspezifischen PM-Hybridisierung in Konkurrenz nicht durch Oberflächeneffekte zustande kommt. Die Messungen in Lösung bestätigen die Ergebnisse der Microarray-Experimente.

### 5.2.3 Wahl des Sequenz-Motivs

#### Experiment

Um auszuschließen, dass der beobachtete Effekt der hochspezifischen Hybridisierung in Konkurrenz nicht nur bei dem in Tab. 5.1 angegebenen Sequenz-Motiv auftritt, wird ein weiteres Sequenz-Motiv entworfen und die Hybridisierung in Konkurrenz untersucht (siehe [96, 97]). Das Motiv beinhaltet eine neue Probe-Sequenz, eine dazu perfekt komplementäre Sequenz PM\* und eine Sequenz mit Fehlstellen MM\*. Das



**Abbildung 5.4:** Messungen der DNA-Konkurrenzhybridisierung in Lösung mittels FRET (durchgeführt von Marc Schenkelberger [96], siehe auch [97]). **A)** Konkurrenzsystem „PM vs. MM2“: Das FRET-Paar besteht aus PM und Probe während MM2 unmarkiert ist. Wird die Konzentration von MM2 auf 1000 nM erhöht und die PM-Konzentration konstant auf 5 nM gehalten, so nimmt das PM-Signal ab. **B)** Konkurrenzsystem „PM vs. MM1“: Das FRET-Paar besteht aus PM und Probe, MM1 ist unmarkiert. Auch ein Konzentrationsverhältnis von  $[PM]:[MM1]=5:1000$  nM hat keinen Einfluss auf das PM-Signal. Die Messungen in Lösung zeigen also das gleiche Verhalten wie die Messungen auf dem Microarray (siehe Abb. 5.1). **C)** Konkurrenzsystem „PM vs. MM2“: Das FRET-Paar besteht hier aus MM2 und Probe, der PM bleibt unmarkiert. Während die Konzentration von MM2 konstant auf 100 nM gehalten wird, wird die Konzentration von PM von 0 nM auf 20 nM erhöht. Unabhängig von der PM-Konzentration erreicht das MM2-Signal immer das gleiche Niveau. **D)** Konkurrenzsystem „PM vs. MM1“: MM1 und Probe bilden das FRET-Paar, während der PM unmarkiert bleibt. Im Gegensatz zu Teil C) bewirkt eine geringe Konzentrationserhöhung des PM einen starken Rückgang des MM1-Signals. Bei Konzentrationsverhältnissen von  $[MM1]:[PM]=100:10$  bzw.  $100:20$  (jeweils in nM) liegt das MM1-Signal unterhalb der Auflösungsgrenze des *Plate Readers*.

**Tabelle 5.3:** Sequenzen und Bindungskonstanten der verwendeten Targets in Abschnitt 5.2.3. Das Target PM\* ist perfekt komplementär zu den Probes auf der Oberfläche. MM\* hat zwei in rot markierte Fehlstellen in der Mitte. Die Sequenzen der Targets orientieren sich an den Molekülen TOI und COM aus [98]. Die angegebenen Bindungskonstanten wurden mittels TIRF bestimmt (durchgeführt von Marc Schenkelberger [96], siehe auch [97]).

Name des Targets	Sequenz (3'-5')	$K$ [1/M] bei 44 °C
PM*	CAT-GAT-AAC-GAC-GGG	$(2 \pm 0.3) \cdot 10^9$
MM*	CAT-GAT- <b>TTC</b> -GAC-GGG	$(2 \pm 0.8) \cdot 10^9$

neue Sequenz-Motiv ähnelt dabei sehr stark den Molekülen TOI und COM, die in [98] verwendet wurden. Tab. 5.3 zeigt die verwendeten Target-Sequenzen und die mit TIRF experimentell bestimmten Bindungskonstanten  $K$  bei 44 °C. Die Fehlstellen von MM\* sind rot gekennzeichnet. Analog zu den vorherigen Experimenten wird die PM\*-Konzentration konstant auf 5 nM gehalten, die Konzentration von MM\* schrittweise von 0 nM auf 1000 nM erhöht und die Duplexkonzentration des PM\* auf der Oberfläche mittels TIRF. Die gewählte Temperatur im Experiment beträgt 44 °C.

### Ergebnisse und Diskussion

Abb. 5.5 zeigt die Duplexkonzentrationen des PM\* auf der Oberfläche nach Normierung auf die Duplexkonzentration des PM\* aus der Einzelhybridisierung ( $[PM^*]=5$  nM,  $[MM^*]=0$  nM). Selbst eine 200-fache Überkonzentration des MM\* (Erhöhung der MM\*-Konzentration auf 1000 nM) hat keinen Einfluss auf die PM\*-Duplexkonzentration. Die PM\*-Spezies hybridisiert auch bei dem geänderten Sequenzmotiv so, als wäre keine Konkurrenzspezies anwesend. Die Vorhersagen von Gl. 5.1 unter Verwendung der Bindungskonstanten aus Tab. 5.3 geben die experimentellen Daten nicht wieder (Symbole hinter den Kurven). Dies bestätigt den Effekt der hochspezifischen Hybridisierung und zeigt gleichzeitig, dass der Effekt nicht nur im Rahmen des gewählten Sequenzmotivs aus Abschnitt 5.1 auftritt.

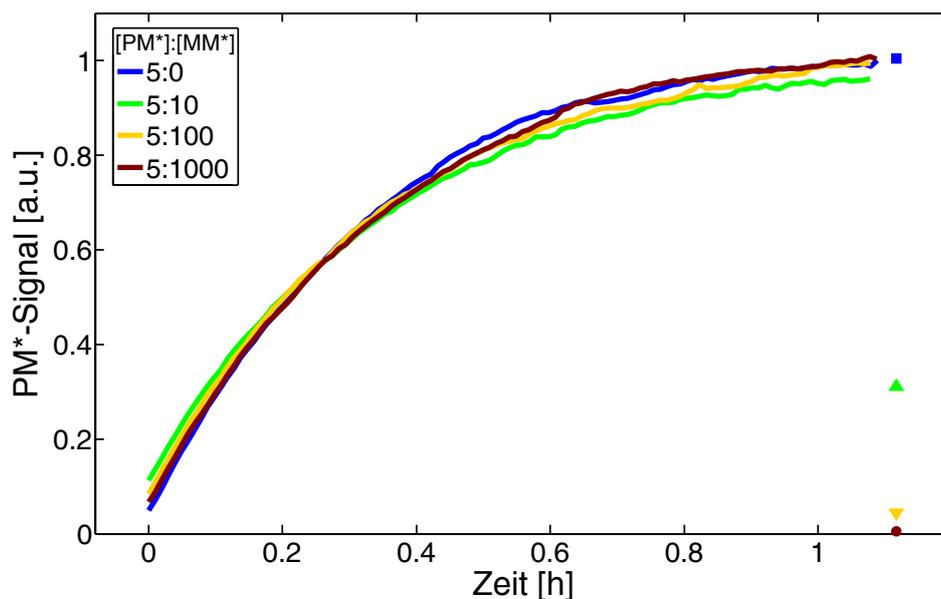
## 5.3 Charakterisierung des Effekts der hochspezifischen Hybridisierung

In den vorherigen Abschnitten wurde die hochspezifische DNA-Hybridisierung in Konkurrenz vorgestellt und durch ergänzende Messungen ausgeschlossen, dass Artefakte in den verwendeten Messmethoden als Ursache für den Effekt verantwortlich sind. Im Folgenden sollen der Effekt der hochspezifischen Hybridisierung durch weitere Untersuchungen charakterisiert und die Bedingungen für sein Auftreten untersucht werden.

### 5.3.1 Unterteilung in Systeme

#### Experiment

Es werden Konkurrenzhybridisierungen zwischen allen in Tab. 5.1 aufgeführten Sequenzen durchgeführt. Dabei wird die Konzentration der Target-Spezies mit der größeren

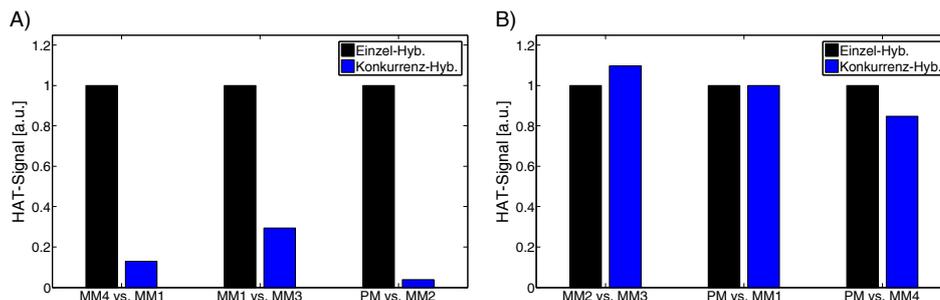


**Abbildung 5.5:** Konkurrenzhybridisierung zwischen PM\* und MM\* (durchgeführt von Marc Schenkelberger [96], siehe auch [97]). Die Abbildung zeigt die PM\*-Signale als Funktion der Zeit für vier verschiedene MM\*-Konzentrationen bei 44 °C. Die Signale sind auf den PM\*-Gleichgewichtswert der Einzelhybridisierung ( $[PM^*]=5$  nM,  $[MM^*]=0$  nM) normiert. Die ausgefüllten Symbole geben die theoretischen Vorhersagen für das PM\*-Signal nach Gl. 5.1 unter Verwendung der Bindungskonstanten aus Tab. 5.3 an. Auch bei verändertem Sequenz-Motiv tritt der Effekt der hochspezifischen Hybridisierung auf.

Bindungskonstante (HAT von engl. *high affinity target*) konstant gehalten, die Konzentration der Target-Spezies mit der kleineren Bindungskonstante (LAT von engl. *low affinity target*) auf 1000 nM eingestellt und die Oberflächenbedeckung des HAT im entsprechenden Konkurrenzsystem der Form „HAT vs. LAT“ mittels TIRF gemessen (siehe [96, 97]). Die HAT-Konzentration ist 10 nM für MM4 im System „MM4 vs. MM1“ und für MM1 im System „MM1 vs. MM3“. In allen anderen Konkurrenzsystemen beträgt die HAT-Konzentration 5 nM. Alle Messungen werden bei 44 °C durchgeführt.

## Ergebnisse und Diskussion

Abb. 5.6 zeigt die Oberflächenbedeckung des HAT bei Einzel- und Konkurrenzhybridisierung in Systemen der Form „HAT vs. LAT“. Die Daten werden jeweils auf den Wert der HAT-Oberflächenbedeckung aus den Einzelhybridisierungen normiert ( $[LAT]=0$  nM). Die Ergebnisse zeigen, dass man eine Unterscheidung in standard- (A)) und hochspezifische (B)) Systeme vornehmen kann. Die hochspezifischen Systeme zeichnen sich dadurch aus, dass die Oberflächenbedeckung des HAT nicht von der Anwesenheit des um ein Vielfaches höher konzentrierten LAT beeinflusst wird, obwohl sich die Bindungskonstanten von HAT und LAT nur geringfügig unterscheiden (siehe Tab. 5.1).



**Abbildung 5.6:** Unterteilung der DNA-Konkurrenzhybridisierung in standard- und hochspezifische Systeme (durchgeführt von Marc Schenkelberger [96], siehe auch [97]). Dargestellt sind die HAT-Signale der Einzelhybridisierung (schwarze Balken,  $[LAT]=0$  nM) und in Konkurrenzsystemen der Form „HAT vs. LAT“ (blaue Balken) für sechs verschiedene Systeme bei einer LAT-Konzentration von 1  $\mu$ M und einer Temperatur von 44 °C (HAT von engl. *high affinity target*, LAT von engl. *low affinity target*). Die HAT-Konzentration beträgt 10 nM in den Systemen „MM4 vs. MM1“ und „MM1 vs. MM3“ bzw. 5 nM in allen anderen. Die Signale sind jeweils auf den Wert der Einzelhybridisierung normiert. Anhand der Daten kann eine Unterscheidung in zwei Systeme getroffen werden. **A)** Standardspezifische Systeme: Gl. 5.1 gibt die HAT-Signale korrekt wieder. **B)** Hochspezifische Systeme: Gl. 5.1 versagt.

Dies entspricht nicht den Vorhersagen durch Gl. 5.1. In standardspezifischen Systemen verhält sich die Oberflächenbedeckung des HAT in Konkurrenz zum LAT entsprechend Gl. 5.1.

Im folgenden Abschnitt wird gezeigt, dass eine Unterteilung der Systeme anhand der Schmelztemperaturen von HAT und LAT vorgenommen werden kann.

### 5.3.2 Schmelztemperaturen

#### Experiment

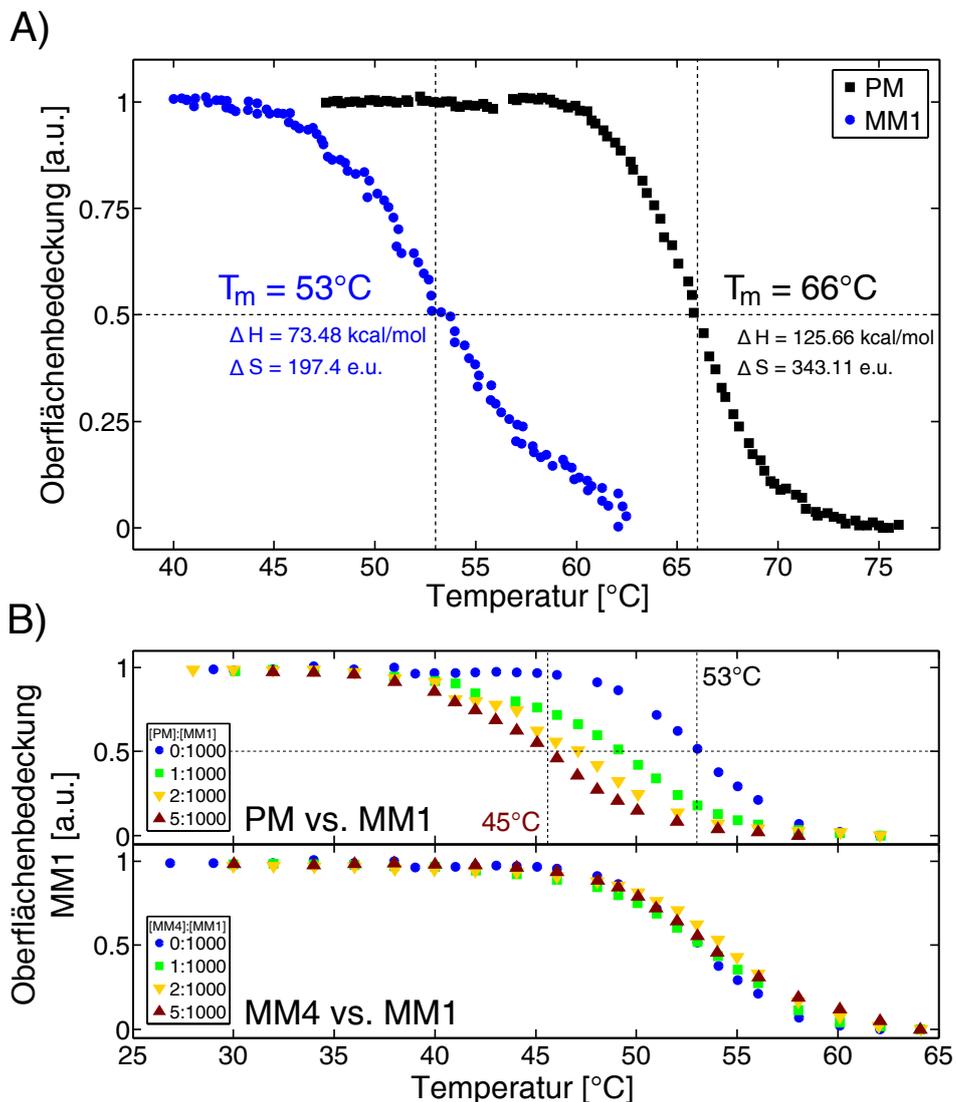
Wird die Hybridisierungstemperatur langsam erhöht, so denaturiert doppelsträngige DNA im thermodynamischen Quasi-Gleichgewicht in zwei Einzelstränge (siehe Abschnitt 2.3.2). Misst man den Anteil hybridisierter Probes als Funktion der Temperatur, so kann die Schmelztemperatur, die Enthalpieänderung  $\Delta H$  und Entropieänderung  $\Delta S$  der Target-Spezies bei Denaturierung bestimmt werden. Die Schmelztemperatur entspricht per Definition dabei gerade der Temperatur, bei der 50 % der Probes noch hybridisiert sind. Mittels TIRF werden so die Schmelztemperaturen von PM und MM1 in Einzelhybridisierungsexperimenten bestimmt und die Werte von  $\Delta H$  und  $\Delta S$  extrahiert (siehe [96, 97]). Die Konzentration der Targets in den Einzelhybridisierungen beträgt 1000 nM. Zusätzlich werden noch die Schmelzkurven von MM1 in den Systemen „PM vs. MM1“ und „MM4 vs. MM1“ bei variierenden HAT-Konzentrationen gemessen. Die MM1-Konzentration in den Konkurrenzsystemen liegt konstant bei 1000 nM.

## Ergebnisse und Diskussion

In Abb. 5.7 A) sind die in Einzelhybridisierungen experimentell bestimmten Schmelzkurven der Targets PM und MM1 mit den eingezeichneten Schmelztemperaturen dargestellt. Die Schmelztemperatur des PM-Probe-Duplexes beträgt 66 °C, die des MM1-Probe-Duplexes 55 °C. Des Weiteren sind aus den Schmelzkurven die jeweiligen Enthalpieänderungen  $\Delta H$  und Entropieänderungen  $\Delta S$  beim Übergang von der Doppelhelix zum denaturierten Zustand extrahiert (für Details siehe [96, 97]). Auffallend ist, dass die Entropieänderung beim PM wesentlich größer ist als beim MM1. Geht man davon aus, dass beide Target-Spezies als Einzelstrang die gleiche Persistenzlänge und damit die gleiche Anzahl von entropischen Freiheitsgraden haben, so ist die Anzahl der Freiheitsgrade im hybridisierten Zustand bei MM1-Probe-Duplex größer als beim PM-Probe-Duplex. Daraus folgt auch, dass die Persistenzlänge des Duplexes, bestehend aus PM und Probe, wesentlich größer ist als die des MM1-Probe-Duplexes. Ein möglicher Grund ist die Fehlstelle des MM1, die genau in der Mitte des Strangs liegt und die eine Ausbildung einer Doppelhelix entlang des gesamten Strangs erschwert. Bereits in [3] wurde beobachtet, dass eine Fehlstellenposition in der Mitte eines kurzen Strangs die Stabilität des Duplexes am stärksten herabsetzt.

Teil B) von Abb. 5.7 zeigt die Schmelzkurven von MM1 in Konkurrenz bei variierenden Konzentrationen der jeweiligen Konkurrenz-moleküle PM bzw. MM4. Die Schmelztemperatur des MM1 im hochspezifischen System „PM vs. MM1“ ist von der Konzentration des PM abhängig: Sie sinkt von 53 °C in der Einzelhybridisierung ( $[PM]=0$  nM,  $[MM1]=1000$  nM) auf ungefähr 45 °C in der Konkurrenzhybridisierung ( $[PM]=5$  nM,  $[MM1]=1000$  nM). Obwohl MM1 200-fach höher konzentriert ist als die PM-Konkurrenzspezies, zeigt sich in den Schmelzkurven eine empfindliche Abhängigkeit der MM1-Hybridisierung von der PM-Konzentration. Im Gegensatz dazu hat eine MM4-Konzentrationserhöhung im standardspezifischen System „MM4 vs. MM1“ keinen Einfluss auf die MM1-Schmelztemperatur. Die Schmelztemperatur liegt auch bei einer MM4-Konzentration von 5 nM immer noch ungefähr beim Wert aus der Einzelhybridisierung von 53 °C. Diese Ergebnisse sind ein Indiz für eine Wechselwirkung zwischen PM und MM1 im hochspezifischen Fall, die so im standardspezifischen Fall nicht auftritt.

Die Schmelztemperaturen der übrigen Target-Moleküle aus den Tabellen 5.1 und 5.3 werden mit dem DINAMelt Server [41] bestimmt. Vergleicht man die experimentell bestimmten Schmelztemperaturen von PM und MM1 aus Abb. 5.7 A) mit den berechneten Werten, so ergibt sich ein Offset der experimentellen Werte von +8 °C. Dieser Unterschied wird von der Markierung der Targets (die eine Duplexstabilisierung bewirkt) und der Präsenz der Oberfläche verursacht.



**Abbildung 5.7:** Experimentelle Bestimmung der Schmelztemperaturen (durchgeführt von Marc Schenkelberger [96], siehe auch [97]). **A)** Gezeigt sind die Schmelzkurven von PM und MM1 in Einzelhybridisierungsexperimenten bei einer Target-Konzentration von 1  $\mu\text{M}$  mit den extrahierten Werten der Enthalpieänderung  $\Delta H$  und Entropieänderung  $\Delta S$  (e.u.=*entropic units*=kcal/(mol·K)).  $\Delta S$  ist im Falle der PM-Hybridisierung wesentlich größer als bei der Hybridisierung mit MM1. Für die Schmelztemperatur gilt:  $T_{PM,m} = 66$  °C,  $T_{MM1,m} = 53$  °C. **B)** Dargestellt sind die Schmelzkurven von MM1 in Präsenz des PM (oberer Teil, hochspezifische Hybridisierung) bzw. des MM4 (unterer Teil, standardspezifische Hybridisierung) für vier verschiedene Konzentrationsverhältnisse (in nM). In Präsenz des PM sinkt die Schmelztemperatur von MM1 von 53 °C auf ca. 45 °C. Im System „MM4 vs. MM1“ bleibt die Schmelztemperatur von MM1 konstant bei 53 °C.

**Tabelle 5.4:** Schmelztemperaturen aller verwendeter Targets bei einer Target-Konzentration von 1  $\mu\text{M}$ . Die Schmelztemperaturen von PM und MM1 wurden experimentell bestimmt (siehe Abb. 5.7), die Schmelztemperaturen der restlichen Targets mit dem DINAMelt-Server [41] für eine 0.57 M monovalente Salzkonzentration (3xSSC-Puffer) berechnet.

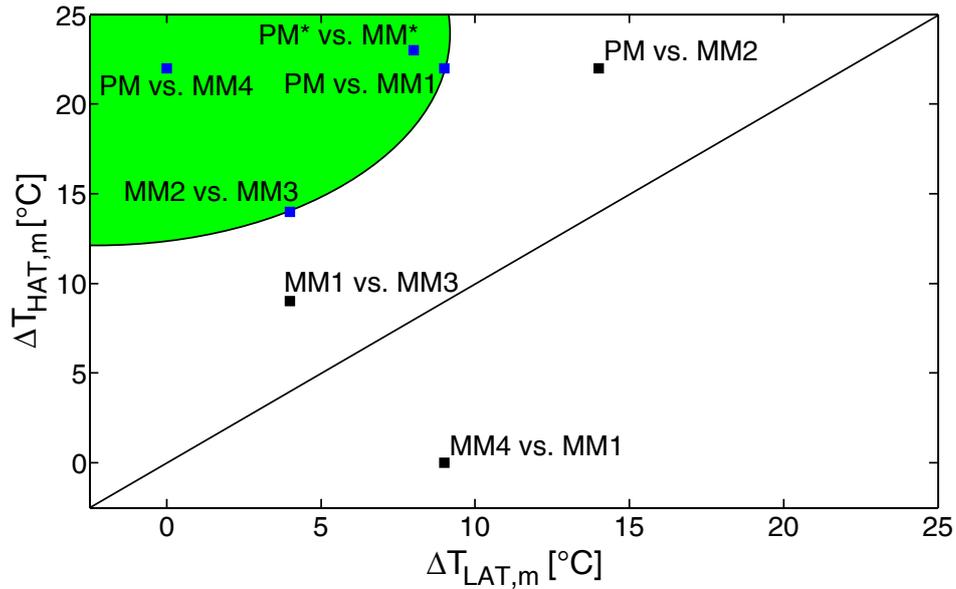
Name des Targets	Schmelztemperatur $T_m$ [°C]
PM	66
MM1	53
MM2	58
MM3	48
MM4	44
PM*	67
MM*	52

In Tab. 5.4 sind die korrigierten Werte der Schmelztemperaturen  $T_m$  aller Targets angegeben. Anhand der Schmelztemperaturen von HAT und LAT eines Konkurrenzsystems „HAT vs. LAT“ lässt sich voraussagen, ob das System standard- oder hochspezifisch ist: Sind die Schmelztemperaturen von HAT und LAT vergleichbar ( $T_{LAT,m} \approx T_{HAT,m}$ ), so ist das entsprechende Konkurrenzsystem standardspezifisch. Unterscheiden sich die Schmelztemperaturen der beiden Targets jedoch um mehr als 10 °C, liegt ein hochspezifisches System vor. Dies ist in Abb. 5.8 dargestellt. Dort sind die Differenzen zwischen den Schmelztemperaturen der HATs und der Hybridisierungstemperatur  $\Delta T_{HAT,m} = T_{HAT,m} - 44$  °C gegen die Differenzen der Schmelztemperaturen der LATs und der Hybridisierungstemperatur  $\Delta T_{LAT,m} = T_{LAT,m} - 44$  °C aufgetragen. Blaue Quadrate entsprechen hochspezifischen, schwarze standardspezifischen Systemen. Für Punkte auf der schwarzen Linie gilt  $\Delta T_{HAT,m} = \Delta T_{LAT,m}$ .

### 5.3.3 FCS-Untersuchungen

#### Experiment

Um die Fluktuationsdynamik der PM-Probe-Duplexe bzw. der MM1-Probe-Duplexe zu untersuchen, werden FCS-Messungen analog den Experimenten aus Abschnitt 4.5.4 in Lösung durchgeführt. Dazu werden die zum PM-Target perfekt komplementären Probes an Position 8 mit Cy3 und die Targets an Position 10 mit Cy5 markiert (Zählrichtung 3'-5'). Um den Beitrag der Diffusion zu eliminieren, werden wiederum entsprechende Targets ohne interne Markierung verwendet (siehe Abschnitt 3.3 und 4.5.4). So können die diffusionsbereinigten Korrelationsfunktionen der Öffnungs- und Schließdy-



**Abbildung 5.8:** Auftragung der Schmelztemperaturen der HATs (von engl. *high affinity target*) gegen die Schmelztemperaturen der LATs (von engl. *low affinity target*) für alle Systeme aus Abb. 5.6 und das System „PM\* vs. MM\*“ aus Abb. 5.5 ( $\Delta T_m = T_m - 44$  °C). Hochspezifische Systeme sind durch blaue, standardspezifische Systeme durch schwarze Quadrate gekennzeichnet. Für Punkte auf der durchgezogenen Linie gilt:  $\Delta T_{HAT,m} = \Delta T_{LAT,m}$ .

namik der beiden Target-Spezies PM und MM1 bei Hybridisierung gemessen werden. Die FCS-Messungen der Einzelhybridisierungen werden mit Lösungen aus 10 nM Probes und 20 nM Targets durchgeführt. Bei den Messungen der Fluktuationsdynamik des MM1 in Präsenz des PM bestehen die Lösungen zusätzlich noch aus 100 pM PM. Die Experimente werden bei Raumtemperatur und bei 44 °C durchgeführt. Das experimentelle Vorgehen ist hier analog zu Abschnitt 4.5.4.

## Ergebnisse und Diskussion

In Abb. 5.9 sind die diffusionsbereinigten Korrelationsfunktionen von PM und MM1 aus den Einzelhybridisierungen bei Raumtemperatur (A)) bzw. 44 °C (B)) gezeigt. In Teil B) ist zusätzlich noch die Korrelationsfunktion von MM1 in Präsenz des PM dargestellt. Der Hintergrund ist jeweils subtrahiert und die Amplituden auf 1 normiert. Die angegebenen Relaxationskonstanten  $\tau$  werden mit Hilfe der Gleichung

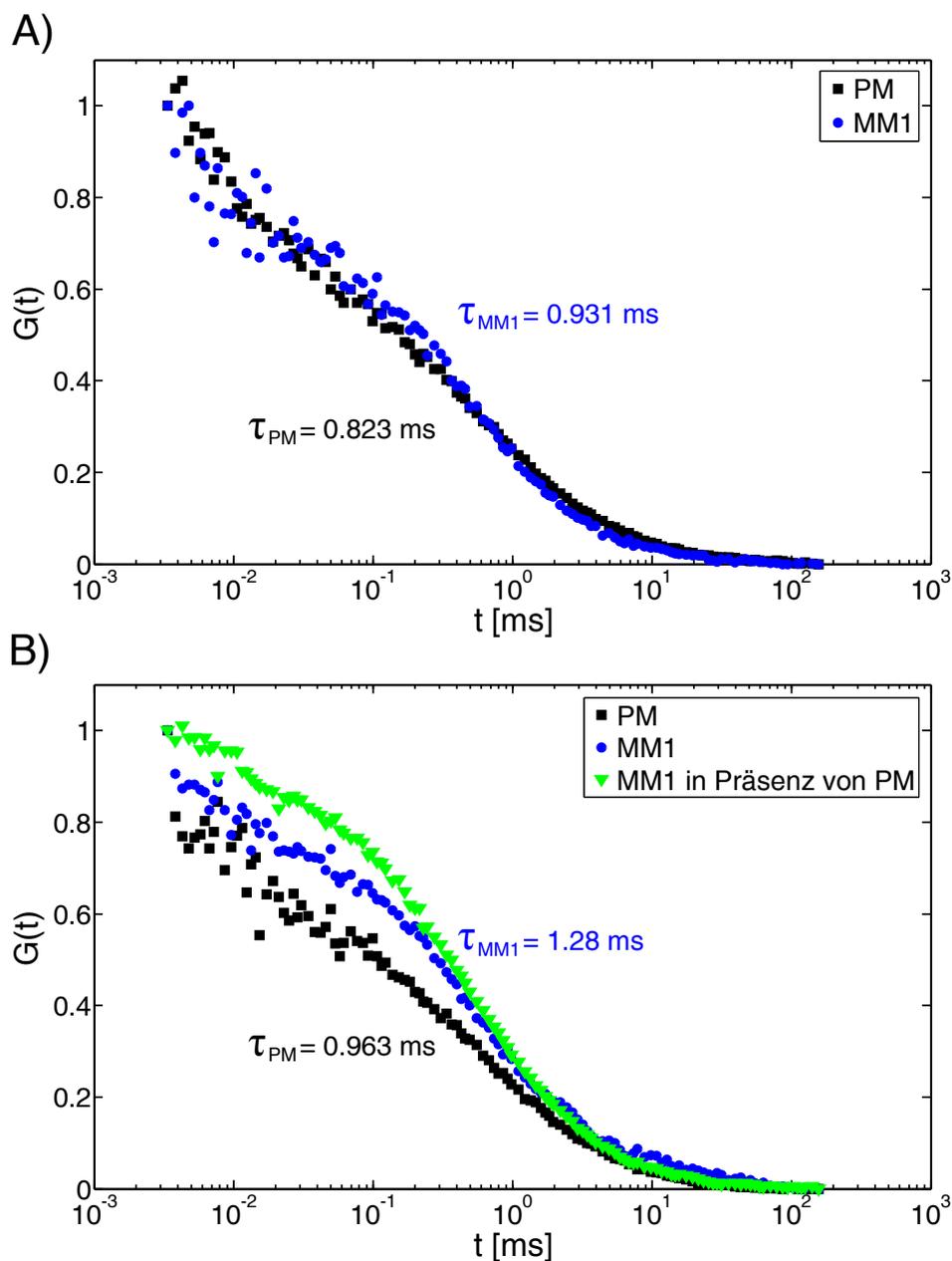
$$G(t) \propto \left(1 + \frac{t}{2\tau}\right) \cdot \operatorname{erfc}\left(\sqrt{\frac{t}{4\tau}}\right) - \sqrt{\frac{t}{\pi\tau}} \cdot e^{-\frac{t}{4\tau}} \quad (5.3)$$

mit  $\operatorname{erfc}(u) = 1 - \frac{2}{\sqrt{\pi}} \int_0^u e^{-x^2} dx$  aus [85] extrahiert (siehe auch Anhang A.3). Die Relaxationskonstanten von PM und MM1 sind bei Raumtemperatur sehr ähnlich

( $\tau_{PM} = 0.823$  ms,  $\tau_{PM} = 0.931$  ms, ähnliche Relaxationsdynamik). Bei 44 °C ist die Relaxationsdynamik des MM1 jedoch wesentlich langsamer als die des PM ( $\tau_{PM} = 0.963$  ms,  $\tau_{PM} = 1.28$  ms). In Präsenz des PM wird die Korrelationsfunktion des MM1 weiter modifiziert. Für diese Korrelationsfunktion wird keine Relaxationskonstante extrahiert, da das Modell aus [85] explizit nur Einzelhybridisierungen behandeln kann. In Übereinstimmung mit den Schmelzkurven aus Abschnitt 5.3.2 (größeres  $\Delta S$  bei PM-Hybridisierung) unterstützen die unterschiedlichen Korrelationsfunktionen von PM und MM1 bei 44 °C die Idee, dass dem MM1 wesentlich mehr entropische Freiheitsgrade im hybridisierten Zustand zur Verfügung stehen als dem PM, wodurch die Relaxationsdynamik verlangsamt wird. Aus der Literatur ist bekannt, dass DNA bei Temperaturen nahe dem Schmelzpunkt in sog. *pre-melted*-Konfigurationen vorliegt [99, 100]. Diese Konfigurationen zeichnen sich im Gegensatz zu den hochkooperativen Bindungskonfigurationen der klassischen Doppel-Helix durch eine niedrigere Kooperativität zwischen benachbarten Basen, einer kleineren Persistenzlänge und einer größeren Anzahl von entropischen Freiheitsgraden aus. Aus den Korrelationsfunktionen bei 44 °C lässt sich schließen, dass die MM1-Spezies in einer teilweise denaturierten, offenen Konfiguration hybridisiert, der PM dagegen in einer eher geschlossenen Struktur. Es wird angemerkt, dass sich die hier vorliegenden experimentellen Bedingungen von denen aus [85] unterscheiden, für die Gl. 5.3 entwickelt wurde. Die hier verwendeten Duplexe sind keine *hairpin loops*, die GC-Klappen auf beiden Seiten des Duplexes fehlen und es wird eine höhere Salzkonzentration von 0.57 M (3xSSC-Puffer) monovalenten Ionen benutzt. Trotzdem wird Gl. 5.3 benutzt um die Relaxationskonstanten zu extrahieren. Ein an die geänderten experimentellen Bedingungen angepasstes Modell für die Fluktuation der Basenpaare ist nicht zentrales Thema dieser Arbeit. Darüber hinaus sind die Unterschiede in den Korrelationsfunktionen so deutlich sichtbar, dass eine Überarbeitung des Modells für die hier getroffenen qualitativen Aussagen nicht nötig ist. Die hier gemessenen Relaxationskonstanten sind größer als die in [85], was sich auf die geänderten experimentellen Bedingungen zurückführen lässt.

#### 5.3.4 Positionsabhängigkeit

In [2, 3] wurde ein Einfluss der Fehlstellenposition auf das Hybridisierungssignal beobachtet: Fehlstellen in der Mitte des Duplexes setzten die Stabilität des Duplexes wesentlich stärker herab als endnahe Fehlstellen. Daher soll im Folgenden untersucht werden, ob die Fehlstellenposition auch für die Unterteilung in standard- und hochspezifische Konkurrenzsysteme eine Rolle spielt.



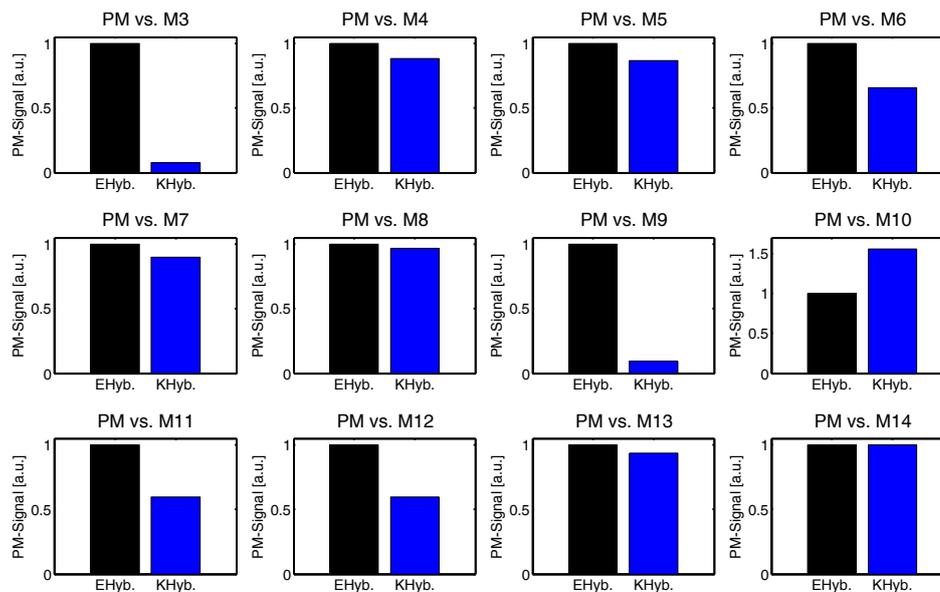
**Abbildung 5.9:** FCS-Messungen der Öffnungs- und Schließdynamik von PM und MM1 in Lösung. Die Korrelationsfunktionen  $G(t)$  werden bei Raumtemperatur (A)) und 44 °C (B)) gemessen. Die Probe-Konzentration beträgt 10 nM, die Target-Konzentration 20 nM. Die angegebenen mittleren Relaxationskonstanten der Fluktuationsmoden werden mit Gl. 5.3 extrahiert, die aus [85] entnommen wurde. Bei Raumtemperatur unterscheiden sich die Relaxationskonstanten von PM und MM1 nur unwesentlich voneinander. Bei 44 °C ergibt sich ein deutlicherer Unterschied. In B) ist gezeigt, dass die Korrelationsfunktion der MM1-Spezies durch die Präsenz von 100 pM PM modifiziert wird (grüne Kurve). Jedoch kann für diesen Fall keine Relaxationskonstante bestimmt werden, da das Modell aus [85] keine Konkurrenzsituationen behandelt.

**Tabelle 5.5:** Sequenzen der Probe und aller verwendeten Targets zur Untersuchung der Positionsabhängigkeit der hochspezifischen Hybridisierung. Target PM ist perfekt komplementär zur Probe, alle anderen Targets besitzen eine Fehlstelle. Die Fehlstelle ist rot markiert und wird von Position 3 bis Position 14 (Zählrichtung 3'-5') variiert. Die Base der Fehlstelle orientiert sich an der Probe-Sequenz: Hat die Probe an der entsprechenden Position die Base X, so wird diese Base auch in der Target-Sequenz übernommen.

Name der Sequenz	Sequenz (3'-5')
Probe	TTC-CTA-GTC-TAG-CAT-T
PM	AAT-GCT-AGA-CTA-GGA-A
M3( $\cong$ MM2)	AA <b>C</b> -GCT-AGA-CTA-GGA-A
M4	AAT- <b>C</b> CT-AGA-CTA-GGA-A
M5	AAT-G <b>G</b> T-AGA-CTA-GGA-A
M6	AAT-GC <b>A</b> -AGA-CTA-GGA-A
M7	AAT-GCT- <b>T</b> GA-CTA-GGA-A
M8( $\cong$ MM1)	AAT-GCT-A <b>C</b> A-CTA-GGA-A
M9	AAT-GCT-AG <b>T</b> -CTA-GGA-A
M10	AAT-GCT-AGA- <b>G</b> TA-GGA-A
M11	AAT-GCT-AGA-C <b>A</b> A-GGA-A
M12	AAT-GCT-AGA-CT <b>T</b> -GGA-A
M13	AAT-GCT-AGA-CTA- <b>C</b> GA-A
M14	AAT-GCT-AGA-CTA-G <b>C</b> A-A

## Experiment

Für die Untersuchung der Positionsabhängigkeit der hochspezifischen Hybridisierung werden zusätzlich zum perfekt komplementären Target PM zwölf weitere Target-Sequenzen verwendet, die im Vergleich zum PM-Target eine Fehlstelle haben. Die Targets unterscheiden sich untereinander durch die Position ihrer Fehlstelle, welche zwischen Position 3 und 14 (Zählrichtung 3'-5') variiert. Die Base der Fehlstelle orientiert sich an der zum PM-Target komplementären Probe-Sequenz: Befindet sich an der entsprechenden Position der Probe die Base X, so wird diese Base auch im Target übernommen (bis auf M3, das die Base Cytosin an der Fehlstellenposition hat). Die verwendeten Target-Sequenzen sind in Tab. 5.5 angegeben. Die Fehlstelle ist rot markiert. M3 bzw. M8 entspricht Target MM2 bzw. MM1 aus Abschnitt 5.1. Es werden Konkurrenzhybridisierungen zwischen der PM-Spezies und jeweils einer Target-Spezies M<sub>x</sub> aus Tab. 5.5 bei 44 °C durchgeführt ([PM]:[M<sub>x</sub>]=1:1000 nM, x steht für die Fehlstellenposition). Die Probes auf der Oberfläche sind dabei komplementär zur PM-Spezies. Vor jeder Konkurrenzhybridisierung wird der Chip zuerst mit 1 nM PM hybridisiert.



**Abbildung 5.10:** Untersuchung der Positionsabhängigkeit der hochspezifischen Hybridisierung. Dargestellt sind die PM-Signale der Einzel- (schwarz,  $[PM]:[Mx]=1:0$  nM) und Konkurrenzhybridisierungen (blau,  $[PM]:[Mx]=1:1000$  nM) für alle Systeme der Form „PM vs. Mx“. Mx steht dabei für ein Target mit Fehlstelle aus Tab. 5.5, wobei x die Fehlstellenposition angibt (Zählrichtung 3'-5'). Die Signale sind jeweils auf den Wert der Einzelhybridisierung normiert. Es lässt sich kein systematischer Zusammenhang zwischen der Fehlstellenposition und dem Auftreten der hochspezifischen Hybridisierung feststellen.

Nach Bilderaufnahme, Ausspülen der alten Target-Lösung und Reinigung des Chips wird die Konkurrenzhybridisierung durchgeführt.

## Ergebnisse und Diskussion

Abb. 5.10 zeigt die jeweils auf die entsprechenden Signale der Einzelhybridisierung normierten PM-Signale. Die jeweiligen Konkurrenzsysteme sind über den Teilabbildungen in der Form „PM vs. Mx“ angegeben. Die Daten lassen keine systematische Abhängigkeit von der Fehlstellenposition erkennen, anhand derer man die Systeme in standard- bzw. hochspezifisch unterteilen könnte. Die in [2, 3] gefundene Abhängigkeit des Hybridisierungssignals von der Fehlstellenposition scheint hier - wenn überhaupt - eine untergeordnete Rolle zu spielen. Die in den vorherigen Abschnitten durchgeführten Experimente reichen also aus um die Bedingungen für das Auftreten der hochspezifischen Hybridisierung vollständig zu charakterisieren.

## 5.4 Erklärungsansatz

Ausgehend von den experimentellen Ergebnissen aus den Abschnitten 5.1 bis 5.3 soll in diesem Abschnitt ein Erklärungsansatz für den Effekt der hochspezifischen Hybridisierung in Konkurrenz gegeben werden.

Die Experimente zur Schmelztemperatur in Abschnitt 5.3.2 (Abhängigkeit der Schmelztemperatur von der Präsenz des Konkurrenten) und die FCS-Untersuchungen in Abschnitt 5.3.3 (Modifikation der Korrelationsfunktion durch Präsenz des Konkurrenten) lassen den Schluss zu, dass die Konkurrenten in einem hochspezifischen System durch thermische Fluktuationen miteinander wechselwirken (siehe Abb. 5.7 und 5.9). Hochspezifische Konkurrenzsysteme treten auf, wenn sich die Schmelztemperaturen der beiden Konkurrenzspezies um mehr als 10 °C unterscheiden. Zusammen mit den Korrelationsfunktionen ergeben sich daraus unterschiedliche Fluktuationsdynamiken der beiden Konkurrenten (z.B. im Konkurrenzsystem „PM vs. MM1“) und damit unterschiedliche Bindungskonfigurationen: Für Temperaturen, die vergleichbar mit der Schmelztemperatur des DNA-Duplexes sind, liegt der Duplex vorwiegend in sogenannten *pre-melted*-Konfigurationen vor, die sich im Gegensatz zur hochkooperativen klassischen Doppelhelix durch eine niedrigere Kooperativität zwischen benachbarten Basen, einer erhöhten Anzahl von entropischen Freiheitsgraden und einer kleineren Persistenzlänge auszeichnen [99, 100]. Der Übergang zwischen den beiden Konfigurationen - Doppelhelix und *pre-melted*-Konfiguration - ist nicht scharf, sodass beide Konfigurationen auf einem Strang nebeneinander koexistieren können. Die beiden Konfigurationen sind in Abb. 6.10 veranschaulicht. Die experimentellen Daten aus den Abb. 5.7 und 5.9 bestätigen, dass dem LAT mehr entropische Freiheitsgrade zur Verfügung stehen als dem HAT. Es wird daher angenommen, dass der LAT vorwiegend in einer flexiblen *pre-melted*-Konfiguration vorliegt, der HAT hingegen aber eine steife Doppelhelix mit der Probe ausbildet.

Durch diese unterschiedlichen Eigenschaften werden die effektiven freien Bindungsenergien der Spezies durch die gegenseitige Wechselwirkung in einer Konkurrenzsituation zugunsten des HAT modifiziert. Dies widerspricht nicht der Thermodynamik, da die Bindung zwischen Probe und Target in Konkurrenz eine andere Situation darstellt, die nicht notwendigerweise durch die mittlere freie Energie der Bindung der Einzelhybridisierung wiedergegeben werden kann. In einer Konkurrenzsituation können die Bindungszustände der konkurrierenden Targets auf antagonistische Weise miteinander kombinieren, sodass die effektive freie Energielandschaft asymmetrisch zugunsten eines Konkurrenten verändert werden kann. Wie in Abschnitt 5.1 dargelegt wurde, ist der Effekt der hochspezifischen Hybridisierung trotz der relativ ähnlichen Bindungskonstanten der Konkurrenten in der Einzelhybridisierung sehr stark ausgeprägt, da selbst

eine 1000-fache MM1-Konzentrationserhöhung keine Auswirkung auf die Oberflächenbedeckung des PM hat. Im Gegensatz dazu verdrängt der MM2 nach Gl. 5.1 mit zunehmender Konzentration den PM von der Oberfläche. Da es sich bei MM1 und MM2 bis auf die kleine Variation der Fehlstellenposition praktisch um identische Moleküle handelt, wird ein hohes Maß an Kooperativität benötigt. Nur so kann bei einer kleinen Störung des Systems eine hohe Nicht-Linearität verursacht werden, die eine Unterscheidung in standard- und hochspezifische System zulässt. Die Hybridisierung von DNA ist bekannt als stark nicht-lineares und kooperatives Phänomen, das noch immer wenig verstanden ist [101]. Es wird angemerkt, dass es auch in standardspezifischen Systemen zu Wechselwirkungen zwischen den beteiligten Konkurrenten kommen kann. Jedoch ändert sich die effektive freie Energielandschaft aufgrund ähnlicher Eigenschaften der Konkurrenten, wie z.B. ähnlicher Fluktuationsdynamik, nicht asymmetrisch, sodass die Bindungsenergien aus den Einzelhybridisierungen auch den Konkurrenzfall vollständig beschreiben.

Es wird angenommen, dass die Interaktion der beiden Konkurrenten über einen sogenannten Triplex-Zustand bestehend aus Probe, HAT und LAT erfolgt (siehe 6.10 C)). Wie bei allen auf engem Raum befindlichen Polymeren kommt es zu einer entropischen Repulsion der Stränge. Bei komplexeren Molekülen wie DNA oder Proteinen können außerdem die Energieniveaus der molekularen Mikrozustände verschoben werden: Bildet in einer Triplex-Konfiguration eine der beiden Target-Spezies eine (steife) Doppel-Helix mit der Probe, wird die Probe gestreckt, wodurch die Entropiebarriere für die Ausbildung einer Helix-Konformation für den Konkurrenten vergrößert wird. Um die erhöhte Entropie-Barriere zu kompensieren, muss eine größere Anzahl von benachbarten Basen eine Doppel-Helix ausbilden. Die erforderliche Kooperativität der Bindung wird somit erhöht. Aufgrund der Fehlstelle sind diese hochkooperativen Bindungszustände für den LAT nicht zugänglich. Dadurch verliert der LAT im Triplex einen Großteil seiner Bindungsaffinität, während die Bindungsaffinität des HAT, der die erhöhte Kooperativität aufbringen kann, praktisch unbeeinflusst bleibt.

Dieser Mechanismus bewirkt letztendlich die hochspezifische Hybridisierung des HAT in Konkurrenz. In Abschnitt 6.2 wird ein Gleichgewichtsmodell vorgestellt, das im Rahmen dieser Arbeit zur Beschreibung der hochspezifischen DNA-Hybridisierung in Konkurrenz entwickelt wurde und das die hier vorgestellten experimentellen Ergebnisse quantitativ reproduziert.

## 5.5 Ordnungsverhalten

Die in den vorherigen Abschnitten vorgestellten Experimente haben gezeigt, dass sich biologische Systeme unter geeigneten Bedingungen aufgrund von molekularer Erkennung selbstständig ordnen können: Bindungen treten nur zwischen perfekt passenden Bindungspartnern auf. Der entscheidende Parameter ist dabei die Temperatur des Systems. So wurde in [96, 97] beobachtet, dass der Effekt der hochspezifischen Hybridisierung im System „PM vs. MM1“ am deutlichsten auftritt, wenn die Hybridisierungstemperatur des Systems gerade der Schmelztemperatur des LAT entspricht und bei sinkender Temperatur abnimmt. Die Ordnung des Systems wird also bei Temperaturerhöhung vergrößert. In den folgenden Experimenten soll untersucht werden, wie sich komplexe mehrkomponentige Systeme bei Temperaturvariation verhalten. Die Grundidee dabei ist, dass die Bindung zwischen Probes und Targets bei niedriger Systemtemperatur aufgrund geringer Fluktuationen relativ unspezifisch ist. Mit steigender Temperatur wird das System geordnet, sodass Targets fast ausschließlich an den zu ihnen perfekt komplementären Probes binden, da die zu einem Target nicht-komplementären Probes von den entsprechenden komplementären Targets mit höherer Wahrscheinlichkeit besetzt werden. Durch diese indirekte Wechselwirkung der Targets untereinander soll das System in eine geordnete Form gezwungen werden. Über die Auswirkungen einer möglicherweise vorliegenden Wechselwirkung zwischen den Targets auf die Hybridisierungseigenschaften des Systems ist wenig bekannt.

### 5.5.1 Vorversuch: 4-Komponenten-System

#### Experiment

In diesem Experiment soll untersucht werden, ob sich ein Microarray-System bei steigender Temperatur generell ordnet und damit die oben beschriebene Grundidee zutrifft. Dazu werden vier Target-Spezies PM, M8, M8-2 und M8-3 verwendet, von denen nur PM mit Cy3 markiert ist. Die Sequenz der Targets unterscheidet sich nur an Position 8 (Zählrichtung 3'-5'). Im Unterschied zu den vorherigen Experimenten, in denen ein Konkurrenzsystem maximal aus zwei Target-Spezies bestand, bestehen die Target-Lösungen hier aus allen vier Spezies. Außerdem befindet sich auf der Chip-Oberfläche nicht nur die Probe-Spezies, die zur PM-Spezies komplementär ist, sondern komplementäre Probes zu allen Target-Spezies. So kann die unspezifische Hybridisierung zwischen dem PM und den nicht-komplementären Probes in Konkurrenz untersucht werden. Es werden drei Target-Lösungen mit verschiedenen Konzentrationen verwendet, mit denen

der Chip nacheinander hybridisiert wird.

- Lösung A: [PM]:[M8]:[M8-2]:[M8-3]=1:0:0:0 nM (Einzelhybridisierung)
- Lösung B: [PM]:[M8]:[M8-2]:[M8-3]=1:1:1:1 nM
- Lösung C: [PM]:[M8]:[M8-2]:[M8-3]=1:100:100:100 nM

Die Temperatur wird jeweils in 5 °C Schritten von 30 °C auf 50 °C erhöht. Bei jedem Schritt wird für 60 min hybridisiert.

## Ergebnisse und Diskussion

Die linke Spalte von Abb. 5.11 zeigt die jeweils auf das stärkste Signal normierten Signale der verschiedenen Features PM, M8, M8-2 und M8-3 für die Target-Lösungen A, B und C (die Features sind zu den Targets mit den entsprechenden Namen perfekt komplementär). Im Einzelhybridisierungsexperiment A) ist gut zu erkennen, dass der PM bei niedrigen Temperaturen auch an den Features MM1, MM2 bzw. MM3 binden kann. Mit steigender Temperatur verliert der PM diese Fähigkeit und bindet fast ausschließlich an seinem komplementären Feature. Die komplementären Features der restlichen Targets liefern keine oder nur noch sehr schwache Signale. Das Signal von Feature M8-2 kann bei Temperaturerhöhung am längsten gemessen werden. Erst ab ca. 40 °C ist auch hier kein Signal mehr messbar. Dies kann mit Hilfe der NN-Parameter [53] erklärt werden. Der Betrag der Differenz  $|\Delta\Delta G_{PM\leftrightarrow M8-2}|$  zwischen der freien Energie des Duplexes bestehend aus PM-Target und Probe MM8-2 und des Duplexes bestehend aus M8-2-Target und Probe M8-2 ist am geringsten: PM und M8-2 sind sich relativ ähnlich, wodurch der PM an der Probe M8-2 wesentlich besser hybridisieren als an den beiden verbleibenden nicht-komplementären Probes. Tab. 5.6 zeigt die NN-Energie der Duplexe an Basenposition 8 (Berücksichtigung von zwei NN-Paaren; nur an Basenposition 8 unterscheiden sich die NN-Parameter der Duplexe aufgrund der Fehlstelle untereinander) und die entsprechenden Differenzen der freien Energie der Duplexe mit Fehlstelle und den Duplexen ohne Fehlstelle bei 44 °C. Negative Werte von  $\Delta G_{Target\leftrightarrow Probe}$  destabilisieren den Duplex.

Wird die Konzentration der Target-Spezies M8, M8-2 und M8-3 auf 1 nM erhöht (Abb. 5.11 B)), so kann der PM bereits bei niedrigen Temperaturen die zu den Konkurrenten komplementären Features nur noch sehr schwer besetzen. Bei einer Konzentration der Konkurrenz-Targets von 100 nM (Abb. 5.11 C)) werden die Features M8 und M8-3 auch bei 30 °C nicht mehr besetzt. Das PM-Feature kann jedoch selbst bei diesen Konzentrationen weiterhin vom PM-Target besetzt werden: Der Effekt der hochspezifischen Hybridisierung tritt also auch in diesem System auf. Um einen gut sichtbaren

**Tabelle 5.6:** Freie Energie der verschiedenen Target-Probe-Paare. Die linke Spalte gibt an, welche Target- und Probe-Spezies den Duplex bilden. In der mittleren Spalte ist die NN-Energie der achten Duplexposition (Berücksichtigung von zwei NN-Paaren) für jeden der Duplexe bei 44 °C angegeben. Die Energieskala ist hier so gewählt, dass negative Werte von  $\Delta G_{Target \leftrightarrow Probe}$  den Duplex destabilisieren. In der rechten Spalte sind die Beträge der Differenzen der NN-Energien angegeben.

Target $\leftrightarrow$ Probe	$\Delta G_{Target \leftrightarrow Probe}$ [kcal/mol]	$ \Delta \Delta G_{Target \leftrightarrow Probe} $ [kcal/mol]
PM $\leftrightarrow$ M8	-0.37	2.97
M8 $\leftrightarrow$ M8	2.6	
PM $\leftrightarrow$ M8-2	-0.51	1.69
M8-2 $\leftrightarrow$ M8-2	1.18	
PM $\leftrightarrow$ M8-3	-1.09	2.79
M8-3 $\leftrightarrow$ M8-3	1.7	

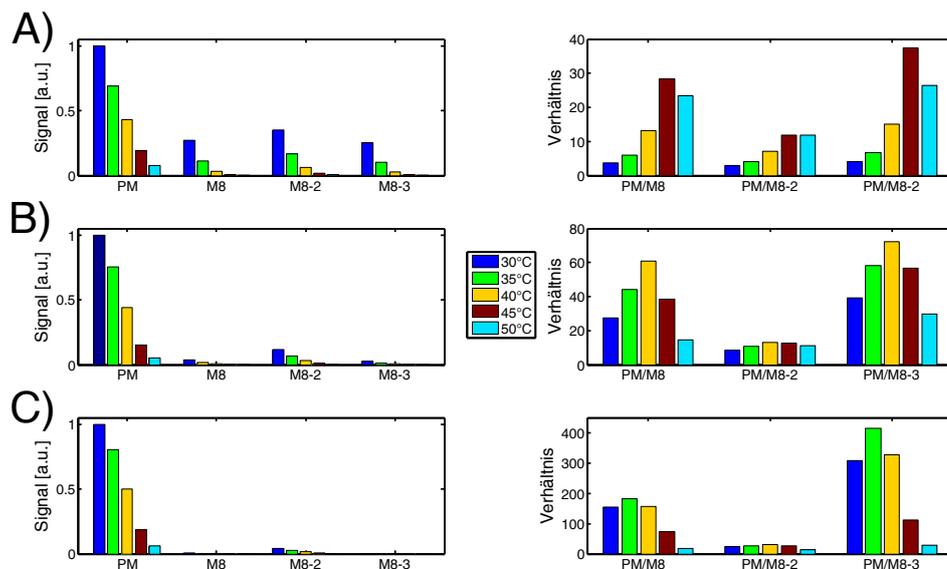
Übergang des Systems vom ungeordneten in den geordneten Zustand zu erzielen, darf die Konzentration der verwendeten Targets nicht zu unterschiedlich sein, da sich das System dann bereits bei sehr niedrigen Temperaturen im geordneten Zustand befindet. Die rechte Spalte von Abb. 5.11 zeigt die Signalverhältnisse der verschiedenen Features. In der Einzelhybridisierung des PM sind die Verhältnisse bei 45 °C maximal. Die Signalunterschiede zwischen dem PM-Feature und den restlichen Features ist hier am größten und das System maximal geordnet. Bei Erhöhung der Konzentration der Konkurrenz-Targets werden die maximalen Signalverhältnisse bei 40 °C (B)) bzw. 35 °C (C)) erreicht. Durch Zugabe von Konkurrenz-Targets erfolgt der Übergang in den geordneten Zustand bei niedrigeren Temperaturen.

Das Experiment zeigt, dass die Grundidee von der Ordnung des Systems als Funktion der Temperatur prinzipiell stimmt. Die Features werden bei steigender Temperatur nur von derjenigen Target-Spezies besetzt, die perfekt komplementär zu ihnen ist.

## 5.5.2 49-Komponenten-System

### Experiment

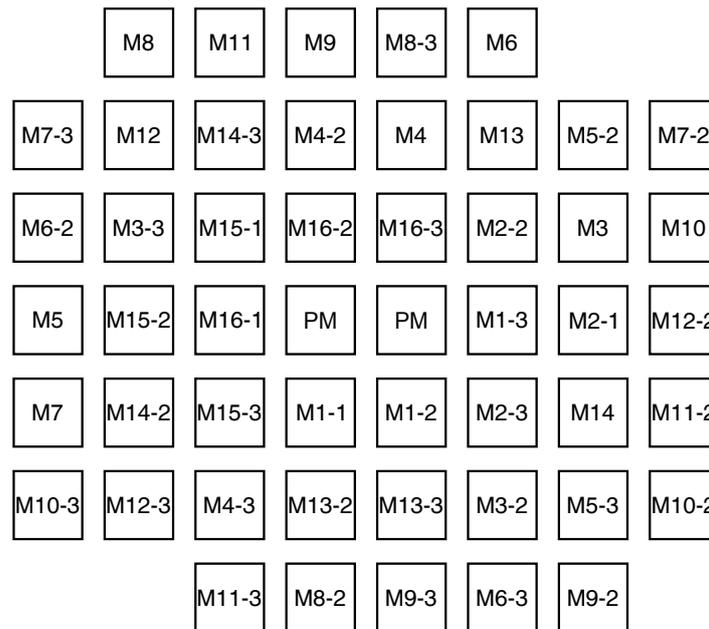
Analog zu Abschnitt 5.5.1 wird in diesem Experiment das Ordnungsverhalten eines Systems untersucht, das aus der PM-Spezies (mit Cy3-Markierung) und aller sich von dieser Sequenz um genau eine Base unterscheidende Spezies besteht. Da die PM-Sequenz eine Länge von 16 Basen hat, ergeben sich so 48 zusätzliche Target-Sequenzen, insgesamt also ein 49-komponentiges System. Durch die zusätzliche Anzahl von Konkurrenten soll die im vorherigen Abschnitt beschriebene indirekte Wechselwirkung zwischen



**Abbildung 5.11:** Hybridisierungssignale und Verhältnisse der Features für drei Target-Lösungen verschiedener Konzentration und fünf Temperaturen. Alle Signale sind auf das jeweilige PM-Signal bei 30 °C normiert. **A)** Der linke Teil der Abbildung zeigt die Signale der Features PM, M8, M8-2 und M8-3 bei Einzelhybridisierung des PM ([PM]:[M8]:[M8-2]:[M8-3]=1:0:0:0 nM). Bei niedrigen Temperaturen kann der PM an allen Features binden. Wird die Temperatur erhöht, so beschränkt sich die Bindung des PM hauptsächlich auf das PM-Feature. Der rechte Teil zeigt die Signalverhältnisse der verschiedenen Features untereinander. Das größte Signalverhältnis ergibt sich bei einer Temperatur von 45 °C. Hier weist das System die größte Ordnung auf. **B)** Hybridisierungssignale bei einer Target-Lösung mit Konzentrationsverhältnissen von [PM]:[M8]:[M8-2]:[M8-3]=1:1:1:1 nM. Hier können die nicht-komplementären Features auch bei niedrigen Temperaturen nur schwer vom PM besetzt werden (linke Seite). Das maximale Signalverhältnis ergibt sich bei einer Temperatur von 40 °C (rechte Seite). **C)** Konzentrationsverhältnis [PM]:[M8]:[M8-2]:[M8-3]=1:100:100:100 nM. Hier sind praktisch keine Signale der nicht-komplementären Features mehr messbar. Jedoch kann der PM trotz der 100-fach erhöhten Konzentration der Konkurrenten das PM-Feature besetzen. Der Effekt der hochspezifischen Hybridisierung tritt also auch hier auf (linke Seite). Die maximale Ordnung des Systems wird bei 35 °C erreicht (rechte Seite).

den Targets verstärkt werden.

Die Targets werden mit Mx-y bezeichnet.  $x$  ( $1 \leq x \leq 16$ ) steht dabei für die Basenposition, an der sich die Sequenz von der PM-Sequenz unterscheidet (Zählrichtung 3'-5').  $y$  ( $1 \leq y \leq 3$ ) bezeichnet eine der 3 Möglichkeiten eine sich an Position  $x$  von der PM-Sequenz unterscheidende Base auszuwählen. Targets mit Namen Mx wurden bereits in Abschnitt 5.3.4 verwendet. Es wird versucht die Features auf dem Chip so anzuordnen, dass sich die Features, die bei Hybridisierung ein vermutlich ähnlich starkes Signal wie das PM-Feature aufweisen, in kleinem Abstand zu diesem befinden. Die Idee dabei ist, dass sich die Hybridisierungseigenschaften der Targets dieser eng beieinanderliegenden Features ähnlich sind und somit eine stärkere Nächste-Nachbar-Wechselwirkung zwi-



**Abbildung 5.12:** Anordnung der Features auf der Microarray-Oberfläche. Jedes Quadrat steht für ein Feature, wobei die Features zum Target mit dem selben Namen komplementär sind.

schen den Konkurrenten vorliegt. Dazu kann die in [3] beobachtete Abhängigkeit des Hybridisierungssignals von der Fehlstellenposition ausgenutzt werden: Mittige Fehlstellenpositionen haben einen größeren Einfluss auf das Hybridisierungssignal als endnahe Positionen. Daraus folgt, dass Targets, die sich an einer mittigen Basenposition von der PM-Sequenz unterscheiden, in einem größerem Abstand von den PM-Features angeordnet werden als Targets mit endnahen Fehlstellenpositionen. Eine mögliche Realisierung der Anordnung der Features ist in Abb. 5.12 gezeigt (die angegebenen Features sind jeweils zu der Target-Spezies mit dem entsprechenden Namen komplementär).

Der Chip wird entweder mit einer 1 nM PM-Lösung (Einzelhybridisierung) oder der 49-komponentigen Lösung hybridisiert, bei der alle Targets in 1 nM Konzentration vorliegen. Nach Einfüllen der entsprechenden Target-Lösung wird der Chip für 2 min auf 80 °C erhitzt und danach auf 23 °C abgekühlt. Nach Erreichen der Temperatur von 23 °C wird die Temperatur in 7 °C Schritten auf 44 °C erhöht und danach wieder in 7 °C Schritten auf 23 °C gesenkt. Bei jeder Temperaturstufe wird 60 min hybridisiert. So soll untersucht werden, ob das System bei Temperaturerhöhung bzw. -Senkung immer wieder in den gleichen Gleichgewichtszustand oder ob die Wechselwirkung der Targets Einfluss auf die Gleichgewichtszustände des Systems hat.

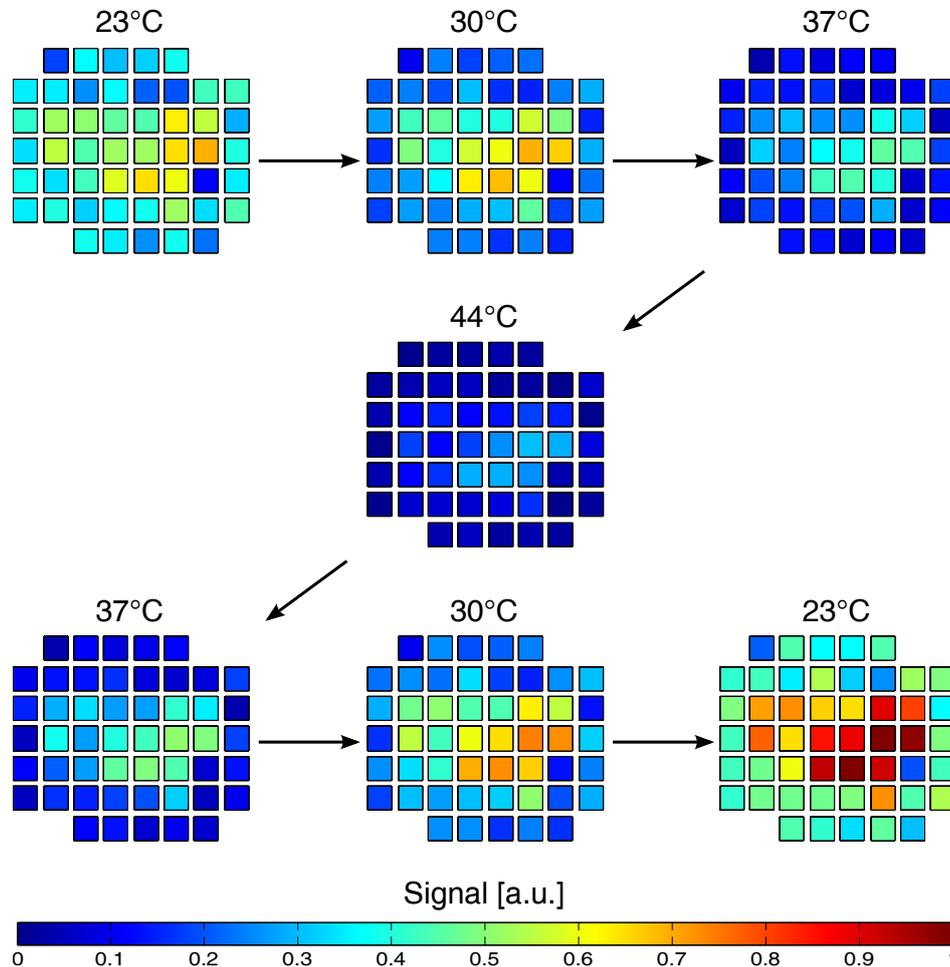
## Ergebnisse und Diskussion

Abb. 5.13 zeigt die Signale der Features bei Hybridisierung mit der 1 nM PM-Lösung

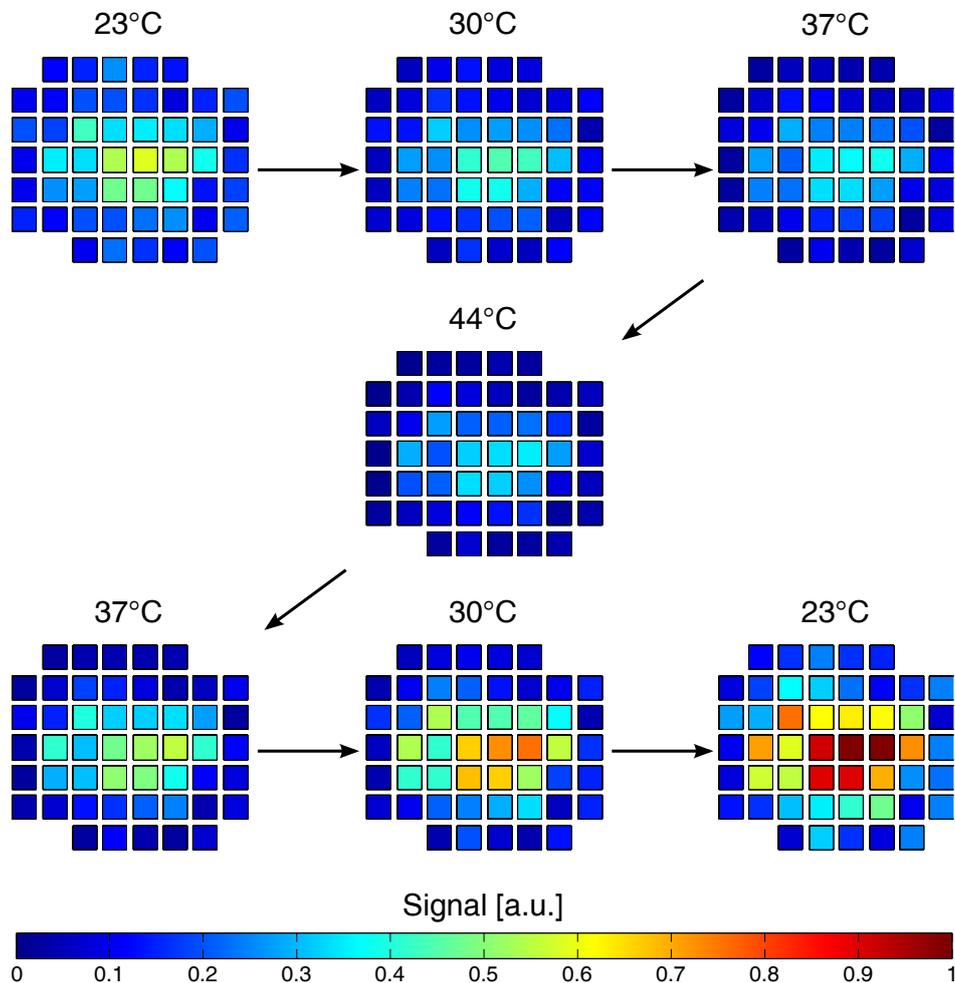
für jeden Schritt der Temperaturrampe nach Normierung der Signale auf das Feature mit dem stärksten Signal. Jedes Quadrat steht für ein Feature mit bestimmter Sequenz. Die Sequenzen der einzelnen Features sind in Abb. 5.12 angegeben. Bei Temperaturerhöhung werden die Signale der Features insgesamt schwächer, erreichen bei 44 °C ihr Minimum und werden bei Temperatursenkung wieder stärker sichtbar. Auch hier setzt sich die Signalvariation aus der Temperaturabhängigkeit des Fluoreszenzfarbstoffs Cy3 und der Anzahl hybridisierter Targets zusammen. Die zum PM-Target perfekt komplementären Features befinden sich genau in der Mitte des Featureblocks. Wie im Design des Experiments geplant, liefern Features, die sich in naher Umgebung der PM-Features befinden (also in der Mitte des Featureblocks) ein ähnlich starkes Signalniveau. Dies bedeutet, dass das PM-Target trotz nicht perfekt komplementärer Sequenz an diesen Features fast genauso gut binden kann wie am PM-Feature. Wird die Temperatur erhöht, so nehmen die Signale der Features in Randgebieten des Featureblocks stärker ab als die Signale der mittigen Features: Die PM-Targets können jetzt nur noch mit mittigen Features Bindungen eingehen. Das System ordnet sich also bei höherer Temperatur.

In Abb. 5.14 sind die entsprechenden Signale bei Hybridisierung mit der 49 komponentigen Target-Lösung gezeigt. Hier kann der PM, wie bereits in Abschnitt 5.5.1 beobachtet, auch bei der niedrigsten Temperatur von 23 °C nur an mittigen Features binden, da vor allem randnahe Features mit höherer Wahrscheinlichkeit von den entsprechenden komplementären Targets besetzt werden. Somit befindet sich das System bereits bei niedrigerer Temperatur im Vergleich zur Einzelhybridisierung in einem geordneteren Zustand.

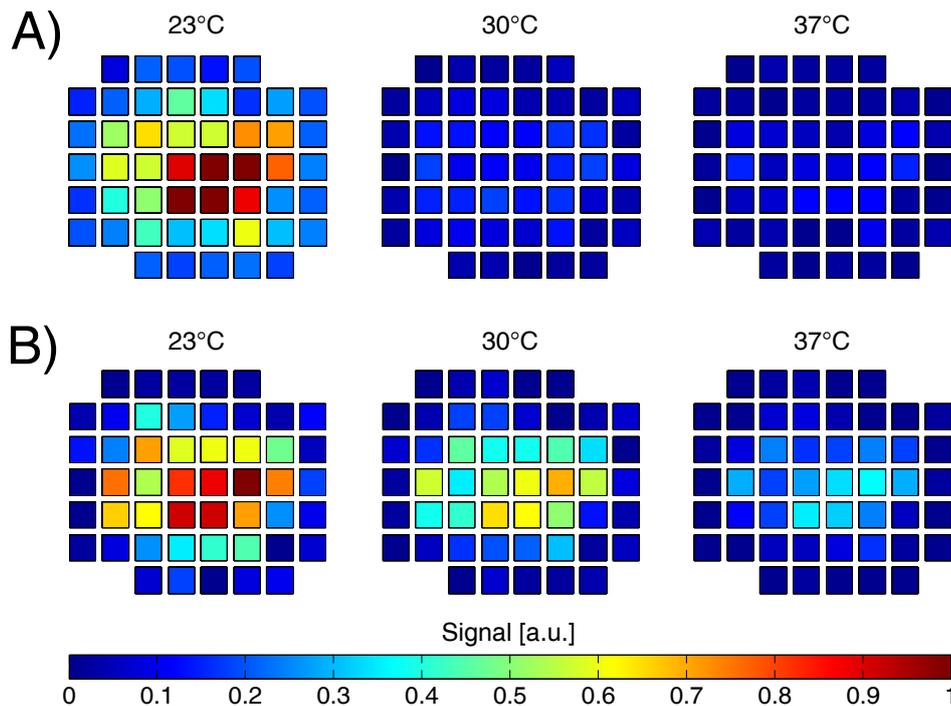
Um zu überprüfen ob das System unabhängig von seiner Vorgeschichte bei einer bestimmten Temperatur immer im gleichen Gleichgewichtszustand endet oder eine Hysterese messbar ist, werden die Differenzen der Signale der einzelnen Features zwischen Temperaturerhöhung und -senkung für drei Temperaturniveaus berechnet (Abb. 5.15). Eine Hysterese in den Signalen der Features wäre ein Hinweis, dass sich das System bei den entsprechenden Temperaturen nicht im Gleichgewicht befindet. Teil A) von Abb. 5.15 zeigt die normierten Signaldifferenzen für die Einzelhybridisierung des PM. Bei einer Temperatur von 23 °C unterscheiden sich die Signale des Systems bei Temperaturerhöhung und -senkung. Die Signalunterschiede für die beiden anderen Temperaturen sind innerhalb des experimentellen Fehlers und damit vernachlässigbar. In der Einzelhybridisierung ist der PM keiner Konkurrenz durch andere Targets ausgesetzt, sodass der Gleichgewichtszustand des Systems nicht durch die Wechselwirkungen der Targets beeinflusst werden kann. Dies bedeutet, dass die Hybridisierungszeit bei einer Temperatur von 23 °C nicht ausgereicht hat, um das Gleichgewicht zu erreichen. In [1] wurde beobachtet, dass die notwendige Zeit bis Erreichen des Gleichgewichts bei den



**Abbildung 5.13:** Normierte Signale der Features aus Abb. 5.12 bei PM-Einzelhybridisierung ( $[PM]=1$  nM) für jeden Schritt der Temperaturrampe. Die Features werden bei jedem Temperaturschritt für 60 min hybridisiert. Die Normierung erfolgt auf das Feature mit dem stärksten Signal. Eine Temperaturerhöhung resultiert in einer Signalabnahme bis zum Minimum bei 44 °C. Umgekehrt werden die Signale bei Temperatursenkung wieder stärker. Kann der PM bei 23 °C noch alle Features besetzen, so werden bei einer zunehmenden Temperaturerhöhung fast ausschließlich mittige Features besetzt: Die Signale randnaher Features des Featuresblocks nehmen stärker ab als die mittleren Features, da die Hybridisierungseigenschaften der Targets, die zu diesen mittleren Features komplementär sind, denen des PM ähneln. Das System ordnet sich mit steigender Temperatur.



**Abbildung 5.14:** Normierte Signale der Features aus Abb. 5.12 bei Hybridisierung mit der 49-komponentigen Target-Lösung für jeden Schritt der Temperaturrampe. Die Features werden bei jedem Temperaturschritt für 60 min hybridisiert. Die Normierung erfolgt auf das Feature mit dem stärksten Signal. Eine Temperaturerhöhung resultiert in einer Signalabnahme bis zum Minimum bei 44 °C. Umgekehrt werden die Signale bei Temperatursenkung wieder stärker. Aufgrund der Konkurrenz durch die anderen Targets kann der PM auch bei 23 °C - im Gegensatz zur Abb. 5.13 - an den randnahen Features des Featureblocks nur noch mäßig binden und besetzt vor allem mittige Features. Das System ist also bereits bei niedrigerer Temperatur in einem geordneteren Zustand.



**Abbildung 5.15:** Differenz der Feature-Signale zwischen Temperaturerhöhung und -senkung für drei Temperaturniveaus. **A)** Dargestellt sind die Signaldifferenzen bei PM-Einzelhybridisierung. Bei den Temperaturniveaus 30 °C bzw. 37 °C ergeben sich im Gegensatz zu 23 °C keine Signalunterschiede bei Erhöhung und Senkung der Temperatur. Bei niedrigen Temperaturen ist die Zeit bis zum Erreichen des Gleichgewichts relativ lang. Die Hybridisierungszeit von 60 min ist anscheinend nicht ausreichend um das Gleichgewicht zu erreichen, sodass sich deutliche Signalunterschiede bei Erhöhung und Senkung der Temperatur ergeben. **B)** Dargestellt sind die Signaldifferenzen bei Hybridisierung mit der 49-komponentigen Target-Lösung. Hier ergeben sich auch bei den Temperaturniveaus von 30 °C und 37 °C Unterschiede bei Erhöhung und Senkung der Temperatur, die in Teil A) nicht aufgetreten sind. Da anhand der Ergebnisse aus Teil A) davon ausgegangen werden kann, dass sich das System bei diesen Temperaturen im Gleichgewicht befindet, kann hier von einer Hysterese gesprochen werden.

hier verwendeten Microarrays neben der Länge der verwendeten Target- und Probe-Sequenzen und der gewählten Temperatur auch davon abhängt, ob die Temperatur erhöht oder abgesenkt wird: Bei Temperaturerhöhung wird das Gleichgewicht schneller erreicht. Das bedeutet, dass die Hybridisierungszeit bei Temperatursenkung auf 23 °C zu kurz ist, sich das System bei 30 °C bzw. 37 °C aber im Gleichgewicht befindet. In Teil B) sind die Signaldifferenzen für die 49-komponentige Target-Lösung dargestellt. Auch hier ergeben sich wie bei der PM-Einzelhybridisierung bei einer Temperatur von 23 °C Signalunterschiede bei Erhöhen und Absenken der Temperatur, die vermutlich auf ein Nicht-Gleichgewicht des Systems zurückgeführt werden können. Jedoch zeigen sich auch deutliche Signalunterschiede bei den anderen beiden Temperaturen, die im Falle der Einzelhybridisierung nicht auftraten. Da die Hybridisierungszeit bei der Einzelhybridisierung und dem 49-komponentigen System bei allen Temperaturniveaus

gleich war und sich das System bei der Einzelhybridisierung des PM bei 30 °C bzw. 37 °C im Gleichgewicht befand, kann im 49-komponentigen System bei diesen Temperaturen von einer Hysterese gesprochen werden, die durch die Wechselwirkung der Targets untereinander zustande kommt: Die Wechselwirkung der Targets verhindert bzw. verzögert das Erreichen des thermodynamischen Gleichgewichts des Systems. Die beobachtete Hysterese fällt vermutlich noch stärker aus, wenn man die Konzentrationen der beteiligten Target-Spezies so einstellt, dass alle Features von Targets besetzt sind und praktisch keine unhybridisierten Targets in Lösung mehr vorliegen: Befindet sich das System bei tiefen Temperaturen im ungeordneten Zustand, in dem die Bindung zwischen Probe und Target relativ unspezifisch ist, können bei Temperaturerhöhung frei werdende Bindungsstellen nicht direkt durch Targets aus der Lösung nahe der Oberfläche besetzt werden und das System geht erst bei hohen Temperaturen in einen geordneten Zustand über. Wenn umgekehrt alle Features im geordneten Systemzustand bei höherer Temperatur ausschließlich mit ihren perfekt komplementären Targets besetzt sind, geht das System erst bei relativ niedrigen Temperaturen in den ungeordneten Zustand über. Dadurch sollte sich in einem gewissen Temperaturbereich ein größerer Signalunterschied zwischen Temperaturerhöhung und -senkung ergeben. Das Einstellen der optimalen Target-Konzentration ist schwierig, da die genaue Anzahl der Probes auf der Oberfläche nicht bekannt ist und nur grob abgeschätzt werden kann. Im verwendeten Microarray-Aufbau wird ein Feature durch 6x6 Pixel dargestellt. Da ein Pixel eine Breite von ca. 4 µm hat, entspricht dies einer Feature-Breite von 24 µm. Geht man davon aus, dass sich die Probes innerhalb eines Features nicht berühren (dies ist bei der verwendeten Probe-Länge von 16 Basen gerechtfertigt), kann man bei einer Probe-Länge von ca. 10 nm ungefähr  $\frac{24 \cdot 10^{-6} \text{ nm}}{10 \cdot 10^{-9} \text{ nm}} = 2.4 \cdot 10^3$  Probes nebeneinander anordnen. Daraus folgt, dass ein Feature aus ungefähr  $1 \cdot 10^7$  Probes besteht. Hat man  $N$  Features pro Target-Spezies, so werden  $N \cdot \frac{1 \cdot 10^7}{6 \cdot 10^{23}}$  mol Targets benötigt um alle Probes mit Targets zu besetzen. Da die Hybridisierungskammer ein Volumen von ca. 0.1 ml hat, entspricht das einer Konzentration der Target-Lösung von  $N \cdot 0.2$  pM pro Target-Spezies. Üblicherweise werden pro Target-Spezies mindestens 4-5 Features synthetisiert, so dass zur Besetzung aller Probes, die zu einer bestimmten Target-Spezies komplementär sind, eine Konzentration von 1 pM notwendig ist. Solch geringe Target-Konzentrationen sind im aktuellen experimentellen Aufbau nicht praktikabel, da die Hybridisierungssignale bei diesen Konzentrationen extrem schwach ausfallen und die Zeitspanne bis zum Erreichen des thermodynamischen Gleichgewichts sehr groß wird. Zusammenfassend kann gesagt werden, dass dieser Versuch die Grundidee des Experiments - die Ordnung des Systems als Funktion der Temperatur - prinzipiell bestätigt. Welchen Einfluss die Wechselwirkungen zwischen den Targets auf das Gleichgewicht des Systems haben, müsste aber noch tiefergehend untersucht werden. Ist der Gleichge-

wichtszustand abhängig von einer möglichen Wechselwirkung der eingesetzten Targets, so hätte dies Auswirkungen auf viele Anwendungsbereiche der DNA-Microarrays wie z.B. der Genexpressionsanalyse, in denen eine Vielzahl von Targets um Bindungsstellen auf der Oberfläche konkurrieren.

## 5.6 Zusammenfassung der Ergebnisse

### 5.6.1 Effekt der hochspezifischen Hybridisierung

- Die Konkurrenzhybridisierungen zwischen zwei Target-Spezies lassen sich in standard- und hochspezifische Systeme unterteilen. In standardspezifischen Systemen können die Oberflächenbedeckungen der beiden Spezies immer mit Gl. 5.1 vorhergesagt werden. In hochspezifischen Systemen ist dies nicht der Fall, da die Hybridisierung der Target-Spezies mit der größeren Bindungskonstante (HAT) unerwartet spezifisch ist: Selbst eine 1000-fache Konzentrationserhöhung des Konkurrenten ändert trotz vergleichbarer Bindungskonstanten der Targets nichts an der Oberflächenbedeckung des HAT.

### 5.6.2 Ausschluss von Artefakten

- Der Effekt der hochspezifischen Hybridisierung tritt im thermodynamischen Gleichgewicht auf.
- Der Effekt ist unabhängig von der verwendeten Messmethode. Die Messungen auf der Oberfläche sowohl mit Hilfe von DNA-Microarrays als auch mittels TIRF ergeben die gleichen Resultate wie die Messungen in Lösung mittels FRET.
- Das Auftreten des Effekts ist unabhängig vom verwendeten Probe-Sequenzmotiv.

### 5.6.3 Charakterisierung des Effekts der hochspezifischen Hybridisierung

- Der Effekt tritt auf, wenn sich die Schmelztemperaturen der konkurrierenden Target-Spezies um mindestens 10 °C unterscheiden und die beiden Moleküle unterschiedliche Fluktuationsdynamiken haben.

- Die Sequenz der in den Konkurrenzhybridisierungen eingesetzten Target-Spezies unterscheidet sich an maximal zwei Positionen. Es konnte keine Abhängigkeit der hochspezifischen Hybridisierung von der Fehlstellenposition festgestellt werden.

#### 5.6.4 Erklärungsansatz

- In hochspezifischen Systemen wechselwirken die verschiedenen Target-Spezies über Triplex-Zustände miteinander.
- Die Triplex-Zustände ändern die effektive freie Energielandschaft asymmetrisch zugunsten des HAT. Dadurch ist die HAT-Hybridisierung unerwartet spezifisch.

#### 5.6.5 Ordnungsverhalten

- Mehrkomponentige Systeme ordnen sich bei steigender Temperatur, da die verwendeten Targets nur noch die zu ihnen komplementären Probes besetzen können.
- Der Gleichgewichtszustand eines Systems kann unter geeigneten Bedingungen durch eine Wechselwirkung der Targets untereinander beeinflusst werden. Dieser Effekt muss in weiterführenden Experimenten noch genauer untersucht werden, da er großen Einfluss auf viele Anwendungsbereiche der DNA-Microarrays haben könnte, in denen viele verschiedene Target-Spezies verwendet werden.

## 6 Theoretische Arbeiten

In diesem Kapitel sollen die experimentellen Resultate aus den Kapiteln 4 und 5 für ein tiefergehendes Verständnis durch Gleichgewichtsmodelle untersucht und reproduziert werden. Im ersten Abschnitt wird der Einfluss von einzelsträngigen Loop-Strukturen in DNA-Duplexen auf das Hybridisierungssignal untersucht. Das im zweiten Abschnitt zur Untersuchung der DNA-Konkurrenzhybridisierung entwickelte Modell zeigt, dass sich trotz vergleichbarer Bindungskonstanten in der Einzelhybridisierung die Spezifität der DNA-Hybridisierung in Konkurrenz im thermodynamischen Gleichgewicht um mehrere Größenordnungen ändern kann.

### 6.1 Stabilität von DNA-Duplexen mit einzelsträngigen Loop-Strukturen

In den Experimenten in Abschnitt 4.1 wurde der Einfluss von einzelsträngigen Loop-Strukturen in der Probe auf das Hybridisierungsverhalten untersucht. Dazu wurde ein Loop verschiedener Länge  $L$  ( $1 \leq L \leq 13$ ) an verschiedenen Positionen  $P$  ( $7 \leq P \leq 26$ ) in die Probe eingebracht. Experimentell wurde beobachtet, dass das Hybridisierungssignal mit steigender Loop-Länge  $L$  monoton abnimmt und das Signal bei einer Loop-Positionsveränderung entlang des Strangs symmetrisch im Bezug zur Mitte des Duplexes variiert. Aufgrund der hohen Stabilität der DNA-Duplexe mit einzelsträngigen Loops haben diese Strukturen einen signifikanten Einfluss auf das Hybridisierungssignal und müssen bei der Dateninterpretation berücksichtigt werden.

Im Rahmen dieser Arbeit wurde ein Gleichgewichtsmodell entwickelt, das den Einfluss von einzelsträngigen Loop-Strukturen in DNA-Duplexen auf die Hybridisierung quantitativ reproduziert [89, 90]. Dazu musste das sog. *double-ended zipper*-Modell, das zur Beschreibung des Einflusses von Einzelfehlstellen (MMs von engl. *mismatches*) auf die Hybridisierung verwendet wurde [1–3], so erweitert werden, dass die zusätzlichen Bindungskonfigurationen aufgrund der Längendifferenz zwischen Probe und Target berücksichtigt werden können. Im Folgenden soll das *Zipper*-Modell noch einmal vorgestellt und die notwendigen Modifikationen erläutert werden.

### 6.1.1 Die Zustandssumme des Zipper-Modells

Nach Abschnitt 2.3 kann die Bindung von DNA-Molekülen an einer Oberfläche mit Hilfe einer Langmuir-Adsorptionsisotherme

$$\theta = \frac{[D]}{[P_0]} = \frac{K \cdot [T_0]}{1 + K \cdot [T_0]} \quad (6.1)$$

beschrieben werden, wobei  $\theta$  der Anteil der hybridisierten Targets,  $[D]$  die Duplexkonzentration der Targets auf der Oberfläche,  $[P_0]$  die Probe-Konzentration,  $K$  die Gleichgewichtsbindungskonstante und  $[T_0]$  die Target-Konzentration in Lösung ist. Da das Fluoreszenzsignal des Microarrays proportional zum Anteil hybridisierter Targets ist, wird  $\theta$  als Hybridisierungssignal bezeichnet. Ein Charakteristikum der Langmuir-Isotherme ist die sehr enge Übergangsregion von niedriger zu hoher Bindungsaffinität. Wie sich bereits in früheren Experimenten gezeigt hat, ist diese Übergangsregion auf den hier verwendeten DNA-Microarrays wesentlich verbreitert. Es konnte gezeigt werden, dass die Verbreiterung durch eine Heterogenität der Bindungsaffinitäten aufgrund unvermeidbarer Synthesefehler während der *in situ*-Synthese zustande kommt [1–3]. Während der Synthese können Streulicht, die nicht hundertprozentige Kopplungseffizienz der Amiditbausteine und andere Effekte zu fehlerhaft aufgebauten Probes führen: Basen können zusätzlich eingebaut werden (*insertions*), fehlen (*deletions*) oder durch falsche Basen ersetzt werden (*substitutions*). Des Weiteren können ganze Probes durch einen alternativen Reaktionsweg von der weiteren Synthese ausgeschlossen werden (*truncations*).

Aufgrund der Heterogenität der Bindungsaffinitäten ist es daher notwendig, die Bindung eines Targets an eine Probe auf der Oberfläche nicht nur durch eine einzelne Bindungskonstante  $K$ , sondern durch eine Verteilung von Bindungskonstanten  $K_i$  zu beschreiben. Für das Hybridisierungssignal einer einzelnen Probe mit Synthesefehlern ergibt sich dann:

$$\theta_i = \frac{K_i \cdot [T_0]}{1 + K_i \cdot [T_0]} \quad (6.2)$$

Nimmt man eine binominale Verteilung der Synthesefehler an mit der Wahrscheinlichkeit  $p$ , dass ein Fehler auftritt, so ist das Hybridisierungssignal  $\theta$  eines Features gegeben durch:

$$\theta = \sum_{k'} x_{k'} \cdot \frac{\sum_{i=1}^{(N')} \theta_i}{\binom{N'}{k'}} = \sum_{k'} x_{k'} \cdot \frac{\sum_{i=1}^{(N')} \frac{K_i \cdot [T_0]}{1 + K_i \cdot [T_0]}}{\binom{N'}{k'}} \quad (6.3)$$

$N'$  ist die Anzahl der Basen in der Probe,  $k'$  die Anzahl der Synthesefehler und  $x_{k'}$  die Wahrscheinlichkeit, dass  $k'$  Synthesefehler in einer Probe der Länge  $N'$  auftreten.

Für  $x_{k'}$  gilt:

$$x_{k'} = \binom{N'}{k'} \cdot p^{k'} \cdot (1-p)^{N'-k'} \quad (6.4)$$

Zur Vereinfachung wird angenommen, dass die unterschiedlichen oben beschriebenen Synthesefehler durch einen einzigen Defektparameter  $\Delta g_{def,syn}$  beschrieben werden können.

Ziel ist es nun, die Bindungskonstanten  $K_i$  als Funktion der Loop-Position  $P$  und der Loop-Länge  $L$  mit Hilfe des *Zipper*-Modells zu berechnen. Im Gegensatz zum NN-Modell (siehe Abschnitt 2.3.3), in dem der DNA-Duplex entweder vollständig denaturiert oder hybridisiert vorliegt, berücksichtigt das *Zipper*-Modell auch Zustände, in denen der Duplex nur teilweise hybridisiert ist. Die Hybridisierung eines DNA-Duplexes wird hier als eine Art doppelseitiger Reißverschluss modelliert. Dabei wird angenommen, dass sich der Duplex nur ausgehend von den hybridisierten Enden, den Reißverschlussgabeln (*zipper forks*), öffnen bzw. schließen kann. Diese Annahme ist gerechtfertigt, da aufgrund der hohen *stacking*-Wechselwirkungen die Öffnung von innen liegenden Basen bei kurzen Strängen sehr unwahrscheinlich ist. Die Zustandssumme  $Z_{zipper}$  des *Zipper*-Modells lautet [102–104]:

$$Z_{zipper} = \sum_{k=0}^{N-1} \sum_{l=k+1}^N \omega_{k,l} = \sum_{k=0}^{N-1} \sum_{l=k+1}^N e^{\Delta G_{k,l}^0/RT} \quad (6.5)$$

$N$  ist die Anzahl der NN-Paare und  $\omega_{k,l}$  das statistische Gewicht des teilweise denaturierten Zustands  $S_{k,l}$ .  $k$  und  $l$  bezeichnen die Positionen der linken und rechten Reißverschlussgabel.  $\Delta G_{k,l}^0$  ist die Summe über die NN-Energien  $\Delta g_i^0$  ( $\Delta g_i^0 > 0$ ) des geschlossenen Duplexbereiches:

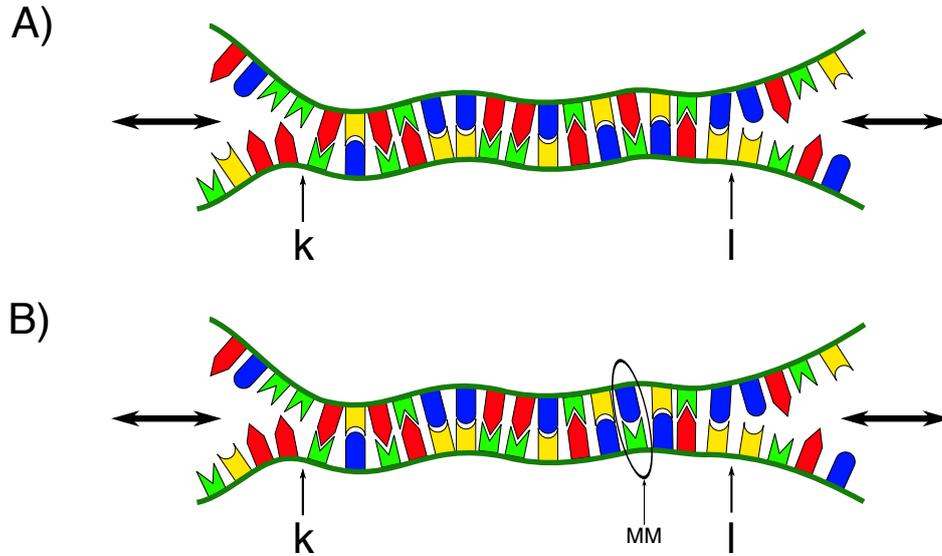
$$\Delta G_{k,l}^0 = \left( \sum_{i=k}^l \Delta g_i^0 \right) + \Delta g_{init} \quad (6.6)$$

$\Delta g_{init} = -4.5$  kcal/mol ist die freie Energie, die zur Duplexinitialisierung überwunden werden muss [3]. Analog zur Berücksichtigung von Synthesefehlern können auch MMs im *Zipper*-Modell berücksichtigt werden, indem die NN-Energien  $\Delta g_i$  durch die entsprechenden MM-Defektparameter ersetzt werden [53].

Nimmt man den vollständig denaturierten Zustand  $S_0$  als Referenzzustand an, so gilt für die Bindungskonstante  $K_i(P, L)$ :

$$K_i(P, L) = Z_{zipper}(P, L) \quad (6.7)$$

Das *Zipper*-Modell und seine Notationen ist in Abb. 6.1 noch einmal veranschaulicht. Dort ist der Duplex zwischen den Positionen  $k$  und  $l$  der Reißverschlussgabeln hybri-



**Abbildung 6.1:** Notationen des *Zipper*-Modells. **A)** Der Duplex ist zwischen den Positionen  $k$  und  $l$  der Reißverschlussgabeln hybridisiert (Zustand  $S_{k,l}$ ). Dadurch wird die freie Energie  $\Delta G_{k,l}$  gewonnen. Die schwarzen Pfeile rechts und links deuten an, dass das Öffnen bzw. Schließen des Duplexes nur von den hybridisierten Enden ausgehen kann. **B)** Hier ist eine Einzelfehlstelle (MM) in den Duplex eingebracht, die im Rahmen des *Zipper*-Modells auch berücksichtigt werden kann.

disiert (Zustand  $S_{k,l}$ ), was die freie Energie  $\Delta G_{k,l}^0$  generiert. Die schwarzen Pfeile auf beiden Seiten deuten an, dass die Öffnung bzw. Schließung des Duplexes nur von den hybridisierten Enden ausgehen kann. In Teil B) ist ein MM in den Duplex integriert, was auch im Rahmen des *Zipper*-Modells berücksichtigt werden kann.

### Loop-Strafenergien

In [3] wurde gezeigt, dass der Einfluss von MMs auf das Hybridisierungssignal mit Hilfe des *Zipper*-Modells und den entsprechenden MM-Defektparametern beschrieben werden kann. Im Folgenden soll untersucht werden, ob sich auch der Einfluss von Loops mit Hilfe des *Zipper*-Modells behandeln lässt. Im Falle der Loops sind keine Defektparameter bekannt. Die als rein entropisch angenommenen Defektparameter werden berechnet, indem die Konformationen eines DNA-Loops als *self-avoiding walk* (SAW) auf einem Gitter aufgefasst werden. Da im *Zipper*-Modell die Duplexöffnung nur von den hybridisierten Enden ausgehen kann und daher die Loops immer geschlossen sind, müssen nur diejenigen SAWs berücksichtigt werden, die wieder zu ihrem Ausgangspunkt zurückkehren. Für die Anzahl der SAWs der Länge  $x$ , die zu ihrem Ausgangspunkt zurückkehren, gilt im Limit  $x \rightarrow \infty$  [32, 105]:

$$\#_{SAW,origin}(x) \propto \sigma \cdot \frac{\mu^x}{x^c} \quad (6.8)$$

$\sigma = 1.75 \cdot 10^{-4}$  ist der sogenannte Kooperativitätsparameter,  $\mu$  die Konnektivitätskonstante und  $c = 2.15$  der *loop closure*-Exponent.  $\sigma$  und  $c$  sind universelle Konstanten, während  $\mu$  von der Geometrie des Gitters abhängt (hier wird  $\mu = 4.684$  benutzt). Für die Gesamtzahl aller möglichen SAW-Konformationen der Länge  $x$  gilt [105, 106]:

$$\#_{SAW} \propto \mu^x \cdot x^{\gamma-1} \quad (6.9)$$

$\gamma = 1.157 \cdot 10^{-3}$  ist der universelle Entropieexponent. Mit Gl. (6.8) und (6.9) ergibt sich die Wahrscheinlichkeit  $\rho(x)$ , dass ein SAW der Länge  $x$  zum Ursprung zurückkehrt zu:

$$\rho(x) = \frac{\#_{SAW,origin}}{\#_{SAW}} \propto \frac{\sigma \cdot \mu^x}{\mu^x \cdot x^{\gamma-1}} \quad (6.10)$$

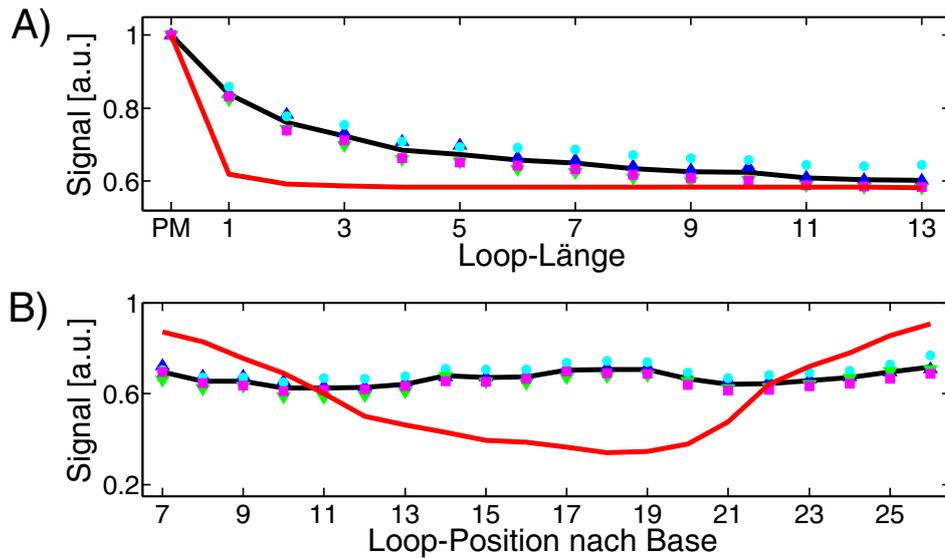
Mit Hilfe von  $\rho(x)$  kann die Entropieänderung  $\Delta S(x)$  bei Schließung eines Loops und der zugehörige Loop-Defektparameter  $\Delta G_{entropy}(x)$  berechnet werden:

$$\Delta S(x) = R \cdot \ln[\rho(x)] \Rightarrow \Delta G_{entropy}(x) = -T \cdot \Delta S(x) \quad (6.11)$$

$R = 1.987 \cdot 10^{-3}$  kcal/(mol·K) ist die universelle Gaskonstante. Die Länge eines DNA-Loops wird durch die Anzahl der Basen  $L$  im Loop und der Internukleotiddistanz  $a_0$  zwischen den einzelnen Basen bestimmt. Die Länge eines *random walks* bezeichnet die Anzahl der Schritte auf dem Gitter mit der Gitterkonstanten  $p_0$ . Beschreibt man die Konformationen eines DNA-Loops als SAW auf einem Gitter, so muss die Persistenzlänge der DNA berücksichtigt werden, die die Anzahl der Schritte des SAW und damit die Gitterkonstante  $p_0$  definiert. Da  $p_0$  und  $a_0$  bei hohen Salzkonzentrationen in der selben Größenordnung liegen [12, 13, 107], werden zur Vereinfachung DNA-Loops der Länge  $L \cdot a_0$  als SAW mit  $L$  Schritten auf einem Gitter mit der Gitterkonstanten  $p_0 \approx a_0$  aufgefasst (die Salzkonzentration in den Experimenten beträgt bei 5xSSPE-Puffer 0.90 M NaCl und 50mM NaH<sub>2</sub>PO<sub>4</sub>). Des Weiteren wurde der Einfluss von  $p_0$  auf die absoluten Werte der Loop-Defektparameter  $\Delta G_{entropy}$  untersucht, wobei keine signifikanten Unterschiede bei Variation von  $p_0$  festgestellt werden konnten. Für einen Loop der Länge  $L$  erhält man somit den Defektparameter:

$$\Delta G_{entropy}(L) = -T \cdot \Delta S(L) \quad (6.12)$$

Abb. 6.2 zeigt den Vergleich zwischen den experimentellen Daten aus Abschnitt 4.1 und den Vorhersagen des *Zipper*-Modells nach Normierung auf das PM-Signal ( $T = 317$  K,  $\Delta g_{def,syn} = -1$  kcal/mol,  $p = 0.084$ ). Teil A) zeigt die Hybridisierungssignale als Funktion der Loop-Länge  $L$  nach Mittelung über alle Loop-Positionen. Teil B) zeigt die Hybridisierungssignale als Funktion der Loop-Position  $P$  nach Mittelung über alle



**Abbildung 6.2:** Vergleich der experimentellen Daten mit den Vorhersagen durch das *Zipper*-Modell. Symbole: Featureblock 1 - blaue Dreiecke; Featureblock 2 - cyanblaue Kreise; Featureblock 3 - grüne Dreiecke; Featureblock 4 - magentarote Vierecke; experimenteller Mittelwert - schwarze Linie; Theorie - rote Linie. Die Daten sind jeweils auf das PM-Signal normiert. Simulationsparameter waren:  $T = 317$  K,  $\Delta g_{def,syn} = -1$  kcal/mol und  $p = 0.084$ . **A)** Hybridisierungssignale als Funktion der Loop-Länge nach Mittelung über alle Loop-Positionen. Der Vergleich zwischen Experiment und Theorie zeigt, dass das *Zipper*-Modell die experimentellen Daten nicht vorhersagen kann. **B)** Hybridisierungssignale als Funktion der Loop-Position nach Mittelung über alle Loop-Längen. Loop-Position  $P$  gibt an, nach welcher Base der Loop in die Ausgangssequenz der Probe eingefügt wird. Auch hier wird keine Übereinstimmung zwischen Theorie und Experiment erzielt.

Loop-Längen. Loop-Position  $P$  gibt an, nach welcher Base der Loop in die Ausgangssequenz der Probe eingefügt wird. Die verschiedenen Symbole kennzeichnen dabei die Signale der einzelnen Feature-Blöcke, die schwarze Linie den Mittelwert der Signale der Feature-Blöcke und die rote Linie die theoretischen Vorhersagen durch das *Zipper*-Modell. Deutlich erkennbar ist, dass die experimentellen Ergebnisse nicht durch dieses einfache *Zipper*-Modell reproduziert werden können. Daher soll im nächsten Abschnitt das Modell so erweitert werden, dass auch der Einfluss von Loops theoretisch vorhergesagt werden kann.

### 6.1.2 Die Zustandssumme des erweiterten Zipper-Modells

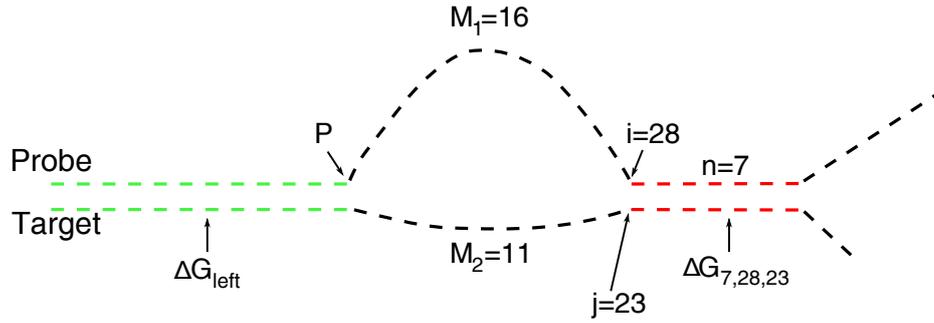
Bislang konnte sich die Öffnung des Duplexes nur von den Enden der hybridisierten Bereiche vollziehen. Wie im Folgenden gezeigt wird, muss die Öffnung des Duplexes an der Loop-Position und die dadurch entstehenden zusätzlichen Bindungszustände zur korrekten Vorhersage der Hybridisierungssignale mit berücksichtigt werden.

Die Zustandssumme  $Z_{on}(P, L)$  eines Duplexes mit Loop der Länge  $L$  an Position  $P$  kann als Summe aus fünf Teilen geschrieben werden:

$$Z_{on}(P, L) = Z_{zipper}(P, L) + Z_{extended, right}(P, L) + Z_{extended, left}(P, L) - Z_{double\ zipper}(P, L) + Z_{non-canonical}(P, L) \quad (6.13)$$

- $Z_{zipper}(P, L)$ : Die Öffnung des Duplexes vollzieht sich von den beiden Enden des hybridisierten Bereichs.  $Z_{zipper}(P, L)$  entspricht der Zustandssumme aus Gl. 6.5.
- $Z_{extended, right}(P, L)$ : Die Stränge von Probe und Target können rechts von der Loop-Position  $P$  jede mögliche Bindungskonfiguration eingehen und unterliegen dabei nicht einer reinen *Zipper*-Bindungsstatistik. So können Loops verschiedener Größe in Probe- und Target-Strang entstehen. Der Duplexbereich links der Loop-Position unterliegt der *Zipper*-Bindungsstatistik und generiert eine freie Energie  $\Delta G_{left}$ . Die Berechnung von  $Z_{extended, right}$  ist in Abb. 6.3 veranschaulicht. Dort bezeichnen die rot und grün gestrichelten Linien hybridisierte Duplexbereiche. Die schwarz gestrichelten Linien repräsentieren denaturierte Bereiche am Ende des Duplexes oder sie bilden einen Loop zwischen Probe und Target (mittlerer Bereich). Im gezeigten Fall bilden Probe und Targets Loops mit einer Länge von 16 bzw. 11 Basen aus. Ab Base 28 des Probe-Strangs und Base 23 des Target-Strangs verbinden sich die beiden Stränge wieder und bilden die nächsten 7 Basen einen Duplex, was eine freie Energie  $\Delta G_{7,28,23}$  generiert.
- $Z_{extended, left}(P, L)$ : wird analog zu  $Z_{extended, right}$  berechnet, nur dass die Seiten vertauscht sind
- $Z_{double\ zipper}(P, L)$ : Beide Duplexbereiche - links und rechts der Loop-Position - verhalten sich wie zwei unabhängige *Zipper*. Um durch die Addition von  $Z_{extended, right}$  und  $Z_{extended, left}$  Zustände nicht doppelt zu zählen, muss diese Zustandssumme in der Berechnung abgezogen werden.
- $Z_{non-canonical}(P, L)$ : Diese Zustandssumme beinhaltet alle nicht-kanonischen Bindungszustände, die gleichzeitig auf beiden Seiten der Loop-Position auftreten. Dieser Term kann aufgrund der nur sehr schwachen Bindungen vernachlässigt werden.

Die Summation über alle möglichen Bindungszustände zwischen Probe und Target rechts der Loop-Position in  $Z_{extended, right}$  hängt von dem aktuellen *Zipper*-Bindungszustand  $S_{k,l}$  des Duplexbereiches links der Loop-Position ab. Dies macht die Berechnung zeitintensiv. Daher wird für die Berechnung von  $Z_{extended, right}$  folgende Ap-



**Abbildung 6.3:** Notation von  $Z_{extended,right}$ . Jede Linie steht für eine einzelne Base. Hier wird ein Loop bestehend aus 9 Basen nach Base 12 in die Probe eingefügt (Loop-Position  $P = 12$ ). Die grünen Linien bezeichnen hybridisierte Duplexbereiche (Watson-Crick Paarung, kanonisch), die schwarzen Linien denaturierte Bereiche (am rechten Ende des Duplexes) bzw. Loop-Bereiche (Mitte). Rechts der Loop-Position  $P$  können Probe und Target alle möglichen Bindungskonfigurationen eingehen, so dass es hier sowohl zu kanonischen als auch nicht-kanonischen Basenpaarungen kommen kann (roter Bereich). Dadurch können in Probe und Target Loops verschiedener Länge entstehen. In diesem Beispiel bilden Probe und Target Loops der Länge  $M_1 = 16$  bzw.  $M_2 = 11$  Basen aus, die sich nach Base  $i = 28$  bzw.  $j = 23$  wieder zu einem Duplex der Länge  $n = 7$  verbinden. Dies generiert die Energie  $\Delta G_{7,28,23}$ . Um die Berechnungszeit zu verkürzen, wird angenommen, dass der grüne Duplexbereich links der Loop-Position immer hybridisiert ist. Die freie Energie  $\Delta G_{left}$  dieses Bereiches wird mit Gl. (6.15) berechnet. Der schwarze denaturierte Bereich am rechten Ende des Duplexes wird in der Berechnung nicht berücksichtigt.

proximation benutzt (siehe Abb. 6.3):

$$Z_{extended,right}(P, L) = \sum_{n=2}^{N-P+1} \sum_{i=P}^{N+L-n+1} \sum_{j=P}^{N-n+1} \omega_{n,i,j}$$

$$\text{mit } \omega_{n,i,j} = e^{\frac{\Delta G_{n,i,j} + \Delta G_{entropy,right} + \Delta G_{left}}{RT}} \quad (6.14)$$

$$\text{und } \Delta G_{n,i,j} = \sum_{r=1}^{n-1} \Delta g_r^{ij}$$

$i$  und  $j$  kennzeichnen die Positionen der jeweiligen Basen in Probe bzw. Target, ab denen  $n$  Basen ( $n - 1$  Basenpaare) im Bereich rechts von  $i$  und  $j$  hybridisiert sind.  $\Delta G_{n,i,j}$  bezeichnet somit die NN-Energie der  $n - 1$  Basenpaare, die ausgehend von der Position  $i$  in der Probe und der Position  $j$  im Target hybridisiert sind.  $\Delta g_r^{ij}$  kennzeichnet die freie Energie eines einzelnen Basenpaares in diesem hybridisierten Bereich. Die freie Energie  $\Delta G_{left}$  des Duplexbereiches links der Loop-Position wird mit Hilfe des *Zipper-*Modells (Gl. (6.5)) genähert. Da  $\Delta G_{left}$  nur eine Funktion der Loop-Position  $P$  ist und nicht vom aktuellen Bindungszustand  $S_{k,l}$  abhängt, kann die Berechnungsdauer deutlich verkürzt werden.

Es gilt:

$$\Delta G_{left} = \Delta G_{left}(P) = RT \cdot \ln \left[ \sum_{k=0}^{P-2} \sum_{l=k+1}^{P-1} \omega_{k,l} \right] \quad (6.15)$$

$\omega_{k,l}$  wurde in Gl. (6.5) definiert.

Im Gegensatz zum einfachen *Zipper*-Modell, in dem die Loops immer wieder zu ihrem Ausgangspunkt zurückkehren mussten, können die Loops im erweiterten Modell in einem Ursprung  $\vec{0}$  beginnen und an einem Ort  $\vec{r}$  enden (siehe auch Abb. 6.3). Daher müssen die (immer noch rein entropischen) Loop-Defektparameter  $\Delta G_{entropy,right}$  nun auf eine andere Weise berechnet werden. Für zwei SAWs mit  $M_1$  bzw.  $M_2$  Schritten hat man jetzt:

$$\begin{aligned} \Delta G_{entropy,right} &= \Delta G_{entropy,right}(M_1, M_2) \\ &= -RT \cdot \ln [\rho(M_1, M_2)] \\ &\text{mit } M_1 = i - P \\ &\text{und } M_2 = j - P \end{aligned} \quad (6.16)$$

$\rho(M_1, M_2)$  bezeichnet die Wahrscheinlichkeit, dass zwei SAWs an einem gemeinsamen Ursprung starten und sich nach  $M_1$  bzw.  $M_2$  Schritten wieder treffen:

$$\begin{aligned} \rho(M_1, M_2) &= \sum_{\vec{r}, \vec{r}'} \delta(\vec{r} - \vec{r}') \cdot \frac{\#(M_1, \vec{r}) \cdot \#(M_2, \vec{r}')}{\#_{SAW}(M_1) \cdot \#_{SAW}(M_1)} \\ &= \sum_{\vec{r}} \frac{\#(M_1, \vec{r}) \cdot \#(M_2, \vec{r})}{\#_{SAW}(M_1) \cdot \#_{SAW}(M_1)} \\ &\text{mit } |\vec{r}| \leq \min(M_1, M_2) \end{aligned} \quad (6.17)$$

$\#(M_i, \vec{r})$  ist die Anzahl der SAWs, die am Ursprung starten und bei Position  $\vec{r}$  enden. In drei Dimensionen gilt [105]:

$$\begin{aligned} \#(M_i, \vec{r}) &\propto \mu^{M_i} \cdot M_i^{\gamma-1-3\nu} \cdot g\left(\frac{r}{M_i^\nu}\right) \\ &\text{mit } g(x) \propto x^\varphi \cdot e^{-\lambda x^\delta}, \lambda > 0, \delta = \frac{1}{1-\nu} \\ &\text{und } \varphi = \frac{\gamma-1}{\nu} \end{aligned} \quad (6.18)$$

Die Konstanten  $\gamma$  und  $\mu$  sind definiert wie bisher.  $\nu = 0.588 \pm 1.5 \cdot 10^{-3}$  ist der (universelle) metrische Exponent.  $\#_{SAW}(M_i)$  bezeichnet wieder die Gesamtzahl aller SAWs mit  $M_i$  Schritten (siehe Gl. (6.9)).

In analoger Weise zu  $Z_{extended,right}$  wird  $Z_{extended,left}$  berechnet:

$$Z_{extended,left}(P, L) = \sum_{n=2}^P \sum_{i=0}^{L+P-n} \sum_{j=0}^{P-n} \omega_{n,i,j}$$

mit  $\omega_{n,i,j} = e^{\frac{\Delta G_{n,i,j} + \Delta G_{entropy,left} + \Delta G_{right}}{RT}}$  (6.19)

und  $\Delta G_{n,i,j} = \sum_{r=1}^{n-1} \Delta g_r^{ij}$

Weiterhin gilt:

$$\Delta G_{right} = \Delta G_{right}(P) = RT \cdot \ln \left[ \sum_{k=P}^{N-1} \sum_{l=k+1}^N \omega_{k,l} \right] \quad (6.20)$$

und

$$\Delta G_{entropy,left}(M_1, M_2) = RT \cdot \ln [\rho(M_1, M_2)]$$

mit  $M_1 = L + P - n - i$  (6.21)

und  $M_2 = P - n - j$

Für  $Z_{double zipper}(P, L)$  ergibt sich

$$Z_{double zipper}(P, L) = \sum_{k,l} \sum_{o,p} \omega_{k,l,o,p}$$

mit  $\omega_{k,l,o,p} = e^{\frac{\Delta G_{k,l} + \Delta G_{entropy,double} + \Delta G_{o,p}}{RT}}$  (6.22)

und

$$\Delta G_{entropy,double}(M_1, M_2) = -RT \cdot \ln [\rho(M_1, M_2)]$$

mit  $M_1 = o - l + L - 1$  (6.23)

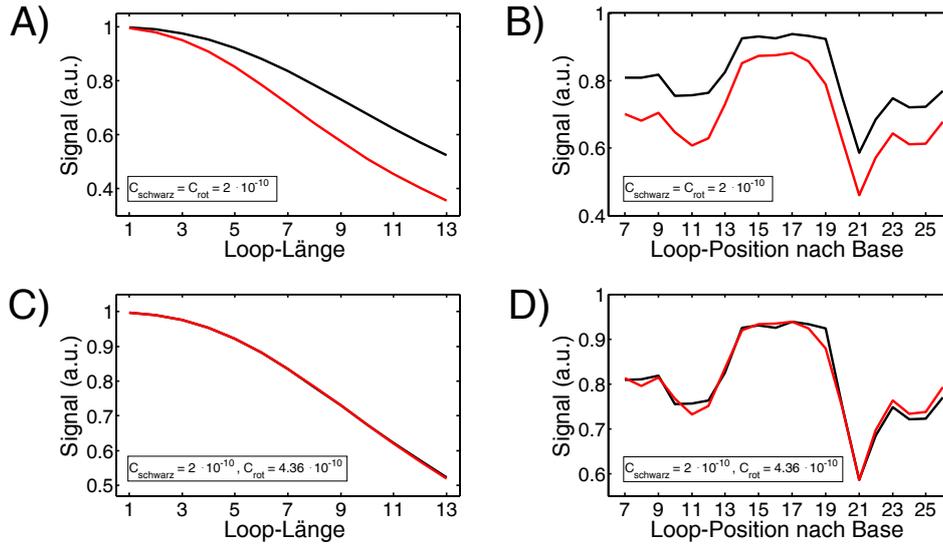
und  $M_2 = o - l - 1$

Die Bindungskonstanten

$$K_i(P, L) = Z_D(P, L) = Z_{zipper} + Z_{extended,right} + Z_{extended,left} - Z_{double zipper} \quad (6.24)$$

werden mit und ohne Näherungen für  $Z_{extended,right}$  (Gl. (6.14)) bzw.  $Z_{extended,left}$  (Gl. (6.19)) für eine einzelne Duplexsequenz berechnet. Um die theoretisch vorhergesagten Signale an die experimentellen Daten anzupassen, muss ein Skalierungsfaktor  $C$  verwendet werden, der die berechneten Bindungskonstanten mit den gemessenen Fluoreszenzintensitäten in Verbindung bringt ( $C$  ist ein freier Parameter):

$$\theta_i = \frac{C \cdot K_i \cdot [T_0]}{1 + C \cdot K_i \cdot [T_0]} \quad (6.25)$$



**Abbildung 6.4:** Vergleich der Signale, die mit bzw. ohne Näherung für  $Z_{extended,right}$  und  $Z_{extended,left}$  berechnet werden. Symbole: schwarze Linie - Berechnung ohne Näherung; rote Linie - Berechnung mit Näherung. **A)+B)** Die berechneten Signale als Funktion der Loop-Länge A)) bzw. -Position B)) unterscheiden sich bei Verwendung des gleichen Skalierungsfaktors  $C_{rot} = C_{schwarz} = 2 \cdot 10^{-10}$  zwar in ihren Absolutwerten, nehmen jedoch einen ähnlichen Verlauf. **C)+D)** Bei Anpassung der Skalierungsfaktoren zu  $C_{rot} = 2 \cdot 10^{-10}$  und  $C_{schwarz} = 4.36 \cdot 10^{-10}$  sind die beiden Kurven praktisch nicht mehr zu unterscheiden. Dies zeigt, dass die in Gl. (6.14) eingeführte Näherung sehr gut ist.

Die Skalierung der Bindungskonstanten  $K_i$  mit dem Faktor  $C$  ist unvermeidbar, da es bekanntermaßen sehr schwierig ist die absoluten Werte von Bindungskonstanten theoretisch vorherzusagen [94, 95].

Abb. 6.4 zeigt, dass die Näherungen für  $Z_{extended,right}$  und  $Z_{extended,left}$  sehr gut sind, wenn der Faktor  $C$  entsprechend angepasst wird. Wird der selbe Faktor  $C$  für die Berechnung mit (rote Kurve) und ohne Näherung (schwarze Kurve) benutzt, so unterscheiden sich die Kurven zwar in ihren Absolutwerten, jedoch ist ihr Verlauf sehr ähnlich (siehe Abb. 6.4 A) und B)). Wird der Faktor  $C$  entsprechend angepasst, so sind die Kurven praktisch nicht zu unterscheiden (siehe Abb. 6.4 C) und D)). Durch die Näherung werden einige Bindungszustände vernachlässigt, wodurch kleinere Absolutwerte der Bindungskonstanten resultieren.

Im Falle gleicher Längen von Probe und Target kann eine erfolgreiche Duplexbindung durch *zipping* nur dann stattfinden, wenn Probe- und Target-Strang perfekt zueinander ausgerichtet sind. Da die Energie  $\Delta g_{init}$  zur Initialisierung des Duplexes als konstant angenommen wurde (siehe Gl. (6.6)), konnte sie in einen Vorfaktor inkludiert und die Bindungskonstanten  $K_i$  als  $K_i = Z_{zipper}$  geschrieben werden.

Im Falle von Duplexen mit Loops ist die Situation anders. Aufgrund der Längendifferenz  $\Delta L$  zwischen Probe und Target, erhöht sich die Anzahl der Probe-Konfigurationen,

die keine Duplexinitialisierung zulassen und die Duplexinitialisierungsenergie  $\Delta g_{init}$  ändert sich entsprechend. Die Anzahl der Kollisionen der beiden Stränge, die nicht zu einer Duplexformation führen, steigt näherungsweise linear mit der Längendifferenz  $\Delta L$  (bei Vernachlässigung der zur Duplexformation notwendigen Entfaltung der beiden Stränge). Dies führt zu einer Änderung der Entropie beim Übergang vom ungebundenen zum gebundenen Zustand:

$$\Delta S_{init} \propto \ln \left( 1 + \frac{\Delta L}{L_0} \right) \quad (6.26)$$

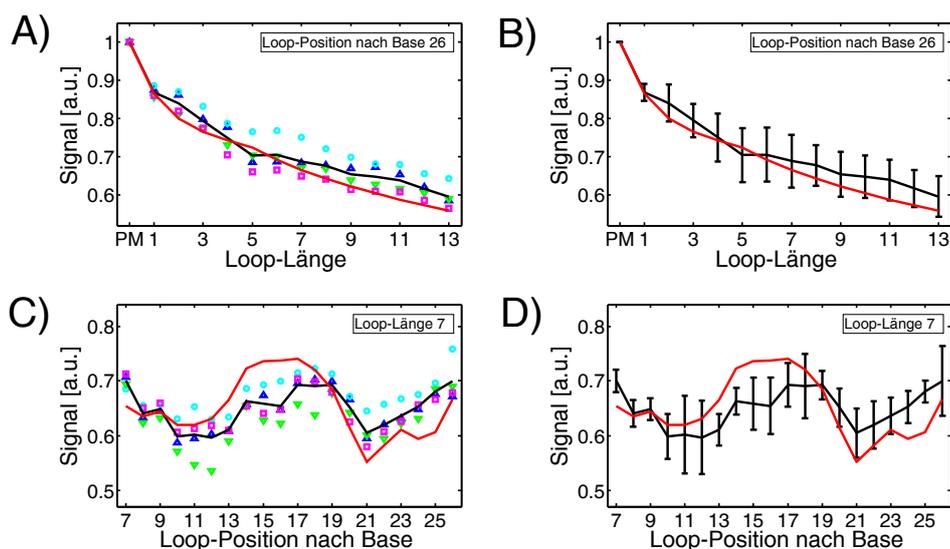
$L_0$  ist die charakteristische Länge des Systems, was gerade der Persistenzlänge einzelsträngiger DNA entspricht (bei den hier vorliegenden experimentellen Bedingungen also einer Base). Im Falle einer kurzen Loop-bildenden Sequenz in der Mitte des Strangs existieren allerdings zwei Positionen, die bei parallelen aber gegeneinander verschobenen Strängen eine Duplexinitialisierung zulassen, so dass der Term  $\Delta L/L_0$  mit dem Faktor  $1/2$  korrigiert werden muss. Befindet sich der Loop am Ende des Strangs, so existiert wiederum nur eine Position, die eine Duplexformation erlaubt und der Korrekturfaktor ist wieder 1. Die Abhängigkeit von  $\Delta S_{init}$  von der Loop-Position  $P$  wird im Folgenden vernachlässigt und durchgängig der Korrekturfaktor  $1/2$  benutzt. Es hat sich gezeigt, dass unterschiedliche Korrekturfaktoren das Ergebnis nicht merklich beeinflussen, wenn der Skalierungsfaktor  $C$  entsprechend angepasst wird.

Die Näherung für  $\Delta S_{init}$  (keine Berücksichtigung der Loop-Position) überschätzt mit steigender Längendifferenz  $\Delta L$  die entsprechende Strafergie der Duplexinitialisierung: Bei großen Längendifferenzen  $\Delta L$  sind die durch die Loop-Sequenz separierten komplementären Bereiche nahezu unabhängig voneinander. Die Initialisierungsenergie nähert sich asymptotisch zwei unabhängigen Hybridisierungsereignissen, wodurch die Abhängigkeit der Initialisierungsenergie von  $\Delta L$  schwächer wird.

Durch die Abhängigkeit der Initialisierungsenergie von  $\Delta L$  berechnet sich die Bindungskonstante  $K_i$  im Gegensatz zu Gl. (6.24) nun mit:

$$K_i = \frac{Z_D(P, L)}{1 + \frac{\Delta L}{2L_0}} \quad (6.27)$$

Die Hybridisierungssignale werden dann mit Hilfe von Gl. (6.25) berechnet. Die Wahl des Nenners in Gl. (6.27), der aus Gl. (6.26) folgt, hat direkten Einfluss auf die berechneten Hybridisierungssignale. Die Anpassung des Modells an die experimentellen Daten (insbesondere die systematische Abweichung der Hybridisierungssignale als Funktion der Loop-Länge, siehe Abb. 6.6) könnte durch eine andere Wahl des Nenners noch weiter verbessert werden. Obwohl die Berechnung der Initialisierungsenergie auch in vielen verwandten Fällen ein Problem darstellt, ist der konkrete Ablauf der DNA-Nukleation noch nicht geklärt. MD-Simulationen könnten helfen den Nukleationsprozess besser zu



**Abbildung 6.5:** Vergleich zwischen den experimentellen Daten und der Vorhersage durch das erweiterte *Zipper*-Modell. Symbole: Featureblock 1 - blaue Dreiecke; Featureblock 2 - cyanblaue Kreise; Featureblock 3 - grüne Dreiecke; Featureblock 4 - magentarote Vierecke; experimenteller Mittelwert - schwarze Linie; Theorie - rote Linie. **A)+C)** Die gemessenen Hybridisierungssignale als Funktion der Loop-Länge bzw. -Position für eine bestimmte Loop-Position bzw. -Länge werden mit den theoretisch berechneten Signalen verglichen. Normierung erfolgt auf den PM. **B)+D)** Vergleich zwischen den entsprechenden 95% Konfidenzintervallen der experimentellen Daten (schwarz) und der Theorie. Da die durch das Modell berechneten Signale zu weiten Teilen innerhalb der Konfidenzintervalle liegen, sind die Vorhersagen statistisch signifikant.

verstehen [108].

Abb. 6.5 zeigt den Vergleich zwischen den experimentellen Daten und der Vorhersage durch das Modell. Die verschiedenen Symbole kennzeichnen die Signale der verschiedenen Featureblöcke als Funktion der Loop-Länge bzw. -Position, die schwarze Linie den experimentellen Durchschnitt und die rote Linie die theoretischen Vorhersagen. In Teil A) sind die Signale als Funktion der Loop-Länge für eine bestimmte Loop-Position, in C) als Funktion der Loop-Position für eine bestimmte Loop-Länge gezeigt. Die Teile B) und D) zeigen außerdem die 95% Konfidenz-Intervalle der experimentellen Daten (schwarz) und den Vergleich mit der Theorie (rot). Wie zu erkennen ist, sind sowohl die experimentellen Daten als auch die Vorhersagen durch das Modell statistisch signifikant.

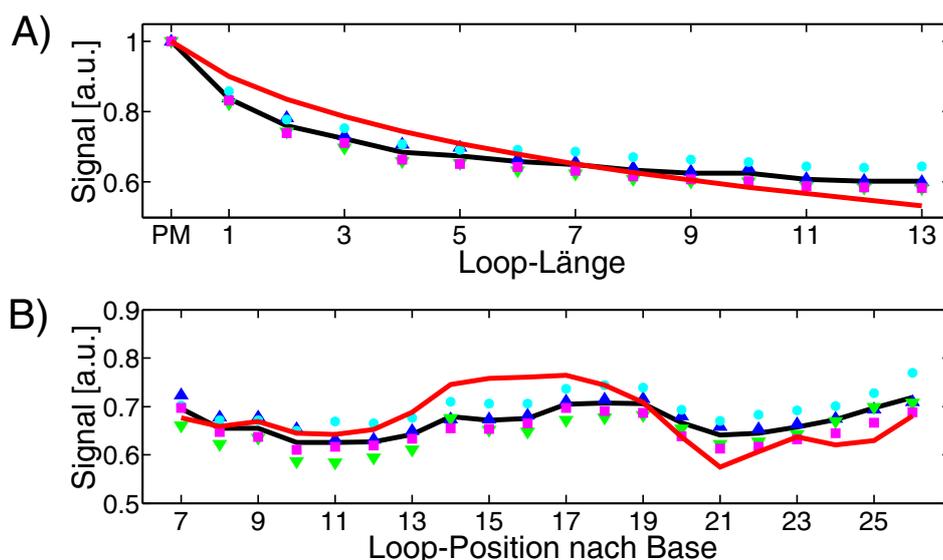
Um die Signalabhängigkeit von der Loop-Länge bzw. der Loop-Position deutlicher darzustellen, wird in Abb. 6.6 A) über alle Loop-Positionen und entsprechend in Abb. 6.6 B) über alle Loop-Längen gemittelt. Die Symbole indizieren die Signale der einzelnen Featureblöcke, die schwarze Linie den experimentellen Mittelwert und die rote Linie die Vorhersagen des Modells. Anhand Abb. 6.5 und 6.6 lässt sich erkennen, dass das erweiterte Modell die experimentellen Ergebnisse gut reproduzieren kann. Die verwen-

deten Simulationsparameter sind neben den (universellen) Parametern für die SAWs [105], die Simulationstemperatur  $T = 317$  K, die Synthesefehlerrate  $p = 0.084$  und der Defektparameter für Synthesedefekte  $\Delta g_{def,syn} = -1$  kcal/mol ( $\Delta g_{syn}$  wurde in [2] bestimmt). Es werden bis zu drei Synthesefehler pro Probe berücksichtigt. Die Berücksichtigung einer größeren Anzahl von Synthesefehlern ist bei der hier verwendeten Sequenzlänge (33 Basen) nicht notwendig, wie bereits in Abschnitt 4.2.3 gezeigt. Die Basen des Loops werden als synthesesehlerfrei behandelt, da diese Basen die meiste Zeit entweder schwach oder gar nicht gebunden sind (es gibt praktisch keine komplementären Basen im Target). Somit ist die Berücksichtigung von Synthesefehlern im Loop nicht notwendig. In der Simulation werden außerdem die NN- und MM-Parameter aus [44, 53] benutzt. Da MM-Defektparameter nur für isolierte MMs verfügbar sind, wird ein weiterer Parameter  $MM_{def} = -2$  kcal/mol für zwei unmittelbar nebeneinanderliegende Fehlstellen im Modell implementiert. Zwei nebeneinanderliegende MMs werden als zwei unabhängige Synthesefehler behandelt:  $MM_{def} = 2 \cdot \Delta g_{def,syn}$ . Die einzigen freien Parameter des Modells sind der Skalierungsfaktor  $C = 1.5 \cdot 10^{-3}$ , der die theoretisch berechneten Bindungskonstanten  $K_i$  mit den gemessenen Fluoreszenzsignalen in Verbindung bringt, und die Wahrscheinlichkeit für das Auftreten von Synthesefehlern  $p = 0.084$ .  $p$  wird allerdings auch dazu benutzt um die Übereinstimmung des Modells mit den bekannten Kopplungs- und Entschützungs effizienzen der Oligonukleotide und des erreichbaren Kontrasts des experimentellen Aufbaus zu überprüfen und ist daher nicht vollkommen frei [70, 73, 74].

Anzumerken bleibt, dass schon  $Z_{double\ zipper}(P, L)$  alleine den ungefähren Verlauf der Hybridisierungssignale als Funktion der Loop-Position in Abb. 6.6 B) wiedergibt. Die aus  $Z_{double\ zipper}(P, L)$  resultierenden Bindungskonstanten sind aber kleiner als die, die mit  $Z_{extended,right}$  und  $Z_{extended,left}$  berechnet werden.  $Z_{extended,right}$  und  $Z_{extended,left}$  reproduzieren den Verlauf und die absoluten Werte der Hybridisierungssignale wesentlich besser. In Abb. 6.7 werden die resultierenden Hybridisierungssignale der beiden Berechnungsarten als Funktion der Loop-Länge nach Mittelung über alle Loop-Positionen A) bzw. der Loop-Position nach Mittelung über alle Loop-Längen B) miteinander verglichen. Die rote bzw. schwarze Linie bezeichnet die Hybridisierungssignale, die mit  $Z_{extended,right} + Z_{extended,left}$  bzw. mit  $Z_{double\ zipper}$  berechnet werden. Die Skalierungsfaktoren  $C$  der beiden Berechnungsarten unterscheiden sich deutlich (siehe Bildunterschrift).

Kleine Unterschiede zwischen den experimentellen Daten und der Vorhersage des Modells im Hinblick auf die Abhängigkeit des Signals von der Loop-Position lassen sich mit einem Blick auf die Duplexsequenz erklären. Hier soll auf zwei Unterschiede eingegangen werden:

- Duplexregion zwischen Loop-Position 14 und 18: In dieser Region befinden sich

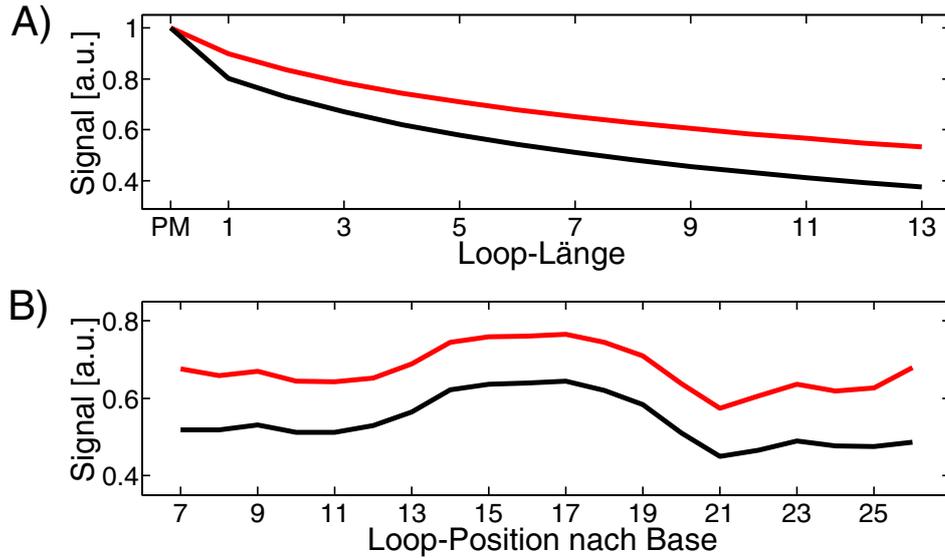


**Abbildung 6.6:** Vergleich der Mittelwerte der experimentellen Daten und der Vorhersage durch das erweiterte *Zipper*-Modell. Symbole: Featureblock 1 - blaue Dreiecke; Featureblock 2 - cyanblaue Kreise; Featureblock 3 - grüne Dreiecke; Featureblock 4 - magentarote Vierecke; experimenteller Mittelwert - schwarze Linie; Theorie - rote Linie. Normierung erfolgt auf das PM-Signal. **A)** Vergleich der experimentellen Daten mit der Theorie als Funktion der Loop-Länge nach Mittelung über alle Loop-Positionen. **B)** Vergleich der experimentellen Daten mit der Theorie als Funktion der Loop-Position nach Mittelung über alle Loop-Längen. Die Ergebnisse zeigen, dass das erweiterte *Zipper*-Modell die experimentellen Daten gut reproduzieren kann.

viele A/T-Basen. Der Abstand zweier C-Basen ist hier außerdem am größten im gesamten Duplex. Die Duplexdestabilisierung durch A/T-reiche Regionen wird sehr wahrscheinlich unterschätzt.

- Loop-Position 21: In dieser Region befinden sich viele C-Basen. Des Weiteren werden die zusätzlichen Loop-Basen unmittelbar hinter zwei C-Basen eingefügt. Wie in [1, 109, 110] gezeigt wurde, können degenerierte Basenpaarungen die Duplexstabilität wesentlich erhöhen. Diese Tatsache ist im Modell nicht implementiert.

In der Literatur werden interne DNA-Loops oder sog. *bubbles* der Länge  $l = l_1 + l_2$  ( $l_1$  ungebundene Basen in der Probe,  $l_2$  ungebundene Basen im Target), die insbesondere bei sehr langen DNA-Sequenzen für die Denaturierung entscheidend sind, oft als SAWs der Länge  $l$  behandelt, die zu ihrem Ausgangsort zurückkehren [32]. Die oben gezeigten experimentellen Daten konnten mit Hilfe dieser Beschreibung nicht reproduziert werden, da die resultierenden Loop-Defektparameter zu groß waren. Die Modellierung eines DNA-Loops als SAW der Länge  $l = l_1 + l_2$ , der zu seinem Ausgangsort zurückkehrt, entspricht nicht der im erweiterten Modell vorgestellten Berechnung: Die Anzahl der Konformationen zweier unabhängiger SAWs der Längen  $l_1$  bzw.  $l_2$ , die am selben Ort starten und sich in einigem Abstand wieder treffen, ist wesentlich kleiner als die



**Abbildung 6.7:** Vergleich der beiden Berechnungsarten. Symbole: resultierende Signale aus  $Z_{extended,right} + Z_{extended,left}$  - rote Linie ( $C = 1.5 \cdot 10^{-3}$ ); resultierende Signale aus  $Z_{double zipper}$  - schwarze Linie ( $C = 3$ ). **A)** Hybridisierungssignale als Funktion der Loop-Länge nach Mittelung über alle Loop-Positionen. **B)** Hybridisierungssignale als Funktion der Loop-Position nach Mittelung über alle Loop-Positionen. Der Verlauf der Hybridisierungssignale ist ähnlich, jedoch unterscheiden sie sich in ihren Absolutwerten. Der Skalierungsfaktor  $C$  musste im Falle der Berechnung mit  $Z_{double zipper}$  deutlich erhöht werden um eine ungefähre Übereinstimmung der Absolutwerte der beiden Berechnungsarten zu erzielen.

Anzahl der Konformationen eines SAW der Länge  $l$ , da im zweiten Fall die Bedingung zu  $l'_1 + l'_2 = l$  für jedes Paar  $l'_1$  und  $l'_2$  abgeschwächt wird. Die beiden Berechnungen könnten unter der Bedingung, dass ein SAW der Länge  $l = l_1 + l_2$  nach  $l_1$  Schritten den Ort  $\vec{r}$  erreicht, die gleichen Resultate liefern [111]. Diese Unterschiede mögen bei Denaturierungsexperimenten nicht wesentlich sein. Jedoch ist der Einfluss der Loop-Defektparameter im hier vorliegenden Fall deutlich größer, da die verwendeten DNA-Stränge wesentlich kürzer sind und somit die freie Bindungsenergie kleiner ist.

Zusammenfassend kann gesagt werden, dass der Einfluss einzelsträngiger Loop-Strukturen in DNA-Duplexen auf die Hybridisierung durch die Erweiterung des *Zipper*-Modells theoretisch reproduziert werden konnte. Dazu musste die Öffnung des Duplexes an der Loop-Position zwingend mit in die Betrachtung aufgenommen werden. Wie bereits in vorherigen Arbeiten gezeigt [3], können auch hier NN-Parameter aus [53] basierend auf Experimenten in Lösung zur Beschreibung der DNA-Hybridisierung auf Microarray-Oberflächen verwendet werden. Die notwendigen Loop-Defektparameter konnten durch die Modellierung der Loops als SAW auf einem Gitter berechnet werden.

## 6.2 Hochspezifische DNA-Hybridisierung in Konkurrenz

In Kapitel 5 wurden Experimente vorgestellt, die die Unterteilung von DNA-Konkurrenzsyste men in standard- und hochspezifische Systeme zulassen. In standardspezifischen Systemen (z.B. „PM vs. MM2“) kann die Hybridisierung vollständig durch die Bindungskonstanten aus den Einzelhybridisierungen charakterisiert und die Oberflächenbedeckung der einzelnen Spezies leicht durch die jeweiligen Konzentrationen eingestellt werden. In hochspezifischen Konkurrenzsyste men hingegen (z.B. „PM vs. MM1“) ist die Oberflächenbedeckung eines dominierenden Moleküls mit der größeren Bindungskonstante (HAT von engl. *high affinity target*) weitestgehend unabhängig von der Konzentration eines Konkurrenz moleküls mit der niedrigeren Bindungskonstante (LAT von engl. *low affinity target*), obwohl die Bindungskonstanten aus den Einzelhybridisierungen in der gleichen Größenordnung liegen.

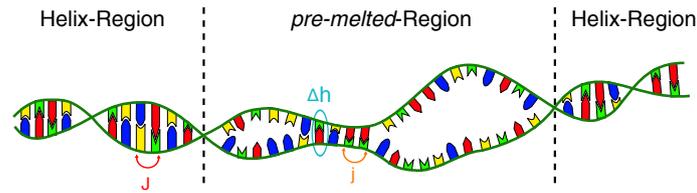
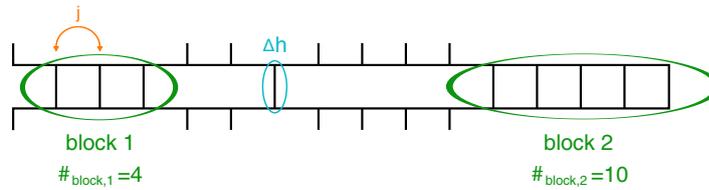
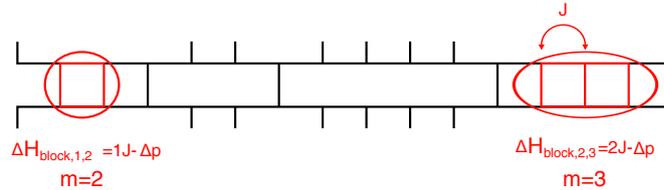
Im Rahmen dieser Arbeit wurde ein Gleichgewichtsmodell entwickelt, mit dem der Effekt der hochspezifischen Hybridisierung eines HAT qualitativ wiedergegeben und die Mechanismen, die für den Effekt verantwortlich sind, charakterisiert werden können [97]. Dazu werden im Modell zwei unterschiedliche Bindungskonfigurationen der DNA unterschieden. In Abb. 6.8 A) sind die beiden möglichen Bindungskonfigurationen in einem Duplex veranschaulicht:

- ***pre-melted*-Konfiguration**

In dieser Konfiguration ist die typische DNA-Doppelhelix aufgebrochen und der Duplex liegt als flexibles Molekül (niedrige Persistenzlänge) mit vielen entropischen Freiheitsgraden vor. Die Bindungsenthalpie zwischen zwei gegenüberliegenden Basen im Duplex beträgt  $\Delta h$  und die Kopplungsenergie zwischen zwei benachbarten gebundenen Basen  $j$ . Die Modellierung dieser Konfiguration wird in Abb. 6.8 B) veranschaulicht.

- **Helix-Konfiguration**

Der DNA-Duplex liegt in der bekannten Doppelhelixstruktur vor. Diese zeichnet sich durch geringe Flexibilität (große Persistenzlänge) und wenige entropische Freiheitsgrade aus. Die Helix-Konfiguration und die *pre-melted*-Konfiguration können in einem DNA-Duplex gleichzeitig vorliegen. Gehen zwei benachbarte gebundene Basen einer *pre-melted*-Konfiguration in eine Helix-Konfiguration über, so wird die Kopplungsenergie  $J$  generiert, was jedoch auf Kosten von entropischen Freiheitsgraden vollzogen wird. Die Modellierung dieser Konfiguration ist in Abb. 6.8 C) dargestellt.

A) Koexistierende *pre-melted*- und Helix-KonfigurationB) *pre-melted*-Konfiguration:  $\vec{v}=(0,1,1,1,0,0,1,0,0,0,0,1,1,1,1,1)$ Anzahl Blöcke  $\#_{\text{blocks}}=2$ Anzahl möglicher Konfigurationen  $k$  der *pre-melted*-Konfiguration  $\vec{v} = \#_{\text{block},1} \cdot \#_{\text{block},2} = 40$ C) eine mögliche Helix-Konfiguration der *pre-melted*-Konfiguration  $\vec{v}$ 

**Abbildung 6.8:** Modell und numerische Beschreibung der verschiedenen DNA-Bindungskonfigurationen. **A)** Modell: Gezeigt ist ein DNA-Duplex, in dem die beiden Bindungskonfigurationen nebeneinander vorliegen. In der Mitte des Duplexes befindet sich eine flexible *pre-melted*-Region, die viele entropische Freiheitsgrade besitzt. Bei Bindung zweier Basen wird die Enthalpie  $\Delta h$  generiert (blauer Kreis). Außerdem existiert eine Kopplungsenergie  $j$  (orange) zwischen benachbarten Basen. Die beiden Endregionen des Duplexes befinden sich in einer steifen Helix-Konfiguration mit wenigen entropischen Freiheitsgraden. Die Kopplungsenergie zwischen benachbarten Basen ist hier  $J$  (rot,  $J > j$ ). **B)+C)** Numerische Berechnung. In B) ist die *pre-melted*-Konfiguration  $\vec{v} = (0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1)$  gezeigt. Die Enthalpie  $\Delta H_{\vec{v}}^p$  aus Gl. (6.28) und die Entropie  $\Delta S^p$  aus Gl. (6.29) ergeben zusammen die freie Energie  $\Delta G_{\vec{v}}^p$  (Gl. (6.30)) dieser Konfiguration. Die Anzahl der Blöcke  $\#_{\text{block}}$  von aufeinanderfolgenden geschlossenen Basen dieser *pre-melted*-Konfiguration  $\vec{v}$  ist 2 (grüne Kreise). Somit ergibt sich nach Gl. (6.31) die Anzahl  $k$  der möglichen Helix-Konfigurationen dieser *pre-melted*-Konfiguration zu  $k = \#_{\text{block},1} \cdot \#_{\text{block},2} = 4 \cdot 10 = 40$ . C) Hier ist eine der möglichen Helix-Konfigurationen der *pre-melted*-Konfiguration  $\vec{v}$  dargestellt. Die roten Bereiche haben jeweils eine Helix ausgeprägt. Die dadurch zusätzlich generierte Enthalpie  $\Delta H_{\text{block},i,m}$  wird anhand Gl. (6.33) berechnet. Die Ausbildung einer Helix-Region in Anwesenheit eines Konkurrenten wird mit der Entropiebarriere  $\Delta p$  bedacht (siehe Abschnitt 6.2.2). Durch Berücksichtigung der Entropieänderung  $S^h$  beim Übergang von der *pre-melted*-Konfiguration  $\vec{v}$  zu dieser bestimmten Helix-Konfiguration kann die freie Energie dieser Helix-Konfiguration  $\Delta G_{\vec{v},k}^h$  berechnet werden.  $\Delta G_{\vec{v},k}^h$  und  $\Delta G_{\vec{v}}^p$  zusammen ergeben die freie Energie  $\Delta G_{\vec{v},k}$  dieser speziellen Bindungskonfiguration.

In den folgenden Abschnitten wird die Beschreibung des Modells durch Zustandssummen erläutert.

### 6.2.1 Freie Energie einer *pre-melted*-Konfiguration

Für die Berechnung der Enthalpie  $\Delta H_{\vec{v}}^p$  einer *pre-melted*-Konfiguration wird eine ISING-Beschreibung benutzt, die der Kollektivität der Basenbindung Rechnung trägt: Ist eine Base des Duplexes geschlossen, so besteht eine erhöhte Wahrscheinlichkeit, dass auch benachbarten Basen geschlossen sind. Somit ergibt sich für  $\Delta H_{\vec{v}}^p$ :

$$\Delta H_{\vec{v}}^p = \Delta H_{\vec{v}}^p(\vec{v}) = \sum_{i=1}^M j \cdot v_i \cdot v_{i+1} + \sum_{i=1}^{M+1} \delta(v_i - 1) \cdot \Delta h \quad (6.28)$$

Dabei steht  $M$  für die maximale Anzahl der NN-Paare des DNA-Duplexes (in den in Kapitel 5 vorgestellten Experimenten beträgt die Länge der DNA 16 Basen, also wäre hier  $M = 15$ ).  $\vec{v}$  ist ein Vektor der Länge  $M + 1$  mit Einträgen 0 und 1. Er repräsentiert eine Bindungskonfiguration zwischen Target und Probe, wobei 1 eine geschlossene Base und 0 eine offene Base indiziert.  $j > 0$  ist die Kopplungsenergie zwischen zwei aufeinanderfolgenden Basen und  $\Delta h$  der Enthalpiegewinn bei Bindung zweier gegenüberliegender Basen des Duplexes der *pre-melted*-Konfiguration (siehe Abb. 6.8 B)). Nimmt man den vollständig denaturierten Zustand als Referenzzustand an, so ist  $\Delta h > 0$ .

Um die freie Energie einer *pre-melted*-Konfiguration bestimmen zu können, muss noch die Entropie berücksichtigt werden. Der Übergang vom vollständig denaturierten Zustand zu einer *pre-melted*-Konfiguration verursacht eine Entropieänderung, da die Persistenzlänge des Duplexes nach Bindung im Vergleich zum denaturierten Zustand höher ist. Es stehen somit weniger entropische Freiheitsgrade zur Verfügung.

Zur Berechnung der Entropieänderung werden wie in Abschnitt 6.1 die möglichen Konformationen des DNA-Strangs als SAW auf einem Gitter aufgefasst. Die Anzahl der Konformationen wird mit Gl. (6.9) berechnet. Analog wird auch hier vereinfachend angenommen, dass eine Base des DNA-Strangs einem Schritt auf dem Gitter entspricht. Somit ergibt sich beim Übergang vom vollständig denaturierten Zustand zu einer *pre-melted*-Konfiguration mit  $y$  geschlossenen Basen eine Entropieänderung  $\Delta S^p(y)$  von:

$$\Delta S^p(y) = R \cdot \begin{cases} \ln \left[ \frac{\#_{SAW}^{(M+1)}}{\#_{SAW}^{(M+1-y)}} \right] & , \text{ wenn } y < M + 1 \\ \ln [\#_{SAW}^{(M+1)}] & , \text{ wenn } y = M + 1 \end{cases} \quad (6.29)$$

Demnach ist die Entropieänderung beim Übergang in die vollständig geschlossene *pre-melted*-Konfiguration am größten.

Die freie Energie  $\Delta G_{\vec{v}}^p$  einer *pre-melted*-Konfiguration  $\vec{v}$  bei der Temperatur  $T$  ergibt sich zu:

$$\Delta G_{\vec{v}}^p = \Delta H_{\vec{v}}^p - T \cdot \Delta S^p \quad (6.30)$$

## 6.2.2 Freie Energie von koexistierenden *pre-melted*- und Helix-Konfigurationen

Befinden sich in einer *pre-melted*-Konfiguration zwei gebundene Basen unmittelbar nebeneinander, so können diese in eine Helix-Konfiguration übergehen und dabei die zusätzliche Energie  $J$  generieren. Um die Anzahl  $k$  aller möglichen Helix-Konfigurationen einer bestimmten *pre-melted*-Konfiguration  $\vec{v}$  zu bestimmen, muss zuerst die Gesamtzahl  $\#_{blocks}$  aller individuellen Blöcke  $i$  mit einer Länge  $l_i \geq 2$  von aufeinanderfolgenden gebundenen Basen bestimmt werden. In der in Abb. 6.8 B) gezeigten *pre-melted*-Konfiguration existieren zwei solcher Blöcke der Länge  $l_1 = 3$  bzw.  $l_2 = 5$ . Die einzelnen Blöcke können unabhängig voneinander eine Helix ausbilden. In Abb. 6.8 C) sind zwei Basen von Block 1 und drei Basen von Block 2 in eine Helix-Konfiguration übergegangen (roter Bereich). Folgen die Blöcke einer *zipper*-Bindungsstatistik (siehe Abschnitt 6.1.1), so ergibt sich für einen einzelnen Block  $i$  der Länge  $l_i$  die Anzahl der möglichen Helix-Konfigurationen  $\#_{block,i}$  zu

$$\#_{block,i} = \frac{l_i \cdot (l_i - 1)}{2} + 1 \quad (6.31)$$

Die Gesamtzahl  $k$  aller möglichen Helix-Konfigurationen einer bestimmten *pre-melted*-Konfiguration  $\vec{v}$  ergibt sich zu:

$$k = \prod_{i=1}^{\#_{block}} \#_{block,i} \quad (6.32)$$

Befinden sich  $2 \leq m \leq l_i$  Basen eines Blocks  $i$  in der Helix-Konfiguration, so wird dadurch die Enthalpie

$$\Delta H_{block,i,m} = (m - 1) \cdot J - \Delta p \quad (6.33)$$

generiert. Wie in Abschnitt 5.4 erläutert, deuten die Experimente zur Konkurrenzhybridisierung im hochspezifischen Fall auf eine verstärkte, asymmetrische Wechselwirkung zwischen den beiden Target-Spezies in Form von kurzlebigen Triplexen hin. Zur Berücksichtigung dieser Wechselwirkung, wird  $\Delta p > 0$  als Entropiebarriere eingeführt.  $\Delta p$  bezeichnet die Entropiestrafe, die für die Ausbildung einer steifen helikalen Duplexregion bei Anwesenheit eines zweiten gebundenen Targets, das sich seiner Deformation und der Deformation der Probe widersetzt, nötig ist. Daraus folgt, dass  $\Delta p$  bei der voll-

ständig ausgeprägten Helix nicht in Erscheinung tritt. Für die weiteren Berechnungen wird  $\Delta p$  als ganzzahliges Vielfaches der Kopplungsenergie  $J$  ausgedrückt. Außerdem wird in Übereinstimmung mit dem NN-Modell der DNA-Hybridisierung  $\Delta p$  im Falle der Einzelhybridisierung bzw. schwacher, symmetrischer Wechselwirkung in Konkurrenzhybridisierungen (standardspezifische Systeme) vernachlässigt [53].

Für die Berechnung der Zustandssumme müssen nur Blöcke berücksichtigt werden, bei denen  $\Delta H_{block,i,m} > 0$  ist. Ist hingegen  $(m - 1) \cdot J \leq \Delta p \Leftrightarrow \Delta H_{block,i,m} \leq 0$ , so fällt der entsprechende Zustand als mögliche Helix-Konfiguration weg. Gl. (6.31) ist also eine Funktion von  $\Delta p$ .

Die zusätzliche Enthalpie  $\Delta H_{\vec{v},k}^h$ , die durch die Ausbildung einer Helix-Konfiguration erhalten wird, erhält man, indem man über alle individuellen Blöcke summiert:

$$\Delta H_{\vec{v},k}^h = \sum_{i=1}^{\#block} \Delta H_{block,i,m} \quad (6.34)$$

Befinden sich keine Basen in der Helix-Konfiguration, so ist  $\Delta H_{\vec{v},k}^h = 0$  und der Duplex generiert seine Stabilität nur aus der freien Energie  $\Delta G_{\vec{v}}^p$  der *pre-melted*-Konfiguration (Gl. (6.30)).

Der Übergang von einer *pre-melted*-Konfiguration zu einer Helix-Konfiguration bewirkt eine weitere Entropieänderung, da die Persistenzlänge des Duplexes in der Helix-Konfiguration größer ist. Für die freie Energie erhält man

$$\Delta G_{\vec{v},k}^h = \Delta H_{\vec{v},k}^h - T \cdot S^h \quad (6.35)$$

Dabei ist  $S^h = S^h(z)$  die oben erwähnte Entropieänderung, die nach Gl. (6.9) berechnet wird.  $z$  beschreibt die Anzahl der Basen in der Helix-Konfiguration. Da der Referenzzustand immer noch der vollständig denaturierte Zustand ist, muss hier  $S^h$  und nicht  $\Delta S^h$  geschrieben werden. Die vollständig ausgeprägte Helix-Konfiguration hat die größte Entropieänderung.

Kombiniert man Gl. (6.30) und Gl. (6.35), so ergibt sich die gesamte freie Energie  $\Delta G_{\vec{v},k}$  einer gebundenen Konfiguration zu:

$$\Delta G_{\vec{v},k} = \Delta G_{\vec{v}}^p + \Delta G_{\vec{v},k}^h \quad (6.36)$$

Damit können nun in den folgenden Abschnitten die Zustandssummen für Einzel- und Konkurrenzhybridisierung und damit die Bindungskonstanten berechnet werden.

### 6.2.3 Einzelhybridisierung

Gl. 2.3 aus Kapitel 2 kann auch durch statistische Überlegungen hergeleitet werden. Ein Ensemble mit  $N_A$  Molekülen der Spezies  $A$  habe  $N$  Bindungsstellen ( $N > N_A$ ). Die einzelnen Bindungsstellen sind unabhängig voneinander und können maximal nur von einem Molekül der Spezies  $A$  besetzt werden. Wird bei der Bindung eines Moleküls die Energie  $\Delta G_A + \mu_A$  generiert ( $\mu_A < 0$  ist das chemische Potential), so lautet die Zustandssumme des Systems

$$Z = \sum_{N_A} \binom{N}{N_A} e^{\beta N_A (\Delta G_A + \mu_A)} = \left[ 1 + e^{\beta (\Delta G_A + \mu_A)} \right]^N \quad (6.37)$$

mit  $\beta = 1/RT$  und  $(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k$ . Der Term in den eckigen Klammern entspricht dabei gerade der Zustandssumme einer einzelnen Bindungsstelle. Die durchschnittliche Anzahl  $\langle N_A \rangle$  besetzter Bindungsstellen durch Spezies  $A$  ergibt sich zu:

$$\langle N_A \rangle = \frac{\partial \ln Z}{\partial (\beta \mu_A)} = N \cdot \frac{e^{\beta (\Delta G_A + \mu_A)}}{1 + e^{\beta (\Delta G_A + \mu_A)}} = N \cdot \frac{K_A \cdot [A]}{1 + K_A \cdot [A]} \quad (6.38)$$

$[A] = e^{\beta \mu_A}$  ist die Konzentration der Spezies  $A$  in Lösung und  $K_A = e^{\beta \Delta G_A}$  die Bindungskonstante. Gl. (6.38) ist abgesehen von dem Vorfaktor  $N$  identisch mit der Langmuir-Isotherme Gl. 2.3 aus Abschnitt 2.3.1 (auch hier wird angenommen, dass die Konzentration der Spezies  $A$  in Lösung so hoch ist, dass eine Reduktion der Moleküle in Lösung durch die Hybridisierungsreaktion vernachlässigbar ist). Hat Molekül  $A$  eine Verteilung von Bindungskonfigurationen, so kann die Bindungskonstante  $K_A$  analog zu Abschnitt 6.1.1 über die Zustandssumme  $Z_A$  der verschiedenen Bindungskonfigurationen berechnet werden:

$$K_A = Z_A \quad (6.39)$$

Im Folgenden soll die Zustandssumme und damit die Bindungskonstante eines *high affinity targets* (HAT) und eines *low affinity targets* (LAT) berechnet werden. Exemplarisch wird hier das System „PM vs. MM1“ modelliert, so dass der HAT den PM und der LAT den MM1 repräsentiert (MM1 hat eine Fehlstelle an Position 8 im Vergleich zum PM, siehe Tab. 5.1). Zur Vereinfachung werden die beiden Targets als Homopolymere angenommen. Die Zustandssumme lautet:

$$Z_t = \sum_{\vec{v}} \sum_{k(\vec{v})} e^{\beta \Delta G_{\vec{v},k}} \quad (6.40)$$

$t$  steht dabei entweder für HAT oder LAT,  $\Delta G_{\vec{v},k}$  wurde in Gl. (6.36) definiert. Die erste Summe von Gl. (6.40) läuft über alle *pre-melted*-Konfigurationen  $\vec{v}$  des Duplexes.

**Tabelle 6.1:** Bindungskonstanten für HAT und LAT nach Gl. (6.40) und ihr Verhältnis bei gegebenem Parametersatz und einer Temperatur von  $T = 317$  K.

$K_{HAT}$	$K_{LAT}$	$K_{HAT}/K_{LAT}$	$j$ [kcal/mol]	$J$ [kcal/mol]	$\Delta h$ [kcal/mol]
$3.1995 \cdot 10^7$	$1.8799 \cdot 10^6$	17.02	0.3	1.7	0.7

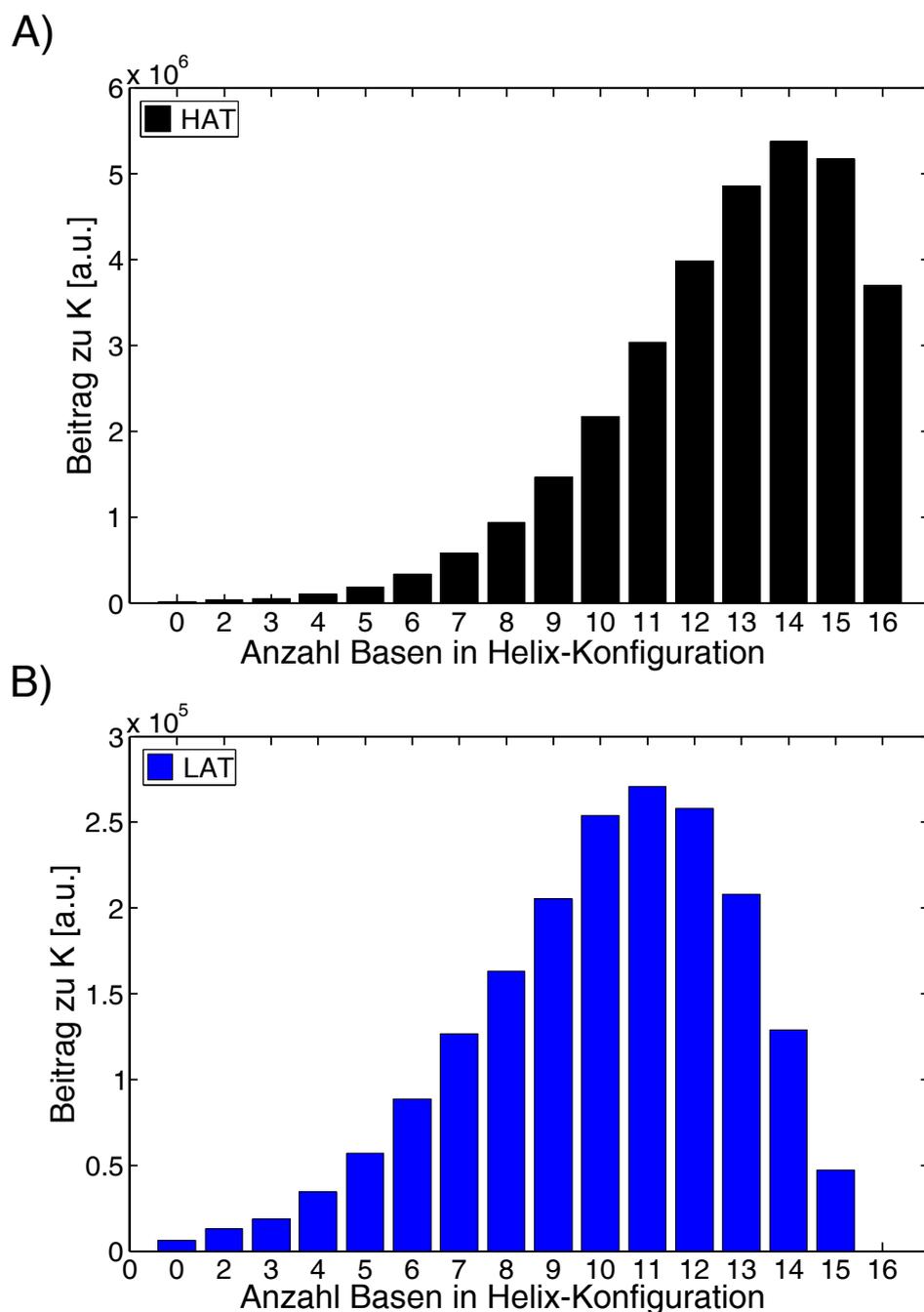
Im Falle des HAT sind das alle  $2^{M+1}$  Kombinationen von gebundenen und ungebunden Basen in einem Strang der Länge  $M+1$ . Im Falle des LAT läuft die Summe nur über  $2^M$  *pre-melted*-Konfigurationen, da sich an Position 8 eine Fehlstelle befindet und die entsprechende Base nie geschlossen sein kann. Die zweite Summe in obiger Gleichung läuft über alle Helix-Konfigurationen  $k$ , die zu einer bestimmten *pre-melted*-Konfiguration  $\vec{v}$  gehören.

Die Werte der Parameter  $j$ ,  $J$  und  $\Delta h$ , die für die Berechnung von  $\Delta G_{\vec{v},k}$  nötig sind, sind nicht bekannt. Sie werden daher als freie Parameter betrachtet, müssen jedoch qualitativ mit den Energien aus dem NN-Modell [53] vergleichbar sein. Die Parameter werden so eingestellt, dass das Verhältnis  $K_{HAT}/K_{LAT}$  ungefähr dem Verhältnis  $K_{PM}/K_{MM1}$  aus den Einzelhybridisierungsexperimenten entspricht ( $K_{PM}/K_{MM1} \approx 15$ , siehe Tab. 5.1). Ein möglicher Parametersatz ist in Tabelle 6.1 gezeigt.

In Abb. 6.9 ist der Anteil der helikalen Konfigurationen an der jeweiligen Bindungsaffinität des HAT bzw. des LAT als Funktion der Anzahl der Basen in der Helix-Konfiguration gezeigt. Ein Helix-Bereich muss aus mindestens zwei Basen bestehen. Der Beitrag zur Bindungsaffinität der *pre-melted*-Konfigurationen ohne Helixausbildung ist mit null Basen gekennzeichnet. Werden alle Beiträge summiert, so ergibt sich wieder die entsprechende Bindungskonstante des HAT bzw. des LAT aus Tab. 6.1. Die Simulation zeigt, dass unter gleichen Bedingungen die Bindungsaffinität des LAT hauptsächlich durch viele schwache Bindungszustände mit hoher Entropie generiert wird (siehe Abb. 6.9 B)), wohingegen der größte Anteil der Bindungsaffinität des HAT von den fast vollständig ausgeprägten Helix-Konfigurationen stammt (siehe Abb. 6.9 A)). Dies legt nahe, dass die Fluktuationsdynamiken von HAT und LAT bei Bindung an die Probe bei der gewählten Temperatur unterschiedlich sind, was experimentell durch die FCS-Messungen in Abschnitt 5.3.3 bestätigt wurde. Die unterschiedlichen thermischen Fluktuationen sind eine der Voraussetzung für das Auftreten der hochspezifischen Hybridisierung wie in Abschnitt 6.2.5 erläutert wird.

## 6.2.4 Konkurrenzhybridisierung ohne Wechselwirkung: standardspezifische Systeme

Im Folgenden wird ein Konkurrenzsystem aus  $N_A$  Molekülen der Spezies  $A$ ,  $N_B$  Mole-



**Abbildung 6.9:** Beitrag der Helix-Konfigurationen zu den Bindungskonstanten des HAT und des LAT (der Maßstab beider Abbildungen ist identisch). Für die Berechnungen werden beide Moleküle als Homopolymere betrachtet. Ein Helix-Bereich besteht aus mindestens 2 Basen. Der Beitrag zur Bindungskonstante der *pre-melted*-Konfigurationen ohne Helix-Ausbildung ist mit null Basen gekennzeichnet. Summiert man über alle Beiträge, so erhält man wieder die Bindungskonstanten aus Tab. 6.1. Die verwendeten Parameter sind:  $j = 0.3$  kcal/mol,  $J = 1.7$  kcal/mol,  $\Delta h = 0.7$  kcal/mol,  $T = 317$  K. **A)** Beitrag der Helix-Konfigurationen zur Bindungskonstanten des HAT. **B)** Beitrag der Helix-Konfigurationen zur Bindungskonstanten des LAT. Während der LAT seine Bindungsaffinität durch viele schwächer gebundene Zustände generiert, stammen die Hauptbeiträge zur Bindungsaffinität des HAT von den Zuständen, die fast vollständig eine Helix ausgeprägt haben.

külen der Spezies  $B$  und  $N$  Bindungsstellen ( $N_A + N_B < N$ ) in einem großkanonischen Ensemble ohne Wechselwirkung ( $\Delta p = 0$ , siehe Abschnitt 6.2.5) zwischen den Spezies betrachtet. Es gibt außerdem keine Wechselwirkung zwischen benachbarten Bindungsstellen und maximal ein Molekül pro Bindungsstelle. Bindet ein Molekül  $A$ , so wird die Energie  $\Delta G_A + \mu_A$  frei, bei Bindung eines Moleküls  $B$  die Energie  $\Delta G_B + \mu_B$ . Die großkanonische Zustandssumme dieses Systems ist:

$$\begin{aligned}
 Z &= \sum_{N_A} \sum_{N_B} \binom{N}{N_A} \binom{N - N_A}{N_B} e^{\beta[N_A(\Delta G_A + \mu_A) + N_B(\Delta G_B + \mu_B)]} \\
 &= \sum_{N_A} \binom{N}{N_A} e^{\beta N_A(\Delta G_A + \mu_A)} \cdot \sum_{N_B} \binom{N - N_A}{N_B} e^{\beta N_B(\Delta G_B + \mu_B)} \\
 &= \sum_{N_A} \binom{N}{N_A} e^{\beta N_A(\Delta G_A + \mu_A)} \cdot [1 + e^{\beta(\Delta G_B + \mu_B)}]^{N - N_A} \\
 &= [1 + e^{\beta(\Delta G_A + \mu_A)} + e^{\beta(\Delta G_B + \mu_B)}]^N
 \end{aligned} \tag{6.41}$$

Der Term in eckigen Klammern in der letzten Zeile entspricht der Zustandssumme einer einzelnen Bindungsstelle. Analog zu Gl. (6.38) ist die durchschnittliche Anzahl  $\langle N_A \rangle$  besetzter Bindungsstellen durch Molekül  $A$  gegeben durch:

$$\langle N_A \rangle = N \cdot \frac{e^{\beta(\Delta G_A + \mu_A)}}{1 + e^{\beta(\Delta G_A + \mu_A)} + e^{\beta(\Delta G_B + \mu_B)}} \tag{6.42}$$

$\langle N_B \rangle$  wird in analoger Weise berechnet. Damit ergibt sich das Verhältnis  $\langle N_A \rangle / \langle N_B \rangle$  zu:

$$\frac{\langle N_A \rangle}{\langle N_B \rangle} = \frac{e^{\beta(\Delta G_A + \mu_A)}}{e^{\beta(\Delta G_B + \mu_B)}} = \frac{[A] \cdot K_A}{[B] \cdot K_B} = \frac{[A]}{[B]} \cdot e^{\beta \Delta \Delta G} \tag{6.43}$$

$K_A$  und  $K_B$  sind die Gleichgewichtsbindungskonstanten aus der Einzelhybridisierung und  $\Delta \Delta G = \Delta G_A - \Delta G_B$  der Unterschied zwischen den Bindungsenergien der Moleküle. Gl. (6.43) besagt, dass das Verhältnis  $\langle N_A \rangle / \langle N_B \rangle$  nur vom Unterschied der Bindungsenergien und der Konzentration der zwei konkurrierenden Moleküle abhängt. Liegen keine (oder vernachlässigbare) Wechselwirkungen zwischen den Spezies vor, so beschreiben also die Bindungskonstanten der Einzelhybridisierung den Konkurrenzfall vollständig. Dies entspricht dem Verhalten der standardspezifischen Systeme (z.B. „PM vs. MM2“), in denen die Oberflächenbedeckungen der beiden Target-Spezies Gl. (6.42) folgen (siehe Abschnitt 5.3.1). In hochspezifischen Systemen kann Gl. (6.43) bei Verwendung der Bindungskonstanten aus den Einzelhybridisierungen das Verhältnis der Oberflächenbedeckungen  $\langle N_A \rangle / \langle N_B \rangle$  wie im nächsten Abschnitt gezeigt nicht beschreiben.

## 6.2.5 Konkurrenzhybridisierung mit Wechselwirkung: hochspezifische Systeme

Wie in [96, 97] experimentell beobachtet, ist die Spezifität der DNA-Konkurrenzhybridisierung dann besonders stark ausgeprägt, wenn für die Schmelztemperaturen  $T_m$  der verschiedenen Spezies gilt:

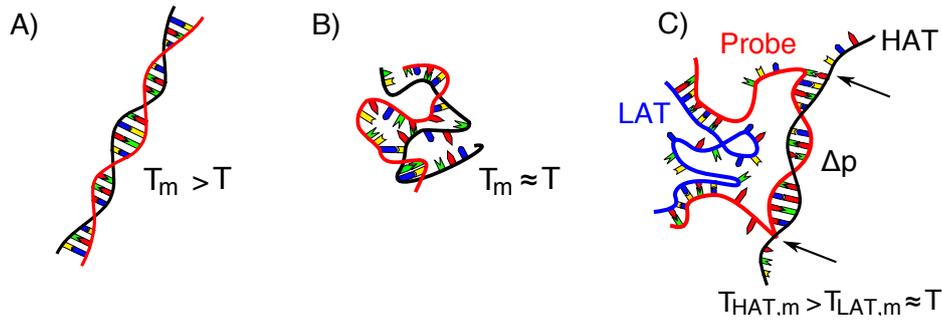
$$T_{HAT,m} > T_{LAT,m} \approx T \quad (6.44)$$

$T_{HAT,m}$  ist die Schmelztemperatur des HAT,  $T_{LAT,m}$  die des LAT und  $T$  die Hybridisierungstemperatur. Wie im Abschnitt 5.4 erläutert, hybridisiert der LAT in diesem Szenario mit erhöhter Wahrscheinlichkeit in einer *pre-melted*-Konfiguration (Abb. 6.10 B)), wohingegen der HAT eine Helix-Konfiguration besetzt (Abb. 6.10 A)). Dies wird durch die unterschiedlichen Fluktuationen von HAT und LAT (siehe FCS-Messungen in Abschnitt 5.3.3) und die theoretischen Ergebnisse aus Abschnitt 6.2.3 (Abb. 6.9) bestätigt.

Wie in Abschnitt 5.4 und 6.2.2 dargestellt, vollzieht sich die Wechselwirkung zwischen den verschiedenen Spezies in hochspezifischen Systemen im einfachsten Fall über kurzlebige Triplex-Konfigurationen - bestehend aus LAT, HAT und Probe -, wodurch die effektive freie Energielandschaft zugunsten eines Konkurrenten geändert werden kann. In einem Triplex wird die Ausbildung einer helikalen Konfiguration zwischen Probe und einer Spezies mit einer Entropiebarriere  $\Delta p > 0$  bestraft, da die hohe Persistenzlänge des Helix-Bereiches die entropischen Fluktuationen einschränkt und sich die ebenfalls gebundene Konkurrenzspezies seiner Deformation und der Deformation der Probe widersetzt (Abb. 6.10 C)).  $\Delta p$  hat direkten Einfluss auf die Anzahl der möglichen Konfigurationen  $\#_{block,i}$  von Block  $i$  in Gl. (6.31) und damit auf die Zustandssumme bzw. die Bindungskonstante. Physikalisch bedeutet eine Erhöhung von  $\Delta p$  eine Erhöhung der Kollektivität der helikalen Bindungskonfigurationen, da zur erfolgreichen Bindung zusätzlich die Barriere  $\Delta p$  überwunden werden muss und somit mehr Basen gleichzeitig in eine Helix-Konfiguration übergehen müssen. In Analogie zu Gl. (6.43) ergibt sich für  $\Delta p > 0$  das Verhältnis der durchschnittlich besetzten Bindungsstellen zu:

$$\frac{\langle N_{HAT} \rangle}{\langle N_{LAT} \rangle} = \frac{K_{HAT,C} \cdot [HAT]}{K_{LAT,C} \cdot [LAT]} \quad (6.45)$$

$K_C$  ist die durch den Einfluss von  $\Delta p$  veränderte Bindungskonstante im Falle wechselwirkender Konkurrenten. Berechnet man  $K_C$  als Funktion von  $\Delta p$  für den gleichen Parametersatz wie in Tabelle 6.1, so ergibt sich eine deutliche Abnahme von  $K_{LAT,C}$  im Vergleich zu  $K_{HAT,C}$  ( $K_{HAT,C} \approx K_{HAT}$ ,  $K_{LAT,C} \ll K_{LAT}$ , siehe Tabelle 6.2 und Abb. 6.11 A)) und damit eine Erhöhung der Differenz der freien Energien



**Abbildung 6.10:** Bindungskonfigurationen der DNA. **A)** Ist die Schmelztemperatur  $T_m$  eines DNA-Duplexes größer als die Hybridisierungstemperatur  $T$  ( $T_m > T$ ), so liegt der Duplex in einer geschlossenen doppelhelikalen Struktur vor, die sich durch eine große Persistenzlänge und wenige entropische Freiheitsgrade auszeichnet. **B)** Ist die Hybridisierungstemperatur vergleichbar mit der Schmelztemperatur des Duplexes,  $T_m \approx T$ , so dominiert die *pre-melted*-Konfiguration. In dieser Konfiguration ist der Duplex teilweise denaturiert und die Bindungsaffinität des Moleküls wird durch viele schwach gebundene Zustände mit niedriger Persistenzlänge generiert. **C)** In einer Konkurrenzsituation zwischen HAT und LAT können die Moleküle gleichzeitig an die Probe binden und eine unvollständige Triplex-Struktur generieren. Die Ausbildung einer Helix zwischen Probe und einem Target ist in solchen Situation erschwert, da sich der jeweilige Konkurrent seiner Deformation und der Deformation der Probe widersetzt. In einer *mean field*-Betrachtung kann dies durch die Entropiebarriere  $\Delta p$  beschrieben werden, die sich sowohl auf HAT und LAT auswirkt. Physikalisch bedeutet die Einführung von  $\Delta p$  eine zur erfolgreichen Ausbildung einer Helix-Konfiguration erhöhte Kooperativität. Hybridisiert der HAT in Konfiguration A) und der LAT in B), so kann der LAT im Gegensatz zum HAT die erhöhte Kollektivität nicht kompensieren. Der HAT hat damit eine höhere Wahrscheinlichkeit in einen Duplex bestehend aus Probe und LAT zu binden als der LAT. Dies führt zu einer asymmetrischen Änderung der effektiven freien Energielandschaft zugunsten des HAT.

$\Delta\Delta G = \Delta G_{HAT,C} - \Delta G_{LAT,C}$  (siehe Abb. 6.11 B)) schon bei einer geringen Erhöhung der Entropiebarriere  $\Delta p$ . Da dem HAT im Vergleich zum LAT hochkollektive Helix-Konfigurationen über der Energiebarriere  $\Delta p$  zur Verfügung stehen, kann die Erhöhung von  $\Delta p$  vom HAT kompensiert werden. Der LAT hat aufgrund seiner Fehlstelle und seiner unterschiedlichen Fluktuationsdynamik keinen Zugang zu diesen kollektiven Konfigurationen. Dies bedeutet auch, dass der HAT mit höherer Wahrscheinlichkeit in einen Duplex bestehend aus Probe und LAT binden kann als ein weiterer LAT. Andererseits ist die erfolgreiche Bindung eines LAT in einen Triplex bestehend aus Probe und HAT sehr unwahrscheinlich. Daraus ergibt sich eine asymmetrische Besetzung der Triplex-Konfiguration, was die effektive freie Energielandschaft zugunsten des HAT ändert. Es soll angemerkt werden, dass sich die Energie der einzelnen Paare Probe-LAT bzw. Probe-HAT natürlich nicht verschoben hat.

**Tabelle 6.2:** Ausgehend vom Parametersatz in Tabelle 6.1 werden die Bindungskonstanten des HAT und des LAT als Funktion von  $\Delta p$  bei einer Temperatur von  $T = 317$  K berechnet ( $j = 0.3$  kcal/mol,  $J = 1.7$  kcal/mol,  $\Delta h = 0.7$  kcal/mol). Schon bei einer geringen Erhöhung von  $\Delta p$  ändert sich das Verhältnis  $K_{HAT,C}/K_{LAT,C}$  zugunsten des HAT deutlich. Dies erklärt die unerwartet hohe Spezifität der Konkurrenzhybridisierung. Für  $\Delta p = 0$  entsprechen die Werte von  $K_{HAT,C}/K_{LAT,C}$  denen der Einzelhybridisierung für  $K_{HAT}/K_{LAT}$ .

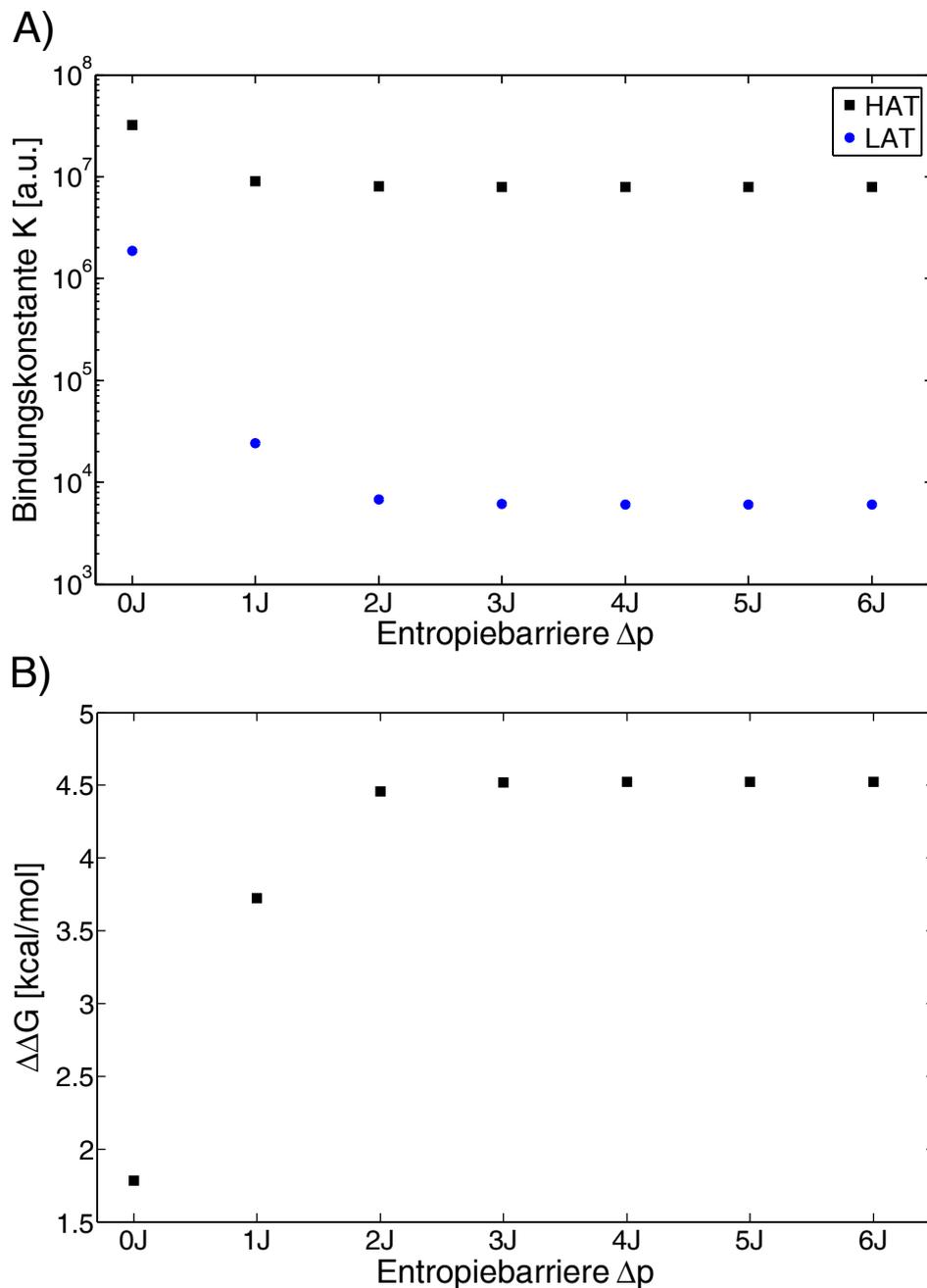
$\Delta p$	$K_{HAT,C}$	$K_{LAT,C}$	$K_{HAT,C}/K_{LAT,C}$
0	$3.1995 \cdot 10^7$	$1.8799 \cdot 10^6$	17.02
1J	$9.0110 \cdot 10^6$	$2.4337 \cdot 10^4$	370.26
2J	$8.0419 \cdot 10^6$	$6.8103 \cdot 10^3$	1180.84
3J	$7.9787 \cdot 10^6$	$6.0953 \cdot 10^3$	1308.99
4J	$7.9745 \cdot 10^6$	$6.0558 \cdot 10^3$	1316.84
5J	$7.9742 \cdot 10^6$	$6.0538 \cdot 10^3$	1317.22
6J	$7.9742 \cdot 10^6$	$6.0537 \cdot 10^3$	1317.24

## 6.2.6 Helmholtz Energie der Oberfläche

Mit diesen Ergebnissen soll nun die gesamte Helmholtzsche freie Energie  $F$  eines Systems bestehend aus  $N$  Bindungsstellen und zwei Teilchenspezies HAT und LAT für alle möglichen Kombinationen der Besetzungen der Bindungsstellen berechnet werden. Jede Bindungsstelle kann entweder leer oder von einer Teilchenspezies besetzt sein. Um der Triplex-Konfiguration Rechnung zu tragen, wird außerdem die Besetzung der Bindungsstellen mit zwei unterschiedlichen Teilchen zugelassen. Zwei Teilchen der selben Spezies können aber nicht gleichzeitig eine Bindungsstelle besetzen. Mit der hier verwendeten Vorzeichenkonvention, ist der Zustand des Systems am wahrscheinlichsten, der  $F$  maximiert. Für die effektive freie Energie  $F$  des Systems gilt:

$$\begin{aligned}
\frac{F}{N} &= U - R \cdot T \cdot S \\
&= c_{HAT} \cdot (\epsilon_{HAT}^0 + \mu_{HAT}) - c_T \cdot (\epsilon_{HAT}^0 + \mu_{HAT} - \epsilon_{HAT}^T - \mu_{HAT}) \\
&+ c_{LAT} \cdot (\epsilon_{LAT}^0 + \mu_{LAT}) - c_T \cdot (\epsilon_{LAT}^0 + \mu_{LAT} - \epsilon_{LAT}^T - \mu_{LAT}) - R \cdot T \cdot S \quad (6.46) \\
&= c_{HAT} \cdot (\epsilon_{HAT}^0 + \mu_{HAT}) - c_T \cdot \tilde{\epsilon}_{HAT} \\
&+ c_{LAT} \cdot (\epsilon_{LAT}^0 + \mu_{LAT}) - c_T \cdot \tilde{\epsilon}_{LAT} - R \cdot T \cdot S
\end{aligned}$$

$R$  ist die ideale Gaskonstante,  $T$  die Temperatur,  $S$  die Entropie und  $\tilde{\epsilon}_{HAT} = \epsilon_{HAT}^0 - \epsilon_{HAT}^T$ .  $\epsilon_{HAT}^0 = R \cdot T \cdot \ln(K_{HAT})$  entspricht der Energie bei Bindung eines Teilchens der HAT-Spezies an eine Bindungsstelle, wobei ein Teilchen der LAT-Spezies nicht an diese Bindungsstelle gebunden ist.  $\epsilon_{HAT}^T = R \cdot T \cdot \ln(K_{HAT}^T)$  ist die Energie bei



**Abbildung 6.11:** Einfluss der Entropiebarriere  $\Delta p$  auf das Bindungsverhalten von HAT und LAT. **A)+B)** Die Bindungskonstanten von HAT und LAT (A)) bzw. die Differenz der freien Energie  $\Delta\Delta G = \Delta G_{HAT,C} - \Delta G_{LAT,C}$  (B)) werden als Funktion der Entropiebarriere  $\Delta p$  dargestellt. Eine geringe Erhöhung von  $\Delta p$  reicht aus um das Verhältnis der Bindungskonstanten zugunsten des HAT zu verschieben.

Bindung eines Teilchens der HAT-Spezies in einer Triplex-Konfiguration (LAT-Spezies gleichzeitig gebunden),  $\mu_{HAT}$  das chemische Potential der Spezies,  $c_{HAT}$  der Anteil an Bindungsstellen, die von der HAT-Spezies besetzt sind ( $0 \leq c_{HAT} \leq 1$ ).  $c_T$  ist der Anteil von  $c_{HAT}$  und  $c_{LAT}$ , die als Triplex vorliegen. Der Term  $c_T \cdot \tilde{\epsilon}_{HAT}$  korrigiert den Term  $c_{HAT} \cdot (\epsilon_{HAT}^0 + \mu_{HAT})$  um die Anzahl der gebundenen HAT-Teilchen im Triplex, die ohne Korrektur einen zu hohen, fehlerhaften Energiebeitrag für die Rechnung liefern würden. Entsprechende Bezeichnungen gelten für ein Teilchen der LAT-Spezies. Für die Entropie  $S$  gilt:

$$S = S_0 - \ln \left[ \binom{N}{N \cdot c_{HAT}} \cdot \binom{N}{N \cdot c_{LAT}} \right] \quad (6.47)$$

$S_0$  ist dabei eine Konstante, deren Absolutwert für die weiteren Berechnungen nicht entscheidend ist. Gl. 6.46 lässt sich umschreiben zu:

$$\begin{aligned} \frac{F}{N} &= c_{HAT} \cdot (\epsilon_{HAT}^0 + \mu_{HAT}) - c_{LAT} \cdot \kappa_{HAT}^T \cdot \tilde{\epsilon}_{HAT} \\ &+ c_{LAT} \cdot (\epsilon_{LAT}^0 + \mu_{LAT}) \cdot \left( 1 - \frac{\kappa_{HAT}^T \cdot \tilde{\epsilon}_{LAT}}{\epsilon_{LAT}^0 + \mu_{LAT}} \right) - R \cdot T \cdot S \end{aligned} \quad (6.48)$$

Hier wurde ausgenutzt, dass für die Anzahl  $c_T$  der Triplexe gilt:

$$c_T = c_{HAT} \cdot \kappa_{LAT}^T = c_{LAT} \cdot \kappa_{HAT}^T \quad (6.49)$$

Dabei ist  $\kappa_{HAT}^T = \kappa_{HAT}^T(c_{HAT})$  die bedingte Wahrscheinlichkeit, dass ein Teilchen der HAT-Spezies im Triplex bindet, wenn bereits ein Teilchen der LAT-Spezies gebunden ist (entsprechendes gilt für  $\kappa_{LAT}^T$ ). In der Berechnung der Energie  $F$  ist  $\kappa_{HAT}^T$  ein freier Parameter (der umgekehrte Fall, in dem  $\kappa_{LAT}^T$  der freie Parameter ist, läuft analog). Für die weiteren Berechnungen wird angenommen, dass  $\kappa_{HAT}^T$  größer als Null ist. Da bei  $c_{HAT} = 0$  keine Triplexe vorkommen können, wird dadurch gleichzeitig eine minimale Konzentration  $c_{HAT,min}$  der HAT-Spezies definiert:

$$c_{HAT,min} = c_{LAT} \cdot \kappa_{HAT}^T \quad (6.50)$$

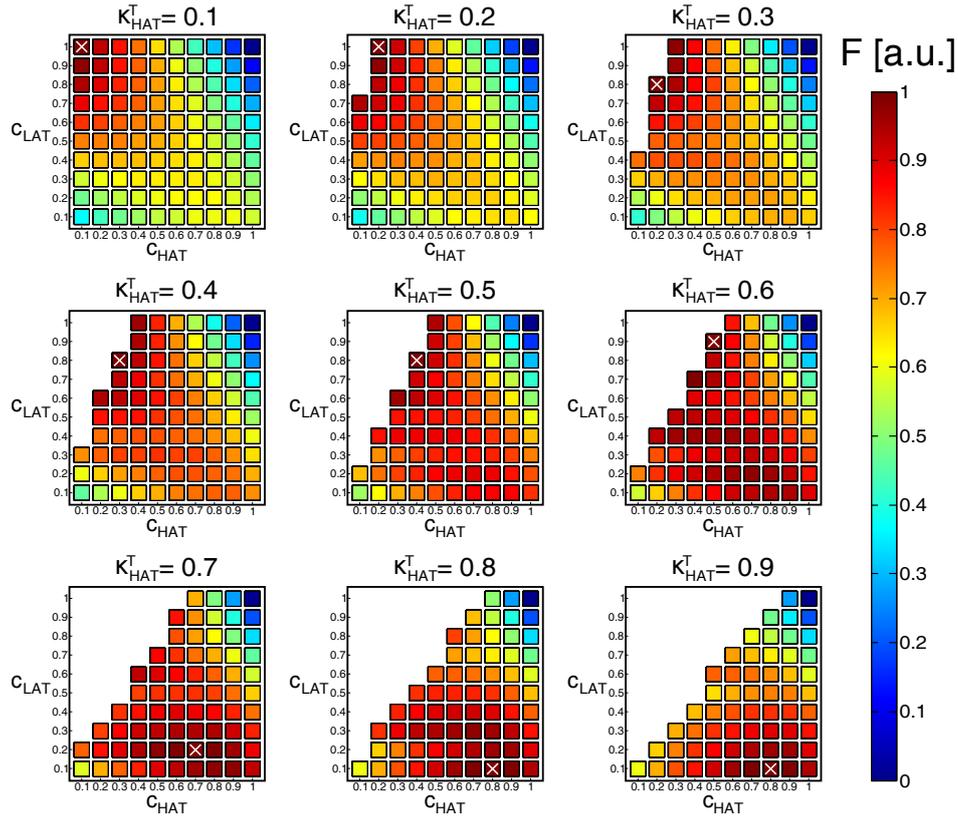
Außerdem muss für  $c_{HAT} = 1$  gelten:  $\kappa_{HAT}^T(c_{HAT} = 1) = 1$ , d.h. die Wahrscheinlichkeit, dass die HAT-Spezies in einem Triplex bei Anwesenheit der LAT-Spezies binden kann, muss 1 sein. Anders ist es bei  $c_{LAT} > 0$  nicht möglich  $c_{HAT} = 1$  zu erreichen. Nimmt man einen linearen Zusammenhang zwischen  $\kappa_{HAT}^T(c_{HAT})$  und  $c_{HAT}$  an, so ergibt sich mit Gl. 6.50:

$$\kappa_{HAT}^T = \frac{c_{HAT,min}}{c_{LAT}} + \frac{c_{LAT} - c_{HAT,min}}{c_{LAT}(1 - c_{HAT,min})} \cdot (c_{HAT} - c_{HAT,min}) \quad (6.51)$$

Im Folgenden soll Teilchensorte HAT der PM-Spezies mit einer Konzentration von 5 nM und Teilchensorte LAT der MM1-Spezies mit einer Konzentration von 1  $\mu$ M aus Kapitel 5 entsprechen (experimentelle Bedingungen im Konkurrenzsystem „PM vs. MM1“). Mit den experimentell bestimmten Bindungskonstanten aus Tab. 5.1 ergeben sich so die Energien bei Einzelbindung von HAT bzw. LAT und die jeweiligen chemischen Potentiale bei einer Temperatur von  $T = 317$  K zu:  $\epsilon_{HAT}^0 = 12.99$  kcal/mol,  $\mu_{HAT} = -12.04$  kcal/mol,  $\epsilon_{LAT}^0 = 11.28$  kcal/mol,  $\mu_{LAT} = -8.7$  kcal/mol.

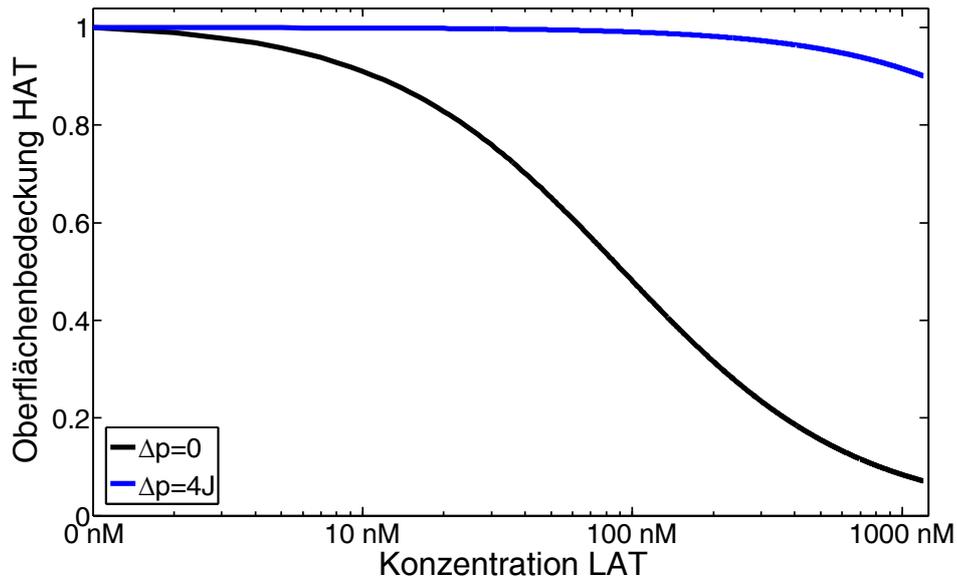
Nach Tab. 6.2 nimmt die Bindungskonstante des HAT bei  $\Delta p = 4J$  um einen Faktor 4.01, die Bindungskonstante des LAT um einen Faktor 310.43 im Vergleich zur Einzelbindung ab. Damit ergeben sich die entsprechenden Energien im Triplex zu  $\epsilon_{HAT}^T = R \cdot T \cdot \ln(K_{HAT}/4.01) = 12.11$  kcal/mol und  $\epsilon_{LAT}^T = R \cdot T \cdot \ln(K_{LAT}/310.43) = 7.67$  kcal/mol und damit  $\tilde{\epsilon}_{HAT} = 0.88$  kcal/mol bzw.  $\tilde{\epsilon}_{LAT} = 3.61$  kcal/mol.  $K_{HAT}$  bzw.  $K_{LAT}$  entsprechen hier den Bindungskonstanten des PM bzw. des MM1 aus Tab. 5.1. Die effektive freie Energielandschaft  $F$  ist in Abb. 6.12 für alle möglichen Kombinationen von  $c_{HAT}$  und  $c_{LAT}$  als Funktion von  $\kappa_{HAT}^T$  nach Normierung des Maximums auf 1 und des Minimums auf 0 dargestellt. In jeder Teilabbildung ist das Maximum der freien Energie mit einem weißem „X“ gekennzeichnet. Da bei  $\kappa_{HAT}^T > 0$  - wie oben beschrieben - ein minimale Konzentration  $c_{HAT,min}$  auf der Oberfläche vorgegeben wird, sind nicht alle Kombinationen von  $c_{HAT}$  und  $c_{LAT}$  möglich. Für niedrige Wahrscheinlichkeiten  $\kappa_{HAT}^T \leq 0.6$  liegt das Maximum der freien Energie bei großen  $c_{LAT}$  und niedrigen  $c_{HAT}$  (z.B. bei  $\kappa_{HAT}^T = 0.3$ :  $c_{LAT} = 0.8$ ,  $c_{HAT} = 0.2$ ). Wird  $\kappa_{HAT}^T$  von 0.6 auf 0.7 erhöht, ändert sich die Energielandschaft sprunghaft zugunsten der HAT-Spezies. Das Maximum liegt dann bei  $c_{LAT} = 0.2$ ,  $c_{HAT} = 0.7$ . Bei höheren Wahrscheinlichkeiten  $\kappa_{HAT}^T$  wird  $F$  bei  $c_{LAT} = 0.1$ ,  $c_{HAT} = 0.8$  maximiert: Das System bevorzugt trotz  $\epsilon_{LAT}^0 > \epsilon_{HAT}^0$ , was durch die wesentlich höhere LAT-Konzentration in Lösung verursacht wird, hohe HAT- und niedrige LAT-Konzentrationen auf der Oberfläche. Ohne dass Triplete in hohem Maße besetzt werden müssen (das Vorkommen von Triplexen wäre erst zwingend notwendig für  $c_{HAT} + c_{LAT} > 1$ ), ändert also bereits die zusätzlich mögliche Triplex-Bindungskonfiguration, bei der die Energie der HAT-Spezies wenig und die der LAT-Spezies stark im Vergleich zur Einzelbindung verändert wird, die Oberflächenbedeckung asymmetrisch zugunsten der HAT-Spezies. Genau dieses Verhalten wird in hochspezifischen Systemen experimentell beobachtet (siehe z.B. Abb. 5.1).

Unter der Annahme, dass die zusätzlich mögliche Triplex-Konfiguration die Energielandschaft des Systems so verändert, dass im Wesentlichen die Bindungsenergien  $\epsilon_{HAT}^T$  und  $\epsilon_{LAT}^T$  die Konkurrenzhybridisierung beschreiben, kann anhand Gl. (2.6) aus Abschnitt 2.3.1 bzw. Gl. (6.42) die HAT-Oberflächenbedeckung in Konkurrenz des LAT berechnet werden. Die Eigenschaften der HAT- und der LAT-Spezies entsprechen nach



**Abbildung 6.12:** Effektive freie Energielandschaft  $F$  als Funktion der Wahrscheinlichkeit  $\kappa_{HAT}^T$ . Minimum und Maximum werden auf 0 bzw. auf 1 normiert. Das Maximum der Energie ist für jeden Teilgraphen durch ein weißes „X“ markiert. Für  $\kappa_{HAT}^T \leq 0.6$  wird die Energie bei hohen LAT- und niedrigen HAT-Konzentrationen auf der Oberfläche maximiert. Die Vergrößerung von  $\kappa_{HAT}^T$  auf 0.7 führt zu einer sprunghaften Änderung der Energielandschaft zugunsten der HAT-Spezies. Das Maximum liegt dann bei  $c_{LAT} = 0.2$  und  $c_{HAT} = 0.7$ . Eine weitere Erhöhung verschiebt das Maximum zu  $c_{LAT} = 0.1$  und  $c_{HAT} = 0.8$ .

wie vor denen des PM und des MM1. Abb. 6.13 zeigt die HAT-Oberflächenbedeckung als Funktion der LAT-Konzentration für zwei Werte von  $\Delta p$ : Im Fall  $\Delta p = 0$  (schwarze Kurve) können die beiden Spezies mit unveränderter Energie in einer Triplex-Konfiguration binden, sodass hier die Bindungskonstanten aus der Einzelhybridisierung  $K_{HAT}$  und  $K_{LAT}$  nach Tab. 5.1 verwendet werden. Bei  $\Delta p = 4J$  (blaue Kurve) ändern sich die Bindungskonstanten zu  $K_{HAT,C} = K_{HAT}/4.01$  und  $K_{LAT,C} = K_{LAT}/310.43$ . Normiert wird jeweils auf die HAT-Oberflächenbedeckung bei einer LAT-Konzentration von 0 nM. Die HAT-Konzentration beträgt 5 nM. Im Falle  $\Delta p = 0$  kann der LAT mit zunehmender Konzentration den HAT leicht von der Oberfläche verdrängen. Bei  $\Delta p = 4J$  bleibt die Oberflächenbedeckung des HAT konstant. Dieses Verhalten wird auch experimentell z.B. im System „PM vs. MM1“ beobachtet (siehe Abschnitt 5.1).



**Abbildung 6.13:** HAT-Oberflächenbedeckung. Anhand Gl. (6.42) wird die HAT-Oberflächenbedeckung als Funktion der LAT-Konzentration für zwei Werte der Entropiebarriere  $\Delta p$  bestimmt. Die Normierung erfolgt jeweils auf den Wert der HAT-Oberflächenbedeckung bei 0 nM LAT-Konzentration. HAT entspricht der PM-Spezies, LAT der MM1-Spezies aus Kapitel 5. Bei  $\Delta p = 0$  (Verwendung der individuellen Bindungskonstanten aus der Einzelhybridisierung aus Tab. 5.1) wird der HAT bei Erhöhung der Konzentration des LAT leicht von der Oberfläche verdrängt. Im Falle  $\Delta p = 4J$  (als Beispiel eines hochspezifischen Systems, Verwendung der Bindungskonstanten  $K_{HAT,C} = K_{HAT}/4.01$  und  $K_{LAT,C} = K_{LAT}/310.43$ ) ändert sich die Oberflächenbedeckung des HAT im experimentell untersuchten LAT-Konzentrationsbereich ( $0 \leq [LAT] \leq 1 \mu\text{M}$ ) praktisch nicht.

#### Bemerkungen:

- i) Die Anzahl von Basenpaaren, die an der Ausbildung einer Helix-Konfiguration beteiligt sind, kann unterschiedlich hoch sein (siehe Anzahl der Helix-Konfigurationen  $\#_{block,i}$  eines individuellen Blocks  $i$  in Gl. (6.31)). Der Parameter  $\Delta p$  wird hier unabhängig von der Anzahl dieser Basenpaare als konstant angenommen. Bei einer kleinen bzw. großen Anzahl beteiligter Basen sollte  $\Delta p$  klein sein und ein Maximum bei einer mittleren Anzahl von Basen haben. Diese Längenabhängigkeit von  $\Delta p$  wurde nicht im Modell implementiert, da sie lediglich die Spezifität des HAT erhöhen würde.
- ii) In diesem Modell wurde die Entropiebarriere  $\Delta p$  nur auf Helix-Konfigurationen angewandt. Eine Verbesserung des Modells könnte darin bestehen, auch auf die *pre-melted*-Konfigurationen eine Entropiebarriere anzuwenden, solange sie Teil eines Triplexes sind. Da dadurch kollektivere *pre-melted*-Konfigurationen bevorzugt würden, hätte diese Entropiebarriere den gleichen Einfluss wie  $\Delta p$  und würde die Spezifität des HAT gegenüber dem LAT weiter erhöhen.

Beide oben genannten Verbesserungen für die Näherung von  $\Delta p$  zielen also auf eine weitere Erhöhung der Spezifität des HAT ab.

Zusammenfassend lässt sich sagen, dass sich das Verhalten hochspezifischer Systeme der DNA-Konkurrenzhybridisierungen mit Hilfe dieses Modells charakterisieren lässt. Dazu wurden zwei unterschiedliche kollektive Bindungszustände und Fluktuationsdynamiken der beteiligten DNA-Spezies benötigt. Obwohl die Wechselwirkung zwischen DNA-Spezies aufgrund thermischer Fluktuationen im Detail sehr schwer zu beschreiben ist, konnte mit dem Modell eine qualitative Übereinstimmung mit den Experimenten erzielt werden (die unterschiedlichen thermischen Fluktuationen wurden im Mittel mit Hilfe des Parameters  $\Delta p$  im Modell abgebildet). Aufwendige MD-Simulationen könnten weitere Einblicke in die komplexe Situation fluktuierender DNA-Moleküle bringen und dazu beitragen, ein detaillierteres Modell zu erarbeiten.

In Abschnitt 6.2.4 wurde die Wechselwirkung durch Triplex-Konfigurationen vernachlässigt. Zwar können auch in standardspezifischen Systemen Triplex-Konfigurationen vorliegen, jedoch würde hier gelten:  $\epsilon_{HAT}^0 \approx \epsilon_{LAT}^0$  und  $\epsilon_{HAT}^T \approx \epsilon_{LAT}^T$ . Das bedeutet, dass eine zusätzlich mögliche Triplex-Konfiguration die Energielandschaft nicht asymmetrisch zugunsten eines Konkurrenten verändert, da beide Spezies aufgrund ähnlicher Fluktuationsdynamik bei Bindung im Triplex in gleichem Maße Bindungsenergie verlieren. Dadurch beschreiben die Bindungskonstanten der Einzelhybridisierung auch die Konkurrenzhybridisierung vollständig (siehe Kapitel 5). Das Auftreten von Triplexen alleine reicht also nicht aus, um in einer Konkurrenzsituation das Gleichgewicht zugunsten einer Spezies zu verändern. Erst durch antagonistische Fluktuationsdynamiken wird die freie Energielandschaft zugunsten einer Spezies verändert. So kann die Spezifität der DNA-Hybridisierung um mehrere Größenordnungen verbessert werden. In einem hochspezifischen System kann also nicht von den Eigenschaften der Einzelhybridisierung auf die eines Konkurrenzsystems geschlossen werden: Das Konkurrenzsystem ist mehr als nur die Summe seiner Teile.

## 7 Zusammenfassung und Ausblick

Die im Rahmen dieser Arbeit durchgeführten Experimente und entwickelten Modelle sind alle im Kontext eines besseren Verständnisses der physikalischen Mechanismen der molekularen Erkennung im Falle der DNA-Hybridisierung zu sehen. Durch Kenntnis der zugrundeliegenden thermodynamischen Gesetzmäßigkeiten können außerdem die auf DNA-Hybridisierung basierenden biotechnologischen Anwendungen auf eine solide Grundlage gestellt und ihr volles Potential ausgeschöpft werden. Die in dieser Arbeit behandelten Aspekte der DNA-Hybridisierung lassen sich in zwei Bereiche unterteilen:

- Stabilität von DNA-Duplexen mit einzelsträngigen Schleifen-Strukturen sog. *Loops*
- DNA-Hybridisierung in Konkurrenz

Der Prozess der molekularen Erkennung wird durch die Ausbildung von Loops erheblich komplizierter, da so neben perfekt komplementären DNA-Strängen auch teilweise nicht-komplementäre Stränge stabile Bindungen eingehen können. In technischen Anwendungen spielen Loops eine wichtige Rolle, da z.B. die Längenverteilung der bei Microarray-Hybridisierungen eingesetzten (unbekannten) Targets aufgrund des Herstellungsprozesses oft nur schwer kontrollierbar ist und nicht unbedingt mit der Länge der immobilisierten Probes übereinstimmen muss. So können sich bei Hybridisierung nicht-kanonische Bindungskonformationen, wie beispielsweise Loops, ergeben, über deren Auswirkungen auf die Stabilität des DNA-Duplexes wenig bekannt ist. Um die Loop-Strukturen zu untersuchen, wurde eine zusätzliche Sequenz in die Ausgangssequenz der Probes auf der Microarray-Oberfläche eingefügt (Abschnitte 4.1 und 4.2). Wird eine solche Probe mit einem zur Ausgangssequenz der Probe komplementären Target-Molekül hybridisiert, bildet sich ein Loop aus. Durch Variation der Position und der Länge der zusätzlichen Sequenz können Position und Länge der Loops auf einfache Weise kontrolliert werden. Es zeigte sich dabei eine monotone Signalabnahme mit zunehmender Länge der bis zu 13 Basen langen eingefügten Sequenz. Auf die maximale Länge der zusätzlichen Sequenz, ab der kein Signal mehr messbar ist, wird weiter unten eingegangen. Bei der Abhängigkeit des Signals von der Loop-Position ergab sich ein symmetrisches „Loop-Defektprofil“ im Bezug zur Mitte der Probe: Das

Signal erreichte für mittige und endnahe Loop-Positionen ein Maximum, während es bei Loop-Positionen in einem Abstand von 3-4 Basen von der Mitte der Probe ein Minimum annahm. In weiterführenden Experimenten konnte weder ein Einfluss der Microarray-Oberfläche auf das Hybridisierungssignal noch ein Einfluss der Sequenz der zusätzlich eingefügten Basen auf die beobachtete Signalabhängigkeit von der Loop-Länge festgestellt werden. Die beobachtete hohe Stabilität aller DNA-Duplexe mit Loop zeigt, dass die Formation von Loops auch im thermodynamischen Gleichgewicht ein wichtiger Aspekt ist, der bei der Analyse der mit Microarrays gewonnenen Daten oder DNA-Hybridisierungen von komplexen Mischungen berücksichtigt werden muss. Weitere Experimente müssten zeigen, inwieweit der Salzgehalt der Puffer-Lösung oder die Hybridisierungstemperatur einen Einfluss auf die Stabilität der DNA-Duplexe mit Loop hat. Da die Persistenzlänge der DNA vom Salzgehalt der Pufferlösung abhängig ist, können bei Variation des Salzgehaltes nicht unbedingt vergleichbare Ergebnisse erwartet werden.

Zum besseren Verständnis der experimentellen Ergebnisse der Loop-Strukturen wurde im Rahmen dieser Arbeit auf Grundlage des in [1] vorgestellten *Zipper*-Modells zur Beschreibung von Einzelbasendefekten ein Gleichgewichtsmodell entwickelt (siehe Abschnitt 6.1). Dazu musste das Modell so erweitert werden, dass nun auch eine Öffnung des Duplexes an der Loop-Position - im Gegensatz zum *Zipper*-Modell - berücksichtigt werden konnte. Es wurde eine erfolgreiche Näherung des erweiterten Modells ausgearbeitet, die die Signalabhängigkeit von der Loop-Länge und -Position mit geringem Rechenaufwand reproduzieren konnte [89].

Da die Experimente zur Stabilität von DNA-Duplexen mit einzelsträngigen Loop-Strukturen gezeigt haben, dass selbst Duplexe mit langem Loop einen signifikanten Beitrag zum Hybridisierungssignal liefern, wurde untersucht, ab welcher Länge der zusätzlich eingefügten Basensequenz keine nennenswerte Hybridisierung mehr auftritt (Grenze der Hybridisierung, siehe Abschnitte 4.3 und 4.4). Aufgrund der Längenbeschränkung der immobilisierten Probes auf der Oberfläche durch die *in situ*-Synthese und damit der Länge der eingefügten Basensequenzen in die Probes, wurden zur Untersuchung der Grenze der Hybridisierung die zusätzlichen Basen in die Targets eingebettet, die keiner Längenbeschränkung unterliegen. Auf diese Weise wurden Duplexe mit bis zu 52 Basen langen Loops generiert, bei denen sich auch eine monotone Signalabnahme mit zunehmender Loop-Länge ergab. Die Extrapolation auf noch größere Loop-Längen erfolgte mit Hilfe eines *power law* der Form  $a \cdot x^{-b}$ , wodurch sich eine Loop-Länge von 79 Basen ergab, ab der keine nennenswerte Hybridisierung mehr zwischen Probe und Target auftritt. Da dies nur für die hier verwendeten Sequenzen gilt, müssten zukünftige Messungen untersuchen, inwieweit dies auch für Ausgangssequenzen unterschiedlicher Länge oder hohem CG-Gehalt gültig ist.

---

Die Hybridisierungssignale der Duplexe mit Loop im Target waren wesentlich stärker als die Signale der Duplexe mit Loop in der Probe. Als Ursache für die unterschiedlichen Signale konnten Synthesefehler, die während der *in situ*-Synthese unvermeidbar auftreten, ausgeschlossen werden. Bei der Untersuchung des Einflusses des Fluoreszenzfarbstoffes zeigte sich in Übereinstimmung mit [93], dass die Position der Markierung (am 3'- oder am 5'-Ende) einen Einfluss auf die Duplexstabilität hat: Die am 5'-Ende markierten Targets lieferten wesentlich stärkere Signale als ihre am 3'-Ende markierten Pendants, was nicht durch eine Wechselwirkung der Fluorophore mit der Oberfläche begründet werden konnte.

Im Abschnitt 4.5 wurde der Einfluss von zwei Loops in einer Probe auf das Hybridisierungsverhalten untersucht. Insbesondere galt hier das Interesse dem Wechselwirkungsabstand der beiden Loops. Durch einfache theoretische Überlegungen und experimentelle Daten (sowohl Microarray- als auch FCS-Messungen) konnte der Wechselwirkungsabstand der Loops auf 5-8 Basen bestimmt werden. Hier müssten weiterführende Experimente zeigen, inwieweit der Wechselwirkungsabstand von verschiedenen Parametern wie der Länge der beiden Einzel-Loops, dem Salzgehalt der Pufferlösung oder der Temperatur abhängt.

Der zweite Teil der Arbeit beschäftigt sich mit DNA-Hybridisierung in Konkurrenz (siehe Kapitel 5). In biologischen Systemen, in denen eine Vielzahl von Molekülen um Bindungsstellen konkurrieren, ist die korrekte molekulare Erkennung von passenden Bindungspartnern von größter Wichtigkeit. Nur so kann die Funktionalität des biologischen Systems, wie beispielsweise einer Zelle, gewährleistet sein. Molekulare Erkennung ist auch bei vielen biotechnologischen Anwendungen von großer Bedeutung. So konkurriert in vielen Microarray-Experimenten eine Vielzahl von Target-Spezies um Bindungsstellen auf der Oberfläche. Die Qualität der Microarray-Daten hängt auch davon ab, ob die Proben nur die zu ihnen komplementären Targets binden oder ob es auch zu unspezifischen Hybridisierungen kommt.

Es wurden verschiedene Konkurrenz-Systeme bestehend aus jeweils zwei Target-Spezies untersucht. In den Konkurrenzsystemen wurde die Oberflächenbedeckung der Spezies mit der größeren Bindungskonstante (HAT von engl. *high affinity target*) bei konstanter Konzentration als Funktion der Konzentration der Spezies mit der niedrigeren Bindungskonstante (LAT von engl. *low affinity target*) gemessen. Die Bindungskonstanten der verwendeten Target-Spezies wurden in Einzelhybridisierungsexperimenten in [96, 97] bestimmt. Zur Beschreibung der Oberflächenbedeckung von DNA wird standardmäßig das Langmuir-Modell verwendet. Danach hängen die Oberflächenbedeckungen der Spezies im thermodynamischen Gleichgewicht nur von dem Verhältnis der Bindungskonstanten aus der Einzelhybridisierung und der jeweiligen Konzentrationen ab. Die Untersuchung der Konkurrenz-Systeme ergab, dass eine Unterschei-

derung in standard- und hochspezifische Systeme getroffen werden muss (siehe Abschnitte 5.1 und 5.3). Im Falle der standardspezifischen Systeme ist eine Beschreibung der Oberflächenbedeckungen der beiden Spezies durch das Langmuir-Modell erfolgreich. Die HAT-Hybridisierung in hochspezifischen Systemen ist hingegen unerwartet spezifisch, da eine im Vergleich zur HAT-Konzentration um ein Vielfaches höhere LAT-Konzentration trotz vergleichbarer Bindungskonstanten keine Auswirkung auf die HAT-Oberflächenbedeckung hat.

In Zusammenarbeit mit Marc Schenkelberger konnten Artefakte in den Messungen als Ursache für die hochspezifische Hybridisierung durch weitere Experimente ausgeschlossen werden (siehe [96, 97], Abschnitt 5.2): Messungen mit TIRF (von engl. *total internal reflection fluorescence*) bestätigten den Effekt der hochspezifischen HAT-Hybridisierung auf einer Oberfläche. Zudem konnte zum einen durch Messungen in Lösung gezeigt werden, dass die hochspezifische Hybridisierung kein Oberflächeneffekt ist und zum anderen nicht von der Wahl des Sequenzmotivs der eingesetzten Targets abhängt. Darüber hinaus wurde verifiziert, dass die hochspezifische Hybridisierung im thermodynamischen Gleichgewicht auftritt.

Zur Charakterisierung der hochspezifischen Hybridisierung waren weitere Experimente notwendig (siehe Abschnitt 5.3). In Zusammenarbeit mit Marc Schenkelberger konnte gezeigt werden, dass hochspezifische Konkurrenzsysteme dann auftreten, wenn sich die Schmelztemperaturen der beiden Target-Spezies um mehr als 10 °C unterscheiden (siehe [96, 97]). Die Abhängigkeit der Schmelztemperatur von der Präsenz des Konkurrenten, die im Falle von hochspezifischen aber nicht bei standardspezifischen Systemen beobachtet werden konnte, deutete auf eine Wechselwirkung zwischen den beiden Target-Spezies hin.

Desweiteren legten die Messungen der Relaxationskonstanten der PM- und der MM1-Spezies in Einzel- und Konkurrenzhybridisierungen mittels FCS zusammen mit den aus den Schmelzkurven extrahierten Entropieänderungen bei PM- bzw. MM1-Hybridisierung nahe, dass ein Duplex bestehend aus Probe und MM1 deutlich flexibler ist als ein Duplex aus Probe und PM.

Basierend auf den experimentellen Ergebnissen wurde in Abschnitt 5.4 ein Erklärungsansatz für den Effekt der hochspezifischen Hybridisierung gegeben: Die Änderung der Schmelztemperatur und die Modifikation der Korrelationsfunktion des LAT in Präsenz des HAT deuten auf eine Wechselwirkung der beiden Konkurrenten hin. Eine einfache Möglichkeit der Wechselwirkung ist die Bildung eines Triplex-Zustandes, in dem HAT und LAT gleichzeitig an die Probe gebunden sind. Um die Auswirkungen dieser hypothetischen Triplex-Konfiguration auf die Energielandschaft zu untersuchen, wurde in Abschnitt 6.2 ein Gleichgewichtsmodell präsentiert, in dem ein DNA-Duplex in Mischzuständen aus zwei unterschiedlichen Bindungskonfigurationen vorliegen kann: einer

---

*pre-melted*- und einer Helix-Konfiguration, die sich durch die Anzahl ihrer entropischen Freiheitsgrade und ihrer Persistenzlänge unterscheiden. Die numerische Berechnung ergab, dass der HAT mit einer deutlich höheren Bindungskonstante im Triplex binden kann als der LAT. Die Analyse der effektiven freien Energielandschaft zeigte, dass das System dadurch eine hohe HAT-Oberflächenbeckung trotz der um ein Vielfaches höheren LAT-Konzentration bevorzugt: Durch Einführen einer Triplex-Konfiguration wird die Energielandschaft drastisch zugunsten des HAT verschoben. Der Effekt der hochspezifischen Hybridisierung zeigt, dass ein DNA-Konkurrenzsyste.ms nicht immer als Summe der Einzelsysteme behandelt werden kann und der Langmuir-Ansatz immer kritisch geprüft werden muss.

Zum Abschluss wurde ein Konkurrenzsystem bestehend aus 49 verschiedenen Target-Spezies untersucht, wobei für jede Target-Spezies komplementäre Probes auf der Oberfläche vorhanden waren (siehe Abschnitt 5.5). Es wurde untersucht, ob sich das 49-komponentige System mit steigender Temperatur selbstständig ordnen kann, d.h. ob die Probes auf der Oberfläche nur die zu ihnen komplementären Targets binden. Bei niedrigen Temperaturen war die Bindung der Targets an die Probes noch relativ unspezifisch, jedoch ordnete sich das System bei Temperaturerhöhung. Eine erneute Temperatursenkung resultierte wieder in einem ungeordneteren System. Unerwarteterweise war dabei eine „Hysterese“ messbar: Bei identischen Temperaturen ergaben sich unterschiedliche Hybridisierungssignale. Da dieser Effekt bei Systemen mit nur 1-2 Target-Spezies nicht beobachtet wurde, deutet dies darauf hin, dass die Wechselwirkung der verschiedenen Target-Spezies das Erreichen des thermodynamischen Gleichgewichts verhindert bzw. verzögert. Hier bedarf es weiterer Experimente, um das Auftreten dieser Hysterese zu verifizieren. Sollten sich die Ergebnisse bestätigen, hätte dies weitreichende Folgen für die Analyse des Hybridisierungsverhaltens von komplexen DNA-Mischungen, da die Wechselwirkung der beteiligten Target-Spezies berücksichtigt werden müsste.

Zusammenfassend lässt sich sagen, dass die vorliegende Arbeit zum besseren Verständnis der molekularen Erkennung im Falle der DNA-Hybridisierung beiträgt. Zahlreiche Effekte, von denen in dieser Arbeit die Ausbildung von Loops und die Vorgabe einer Konkurrenzsituation untersucht wurden, haben Einfluss auf die Genauigkeit der molekularen Erkennung. So kann beispielsweise der Effekt der hochspezifischen Hybridisierung in Konkurrenz die Genauigkeit der molekularen Erkennung im thermodynamischen Gleichgewicht um mehrere Größenordnungen verbessern. Dies zeigt, dass aufgrund kollektiver Effekte die Eigenschaften eines Konkurrenzsystems nicht notwendigerweise durch die Eigenschaften der beteiligten Einzelsysteme vorgegeben sein müssen. Dies muss nicht zwangsläufig nur für die molekulare Erkennung im Falle der DNA-Hybridisierung gültig sein, sondern kann auch in anderen Bereichen eine Rolle spielen, wie beispielsweise bei der Wechselwirkung von Antikörpern und Antigenen oder bei

der Bindung von Proteinen an DNA. Daher wird die Untersuchung anderer Systeme notwendig sein, um ein umfassendes Bild der molekularen Erkennung zu erhalten.

# Literaturverzeichnis

- [1] T. Naiser. *Characterization of Oligonucleotide Microarray Hybridization*. Dissertation, Universität Bayreuth, 2007.
- [2] T. Naiser, J. Kayser, T. Mai, W. Michel, and A. Ott. Position dependent mismatch discrimination on DNA microarrays - experiments and model. *BMC Bioinformatics*, 9:509, 2008.
- [3] T. Naiser, J. Kayser, T. Mai, W. Michel, and A. Ott. Stability of a surface-bound oligonucleotide duplex inferred from molecular dynamics: A study of single nucleotide defects using DNA microarrays. *Phys. Rev. Lett.*, 102, 218301, 2009.
- [4] A. Brazma and D. Hingamp. Minimum information about a microarray experiment (MIAME) - towards standards for microarray data. *Nat. Genet.*, 29, 2001.
- [5] Y. Dharmadi and R. Gonzalez. DNA microarrays: Experimental issues, data analysis, and application to bacterial systems. *Biotechnol. Prog.*, 20:1309–1324, 2004.
- [6] H. C. King and A. A. Sinha. Gene expression profile analysis by DNA microarrays: Promise and pitfalls. *JAMA*, Vol 286, No. 18 (Reprinted), 2001.
- [7] T. Wilkes, H. Laux, and C. A. Foy. Microarray data quality - review of current developments. *Omic*s, 11(1):1–13, 2007.
- [8] C. L. Yauk and M. L. Berndt. Review of the literature examining the correlation among DNA microarray technologies. *Environ. Mol. Mutagen*, 48(5):380–394, 2007.
- [9] N. A. Campbell and J. B. Reece. *Biologie*. Spektrum, 2003.
- [10] P. A. Levene. The structure of yeast nucleic acid. *J. Biol. Chem.*, 40:415–424, 1919.
- [11] J. D. Watson and F. H. C. Crick. Molecular structure of nucleic acids. *Nature*, 171, 1953.

- [12] B. Tinland, A. Pluen, J. Sturm, and G. Weill. Persistence length of single-stranded DNA. *Macromolecules*, 30:5763–5765, 1997.
- [13] M. C. Murphy, I. Rasnik, W. Cheng, T. M. Lohman, and T. Ha. Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophys. J.*, 86:2530–2537, 2004.
- [14] M. Kohandel and B. Y. Ha. Thermal denaturation of double-stranded DNA: Effect of base stacking. *Phys. Rev. E*, 73(1):011905, 2006.
- [15] P. J. Hagerman. Flexibility of DNA. *Annu. Rev. Biophys. Bio.*, 17:265–286, 1988.
- [16] E. T. Kool. Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annu. Rev. Biophys. Biomol. Struct.*, 30:1–22, 2001.
- [17] P. Yakovchuk, E. Protozanova, and M. D. Frank-Kamenetskii. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucl. Acids Res.*, 34(2), 2006.
- [18] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular Biology of the Cell*. Garland Science, 2002.
- [19] H. M. Temin and S. Mizutani. Viral RNA-dependent DNA polymerase: RNA-dependent DNA polymerase in virions of rous sarcoma virus. *Nature*, 226:1211–1213, 1970.
- [20] F. Crick. On protein synthesis. *Symp. Soc. Exp. Biol. XII*, pages 139–163, 1958.
- [21] F. Crick. Central dogma of molecular biology. *Nature*, 227(5258):561–3, 1970.
- [22] S. Spiegelman, I. Haruna, I. B. Holland, G. Beaudreau, and D. Mills. The synthesis of a self-propagating and infectious nucleic acid with a purified enzyme. *Proc. Nat. Acad. Sci. USA*, 54 (3):919–927, 1965.
- [23] D. Thieffry and S. Sarkar. Forty years under the central dogma. *Trends Biochem.*, 23(8):312–316, 1998.
- [24] The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489:57–74, 2012.
- [25] Encyclopedia of DNA Elements, 2013. <http://www.encodeproject.org/ENCODE/index.html>.
- [26] D. Pörschke and M. Eigen. Co-operative non-enzymatic base recognition III. kinetics of the helix-coil transition of the oligoribouridylic-oligoriboadenylic acid

- system and of oligoriboadenylic acid alone at acidic pH. *J. Mol. Biol.*, 62, 2:361–364, 1971.
- [27] I. Langmuir. The constitution and fundametal properties of solids and liquids, part I solids. *J. Am. Chem. Soc.*, 38:2221–95, 1916.
- [28] 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. *Biophys. J.*, 92:999–1004, 2007.
- [29] R. Thomas. Recherches sur la dénaturation des acides désoxyribonucléiques. *Biochim. Biophys. Acta*, 14:231–240, 1954.
- [30] S. A. Rice and P. Doty. The thermal denaturation of desoxyribose nucleic acid. *J. Am. Chem. Soc.*, 79 (15):3937–3947, 1957.
- [31] R. M. Wartell and A. S. Benight. Thermal denaturation of DNA molecules: A comparsion of theory with experiment. *Phys. Rep.*, 126:67–107, 1985.
- [32] R. Blossey and E. Carlon. Reparametrizing the loop entropy weights: Effect on DNA melting curves. *Phys. Rev. E*, 68, 061911, 2003.
- [33] M. Baiesi, E. Carlon, Y. Kafri, D. Mukamel, E. Orlandini, and A. L. Stella. Interstrand distance distribution of DNA near melting. *arXiv:cond-mat/0211236v1*, 2002.
- [34] J. Applequist and V. Damle. Thermodynamics of helix-coil equilibrium in oligoadenylic acid from hypochromicity studies. *J. Am. Chem. Soc.*, 87(7):1450–1458, 1965.
- [35] J. G. Duguid, V. A. Bloomfield, J. M. Benevides, and G. J. Thomas Jr. DNA melting investigated by differential scanning calorimetry and raman spectroscopy. *Biophys J.*, 71(6):3350–336, 1996.
- [36] J. Kypr, I. Kejnovská, D. Renčiuk, and M. Vorlíčková. Circular dichroism and conformational polymorphism of DNA. *Nucl. Acids Res.*, 37, No. 6 1:713–1725, 2009.
- [37] D. J. Patel and L. Canuel. Nuclear magnetic resonance studies of the helix-coil transition of poly(dA-dT) in aqueous solution. *Proc. Nat. Acad. Sci. USA*, 73, No. 3:674–678, 1976.
- [38] R. Owczarzy. Melting temperatures of nucleic acids: Discrepancies in analysis. *Biophys. Chem.*, 117(3):207–215, 2005.

- [39] R. B. Wallace, J. Shaffer, R. F. Murphy, J. Bonner, T. Hirose, and K. Itakura. Hybridization of synthetic oligodeoxyribonucleotides to phi-chi-174 DNA: the effect of single base pair mismatch. *Nucl. Acids Res.*, 6, 1979.
- [40] J. G. Wetmur. DNA probes: Applications of the principles of nucleic-acid hybridization. *Crit. Rev. Biochem. Mol. Biol.*, 26(3-4):227–259, 1991.
- [41] N. Markham and M. Zuker. DINAmelt web server for nucleic acid melting prediction. *Nucl. Acids Res.*, 33:W577–W581, 2005.
- [42] D. M. Crothers and B. H. Zimm. Theory of the melting transition of synthetic polynucleotides: Evaluation of the stacking free energy. *J. Mol. Biol.*, 9, 1:1–9, 1964.
- [43] O. Gotoh and Y. Tagashira. Stabilities of nearest-neighbor doublets in doublehelical DNA determined by fitting calculated melting profiles to observed profiles. *Biopolymers*, 20(5):1033–1042, 1981.
- [44] J. SantaLucia Jr. and D. Hicks. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, 33, 2004.
- [45] M. Petersheim and D. H. Turner. Basestacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry*, 22:256–63, 1983.
- [46] M. G. Gore, editor. *The use of spectroscopic techniques in the study of DNA stability. In Spectrophotometry and Spectrofluorimetry: A Practical Approach.* Oxford Univ. Press, 2000.
- [47] A. V. Vologodskii, B. R. Amirikyan, Y. L. Lyubchenko, and M. D. Frank-Kamenetskii. Allowance for heterogeneous in the DNA helix-coil transition theory. *J. Biomol. Struct. Dyn.*, 2:131–148., 1954.
- [48] K. J. Breslauer, R. Franks, H. Blockers, and L. A. Marky. Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA*, Vol. 83:3746–3750, 1986.
- [49] S. G. Delcourt and R. D. Blake. Stacking energies in DNA. *J. Biol. Chem.*, 266(23):15160–9., 1991.
- [50] M. J. Doktycz, R. F. Goldstein, T. M. Paner, F. J. Gallo, and A. S. Benight. Studies of DNA dumbbells I: melting curves of 17 DNA dumbbells with different

- duplex stem sequences linked by T4 endloops. evaluation of the nearest-neighbor stacking interactions in DNA. *Biopolymers*, 32:849–864, 1992.
- [51] J. SantaLucia Jr., H. T. Allawi, and P. A. Seneviratne. Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry*, 35:3555–3562, 1996.
- [52] N. Sugimoto, S. Nakano, M. Yoneyama, and K. Honda. Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucl. Acids Res.*, 24 (22):4501–4505, 1996.
- [53] J. SantaLucia Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. USA*, Vol. 95:1460–1465, 1998.
- [54] H. T. Allawi and J. SantaLucia Jr. Thermodynamics and NMR of internal G-T mismatches in DNA. *Biochemistry*, 36:10581–10594, 1997.
- [55] R. Owczarzy, Y. You, B. G. Moreira, J. A. Manthey, L. Y. Huang, M. A. Behlke, and J. A. Walder. Effects of sodium ions on DNA duplex oligomers: Improved predictions of melting temperatures. *Biochemistry*, 43(12):3537–3554, 2004.
- [56] K. Kleppe, E. Ohtsuka, R. Kleppe, and I. Molineux and H. G. Khorana. Studies on polynucleotides: Xcvi. repair replication of short synthetic DNA's as catalyzed by DNA polymerases. *J. Mol. Biol.*, 56(2):341–361, 1971.
- [57] J. M. S. Bartlett and D. Stirling. *PCR Protocols: Methods in Molecular Biology*, chapter A Short History of the Polymerase Chain Reaction, pages 3–6. Springer, 2003.
- [58] A. Chien, D. B. Edgar, and J. M. Trela. Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J. Bacteriol.*, 127 (3):1550–1557, 1976.
- [59] R. B. Stoughton. Applications of DNA microarrays in biology. *Annu. Rev. Biochem.*, 74:53–82, 2005.
- [60] X. L. Gao, E. Gulari, and X. C. Zhou. In situ synthesis of oligonucleotide microarrays. *Biopolymers*, 73(5):579–596, 2004.
- [61] C. A. Harrington, C. Rosenow, and J. Retief. Monitoring gene expression using DNA microarrays. *Curr. Opin. Microbiol.*, 3:285–291, 2000.

- [62] P. Hofman. DNA microarrays: A practical approach from a pathologist's standpoint. *Nephron. Physiol.*, 99:p85–p89, 2005.
- [63] D. Gresham, M. J. Dunham, and D. Botstein. Comparing whole genomes using DNA microarrays. *Nat. Rev. Genet.*, 9:291–302, 2008.
- [64] D. Gresham, D. M. Ruderfer, S. C. Pratt, J. Schacherer, M. J. Dunham, D. Botstein, and L. Kruglyak. Genome-wide detection of polymorphisms at nucleotide resolution with a single DNA microarray. *Science*, 311:1932–1936, 2006.
- [65] V. Trevino, F. Falciani, and H. A. Barrera-Saldaña. DNA microarrays: a powerful genomic tool for biomedical and clinical research. *Mol. Med.*, 13(9-10):527–541, 2007.
- [66] B. E. Cleven, M. Palka-Santini, J. Gielen, S. Meembor, M. Krönke, and O. Krut. Identification and characterization of bacterial pathogens causing bloodstream infections by DNA microarray. *J. Clin. Microbiol.*, 44(7):2389–2397, 2006.
- [67] S. F. Al-Khaldi, M. M. Mossoba, M. M. Allard, E. K. Lienau, and E. D. Brown. Bacterial identification and subtyping using DNA microarray and DNA sequencing. *Methods Mol. Biol.*, 881:73–95, 2012.
- [68] A. Hasan, K. P. Stengele, H. Giegrich, P. Cornwell, K. R. Isham, R. A. Sachleben, W. Pfeleiderer, and R. S. Foote. Photolabile protecting groups for nucleosides: Synthesis and photodeprotection rates. *Tetrahedron*, 53(12):4247–4264, 1997.
- [69] M. Beier and J. D. Hoheisel. Versatile derivatisation of solid support media for covalent bonding on DNA-microchips. *Nucl. Acids Res.*, 27:197:1970–1977,, 1999.
- [70] T. Naiser, T. Mai, W. Michel, and A. Ott. A versatile maskless microscope projection photolithography system and its application in light-directed fabrication of DNA microarrays. *Rev. Sci. Instrum.*, 77, 063711, 2006.
- [71] A. H. Krotz, C. Rentel, D. Gorman, P. Olsen, H. J. Gaus, J. V. McArdle, and A. N. Scozzari. Solution stability and degradation pathway of deoxyribonucleoside phosphoramidites in acetonitrile. *Nucleos. Nucleot. Nucl.*, 23(5):767–775, 2004.
- [72] V. Le Berre, E. Trevisiol, A. Dagkessamanskaia, S. Sokol, A. M. Caminade, J. P. Majoral, B. Meunier, and J. Francois. Dendrimeric coating of glass slides for sensitive DNA microarrays analysis. *Nucl. Acids Res.*, 31(16):e88, 2003.
- [73] E. F. Nuwaysir, W. Huang, T. J. Albert, J. Singh, K. Nuwaysir, A. Pitas, T. Richmond, T. Gorski, J. P. Berg, J. Ballin, M. McCormick, J. Norton, T. Pollock,

- T. Sumwalt, L. Butcher, D. Porter, M. Molla, C. Hall, F. Blattner, M. R. Sussman, R. L. Wallace, F. Cerrina, and R. D. Green. Gene expression analysis using oligonucleotide arrays produced by maskless photolithography. *Genome Res.*, 12:1749–1755, 2002.
- [74] P. B. Garland and P. J. Serafinowski. Effects of stray light on the fidelity of photodirected oligonucleotide array synthesis. *Nucl. Acids Res.*, 30 No. 19 e99, 2002.
- [75] S. Walbert, W. Pfeiderer, and U. E. Steiner. Photolabile protecting groups for nucleosides: Mechanistic studies of the 2-(2-nitrophenyl)ethyl group. *Helv. Chim. Acta*, 84, 2001.
- [76] C. Trapp. Kodierung und DNA-Microarrays. Diplomarbeit, Universität des Saarlandes, 2009.
- [77] T. Förster. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik*, 2, 6, 1948.
- [78] C. P. Berney. *Quantitative Fluorescence Resonance Energy Transfer Microscopy*. Dissertation, Swiss Federal Institute of Technology Zürich, 2003.
- [79] J. R. Lakowicz. *Principles of Fluorescence Spectroscopy*. Springer, 2006.
- [80] B. Valeur. *Molecular Fluorescence*. Wiley-VCH, 2001.
- [81] X.F. Wang and B. Herman. Fluorescence imaging spectroscopy and microscopy. *Chemical Analysis*, 137, 1996.
- [82] A. Hillisch, M. Lorenz, and S. Diekmann. Recent advances in FRET: distance determination in protein-DNA complexes. *Curr. Opin. Struct. Biol.*, 11:201–207, 2001.
- [83] D. Magde, E. Eison, and W. W. Webb. Thermodynamic fluctuations in a reacting system - measurement by fluorescence correlation spectroscopy. *Phys. Rev. Lett.*, 29:705–708, 1972.
- [84] R. Rigler, Ü. Mets, J. Widengren, and P. Kask. Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion. *Eur. Biophys. J.*, 22:169–175, 1993.
- [85] G. Altan-Bonnet, A. Libchaber, and O. Krichevsky. Bubble dynamics in double-stranded DNA. *Phys. Rev. Lett.*, 90 138101, 2003.

- [86] O. Krichevsky and G. Bonnet. Fluorescence correlation spectroscopy: the technique and its applications. *Rep. Prog. Phys.*, 65:251–297, 2002.
- [87] R. Pecora, editor. *Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy*. Springer, 1985.
- [88] T. Naiser, O. Ehler, J. Kayser, T. Mai, W. Michel, and A. Ott. Impact of point-mutations on the hybridization affinity of surface-bound DNA/DNA and RNA/DNA oligonucleotide-duplexes: Comparison of single base mismatches and base bulges. *BMC Biotechnology*, 8:48, 2008.
- [89] C. Trapp, M. Schenkelberger, and A. Ott. Stability of double-stranded oligonucleotide DNA with a bulged loop: a microarray study. *BMC Biophysics*, 4:20, 2011.
- [90] A. Harrison, H. Binder, A. Buhot, C. J. Burden, E. Carlon, C. Gibas, L. J. Gamble, A. Halperin, J. Hooyberghs, D. P. Kreil, R. Levicky, P. A. Noble, A. Ott, B. M. Pettitt, D. Tautz, and A. E. Pozhitkov. Physico-chemical foundations underpinning microarray and next-generation sequencing experiments. *Nucl. Acids Res.*, 41(5):2779–2796, 2013.
- [91] A. Pozhitkov, P. A. Noble, T. Domazet-Loso, A. W. Nolte, R. Sonnenberg, P. Staehler, M. Beier, and D. Tautz. Tests of rRNA hybridization to microarrays suggest that hybridization characteristics of oligonucleotide probes for species discrimination cannot be predicted. *Nucl. Acids Res.*, Vol. 34, No. 9 e66, 2006.
- [92] R. N. Nurtdinov, M. O. Vasiliev, A. S. Ershova, I. S. Lossev, and A. S. Karyagina. Plandbaffy: probe-level annotation database for affymetrix expression microarrays. *Nucl. Acids Res.*, 38:D726–D730, 2010.
- [93] B. G. Moreira, Y. You, M. A. Behlke, and R. Owczarzy. Effects of fluorescent dyes, quenchers, and dangling ends on DNA duplex stability. *Biochem. Biophys. Res. Commun.*, 327:473–484, 2005.
- [94] D. L. Mobley, A. P. Graves, J. D. Chodera, A. C. McReynolds, B. K. Shoichet, and K. A. Dill. Predicting absolute ligand binding free energies to a simple model site. *J. Mol. Biol.*, 371:1118–1134, 2007.
- [95] D. R. Ripoll, I. V. Khavrutskii, S. Chaudhury, J. Liu, R. A. Kushner, A. Wallqvist, and J. Reifman. Quantitative predictions of binding free energy changes in drug-resistant influenza neuraminidase. *PLoS Comput. Biol.*, 8(8), 2012.

- [96] M. Schenkelberger. *Competitive Biomolecular Binding*. Dissertation, Universität des Saarlandes, 2013.
- [97] M. Schenkelberger, C. Trapp, T. Mai, and A. Ott. Competition for binding entails ultrahigh molecular recognition specificity. *in Vorbereitung*, 2013.
- [98] J. Bishop, A. M. Chagovetz, and S. Blair. Kinetics of multiplex hybridization: Mechanisms and implications. *Biophys. J.*, 94:1726–1734, 2008.
- [99] D. Xu, K. O. Evans, and T. M. Nordlund. Melting and premelting transitions of an oligomer measured by DNA base fluorescence and absorption. *Biochemistry*, 33:9592–9599, 1994.
- [100] H. Ma, C. Wan, A. Wu, and A. H. Zewail. DNA folding and melting observed in real time redefine the energy landscape. *Proc. Nat. Acad. Sci.*, 104(3):712–716, 2007.
- [101] E. J. Sambriski, D. C. Schwartz, and J. J. de Pablo. Uncovering pathways in DNA oligonucleotide hybridization via transition state analysis. *Proc. Natl. Acad. Sci.*, 106:43:18125–18130, 2009.
- [102] J. H. Gibbs and E. A. DiMarzio. Statistical mechanics of helix-coil transitions in biological macromolecules. *J. Chem. Phys.*, 30, No. 1, 1959.
- [103] C. Kittel. Phase transition of a molecular zipper. *Am. J. Phys.*, 37 (9), 1969.
- [104] J. M. Deutsch, S. Liang, and O. Narayan. Modeling of microarray data with zippering. *arXiv:q-bio/0406039v1*, 2004.
- [105] C. Vanderzande. *Lattice models of polymers*. Cambridge University Press, 1998.
- [106] D. Randall and A. Sinclair. Self-testing algorithms for self-avoiding walks. *J. Math. Phys.*, 41(3), 2000.
- [107] K. Rechendorff, G. Witz, J. Adamcik, and G. Dietler. Persistence length and scaling properties of single-stranded DNA adsorbed on modified graphite. *J. Chem. Phys.*, 131, 095103, 2009.
- [108] E. J. Sambriski, D. C. Schwartz, and J. J. de Pablo. A mesoscale model of DNA and its renaturation. *Biophys J.*, 96:1675–1690, 2009.
- [109] S. H. Ke and R. M. Wartell. Influence of neighboring base pairs on the stability of single base bulges and base pairs in a DNA fragment. *Biochemistry*, 34:4593–4600, 1995.

- [110] B. M. Znosko, S. B. Silvestri, H. Volkman, B. Boswell, and M. J. Serra. Thermodynamic parameters for an expanded nearest-neighbor model for the formation of RNA duplexes with single nucleotide bulges. *Biochemistry*, 41:10406–10417, 2002.
- [111] A. Hanke, M. G. Ochoa, and R. Metzler. Denaturation transition of stretched DNA. *Phys. Rev. Lett.*, 100, 018106, 2008.

# Anhang A

## Experimentelle Daten



## A.1 Loops anderer Sequenzen

Verwendete Probe-Sequenzen in Abschnitt 4.2.2 - Loops anderer Sequenzen. Die zusätzlich eingefügten Basen sind rot gekennzeichnet. Im Fall B) wurden die eingefügten Sequenzen aus der Sequenz „CAGT“ gewonnen.

### a) poly-C-Loops

Loop-Position	Loop-Länge	Probe-Sequenz (3'-5')	Länge (Basen)
PM Probe		TTCAATACTACTCATAATTACCAACAACATTAC	33
9	1	TTCAATACTC <b>C</b> ACTCATAATTACCAACAACATTAC	34
9	2	TTCAATACT <b>CC</b> ACTCATAATTACCAACAACATTAC	35
⋮	⋮	⋮	⋮
⋮	⋮	3'-TTCAATACT <b>CC</b> ...ACTCATAATTACCAACAACATTAC	⋮
9	13	TTCAATACT <b>CCCCCCCCCCCC</b> ACTCATAATTACCAACAACATTAC	46
16	1	TTCAATACTACTCATA <b>C</b> ATTACCAACAACATTAC	34
16	2	TTCAATACTACTCATA <b>CC</b> ATTACCAACAACATTAC	35
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
16	13	TTCAATACTACTCATA <b>CCCCCCCCCCCC</b> ATTACCAACAACATTAC	46
23	1	TTCAATACTACTCATAATTACCA <b>C</b> ACAACATTAC	34
23	2	TTCAATACTACTCATAATTACCA <b>CC</b> ACAACATTAC	35
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
23	13	TTCAATACTACTCATAATTACCA <b>CCCCCCCCCCCC</b> ACAACATTAC	46

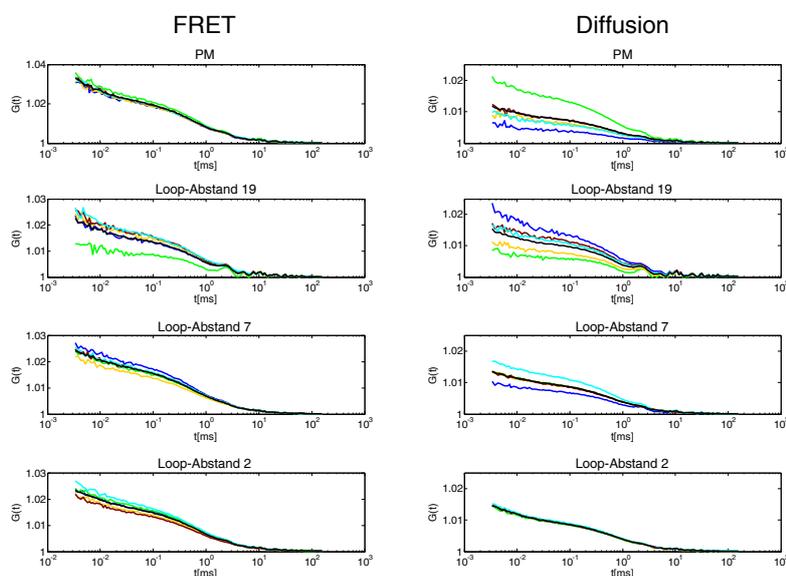
### b) beliebige Loops

Loop-Position	Loop-Länge	Probe-Sequenz (3'-5')	Länge (Basen)
PM Probe		TTCAATACTACTCATAATTACCAACAACATTAC	33
9	1	TTCAATACT <b>C</b> ACTCATAATTACCAACAACATTAC	34
9	2	TTCAATACT <b>CA</b> ACTCATAATTACCAACAACATTAC	35
9	3	TTCAATACT <b>CAG</b> ACTCATAATTACCAACAACATTAC	36
9	4	TTCAATACT <b>CAGT</b> ACTCATAATTACCAACAACATTAC	37
9	5	TTCAATACT <b>CAGTC</b> ACTCATAATTACCAACAACATTAC	38
⋮	⋮	⋮	⋮
⋮	⋮	3'-TTCAATACT <b>CAGTC</b> ...ACTCATAATTACCAACAACATTAC-5'	⋮
9	13	TTCAATACT <b>CAGTCAGTCAGTC</b> ACTCATAATTACCAACAACATTAC-5'	46
16	1	TTCAATACTACTCATA <b>C</b> ATTACCAACAACATTAC	34
16	2	TTCAATACTACTCATA <b>CA</b> ATTACCAACAACATTAC	35
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
16	13	TTCAATACTACTCATA <b>CAGTCAGTCAGTC</b> ATTACCAACAACATTAC	46
23	1	TTCAATACTACTCATAATTACCA <b>C</b> ACAACATTAC	34
23	2	TTCAATACTACTCATAATTACCA <b>CA</b> ACAACATTAC	35
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
23	13	TTCAATACTACTCATAATTACCA <b>CAGTCAGTCAGTC</b> ACAACATTAC	46

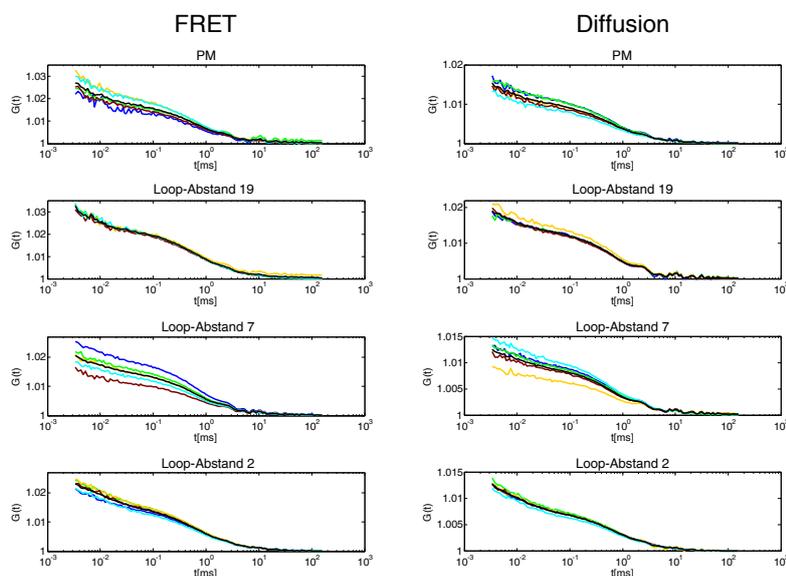
## A.2 Probes mit zwei Loops: FCS-Rohdaten

Diese Daten werden benutzt um die in Abschnitt 4.5.4 gezeigten Mittelwerte der Korrelationsfunktionen zu berechnen. Die farbigen Kurven bezeichnen die Daten der fünf Messungen. Die schwarze Kurve ist der daraus berechnete Mittelwert. Normalerweise werden für jeden Teilgraphen 5 Messungen durchgeführt. Sind weniger Messungen angegeben, so war die entsprechende Messung fehlerhaft.

### A.2.1 Messungen bei 44 °C



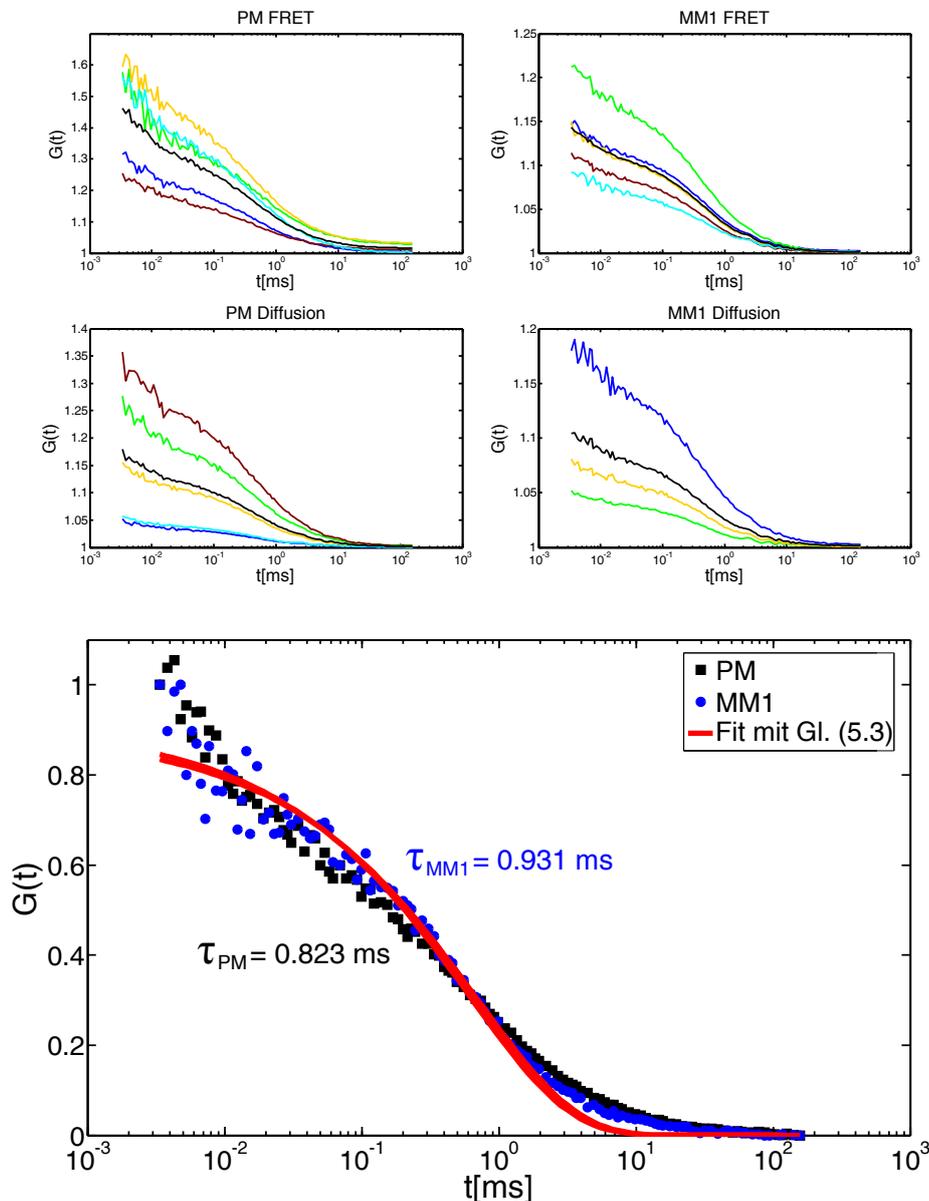
### A.2.2 Messungen bei 55 °C



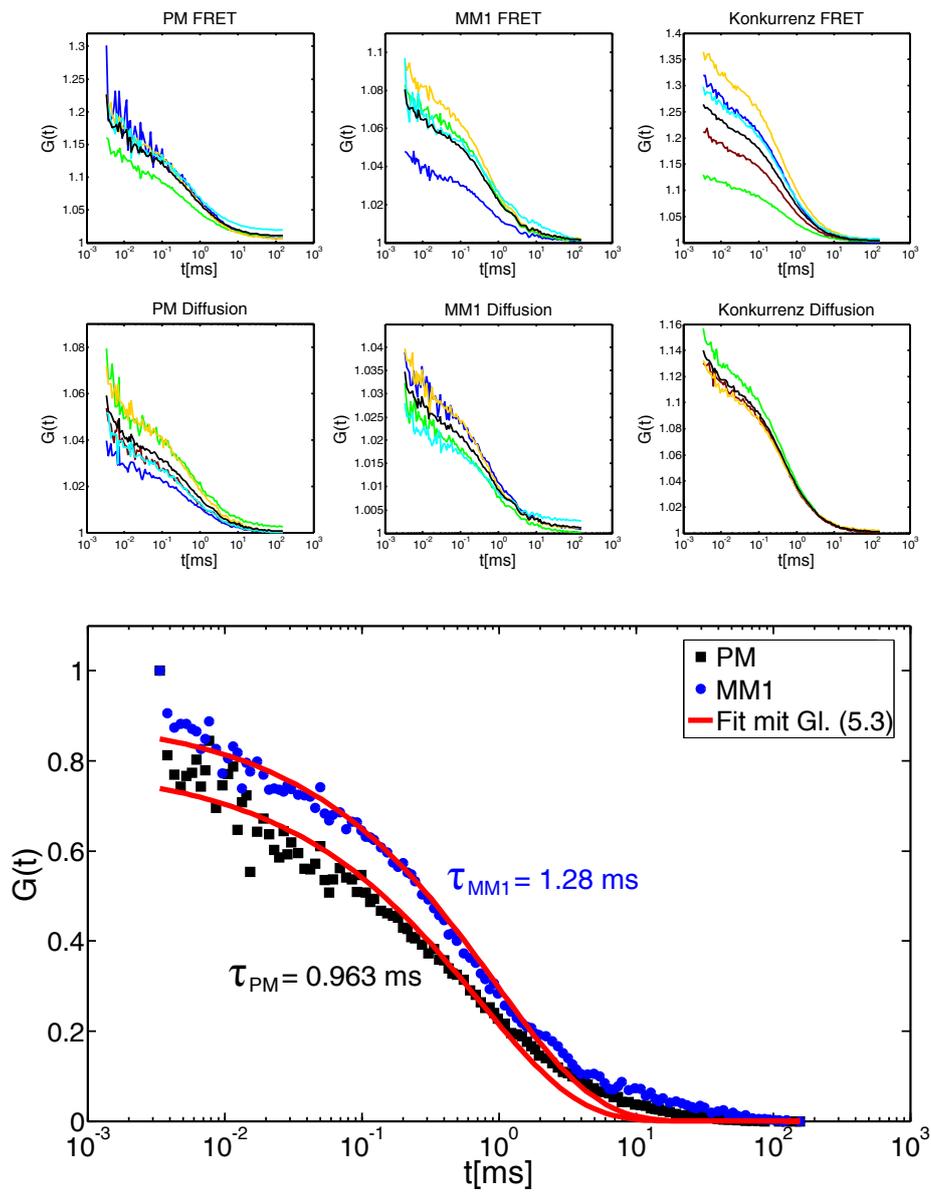
## A.3 DNA-Konkurrenzhybridisierung: FCS-Rohdaten und Fits

Diese Daten werden benutzt um die in Abschnitt 5.3.3 gezeigten Mittelwerte der Korrelationsfunktionen zu berechnen. Die farbigen Kurven bezeichnen die Daten der fünf Messungen. Die schwarze Kurve ist der daraus berechnete Mittelwert. Normalerweise werden für jeden Teilgraphen 5 Messungen durchgeführt. Sind weniger Messungen angegeben, so war die entsprechende Messung fehlerhaft. Die Mittelwerte werden mit dem Modell aus [85] gefittet.

### A.3.1 Messungen bei Raumtemperatur



### A.3.2 Messungen bei 44 °C



# **Anhang B**

## **Ergänzende Informationen**



## B.1 Verwendete Target-Sequenzen

In unten stehender Tabelle sind die im Rahmen dieser Arbeit verwendeten Target-Sequenzen aufgeführt. Die im Text genannten Probe-Sequenzen des gleichen Namens sind zu den hier aufgeführten Sequenzen komplementär. Wenn nicht anders angegeben, sind alle Target-Sequenzen am 3'-Ende mit Cy3 markiert.

Name	Sequenz (3'-5')	Länge
Map2k18833c	AAG TTA TGA TGA GTA TTA ATG GTT GTT GTA ATG	33
Map2k18833c-reverse	GTA ATG TTG TTG GTA ATT ATG AGT AGT ATT GAA	33
Map2k18833c-rev-low	TAT GAG TAG TAT TGA A	16
Map2k18833c-rev-up	GTA ATG TTG TTG GTA AT	17
Map2k18817	ATG AGT ATC AAT GGT TG	17
LT-PM-3Cy3	TTC AAT ACT ACT CAT AAT TAC CAA CAA CAT TAC	33
LT-PM-5Cy3	TTC AAT ACT ACT CAT AAT TAC CAA CAA CAT TAC-Cy3	33
LT-13-3Cy3	TTC AAT ACT ACT CAT A Loop <sub>13</sub> ATT ACC AAC AAC ATT AC, Loop <sub>13</sub> =13·T	46
LT-13-5Cy3	TTC AAT ACT ACT CAT A Loop <sub>13</sub> ATT ACC AAC AAC ATT AC -Cy3, Loop <sub>13</sub> =13·T	46
LT-26-3Cy3	TTC AAT ACT ACT CAT A Loop <sub>26</sub> ATT ACC AAC AAC ATT AC, Loop <sub>13</sub> =26·T	59
LT-26-5Cy3	TTC AAT ACT ACT CAT A Loop <sub>26</sub> ATT ACC AAC AAC ATT AC-Cy3, Loop <sub>13</sub> =26·T	59
LT-52-3Cy3	TTC AAT ACT ACT CAT A Loop <sub>52</sub> ATT ACC AAC AAC ATT AC, Loop <sub>13</sub> =52·T	85
LT-52-5Cy3	TTC AAT ACT ACT CAT A Loop <sub>52</sub> ATT ACC AAC AAC ATT AC-Cy3, Loop <sub>13</sub> =52·T	85
PM	AAT-GCT-AGA-CTA-GGA-A	16
M1-1	CAT-GCT-AGA-CTA-GGA-A	16
M1-2	GAT-GCT-AGA-CTA-GGA-A	16
M1-3	TAT-GCT-AGA-CTA-GGA-A	16
M2-1	ACT-GCT-AGA-CTA-GGA-A	16
M2-2	AGT-GCT-AGA-CTA-GGA-A	16
M2-3	ATT-GCT-AGA-CTA-GGA-A	16
M3(=MM2)	AAC-GCT-AGA-CTA-GGA-A	16
M3-2	AAA-GCT-AGA-CTA-GGA-A	16
M3-3	AAG-GCT-AGA-CTA-GGA-A	16
M4	AAT-CCT-AGA-CTA-GGA-A	16
M4-2	AAT-ACT-AGA-CTA-GGA-A	16
M4-3	AAT-TCT-AGA-CTA-GGA-A	16
M5	AAT-GGT-AGA-CTA-GGA-A	16
M5-2	AAT-GAT-AGA-CTA-GGA-A	16
M5-3	AAT-GTT-AGA-CTA-GGA-A	16
M6	AAT-GCA-AGA-CTA-GGA-A	16
M6-2	AAT-GCC-AGA-CTA-GGA-A	16
M6-3	AAT-GCG-AGA-CTA-GGA-A	16
M7	AAT-GCT-TGA-CTA-GGA-A	16
M7-2	AAT-GCT-CGA-CTA-GGA-A	16
M7-3	AAT-GCT-GGA-CTA-GGA-A	16
M8(=MM1)	AAT-GCT-ACA-CTA-GGA-A	16
M8-2	AAT-GCT-ATA-CTA-GGA-A	16
M8-3	AAT-GCT-AAA-CTA-GGA-A	16
M9	AAT-GCT-AGT-CTA-GGA-A	16
M9-2	AAT-GCT-AGC-CTA-GGA-A	16
M9-3	AAT-GCT-AGG-CTA-GGA-A	16
M10	AAT-GCT-AGA-GTA-GGA-A	16
M10-2	AAT-GCT-AGA-ATA-GGA-A	16
M10-3	AAT-GCT-AGA-TTA-GGA-A	16
M11	AAT-GCT-AGA-CAA-GGA-A	16
M11-2	AAT-GCT-AGA-CCA-GGA-A	16
M11-3	AAT-GCT-AGA-CTA-GGA-A	16
M12	AAT-GCT-AGA-CTT-GGA-A	16
M12-2	AAT-GCT-AGA-CTC-GGA-A	16

M12-3	AAT-GCT-AGA-CTG-GGA-A	16
M13	AAT-GCT-AGA-CTA-CGA-A	16
M13-2	AAT-GCT-AGA-CTA-AGA-A	16
M13-3	AAT-GCT-AGA-CTA-TGA-A	16
M14	AAT-GCT-AGA-CTA-GCA-A	16
M14-2	AAT-GCT-AGA-CTA-GAA-A	16
M14-3	AAT-GCT-AGA-CTA-GTA-A	16
M15-1	AAT-GCT-AGA-CTA-GGC-A	16
M15-2	AAT-GCT-AGA-CTA-GGG-A	16
M15-3	AAT-GCT-AGA-CTA-GGT-A	16
M16-1	AAT-GCT-AGA-CTA-GGA-C	16
M16-2	AAT-GCT-AGA-CTA-GGA-G	16
M16-3	AAT-GCT-AGA-CTA-GGA-T	16

## B.2 Hybridisierungs-Puffer

- 5xSSPE
  - Nukleasefreies Wasser
  - 0.75 M NaCl
  - 50 mM NaH<sub>2</sub>PO<sub>4</sub>
  - 5 mM EDTA
  - Zugabe von 0.01 % (v/v) Tween-20<sup>TM</sup>
  - Einstellen des pH-Werts von 7.4 mit NaOH

Die Zugabe von Tween-20<sup>TM</sup> ( Polyoxyethylen-(20)-sorbitanmonolaurat) ist notwendig um eine starke unspezifische Bindung der Targets an die Oberfläche zu verhindern. EDTA (Ethylendiamintetraacetat) wird zur Hemmung der Nuklease-Aktivität hinzugefügt.

- 3xSSC
  - Nukleasefreies Wasser
  - 0.45 M NaCl
  - 45 mM C<sub>6</sub>H<sub>5</sub>Na<sub>3</sub>O<sub>7</sub> · 2 H<sub>2</sub>O
  - Einstellen des pH-Werts von 7.4 mit NaCl

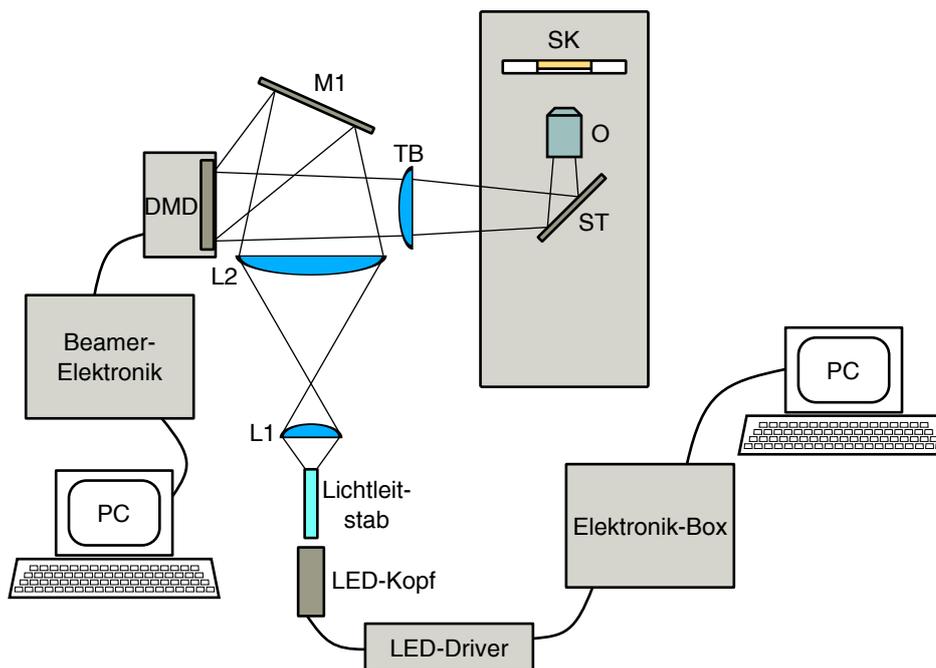
## B.3 Protokoll für die Herstellung der Dendrimer-Oberflächen

- Waschen der Deckglasoberflächen in Deconex (5%, 30 min) im Ultraschallbad
- Waschen in Milipore-Wasser
- Trocknen im N<sub>2</sub>-Gasstrom
- Reinigung im Plasmareiniger (10 min)
- Silanisierung: 10% (v/v) APTES + 95% EtOH (12 h)
  1. 180 ml 95% EtOH (171 ml EtOH, 9 ml Milipore Wasser) + 20 ml APTES, Endvolumen 200 ml
  2. APTES direkt mit Spritze aus der Flasche ziehen (SureSeal), Einleitung von Argon 4.8 mittels zweiter Kanüle
  3. Lagerung von APTES im Gefahrstoffschrank
  4. Fixierung eines Standzylinders (500 ml) auf einem Schüttler (Einstellung 125) unter Abzug
- Waschvorgang je zweimal mit EtOH (250 ml absolut) und Milipore-Wasser (500 ml)
  1. Vorbereitung von vier Bechergläsern, jeweils zwei Bechergläser mit EtOH (absolut, 250 ml) und Milipore-Wasser (500 ml)
  2. Halter mit Deckgläschen aus Silanlösung (Standzylinder) und jeweils nacheinander in die bereitgestellten Bechergläser stellen
  3. Säubern mit Spritze im ersten Becher, Behandlungsdauer etwa eine Minute pro Becherglas
- Trocknen im N<sub>2</sub>-Gasstrom
- Hitzebehandlung (120 °C, 3 h) eventueller Zwischenschritt (falls Dendrimerlösung älter als vier Wochen)
- neue 0.1% (w/v) Dendrimerlösung
  1. 200 mg PMMH Gen. 2.5 oder 4.5 (Lagerung im Gefahrstoffschrank)
  2. 200 ml Dichlormethan

Fortfahren mit Herstellung der Dendrimer-Oberflächen

- 8% KOH Lösung (16 g KOH in 200 ml Milipore-Wasser) (5 min)
- Dreimaliges Waschen in Milipore-Wasser (5-10 min pro Becher)
- Trocknen im N<sub>2</sub>-Gasstrom
- Dendrimerlösung (7 h, Standzylinder, 500 ml, Schüttler (Einstellung 125), Abzug)
- Waschvorgang, je zweimal mit Dichlormethan und EtOH (absolut)
- Trocknen im N<sub>2</sub>-Gasstrom
- 0.35% (w/v) Borhydrid (1 g NaBH<sub>4</sub> in 300 ml Milipore-Wasser (3 h, Schüttler (Einstellung 125), Abzug)
- Waschen in Milipore-Wasser
- Trocknen im N<sub>2</sub>-Gasstrom
- Lagerung bei 4 °C

## B.4 Optischer Aufbau des Synthesizers



- LED: UV-LED module LC-L2, Hamamatsu
- Lichtleitstab: 8 mm Fused Silica Hexagonal Light Pipe, Edmunds
- L1-Lichtleitstab: 30 mm
- L1: plankonvexe Linse,  $f=30$  mm,  $\varnothing=25.4$  mm, UV Fused Silica, Thorlabs
- L1-L2: 295 mm
- L2: plankonvexe Linse,  $f=150$  mm,  $\varnothing=75$  mm, UV Fused Silica, Thorlabs
- L2-M1: 70 mm
- M1: Metallspiegel
- M1-DMD: 80 mm
- DMD-TB: ca. 164.5 mm
- TB: Tubuslinse,  $f=164.5$  mm, Carl Zeiss

## Liste der Publikationen

- C. Trapp, M. Schenkelberger and A. Ott. Stability of double-stranded oligonucleotide DNA with a bulged loop: a microarray study. *BMC Biophysics* 4 (2011) 20
- Andrew Harrison, Hans Binder, Arnaud Buhot, Conrad J. Burden, Enrico Carlon, Cynthia Gibas, Lara J. Gamble, Avraham Halverin, Jef Hooyberghs, David P. Kreil, Rastislav Levicky, Peter A. Noble, B. Montgomery Pettitt, Diethard Tautz, Alexander E. Pozhitkov and A. Ott. Physico-chemical foundations underpinning microarray and next-generation sequencing experiments. *Nucl. Acids Res.*, 41(5):2779–2796, 2013.
- M. Schenkelberger, C. Trapp, T. Mai and A. Ott. Competition for binding entails ultrahigh molecular recognition specificity. *in Vorbereitung*



## Danksagung

Besonderer Dank gilt meinem Betreuer Prof. Dr. Albrecht Ott, der durch seine offene Art und seine stetige Diskussionsbereitschaft bei auftretenden physikalischen Problemen die Arbeit am Lehrstuhl zu einer sehr interessanten und lehrreichen Zeit gemacht hat.

Für die Übernahme des Zweit- bzw. Drittgutachtens meiner Dissertation möchte ich mich bei Prof. Dr. Karin Jacobs und Prof. Dr. Joachim Rädler bedanken.

Dank gilt auch meinen beiden Bürokolleginnen Eva Wollrab und Sabrina Scherer für die angenehme Büroatmosphäre, meinem persönlichen Coach Maschinenmensch Emanuel Worst für die vielen Muskelkater, Marc Schenkelberger für die erfolgreiche Zusammenarbeit sowie den restlichen Mitgliedern der Arbeitsgruppe.

Für das Korrekturlesen der Arbeit möchte ich mich bei Jennifer Collet bedanken, die doch sehr zur besseren Lesbarkeit der Arbeit beigetragen hat.

Mein größter Dank jedoch gilt meinen Eltern und meiner Freundin Lisa, die mich in schwierigen Phasen - nicht nur während der Promotion - unterstützt haben und immer ein offenes Ohr für meine Sorgen und Probleme hatten.

Vielen Dank Lisa, Du bist die Beste!