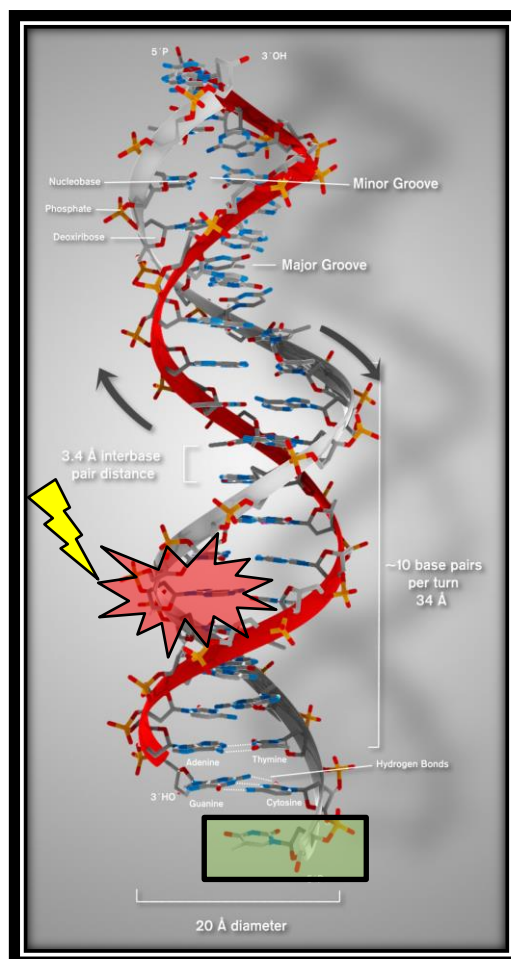




Dissertation

Untersuchung von Nukleinsäuren mit Hilfe von neuen Simulationsmethoden

Mahmut Kara





TECHNISCHE UNIVERSITÄT MÜNCHEN
Physik Department
Lehrstuhl für Molekulardynamik T38



Untersuchung von Nukleinsäuren mit Hilfe von neuen Simulationsmethoden

Dipl.-Bioinf.
Mahmut Kara

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

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Torsten Hugel
Prüfer der Dissertation: 1. Univ.-Prof. Dr. Martin Zacharias
2. Univ.-Prof. Dr. Iris Antes

Die Dissertation wurde am 29.10.2013 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 09.01.2014 angenommen.

Die Wissenschaft fängt eigentlich erst
da an interessant zu werden, wo sie
aufhört

Justus von Liebig (1803-73)

Abstract

The mechanisms behind the folding of nucleic acid structures into defined three-dimensional structures and the recognition of damaged DNA structures by repair enzymes are not well understood. DNA and RNA molecules are highly flexible and their conformation depends on the environment (temperature, pH, salt concentration, interaction with other molecules, etc.). One possible way to explore the dynamics of nucleic acids at atomic resolution is the use of Molecular Dynamics (MD) simulations based on a classical force field. In this dissertation I will present efficient approaches employing MD simulations to study the structure formation of nucleic acids revealing the underlying driving forces. For this purpose, a new, advanced sampling replica-exchange method has been developed to specifically enhance the sampling of conformational sub-states in nucleic acids during MD simulations. The approach, called BP-REMD, employs specific biasing potentials acting on pairs of pseudodihedral angles of the nucleic acid backbone that are added in the replica simulations to promote transitions between the most common sub-states of the nucleic acid backbone. The application to 7,8-dihydro-8oxo-guanosine and to O^6 -Methylguanine indicated a better convergence of sampled states in comparison to longer standard MD simulations for both cases. The newly developed method is well suited to systematically study the fine structure and dynamics of large nucleic acids under realistic conditions, including explicit solvent and ions. The biasing potential-replica exchange MD simulations indicated significant differences in the population of nucleic acid backbone sub-states in the case of modified guanosine compared to a regular guanosine in the same sequence context. Such differences may play a crucial role in the initial recognition of damaged DNA by repair enzymes. In all presented studies the results obtained from BP-REMD were compared to results obtained from classical methods as well as to experimental results.

In addition to DNA damage structures, non-canonical and non-regular structural motifs (RNA bulge structures) were examined by BP-REMD and classical MD simulations. The BP-REMD method showed improved sampling of conformational substates compared to conventional MD simulations. It could be shown that the BP-REMD models global conformational changes more realistically. Furthermore, Umbrella-Sampling free energy simulations of dangling cytosine and guanine nucleotides at the end of duplex and single stranded RNA and DNA molecules were used to investigate the molecular origin of dangling end effects. Calculated free energy contributions associated to the presence of

dangling nucleotides were in reasonable agreement with experiments predicting the general trend of a more stabilizing effect of purine dangling ends on comparison to pyrimidine dangling ends. Additionally, these calculations indicate a stronger stabilizing effect of dangling ends at the 5'-end than at the 3'-end in case of DNA and the opposite trend in the case of RNA.

Zusammenfassung

Nukleinsäuren sind hochflexible Moleküle, deren Konformationen abhängig von der Umgebung ist, die wiederum von einer Vielzahl von Faktoren abhängig ist (z.B. Temperatur, pH-Wert, Salzkonzentration, Interaktion mit anderen Biomolekülen usw.). In dieser Arbeit wurde ein Verfahren entwickelt mit dem der Konformationsraum von Nukleinsäuren effizient abgetastet werden kann. Das entwickelte Verfahren basiert auf der Destabilisierung von Pseudo-Diederwinkeln, die entlang des Zucker-Phosphatrückgrats definiert sind. Anhand der Simulationen können Aussagen über die strukturelle Vielfalt einer Nukleinsäure getroffen werden. Außerdem ist eine vergleichende Studie von Nukleinsäuren möglich, die sich nur geringfügig unterscheiden. Die Bereiche, in denen dieses Verfahren angewandt wurde, sind sehr vielseitig. Neben unterschiedlichen DNA-Schäden wurden auch Konformationsräume von ungepaarten Nukleotiden innerhalb der RNA Doppelhelix (RNA-Bulge-Strukturen) untersucht. Es wurde untersucht, inwieweit die Mutationschäden, ausgehend von einer modellierten B-DNA Struktur, sich auf lokale und globale Eigenschaften einer DNA-Doppelhelix auswirken. Hierbei wurden im speziellen oxidative Schäden (8-oxoguanin) und Alkylierungsschäden (O⁶-Methylguanin) in freier Lösung simuliert. Das Ziel, strukturelle Unterschiede bezüglich der intakten DNA zu bestimmen, wurde mittels des entwickelten Verfahrens verwirklicht. Annahme ist hierbei, dass die strukturellen Unterschiede als Erkennungsmerkmal für Reparaturenzyme dienen. In den vorliegenden Studien wurde stets darauf geachtet, dass ein Vergleich zu klassischen Verfahren und zu experimentellen Ergebnissen stattfindet. In den hier vorgestellten Simulationen konnte gezeigt werden, dass ein effizienteres Abtasten des Konformationsraumes mit dem neu entwickelten Verfahren im Vergleich zu klassischen Verfahren möglich ist.

Neben Mutationsschäden und RNA-Bulge-Strukturen wurden freie Energierechnungen an Dangling-End-Strukturen durchgeführt. Dangling-End Basen sind überhängende Basen am Terminus einer Nukleinsäure, die keine Watson-Crick Partner besitzt. Experimentell wurde bereits nachgewiesen, dass Dangling-End-Strukturen einen stabilisierenden Einfluss auf den Doppelstrang haben. Die Ursachen dieses Einflusses wurden in dieser Arbeit systematisch für GC-haltige Sequenzen untersucht.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Einführung in die Problemstellung	1
1.2	Aufbau und Stabilität von Nukleinsäuren	3
1.3	Sekundärstrukturen von Nukleinsäuren	7
1.3.1	RNA-Bulge	7
1.3.2	Dangling-End	9
1.4	DNA-Schäden und ihre Auswirkungen auf die Feinstruktur der DNA	11
1.4.1	7,8-dihydro-8-oxoguanin	14
1.4.2	O ⁶ -Methylguanin	16
1.5	Untersuchung von Nukleinsäuren mit Hilfe von Computermethoden	19
1.6	Motivation der Arbeit	21
2	Theorie, Methoden und Modelle	22
2.1	Molekulardynamik	22
2.1.1	Kraftfelder	24
2.2	DNA-Konformationsanalyse	26
2.2.1	Konformationen des Zucker-Phosphat-Rückgrats	27
2.2.2	Helikale Parameter	30
2.3	Berechnung der Freien-Energie-Differenz	34
2.3.1	Umbrella-Sampling Simulationen	35
2.3.2	MM/PBSA Rechnungen	37
2.4	Hamiltonian Replica Exchange Methode	39
2.5	Modellierung der Systeme	45
2.5.1	7,8-dihydro-8-oxoguanin	45
2.5.2	O ⁶ -Methylguanin	49
2.5.3	Dangling-End-Strukturen	51
2.5.4	RNA-Bulges	54
3	Ergebnisse und Diskussion	57
3.1	Guanin und 8-oxoguanin	57

3.1.1	Resultate der klassischen MD	58
3.1.2	Resultate der BP-REMD	65
3.1.3	Stereochemische Ursachen der unterschiedlichen Konformationszustände.....	67
3.1.4	Solvatisierung.....	70
3.1.5	Zusammenfassung und Schlussfolgerung	73
3.2	Vergleichende Modellierung von O ⁶ -MeG:C, O ⁶ -MeG:T und G:C	74
3.2.1	Stabilität der Strukturen anhand der RMSD-Werte.....	76
3.2.2	Vergleich zwischen klassischen MD und BP-REMD	78
3.2.3	Bestimmung helikaler Parameter	86
3.2.4	Clusteranalyse	88
3.2.5	Zusammenfassung und Schlussfolgerung	90
3.3	Einfluss von Dangling-Ends auf die Stabilität von Nukleinsäuren	91
3.3.1	Optimierung von Dangling-End Konformationen in freien Simulationen.....	92
3.3.2	Berechnung der Freien Energie mittels Umbrella-Sampling Simulationen.....	94
3.3.3	Analyse der Zusammensetzung der Energiebeiträge	99
3.3.4	Konformationsanalyse der Umbrella-Sampling Simulationen	101
3.3.5	Zusammenfassung und Schlussfolgerung	104
3.4	RNA-Bulge	105
3.4.1	Gesamtstabilität der Strukturen während der Simulation	106
3.4.2	Vergleich zwischen BP-REMD und kMDS Konformationen.....	108
3.4.3	Zusammenfassung und Diskussion	115
4	Abschließende Diskussion und Ausblick.....	116
	Danksagung.....	120
	Publikationsliste.....	121
	Anhang.....	122
	Literaturverzeichnis	125

Abbildungsverzeichnis

Abbildung 1.1 Zusammensetzung von Nukleotiden.....	3
Abbildung 1.2 Basenpaarungen in DNA Strukturen.....	4
Abbildung 1.3 Modell einer B-DNA Struktur	6
Abbildung 1.4 A-DNA, B-DNA und Z-DNA Konformationen.....	7
Abbildung 1.5 Elemente von RNA Sekundärstrukturen	7
Abbildung 1.6 Schematische Darstellung von <i>Dangling-End</i> Basen	10
Abbildung 1.7 Ein genereller Überblick möglicher DNA Schäden und deren Folgen.....	12
Abbildung 1.8 Reparaturmechanismus am Beispiel des 8oxoG.....	13
Abbildung 1.9 Basenexzisionsreparatur Mechanismus der 8oxoG-Base.....	15
Abbildung 1.10 Mögliche Erkennungsmechanismen von modifizierten DNA Basen.....	16
Abbildung 1.11 Beispiel einer Direkten-Reparatur am Beispiel der O ⁶ -MeG Base O ⁶ -MeG O ⁶ -MeG.....	17
Abbildung 1.12 Konformationszustände von alkylierten Guanin Basen.....	19
Abbildung 2.1 Vorgehensweise bei einer Molekulardynamik Simulation.....	23
Abbildung 2.2 Beschreibung der Kraftfeldterme.....	25
Abbildung 2.3 Klassifizierung der Konformationen anhand eines Pseudorotations-Rad.....	28
Abbildung 2.4 Diederwinkel des DNA-Phosphatrückgrats.....	29
Abbildung 2.5 Darstellung eines idealen Basen-Paar Parameters	31
Abbildung 2.6 Helikale Parameter für Nukleinsäuren.....	33
Abbildung 2.7 Thermodynamischer Kreisprozess bei einer MM/PBSA-Rechnung	38
Abbildung 2.8 Schematische Darstellung der BP-REXMD.....	42
Abbildung 2.9 Das 2D-Biasing Potential	43
Abbildung 2.10 Veranschaulichung der verwendeten Pseudo-Diederwinkel η und θ	44
Abbildung 2.11 Selbstkomplementäre Struktur mit zwei 8oxoG-Basen	46
Abbildung 2.12 α/γ and ε/ζ Konformationsbereiche von 96 experimentellen 8oxoG-Basen.....	48
Abbildung 2.13 B-DNA Modell der O ⁶ -MeG:C Fehlpaarung.....	51
Abbildung 2.14 Reaktionskoordinate vom gestackten zum völlig solvatisierten Zustand.....	52
Abbildung 2.15 Pseudo-Diederwinkel des RNA-Bulge Modells.....	55
Abbildung 2.16 Primär und Tertiärstruktur der RNA-Bulge-Strukturen	56
Abbildung 3.1 Zusammenhang zwischen RMSD und Diederwinkel.....	59
Abbildung 3.2 Zusammenhang zwischen Diederwinkel und die Auswirkungen des parm99 Kraftfeldes.....	60
Abbildung 3.3 PMF-Kurve des ε -Diederwinkel.....	62
Abbildung 3.4 Konformationsverteilungen der α/γ -Diederwinkel im Vergleich.....	63
Abbildung 3.5 Konformationsverteilungen der ε/ζ -Diederwinkel im Vergleich.....	64
Abbildung 3.6 Verteilung der χ -Diederwinkel.....	68
Abbildung 3.7 Korrelation zwischen ε und dem χ -Diederwinkel.....	69
Abbildung 3.8 Zusammenhang zwischen BI/BII und den α/γ Zuständen.....	69
Abbildung 3.9 Korrelation der Winkel und der O8-O4 bzw. H8-O4 Distanzen.....	70
Abbildung 3.10 Vergleich der radialen Verteilungsfunktion von 8oxoG und Guanin.....	71
Abbildung 3.11 Wasserstoffbrücken-Netzwerk der 8oxoG Base.....	72
Abbildung 3.12 Verteilung von Wassermolekülen im BI bzw. BII Zustand.....	72
Abbildung 3.13 RMSD Werte der drei Modelle: O ⁶ -MeG:C, O ⁶ -MeG:T und G:C.....	76
Abbildung 3.14 Superpositionierung der Kristall-Struktur 1T38 und der modellierten Struktur.....	78
Abbildung 3.15 Verteilung des δ -Diederwinkel im O ⁶ -MeG:C, O ⁶ -MeG:C und G:C Modell.....	80

Abbildung 3.16 Konformationsverteilungen der α/γ -Diederwinkel im Vergleich.....	82
Abbildung 3.17 α/γ and ε/ζ Konformationsbereiche von experimentellen O^6 -MeG-Basen.....	83
Abbildung 3.18 Verteilung des χ -Diederwinkel im O^6 -MeG:C, O^6 -MeG:T und G:C Modell.	84
Abbildung 3.19 Helikale Parameter der O^6 -MeG:C, O^6 -MeG:T und G:C Modelle im Vergleich.....	87
Abbildung 3.20 Repräsentative Strukturen aus der Clusteranalyse der kMDS-Trajektorie.....	89
Abbildung 3.21 Repräsentative Strukturen aus der Clusteranalyse des O^6 -MeG:T Modells (kMDS).	90
Abbildung 3.22 Superpositionierung von Average-Strukturen und Kristallstrukturen.	93
Abbildung 3.23 Superpositionierung der 5'-Cystein Dangling-End Base.	95
Abbildung 3.24 Vergleich zwischen den RNA und DNA Energieprofilen.	97
Abbildung 3.25 Thermodynamischer Kreis zur Berechnung des Freien Energiebeitrages zur Doppelstrang Stabilität.	98
Abbildung 3.26 Vergleich der Zwischenstände während der Umbrella-Sampling Simulation.....	102
Abbildung 3.27 Charakteristische Zwischenzustände von RNA Dangling-End-Strukturen während der Umbrella-Sampling Simulation.	102
Abbildung 3.28 Verteilung der Phosphatrückgrat Diederwinkel	103
Abbildung 3.29 RMSD Werte der A-Bulge und U-Bulge Modelle im Vergleich.	107
Abbildung 3.30 Charakteristische Zustände des U-Bulges während der kMDS.	108
Abbildung 3.31 Experimentell ermittelte Konformationszustände des A- und U-Bulges.	109
Abbildung 3.32 η/θ -Konformationszustände im Vergleich.....	110
Abbildung 3.33 α/γ und ε/ζ Konformationszustände im Vergleich.	111
Abbildung 3.34 Clusteranalyse der kMDS und BP-REMD Trajektorie.	113
Abbildung 3.35 Vergleich des U- und A-Bulges.....	114

Kapitel 1

1 Einleitung

1.1 Einführung in die Problemstellung

Das grundlegende Verständnis über Struktur und Funktion von Nukleinsäuren nahm im Jahre 1953 mit der Auflösung der drei-dimensionalen Struktur der Desoxyribonukleinsäure (DNS, engl. deoxyribonucleic acid (DNA))-Doppelhelix durch James Watson und Francis Crick ihren Anfang (1). Crick beschreibt in späteren Arbeiten das zentrale Dogma der Molekularbiologie, welches besagt, dass die Information von DNA über RNA hin zum Protein übertragen wird (2). Die DNA übernimmt hierbei die zentrale Funktion der Informationsspeicherung für ein spezifisches Genprodukt ein. Neben der Strukturauflösung war die Entschlüsselung des menschlichen Genoms ein weiterer Schritt zum Verständnis der Funktionsweise des menschlichen Erbguts. Die DNA-Sequenz, auch als Primärstruktur der DNA bekannt, beschreibt die Abfolge der vier Basen (Adenin, Thymin, Cytosin und Guanin) und kann daher nur eine begrenzte Information über Form und Funktion der Struktur liefern. Da die DNA jedoch ein hochflexibles und dynamisches Molekül ist, welches mit einer Vielzahl von Biomolekülen interagieren kann, ist die Sekundär- und Tertiärstruktur, also die räumliche Anordnung der Basenpaare, und die Dynamik dieser Struktur von großem Interesse. Enorme Fortschritte in experimentellen Methoden wie der Kristallstrukturanalyse oder der Kernspinresonanzspektroskopie ermöglichten hierbei einen Einblick in die strukturelle Vielfalt der Nukleinsäuren, wobei es schwierig ist Aussagen über die Dynamik einer Struktur mit Hilfe der Kristallstrukturanalyse zu treffen, da diese von den experimentellen Bedingungen abhängig sind.

Um die Dynamik einer Struktur zu bestimmen, können die experimentell ermittelten Informationen für theoretische Studien herangezogen werden. Durch die neuen Möglichkeiten, die die Informationstechnologie zur Verfügung stellt, ist es nun möglich Ergebnisse zu liefern, die für die Forschung interessant sind. Sie ermöglicht eine Vielzahl von Informationen über die Struktur zu erhalten und

dient als Ergänzung und Hilfe zur Interpretation von experimentellen Ergebnissen. Ziel dieser Arbeit ist die Beschreibung und Analyse von diversen Sequenz-Motiven und Konformationen von Nukleinsäuren hinsichtlich ihrer Dynamik und Stabilität. Zur Untersuchung der Fragestellungen wurden Molekulardynamik-Simulationen (MDS) verwendet. Anhand der MDS können Erkenntnisse über die Dynamik und Konformationsänderungen im Gleichgewichtszustand gewonnen werden. Um diesen Gleichgewichtszustand bestimmen zu können, wurden sowohl *klassische MDS* (kMDS), als auch *Hamiltonian Replica Exchange Molekulardynamik Simulationen* (H-REMD) durchgeführt. Bei der verwendeten H-REMD Methode handelt es sich um eine Simulationstechnik, bei dem parallel Simulationen (Replikate), ausgehend von der gleichen Startstruktur, jedoch mit unterschiedlichen Kraftfeldtermen, durchgeführt werden. Benachbarte Replikate haben die Möglichkeit Konformationen auszutauschen. Die Replikate unterscheiden sich bei der verwendeten Methode anhand unterschiedlicher Terme für die Berechnung bestimmter Diederwinkel. Ziel hierbei ist, Energiebarrieren für die definierten Diederwinkel zu reduzieren und somit einen größeren Konformationsraum abzusuchen. Beide Methoden wurden hinsichtlich ihrer Effizienz miteinander verglichen.

Ein besonderes Augenmerk galt dem Vergleich zwischen mutierter DNA und dem entsprechenden Wildtyp. Dabei sollten vor allem Konformationsänderungen, welche durch die Basenmodifikation hervorgerufen werden, charakterisiert werden. Hierzu wurden Basenmodifikationen, die durch Alkylierungsschäden und Oxidationen entstehen, genauer untersucht. Da der Erkennungsmechanismus dieser Schäden unbekannt ist, liegt der Fokus dieser Arbeit bei einem Vergleich zwischen geschädigter und intakter DNA in Lösung. Die Annahme ist, dass die Mutation zu lokalen Veränderungen der DNA führt, sodass spezifische Konformationen eingenommen werden, welche vom Reparaturenzym erkannt werden.

Diese Arbeit befasst sich neben dem Einfluss von Basenmodifikationen und deren Einfluss auf die Feinstruktur der DNA mit Sekundärstrukturen von Nukleinsäuren. So wurden sowohl *Dangling-End* Motive, und deren Einfluss auf die Gesamtstabilität der Helix, als auch *RNA-Bulge-Strukturen* mit Hilfe von kMDS und H-REMD untersucht. Im Falle der *Dangling-End* Motive wurden Umbrella-Sampling Simulationen und MM/PBSA Rechnungen durchgeführt, um den Einfluss genauer zu untersuchen. Ziel der RNA-Bulge Simulationen war der Vergleich und die Konformationsanalyse unterschiedlicher Bulge-Basen im gleichen Sequenzkontext.

1.2 Aufbau und Stabilität von Nucleinsäuren

Nucleinsäuren sind Biomoleküle die sich aus einzelnen Monomeren, den Nucleotiden, zu einer Polynucleotidkette zusammensetzen. Ein Nucleotid besteht aus aus einem Zucker, einer heterozyklen, stickstoffhaltigen Base und einer Phosphatgruppe. Handelt es sich bei dem Zucker um eine Ribose, so spricht man von Ribonucleotiden. Ist dieser Baustein ein Derivat der Ribose, die Desoxyribose, so handelt es sich um Desoxyribosenucleotide. Die stickstoffhaltige Base wird in die Familie der Purine bzw. Pyrimidine unterteilt. Zu den Purinen gehören Adenin (A) und Guanin (G). Die Vertreter der Pyrimidine sind Cytosin (C), Thymin(T) und Uracil(U) (3).

Die Basen sind kovalent an den Zucker gebunden, der wiederum an die Phosphatgruppe gebunden ist. Die Phosphatgruppe eines Nucleotides ist kovalent an die OH-Gruppe des C3-Kohlenstoff Atoms der nächsten Zucker gebunden, wodurch eine zweite Phosphoesterbindung entsteht. Desoxyribonucleinsäuren (DNA) und Ribonucleinsäuren (RNA) unterscheiden sich in der Zusammensetzung dieser Bausteine. Neben dem unterschiedlichen Zucker besitzen DNA-Moleküle A, C, G und T, wohingegen die RNA Moleküle die Basen A, C, G und U (3) (siehe Abbildung 1.1) enthalten. Wie aus Abbildung 1.2 zu entnehmen ist, sind die einzelnen Nucleotide durch eine 3'-5'-Phosphodiesterbindung miteinander verknüpft. Das Nucleotid am 5'-Ende ist durch eine freie Phosphorylgruppe charakterisiert, wohingegen das Nucleotid am 3'-Ende eine Hydroxylgruppe besitzt.

Ein DNA-Molekül besteht aus zwei Polynucleotidketten, die anti-parallel zueinander verlaufen und über Wasserstoffbrücken miteinander interagieren.

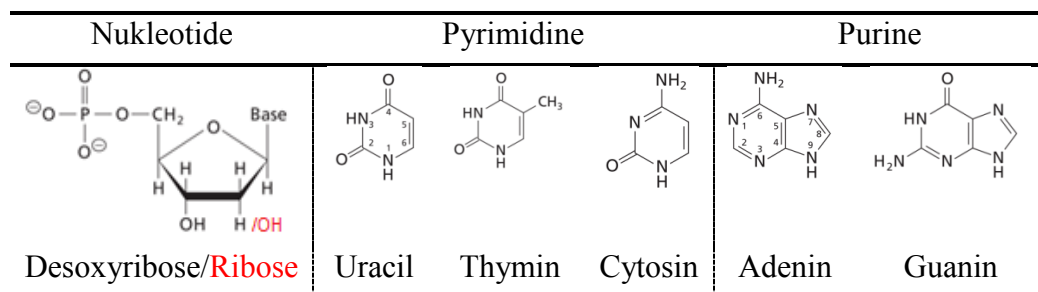


Abbildung 1.1 Zusammensetzung von Nucleotiden. Nucleotide setzen sich zusammen aus einem Zucker, einer stickstoffhaltigen Base und einer Phosphatgruppe. RNA Strukturen unterscheiden sich hinsichtlich des Zuckers (Ribose statt Desoxyribose) und der Basenpaarung (A:U statt A:T)

Sie bilden dabei eine helikale Struktur aus. Zusätzlich ist zu erwähnen, dass beide Stränge komplementär zueinander sind. In einem komplementären DNA-Doppelstrang bilden G:C und A:T (A:U im Falle von RNA) ein Watson-Crick Basenpaar (4).

Da ein DNA-Molekül aus zwei nicht-kovalent miteinander verbundenen Polynukleotidketten besteht, stellt sich die Frage, wodurch die Stabilität eines DNA Moleküls gewährleistet ist. Wie bereits erwähnt bildet sich ein komplementärer DNA-Strang zu einer Helix aus. Diese Helix ist jedoch nicht nur durch Wasserstoffbrücken stabilisiert. Die π - π Wechselwirkungen der Basen führt dazu, dass eine hydrophobe Umgebung zwischen den Basen entsteht, sodass der hydrophile Zucker-Phosphat-Rückgrat der DNA weggedreht wird und mit der wässrigen Umgebung interagieren kann. Die gestapelt übereinander liegenden (engl. stacking), hydrophoben Basen sind dadurch schwer zugänglich für Wasser, wodurch die Wasserstoffbrücken zwischen den Basen zusätzlich stabilisiert werden.

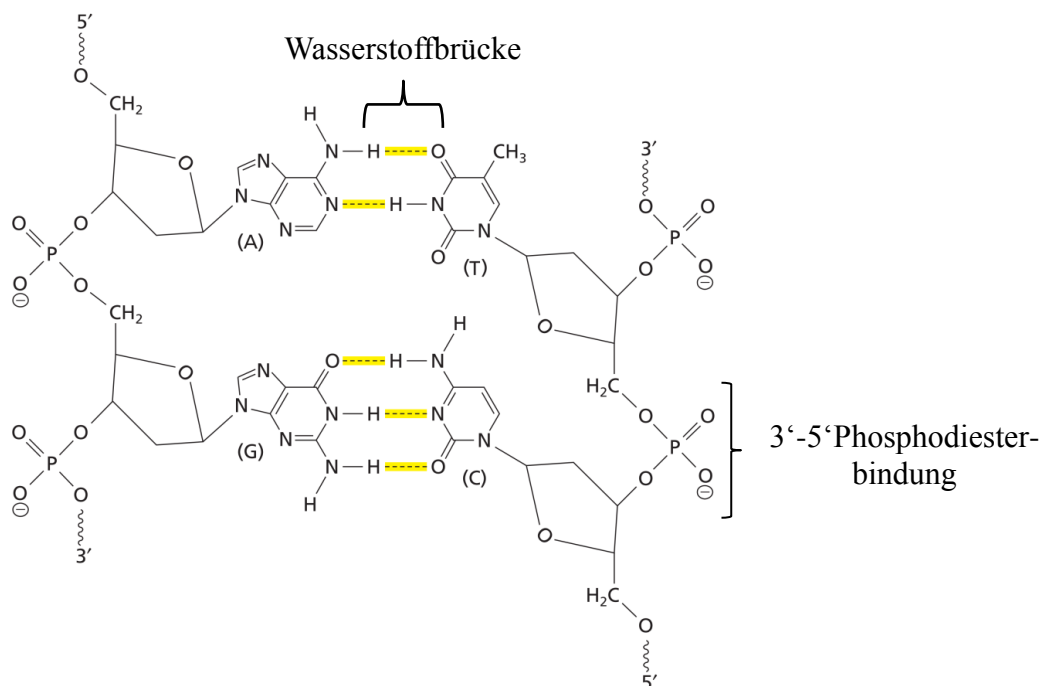


Abbildung 1.2 Basenpaarungen in DNA Strukturen. Die beiden Einzelstränge verlaufen in entgegengesetzte Richtungen. Die Adenin: Thymin Basen und die Guanin: Cytosin Basen bilden jeweils ein Watson-Crick Paar. Durch die Basenpaarung entsteht die optimale Konstellation für Wasserstoffbrücken. Zudem sind Nukleotide über eine Phosphodiesterbindung miteinander verknüpft. Quelle: modifiziert nach Moran, Horton, Scrimgeour, Perry, 2012(5)).

Aus dem Zusammenspiel dieser verschiedenen Kräfte ergibt sich die Form und Geometrie der DNA-Doppelhelix. So entstehen auf der Helixoberfläche zwei unterschiedliche Furchen, die sich in ihrer Spannweite unterscheiden (engl. minor und major groove). Aus den π - π Wechselwirkungen ergeben sich spezifische Distanzen zwischen benachbarten Basen (siehe Abbildung 1.3). Die Zusammensetzung der Basen hat aufgrund der unterschiedlichen Anzahl von gebildeten Wasserstoffbrücken einen Einfluss auf die Stabilität der DNA-Helix. G:C reiche Sequenzen bzw. Strukturen weisen eine höhere Stabilität auf als A:T reiche Strukturen.

Die elektrostatische Abstoßung durch die negativ geladenen Phosphatgruppen verursacht eine Destabilisierung der DNA-Helix, die durch Gegen-Ionen wie z.B. Mg^{2+} reduziert wird. Somit entsteht eine Abhängigkeit zwischen der Stabilität der DNA-Struktur, der Polarität und dem Salzgehalt der Umgebung.

Bei der hier beschriebenen Form der DNA-Struktur handelt es sich um B-DNA Strukturen. Jedoch existiert eine Abhängigkeit zwischen der Struktur der DNA und ihrer Umgebung. Ist die DNA dehydriert so nimmt sie eine A-DNA Form an. Auch die Basen-Zusammensetzung hat einen direkten Einfluss auf die Form der DNA. In einigen G:C-reichen Sequenzen wird eine Z-Form bevorzugt (siehe Abbildung 1.4) (6).

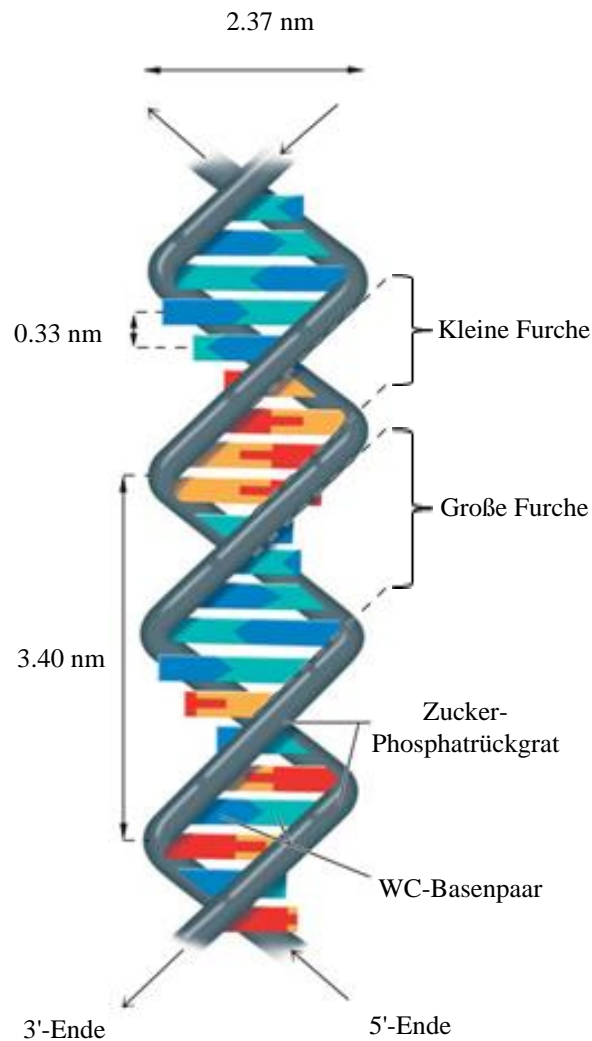


Abbildung 1.3 Modell einer B-DNA Struktur . Wie aus dem Modell schematisch dargestellt ist, befinden sich die Basen innerhalb der Helix, wohingegen das Zucker-Phosphat-Rückgrat die Basen umschließt. Durch die Helix-Windung und die Stacking-Interaktionen der Basen entstehen eine kleine und eine große Furche. Der Anstieg pro Windung beträgt 3.4 nm. Die Distanz zwischen zwei Basen beträgt etwa 0.33 nm und der Gesamtdurchmesser 2.37 nm (Quelle: modifiziert nach Moran, Horton, Scrimgeour, Perry, 2012(5)).

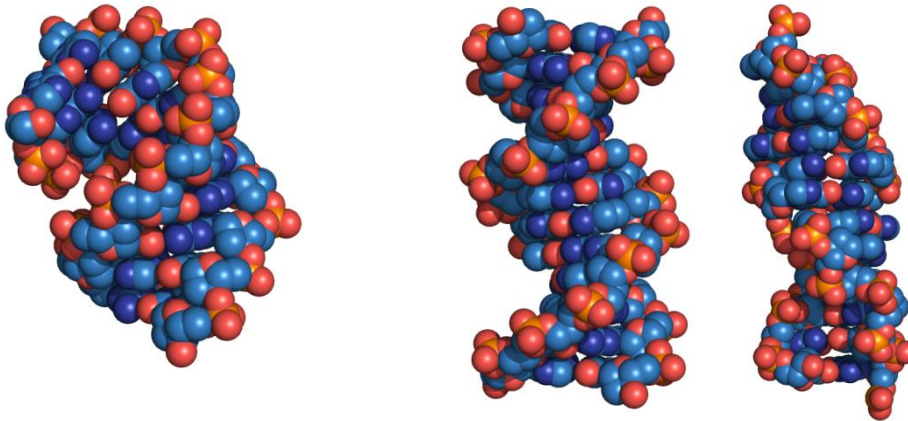


Abbildung 1.4 A-DNA, B-DNA und Z-DNA Konformationen. Die A-DNA Konformation ist bevorzugt in einem dehydrierten Zustand (Links, PDBid 401D(7)). B-DNA (Mitte, PDBid 423D(8)) ist die natürlich vorliegende Konformation. Die Z-DNA Konformation (Rechts, PDBid 279D(9)) wird in einigen G/C-reichen Sequenzen bevorzugt.

1.3 Sekundärstrukturen von Nucleinsäuren

1.3.1 RNA-Bulge

RNA Moleküle können unterschiedliche Sekundärstrukturen ausbilden (Haarnadel-Schleife, Innere-Schleife, Ausbeulung; engl. *Hairpin-Loop*, *Internal-Loop*, *Bulges*) und somit sehr komplexe drei-dimensionale Strukturen einnehmen (10). Zudem besitzen diese Sekundärstrukturelemente spezifische biologische Funktionen (11–13). Durch diese Elemente nehmen RNA Strukturen einen größeren Konformationsraum als DNA Strukturen ein. Die enorme strukturelle Diversität, hervorgerufen durch nicht-helikale Motive, erschwert die Vorhersage der 3D-Strukturen von RNA Molekülen, ausgehend von ihrer Primärsequenz (14–16).



Abbildung 1.5 Elemente von RNA Sekundärstrukturen. Vier Beispiele für Sekundärstruktur-Elemente in RNA Strukturen: (a) Helix, (b) Bulge-Base (c) Fehlpaarungen und (d) Hairpin-Strukturen

Besonders häufig treten hierbei Bulge-Motive in RNA-Strukturen auf, die zudem für die molekulare Erkennung und Funktion der Struktur wichtig sind (17, 18). Bulge Strukturen sind ungepaarte Nukleotide innerhalb der RNA-Doppelhelix. Die Anzahl der ungepaarten Basen kann variieren. Die strukturelle Diversität einer RNA-Doppelhelix, die ausschließlich kanonische Watson-Crick Paarungen ausweist, ist limitiert durch die Anzahl der möglichen Wasserstoffbrückenbindungen und dem π - π -Stacking Interaktionen der aromatischen Ringe (19). Diese Form der Interaktion ist bei einer Bulge-Struktur nicht zwingend vorhanden. Eine Bulge-Struktur kann sowohl inter-, als auch intra-helikale Strukturen einnehmen. Interhelikale Bulge-Basen können die reguläre A-Form der RNA verändern und eine Krümmung der Gesamtstruktur hervorrufen (20). Stabilisiert wird diese Struktur vor allem durch hydrophobe Wechselwirkungen zwischen den Basen. Liegt die Bulge-Struktur extra-helikal vor (*looped out*-Konformation) so nimmt die RNA eine A-Form ähnliche Konformation ein (20). Hierbei sind die Bulge Basen völlig solvatisiert. Die bevorzugte Konformation ist schwierig zu ermitteln und abhängig von vielen Parametern (benachbarte Basenpaare, Ionenkonzentration, Solvatationseffekten usw.) (21–23). Diese Abhängigkeit ist ausschlaggebend für die Limitierung durch experimentelle Methoden. In NMR Studien wurde nachgewiesen, dass Bulge-Strukturen mit nur einem Adenin (Single-Bulge) generell interhelikal vorliegen. In Kristallstrukturen liegen sie aufgrund der Salzkonzentration generell im looped-out Zustand vor (20, 24–26).

Um die biologische Relevanz von *RNA-Bulge-Motiven* zu untermauern werden im Folgenden einige Beispiele-Motive diskutiert:

- Das *Trans-activation response element* (TAR-Element) ist ein regulatorischer Abschnitt der TAR-RNA und Bestandteil des genetischen Codes von HIV-Viren. Die TAR-RNA setzt sich zusammen aus einem Stamm (stem), einer Ausstülpung aus drei Nukleotiden (Bulge) am Stamm und einer Schleife (Loop) und dient als Bindestelle für den Transkriptionsfaktor Tat und zelluläre Transkriptionsfaktoren. Die Bulge-Sequenz besitzt hierbei eine besondere Funktion. Das HIV-1 Tat Protein bindet an die Bulge-Sequenz, wodurch die Transkripte stabilisieren und der Elongationsprozess der Transkription fortgesetzt wird (27–29).

- In allen RNA P4 Elementen der bakteriellen RNaseP befindet sich ein U-Bulge. Es ist Bestandteil einer Pseudoknoten-Struktur und befindet sich im Zentrum der katalytischen Domäne (30–33).
- Das Hüllprotein der Bakteriophage MS2 bindet über einen kleinen Loop-Bereich der RNA das genetische Material des Virus. Als Ankerpunkt zwischen dem Loop und dem Protein dient ein Adenin-Bulge, welches an die hydrophobe Bindetasche des Proteins bindet (34, 35).

Ziel dieser Arbeit ist es den Konformationsraum eines U-Bulges und eines A-Bulges im gleichen Sequenzkontext unter identischen Bedingungen miteinander zu vergleichen. Hieraus können generelle Eigenschaften für Bulge Motive abgeleitet werden. Insbesondere sollen, mit Hilfe von effizienten Simulationsmethoden, wie der H-REMD Methode, Intermediate-Zustände ermittelt und miteinander verglichen werden.

1.3.2 Dangling-End

Die Faltung und Hybridisierung von Nukleinsäuren ist ein komplexer Prozess, der durch eine Vielzahl von Faktoren beeinflusst wird. Ausschlaggebend ist jedoch die spezifische Interaktion der Watson-Crick Partner (WC-Partner). Falls jedoch die Einzelstränge nicht perfekt zusammenpassen, oder sich die Einzelstränge in ihrer Länge unterscheiden, können überhängende Basen am Ende der Doppelhelix entstehen. Diese überhängenden Basen, werden im Folgenden als *Dangling-End* bezeichnet. In solch einem Fall ist die perfekte Bindung der Einzelstränge nicht möglich bzw. notwendig (36–39).

Bei den *Dangling-Ends* handelt es sich um Sekundärstrukturen, die sowohl in DNA als auch in RNA auftreten können. Sie spielen speziell für RNA Strukturen in diversen biologischen Prozessen eine wichtige Rolle. Beispiele sind unter anderem Codon-Anticodon Interaktionen zwischen tRNA und mRNA oder die Erkennung der tRNA (40–42). Um diese spezifischen Interaktionen sowie Faltungs- und Hybridisierungsprozesse besser zu verstehen, ist eine detaillierte Betrachtungsweise der strukturellen und thermodynamischen Eigenschaften von *Dangling-End* Basen notwendig. Experimentell wurde bereits nachgewiesen, dass *Dangling-End* Basen einen stabilisierenden Einfluss auf die Gesamtstruktur haben.

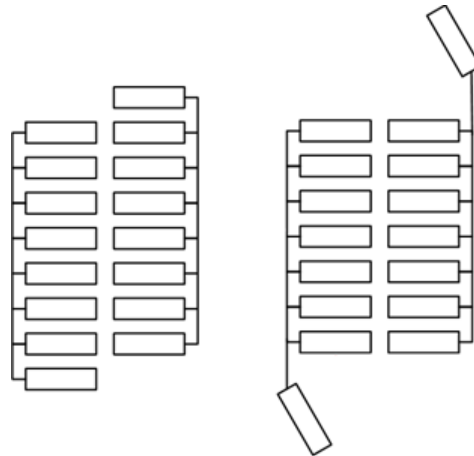


Abbildung 1.6 Schematische Darstellung von *Dangling-End* Basen. Die *Dangling-End* Base kann mit den terminalen Basen in einem gestackten Zustand (links) oder völlig solvatisiert (rechts) vorliegen.

Im Allgemeinen haben RNA-Dangling-End-Strukturen einen stärkeren Einfluss als DNA-Dangling-End-Strukturen (39). Weiterhin spielt die Orientierung und Position der Dangling-End-Struktur eine wichtige Rolle. Im Falle von RNA wirkt ein Dangling-End am 3'-Ende des Stranges stabilisierender als ein Dangling-End am 5'-Ende (43, 44). Interessanterweise verhält es sich bei DNA Strukturen entgegengesetzt (39, 45). Ursache hierfür können unterschiedliche geometrische Einschränkungen sein, die ein cross-stacking ermöglichen bzw. verhindern (46). Einen Hinweis dafür lieferten Femtosekunden-Spektroskopie Messungen mit RNA Strukturen (47). So konnte für Purine Basen am 3'-Ende eine starke Interaktion mit dem terminalen Basenpaar nachgewiesen werden. Im Vergleich hierzu zeigen NMR Untersuchungen, dass Dangling-End-Strukturen am 5'-Ende einen größeren Konformationsraum einnehmen können und häufig keine Interaktion mit den terminalen Basen eingehen (48, 49, 46). Diesen vollständig solvatisierten Zustand bezeichnet man als unstacked-Zustand (siehe Abbildung 1.6). Neben der Position der *Dangling-Ends* (3' vs.- 5') sind auch die Basenzusammensetzung (Purine vs. Pyrimidin), die zum *Dangling-End* benachbarten Basen, die Salzkonzentration sowie die Anzahl der Basen des Dangling-End entscheidend für die Stabilität der Gesamtstruktur (37, 50, 51). So kann ein Dangling-End Base abhängig von seiner Form und Größe einen Schutz vor dem Lösungsmittel für die terminalen Basen bieten und somit ihre WC-Interaktion stabilisieren (52, 53). Basen mit einer großen Oberfläche (A und G) wirken daher stabilisierender als T und C. Weitere Untersuchungen zeigen einen sequenzspezifischen Einfluss von *Dangling-End-Strukturen*. Beispielsweise

besitzen ein 5'-AC/ 3'-G, wobei das A ein 5'-*Dangling-End* und das C der unmittelbare Nachbar am gleichen Strang ist unterschiedliche thermodynamische Eigenschaften als ein 5'-AG/ 3' C-Paar (45).

Die spezifischen Interaktionen zwischen *Dangling-End* Basen und Terminalen Basen sowie die Konformationsräume von 3'- und 5'-*Dangling-End-Strukturen* werden in dieser Arbeit für RNA und DNA Strukturen miteinander verglichen. Ziel ist es hierbei mit Hilfe von Molekulardynamik Simulationen einen besseren Einblick in die Thermodynamik und strukturelle Flexibilität der *Dangling-End-Strukturen* zu erhalten.

1.4 DNA-Schäden und ihre Auswirkungen auf die Feinstruktur der DNA

Die DNA ist das einzige zelluläre Makromolekül, welches von Enzymen erkannt und repariert werden kann (5). Man unterscheidet hierbei verschiedene Typen von DNA Schäden wie z.B. Basenmodifikationen, Fehlpaarungen, Cross-Linking von DNA Strängen oder Brüche des Phosphatrückgrats (54–60). Werden diese Schäden nicht erkannt, kann dies zu Zelltod, Gewebedegeneration oder zur Tumorbildung führen. Während einige DNA Schäden durch endogene Prozesse oder Umweltwelteinflüsse hervorgerufen werden (siehe Abbildung 1.7), entsteht der Großteil der Schäden durch Fehlpaarungen während der DNA-Replikation (61). Basenmodifikationen die trotz des Proofreading-Prozesses der DNA-Polymerase bestehen bleiben, können anhand von spezifischen Enzymen erkannt und repariert werden (62, 61). Diese Enzyme unterscheiden sich anhand ihrer Substratspezifität und ihrer Kinetik. Es wird auf folgende Reparaturmechanismen eingegangen: Direkte Reparatur, Basen-Exzisionsreparatur und Nukleotid-Exzisionsreparatur.

Die direkte Reparatur kommt ohne ein Aufbrechen des Phosphatrückgrats aus. Durch eine enzymatische Reaktion wird die kovalente Bindung einer Läsion mit der DNA aufgebrochen, sodass die Base in ihrer ursprünglichen Form wieder vorliegt. Beispiele für eine direkte Reparatur sind unter anderem die Reparatur von Photodimerisierungs-Prozessen (z.B. Thymin-Dimer) oder Alkylierungsschäden (z.B. O⁶-Methylguanin) (56, 63).

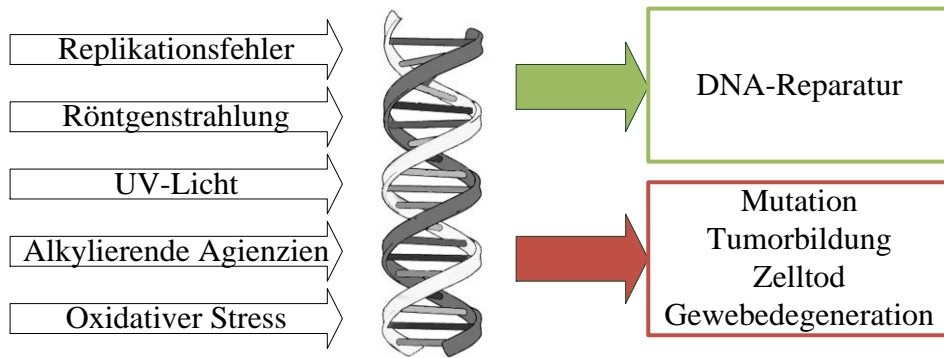


Abbildung 1.7 Ein genereller Überblick möglicher DNA Schäden und deren Folgen. DNA Schäden können durch eine ganze Reihe von Faktoren hervorgerufen werden. Auf der Abbildung sind links einige dieser Faktoren beispielhaft als Pfeile dargestellt. Wird der Schaden nicht erkannt und repariert, können die Folgen gravierend sein (Mutation, Tumorbildung, Zelltod, Gewebedegeneration).

Ionisierende Strahlung oder oxidativen Stress können Basenmodifikationen hervorrufen. Diese Basenmodifikationen werden im Allgemeinen über einen Exzisions-Reparatur-Mechanismus behoben. Die Basen-Exzisions-Reparatur wird am Beispiel eines oxidativen Schaden am Guanin, dem 8-Hydroxydesoxyguanin (8oxoG) diskutiert, da es sich hierbei um einen der am häufigsten auftretenden oxidativen Schäden handelt. Ursache hierfür ist das geringe Oxidationspotential des Guanins im Vergleich zu den anderen Basen. 8oxoG entsteht aufgrund von Reaktionen zwischen Hydroxyradikalen mit Kohlenstoff C8 des Guanins. Behoben wird dieser Schaden durch DNA-Glykolyasen. Hierbei wird die N-glykosidische Bindung des 8oxoG hydrolysiert, wobei eine abasische Seite (AP-Site) entsteht. Die AP-Site wird wiederum von einer AP-Endonuklease erkannt und entfernt die Desoxyribose (64, 65, 55, 66, 59). Die entstandene Lücke wird von der DNA-Polymerase β erkannt und mit Hilfe einer Ligase geschlossen (siehe Abbildung 1.8).

Wie schon bei Basen-Exzisionsreparatur handelt sich bei der Nukleotid-Exzisionsreparatur um einen Prozess bei dem eine Reihe von Enzymen beteiligt ist.

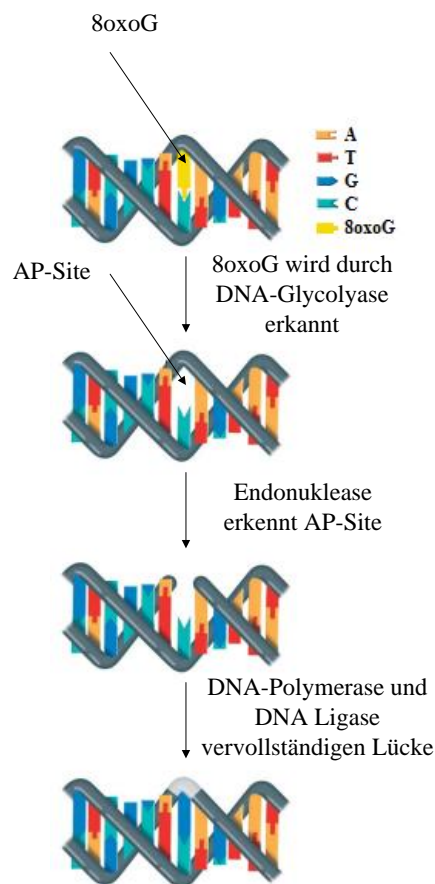


Abbildung 1.8 Reparaturmechanismus am Beispiel des 8oxoG: Für die vollständige Ersetzung der 8oxoG Base zu dem ursprünglichen Guanin werden mehrere Enzyme und Schritte benötigt. Die Erkennung der Base findet durch DNA-Glykolyasen statt, welches die Bindung zwischen Zucker und Base aufbricht, sodass eine AP-Site entsteht. Diese wird wiederum durch Endonukleasen erkannt, welche das Nucleotid aus dem Doppelstrang vollständig entfernt. Die entstandene Lücke wird durch die DNA-Polymerase und DNA-Ligase geschlossen. (Quelle: modifiziert nach Moran, Horton, Scrimgeour, Perry, 2012(5)).

Man unterscheidet grundsätzlich zwischen zwei Mechanismen bei der Nucleotid-Exzisionsreparatur (67–69):

- Global genomische Reparatur (engl. Global genom repair, GCR). Hierbei wird der Schaden durch das Heterodimer XPC-GR23B, DDB1-DDB2-Komplex und dem Gadd45 α erkannt (70).
- Transkriptionsabhängige-Reparatur (engl. Transcription coupled repair, TCR). Eingeleitet wird dieser Prozess durch die Proteine CSA und CSB, welche in der Lage sind die RNA-Polymerase II vom DNA-Schaden zu entfernen. Dadurch wird es den TCR-Proteinen ermöglicht an die DNA zu binden (71).

Die folgenden Schritte der DNA-Reparatur sind bei beiden Mechanismen äquivalent. Die Proteine XPA, RPA und der TFIIH-Komplex entwinden im Bereich des DNA-Schadens die Helix und ermöglichen den Endonukleasen XPG und XPF/ERCC1 das Entfernen eines 20-30 Nukleotid-langen DNA-Molekül. Die entstandene Lücke wird durch die DNA-Polymerase δ und ϵ neu hybridisiert und mit Hilfe der DNA-Ligase I mit dem DNA Strang verbunden (67).

1.4.1 7,8-dihydro-8-oxoguanin

Oxidative Schäden können zu DNA-Mutationen und zur Krebsentstehung führen (61, 54, 72). Ursache für die Modifikation sind freie Radikale oder reaktive Sauerstoffspezien, die mit den Nukleinsäure-Basen reagieren können (72). Das Verständnis über die Ursachen für einen oxidativen Schaden, sowie die detaillierte Untersuchung des Reparaturmechanismus, sind von entscheidender Bedeutung um Strategien gegen Mutationskrankheiten zu entwickeln.

Besonders häufig tritt ein oxidativer Schaden an der Base Guanin auf. Hierbei kommt es zu einer Oxidation vom Kohlenstoff C8 und dem Stickstoff N7. Als Resultat dieser Reaktion entsteht das 7,8-dihydro-8-oxoguanine (8oxoG). 8oxoG besitzt die besondere Eigenschaft mit unterschiedlichen Basen eine Interaktion einzugehen zu können. Neben der Watson-Crick Paarung mit einem Cytosin, kann 8oxoG unter anderem auch mit Adenin, in der sogenannten Hoogsteen Konformation, vorliegen (73–77). Dies führt, falls es von keinem Reparaturenzym erkannt wird, zu einer Mutation von G:C \rightarrow A:T (siehe Abbildung 1.9). Um solch eine Mutation zu vermeiden, existieren Glykolyasen, wie das bakterielle MutM oder die humane variante 8oxoG DNA Glykolyase (hOgg1) (78, 59, 79).

Es existieren bereits eine Vielzahl von Studien bezüglich dem Oxidations- und Reparaturprozess (59, 76, 77, 79–81). Strukturelle Untersuchungen zeigen, dass die Wasserstoffbrückenbindung zwischen 8oxoG und der gegenüberliegenden Base aufgebrochen werden muss, damit 8oxoG-Base in die aktive Region des Enzyms herumklappen kann. Anschließend wird mit Hilfe des Enzyms die Bindung zwischen der Base und dem Zucker aufgebrochen, sodass eine abasische Stelle (AP-Site) entsteht. Die Glykolyase dissoziiert sich von der DNA, sodass die AP-Site von einer Endonuklease erkannt werden kann. Es ist jedoch weiterhin unklar, wie das Enzym den DNA Schaden erkennt.

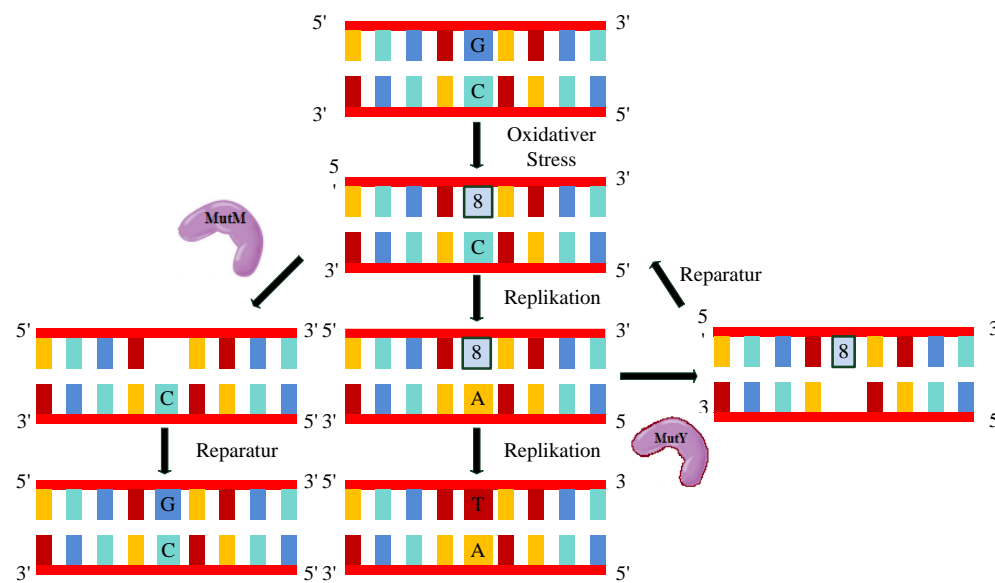


Abbildung 1.9 Basenexzisionsreparatur Mechanismus der 8oxoG-Base. Durch oxidativen Stress kann die Guanin-Base zu 8-oxoguanine modifiziert werden. Diese 8oxoG:C Fehlpaarung kann, falls sie nicht erkannt wird, zu einer irreversiblen T:A Mutation führen (mittlerer Pfad). Die humane DNA-Glykolyase hOGG1 bzw. die bakterielle MutM Variante sind in der Lage die 8oxoG Base zu erkennen. Für den weiteren Verlauf dieses Pfads sei an dieser Stelle auf Abbildung 1.8 verwiesen. Wird die 8oxoG:C Fehlpaarungen nicht erkannt, entsteht nach dem ersten Replikationsschritt eine 8oxoG:A Paarungen. Diese wieder kann durch das Enzym MutYh bzw. durch die bakterielle Variante MutY erkannt werden. Hierbei wird jedoch die Adenin Base entfernt. Der weitere Verlauf entspricht Abbildung 1.8. In beiden Fällen entsteht eine AP-Site die wiederum durch Endonukleasen erkannt wird. Das 8oxoG:C Paar kann anschließend wieder durch hOGG1 bzw. MutM erkannt werden.

Eine Möglichkeit ist, dass die geschädigte Base in einem extra-helikalen Zustand vorliegt und dadurch erkannt wird. Alternativ findet ein intra-helikaler Erkennungsprozess statt. Dies setzt einen aktiven Prozess voraus, in dem die Watson-Crick Paarung der modifizierten Base und der Nachbar Base aufgelöst wird und die Base aus der Helix herausgedreht wird. Dieser Zustand wird im Folgenden als *looped-out* Zustand bezeichnet. Das Enzym muss also in der Lage sein, während des Bindeprozesses die Base effizient zu erkennen und herauszudrehen. Annahme ist, dass bereits in der Feinstruktur der oxidierten Base strukturelle Unterschiede zu Guanin vorhanden sind.

Diese strukturellen Unterschiede könnten Hinweise für das Reparaturenzym liefern. Verdin *et. al* lieferten für diese Hypothese erste Hinweise (78, 66).

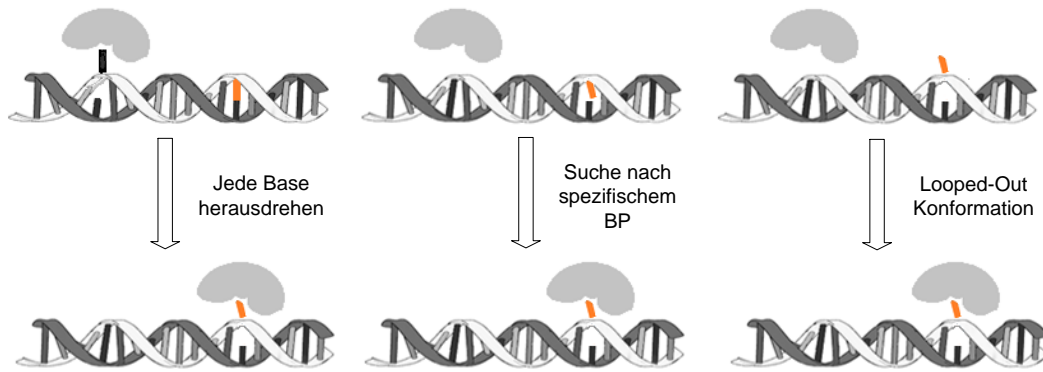


Abbildung 1.10 Mögliche Erkennungsmechanismen von modifizierten DNA Basen.

So konnten sie bei DNA Strukturen in Komplex mit dem Enzym nachweisen, dass bereits in dem inter-helikalen Zustand (*looped-in*) strukturelle Unterschiede den Konformationen des Zucker (Zucker-Pucker, engl. Sugar-Pucker) und Phosphatrückgrat zwischen Guanin und 8oxoG vorliegen. Experimentelle DNA Strukturen, die 8oxoG als Baustein besitzen, weisen jedoch keine Unterschiede in der Basenpaarungsgeometrie, Wasserbrückenbindung oder Stacking-Energien zwischen 8oxoG:C und G:C auf (82).

Die Charakterisierung möglicher spezifischer Zustände von 8oxoG ist daher von enormer Bedeutung, da sie einen Hinweis für die Erkennung durch das Reparaturenzym liefern können. Diese spezifischen Zustände sollen in dieser Arbeit mit Hilfe von Molekulardynamik Simulationen ermittelt werden.

1.4.2 O⁶-Methylguanin

Neben oxidativen Agenzien können auch alkylierende Verbindungen Mutationsschäden hervorrufen (83). Es handelt sich hierbei um elektrophile Substanzen, die Ringstickstoffe, Ringsauerstoffe oder die Phosphatgruppe angreifen. Es existieren zwölf potentielle Reaktionszentren zwischen DNA-Basen und alkylierenden Substanzen (84, 85). Bei dieser Reaktion (nukleophilen Substitution) unterscheidet man zwischen S_N1 und S_N2 Reaktionen.

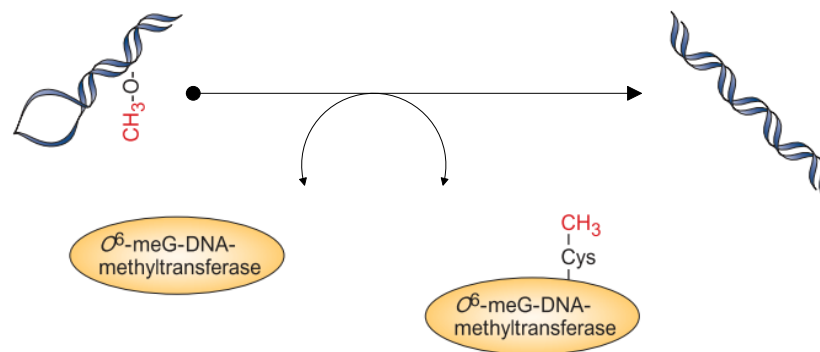


Abbildung 1.11 Beispiel einer Direkten-Reparatur am Beispiel der O⁶-MeG Base. O⁶-MeG Base (dargestellt als CH₃-O) wird direkt demethyliert durch die O⁶-MeG-DNA-Methyltransferase. Das aktive Protein überträgt die Methylgruppe auf ein Cystein-Rest im Aktiven Zentrum und geht aber in einen irreversiblen inaktiven Zustand über (modifizierte Abbildung von (85)).

S_N1 Verbindungen wie N-methyl-N'-nitro-N-nitrosoguanidine reagieren mit dem O₆-Sauerstoff des Guanin zu O⁶-Methylguanin (O⁶-MeG) (85). Wird dieses Produkt nicht von Reparaturenzymen erkannt, kann es zu einer G:C → A:T Mutation, oder einem DNA-Strangbruch kommen (86). S_N2 Verbindungen, wie zum Beispiel Methylmethanesulphonat reagieren mit Adenin zu 1-Methyladenine oder mit Cytosin zu 3-Methylecytosin. Diese Art von Schäden treten häufig in einer Einzelstrang-DNA oder RNA Strukturen auf (56).

O⁶-Methylguanin (O⁶-MeG)-Basen haben einen schwerwiegenden Einfluss auf die Funktionsfähigkeit einer Zelle und können zum Zelltod oder Tumorbildung führen (87–90). Die Methylierung des Guanins hat zur Folge, dass die Doppelbindung zwischen dem C₆ Atom und dem O₆ Atom zu einer Einfachbindung wird, sodass die gesamte Tautomerie des Guanins verändert wird. NMR-Studien an einer O⁶-MeG:C Fehlpaarung in einem Oligomer-Duplex haben gezeigt, dass für beide Basen die glykosidische Bindung in einer *anti*-Konformation vorliegt (91). Die B-DNA Form bleibt weiterhin erhalten, jedoch mit einer *wobble*-Orientierung der Wasserstoffbrücken und ausschließlich einer Wasserstoffbrücke zwischen den Atomen N₂ (O⁶-MeG) und O₂ (C). Die Kristallstruktur einer Z-DNA mit dieser Fehlpaarung zeigt eine Anordnung, die kompatibler mit Standard-Watson-Crick-Wasserstoffbrücken-Abstände ist (92). Für die O⁶-MeG:T Fehlpaarung in Kristallstrukturen wurde ein Watson-Crick ähnliches Wasserstoffbrücken-Muster beobachtet (93, 94). NMR Studien zeigten hingegen, dass nur eine Wasserstoffbrücke zwischen diesen Basen besteht (95). Es ist an dieser Stelle anzumerken, dass sich der Sequenzkontext zwischen

Kristallstrukturen und der NMR Studie unterscheidet, sodass angenommen werden kann, dass der Sequenzkontext Ursache für die unterschiedlichen Wasserstoffbrücken-Muster ist. Die generelle Struktur des O⁶-MeG:T Basenpaares zeigt sowohl in der NMR- als auch in der Kristallstruktur Ähnlichkeiten zu einem G:C Paar, während für O⁶-MeG:C eine *wobble*-Orientierung beobachtet wird. Dieser Unterschied erklärt jedoch nicht, warum eine schwache Präferenz eines O⁶-MeG:T Basenpaaren Einbaus, während der DNA-Replikation besteht

Die Erkennung und Reparatur einer O⁶-MeG-Base kann unterschiedlich erfolgen:

- direkte oxidative Desalkylierung → Die O⁶-MeG-Base kann durch Oxygenasen der AlkB-Familie direkt entfernt werden (96).
- Basenexzisionsreparatur-Glykosylasen → O⁶-MeG-Base wird aus der DNA-Doppelhelix entfernt (97).
- O⁶-Alkylguanin-DNA-Transferasen (AGTs) → Die aktive O⁶-Methylguanin-DNA-Methyltransferase ist in der Lage die O⁶-MeG-Base zuerkennen und transferiert die Alkyl-Gruppen auf ein Cystein-Rest in der aktiven Bindetasche (60).

Da der O⁶-MeG Schaden am häufigsten über AGTs repariert wird, wird im Folgenden näher auf diesen Mechanismus eingegangen (85). Die Alkyl-Gruppe der O⁶-MeG Base wird auf einen Cystein-Rest (Cys145) in seinem aktiven Zentrum des Proteins transferiert. Das Protein geht dabei in einen inaktiven irreversiblen Zustand über und zerfällt anschließend durch Ubiquitinierung.

Die aktive O⁶-Methylguanin-DNA-Methyltransferase besitzt ein Helix-Turn-Helix (HTH)-Motiv, welches mit der kleinen Furche des DNA Doppelstranges interagiert (60). Diese Interaktion hat eine Konformationsänderung der DNA zur Folge, sodass sich die Spannweite der kleinen Furche der DNA vergrößert und die Interkalierung eines Arginin-Fingers in die Helix ermöglicht (98). Die alkylierte Base wird anschließend in die aktive Tasche herausgedreht. Ein konservierter Tyrosinrest induziert zusätzlich über sterische Hinderung und Ladungsabstoßung die Rotation des 3'-Phosphatrestes der alkylierten Base. Die nun ungepaarte Cytosin-Base interagiert mit einem Arginin wodurch die O⁶-MeG Base zusätzlich stabilisiert wird (99–102).

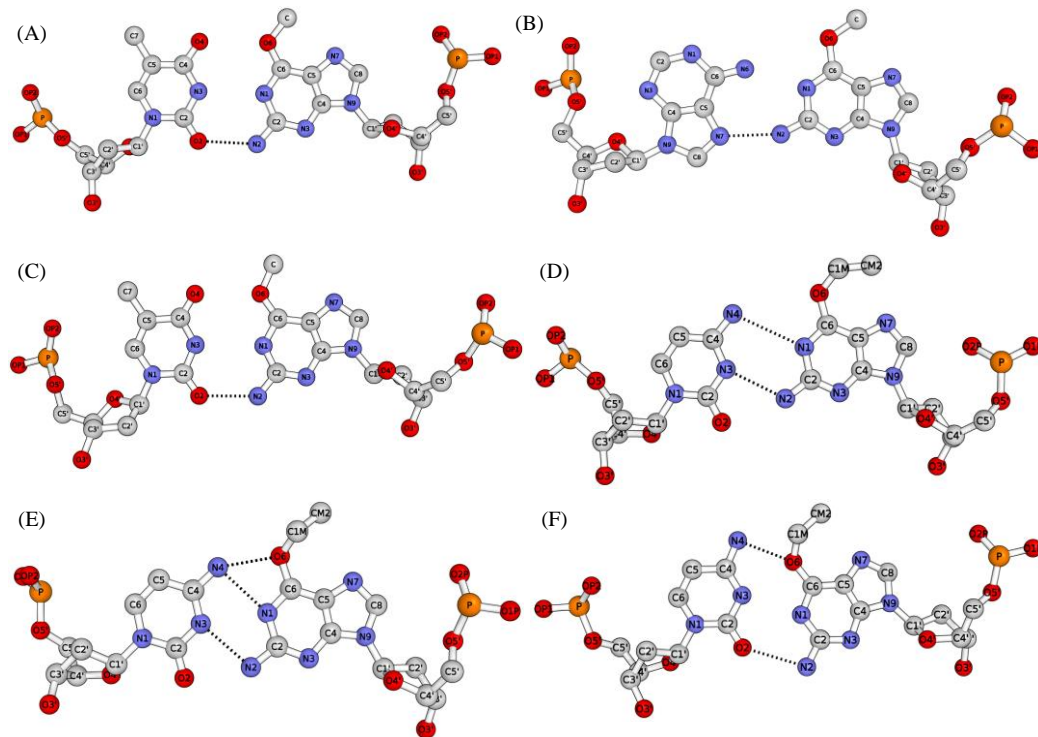


Abbildung 1.12 Konformationszustände von alkylierten Guanin Basen. (A) PDBid 1D27(94) (B) 153D(103) (C) 218D(93) (E) 1D95 (F) 144D(104).

1.5 Untersuchung von Nukleinsäuren mit Hilfe von Computermethoden

Nukleinsäuren sind hochflexible Biomoleküle, deren strukturelle Dynamik nicht direkt aus der Sequenz abgeleitet werden kann (105). Experimentelle Methoden wie die Röntgenstrukturanalyse (engl. X-ray crystallography(X-ray)) oder die Kernspinresonanzspektroskopie (NMR-Spektroskopie, vom engl. nuclear magnetic resonance (NMR)) geben hierbei einen Einblick in die strukturelle Vielfalt von Nukleinsäuren und deren Interaktionen mit diversen Biomolekülen. Der drastische Anstieg der durch diese Methoden bestimmten Strukturen hilft bei der Analyse von spezifischen Interaktionen innerhalb von Nukleinsäuren, Nukleinsäure-Protein-Wechselwirkungen und Nukleinsäure-Ligand-Wechselwirkungen. So konnten eine ganze Reihe von kleinen organischen Strukturen entwickelt werden, die ganz spezifisch an bestimmte Regionen innerhalb der Nukleinsäure binden (106–109). Jedoch liefern diese experimentellen Methoden nur geringfügige Informationen über die treibenden

Kräfte bei Prozessen wie der Faltung von Nukleinsäuren, die Dynamik von Konformationsübergängen oder Assoziation/Dissoziationsprozessen (110–113).

Um ein besseres Verständnis über Konformationen und Konformationsübergängen zu erhalten kann die Methodik der MD-Simulation herangezogen werden. MD-Simulationen ermöglichen einen Einblick in die Konformationsvielfalt von Nukleinsäuren und geben Aufschluss über die strukturellen Ursachen von Konformationsübergängen. Es wurden bereits eine Vielzahl von Fragestellungen bezüglich Nukleinsäuren mit Hilfe von MDS behandelt:

- Flexibilität des Zucker-Phosphatrückgrats (α/γ -Übergänge, BI/BII-Übergänge, Zusammenhang zwischen helikalen Parametern und Diederwinkel (114–116))
- Zusammenhang zwischen Sequenz und Dynamik (117, 118, 78, 119, 114)
- Faltung- und Entfaltungsprozesse (120–122, 111, 123)
- Spezifische Nukleinsäure-Ionen Interaktion (124, 125, 110, 25, 126, 108)
- Einfluss des Lösungsmittels (127, 109, 128)
- Nukleinsäure-Ligand Interaktion (106, 129, 130)
- Strukturverfeinerung und Vorhersage durch coarse-grained Modelle (131–134)
- Globale Bewegungen in Nukleinsäuren (global kink turn motions) (135–138)
- Assoziation und Dissoziation von Nukleinsäuren (139)
- DNA-Oberflächen Interaktion (140, 141)

1.6 Motivation der Arbeit

In der vorliegenden Arbeit sollen strukturelle und thermodynamische Eigenschaften von Nukleinsäuren mit Hilfe von Molekulardynamik-Simulation bestimmt werden. Es handelt sich um Sekundärstrukturelemente wie Dangling-End oder RNA-Bulges. Ferner soll der Einfluss von Basenmodifikationen auf die Feinstruktur der DNA untersucht werden. Hierbei sollen folgende Ziele verwirklicht werden:

- Implementierung und Anwendung von effizienten Simulationsmethoden (siehe Abschnitt 2.4).
- Vergleichende Modellierung von 8oxoG:C und G:C Paarungen sowie die Anwendung von BP-REMD Simulationen und deren direkter Vergleich zu kMDS (siehe Abschnitt 3.1).
- Vergleichende Modellierung von O^6 -MeG:C, O^6 -MeG:T und G:C Paarungen und Bestimmung von spezifischen O^6 -MeG Konformationszuständen mit Hilfe der von BP-REMD und kMDS (siehe Abschnitt 2.5.2 und Abschnitt 3.2).
- Modellierung und Bestimmung von Thermodynamischen Parametern von RNA und DNA Dangling-End-Strukturen (siehe Abschnitt 2.5.3 und Abschnitt 3.3).
- Untersuchung des Konformationsbereiches unterschiedlicher RNA-Bulge-Strukturen mit Hilfe unterschiedlicher Simulationsmethoden (siehe Abschnitt 3.4).
- Für alle Anwendungen sollen zudem Vergleiche zwischen experimentellen Strukturen und berechneten Strukturen erstellt werden.

Kapitel 2

2 Theorie, Methoden und Modelle

2.1 Molekulardynamik

Molekulardynamik-Simulationen sind eine weit verbreitete und etablierte Methode zur Beschreibung und Analyse von Biomolekülen (142). Die Anwendungsgebiete sind sehr vielseitig. So können Molekulardynamik-Simulationen eingesetzt werden um Protein-Faltungsprozesse zu untersuchen (143). Es besteht weiterhin die Möglichkeit den Effekt von Solvensmolekülen, die Stabilität von Strukturen, und Eigenschaften des biomolekularen Systems, wie Interaktionsenergien zwischen einzelnen Atomen, sowie entropische Eigenschaften zu ermitteln (144). Innerhalb der letzten 25 Jahre hat sich die Methode der Molekulardynamik-Simulation sehr rasch entwickelt. Die erste Protein Simulation wurde von McCammon *et al.* 1977 durchgeführt. Es handelte sich dabei um die Simulation eines Trypsin Inhibitors mit 58 Aminosäureresten im Vakuum mit einer Simulationsdauer von 8.8 psec. Die ersten Nukleinsäure-Simulation wurden von Levitt *et. al* 1983 durchgeführt (145). Einige Jahre später simulierten McCammon und Harvey die tRNA^{PHE} (76 Nukleotide) (146). Inzwischen ist man in der Lage Systeme mit bis zu 10⁶ Atomen und Simulationszeiten im Mikrosekunden-Bereich zu erreichen (147–149).

Molekulardynamik-Simulationen basieren auf den Gesetzen der klassischen Mechanik. Die Dynamik eines Systems ist durch die numerische Integration der Newton'schen Bewegungsgleichung definiert. In der klassischen Mechanik ist, unter der Annahme, dass alle Informationen bezüglich der Koordinaten der Atome und deren Impulse/Kräfte vorliegen, die zeitliche Entwicklung eines abgeschlossenen Systems eindeutig definiert. Man erhält für ein System mit N Teilchen die zeitliche Entwicklung gemäß der Newton'schen Bewegungsgleichung wie folgt (142):

$$m_i \frac{d^2 r_i(t)}{dt^2} = F_i \quad (2.1)$$

$$F_i = -\frac{\partial \mu(r_1, \dots, r_N)}{\partial r_i} \quad (2.2)$$

Wobei r_1, \dots, r_N die kartesischen Koordinaten, $\mu(r_1, \dots, r_N)$ die durch die Koordinaten festgelegte Potentialfunktion und F_i , die auf Teilchen i wirkende Kraft darstellen. Zur Bestimmung der Newton'schen Bewegungsgleichung werden also Potentialfunktionen (Kraftfelder) und Algorithmen zur Integration der Bewegungsgleichung benötigt. Bewährte Algorithmen hierfür sind der Verlet-Algorithmus (150) und der Leap-Frog-Algorithmus (151). Die allgemeine Vorgehensweise einer Molekulardynamik Simulation hat folgendes Schema:

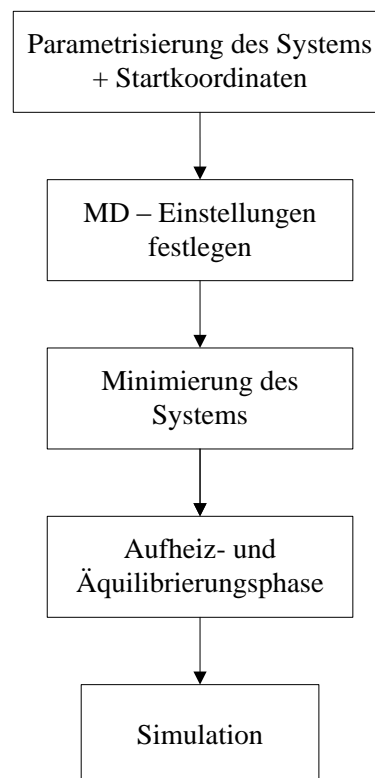


Abbildung 2.1 Vorgehensweise bei einer Molekulardynamik Simulation. Zunächst müssen die Moleküle des zu simulierenden Systems (z.B. Nukleinsäure, Protein, Ligand, Wasser und Ionen) parametrisiert werden. Anschließend legt man die Dauer und die Bedingungen der Minimierung, Equilibrierung und der Simulation fest. Bei der Minimierung ist zu beachten, dass zunächst die Wassermoleküle und anschließend Nukleinsäure minimiert werden. Innerhalb der Aufheizphase wird die Temperatur von 0 K auf die allgemeine Simulationstemperatur von 300 K schrittweise erhöht. Die Equilibrierung findet anschließend statt, sodass es zu einem Ausgleich der kinetischen und potentiellen Energie kommt. Sobald sich das System im Gleichgewicht befindet beginnt die Simulation.

Um eine akzeptable Startgeometrie zu ermitteln und sterische Hinderungen zu vermeiden, wird das System zunächst einer Energieminimierung unterzogen. Dabei wird meist nur ein lokales Energieminimum erreicht, da es äußerst schwierig ist für komplexe Systeme wie Proteine oder Nucleinsäuren das globale Minimum zu ermitteln. Zur Suche des Energieminimums werden Methoden des steilsten Abstiegs und des konjugierten Gradienten herangezogen. Wie in der Beschreibung der Abbildung 2.1 erläutert, werden zunächst die Wassermoleküle minimiert um ungünstige van der Waals Kontakte zwischen Wassermolekülen zu beseitigen. Anschließend findet eine Minimierung des Soluts statt.

In der Aufheizphase wird die Temperatur des Systems schrittweise von 0 K auf 300 K erhöht. Die Nucleinsäure wird hierbei mit Hilfe von harmonischen Rückstellkräften fixiert. Während der Equilibrierungsphase werden die Rückstellkräfte entfernt. Sobald sich die totale Energie im Gleichgewicht befindet, beginnt die Simulation. Simulationsdauer und Simulationsbedingungen sind hierbei abhängig von der jeweiligen Fragestellung.

2.1.1 Kraftfelder

Die Entwicklung von Kraftfeldern ist nach wie vor ein schwieriges Problem, da das Ziel der Modellierung eine möglichst detaillierte Beschreibung des Systems in seiner natürlichen Umgebung ist (152–154). Zu den am weitesten verbreiteten Kraftfeldern gehören unter anderem CHARMM (155), AMBER (156, 153) und das GROMOS (157) Kraftfeld. Viele diese Kraftfelder besitzen eine ähnliche Form, unterscheiden sich jedoch in ihrer Parametrisierung. Die Potentialfunktion des AMBER-Kraftfeldes ist wie folgt definiert:

$$\begin{aligned}
 E_{total} = & \sum_{\text{Bindungen}} K_r (r - r_{eq})^2 + & (2.3) \\
 & \sum_{\text{Winkel}} K_\theta (\theta - \theta_{eq})^2 + \\
 & \sum_{\text{Diederwinkel}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]^2 + \\
 & \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]
 \end{aligned}$$

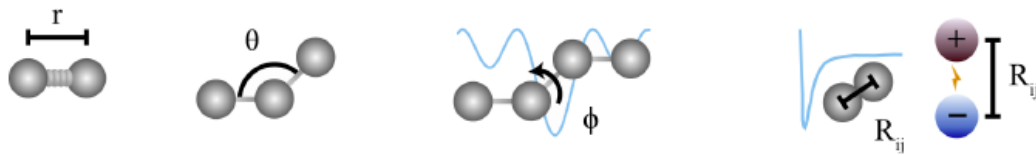


Abbildung 2.2 Beschreibung der Kraftfeldterme. Die Interaktion von Molekülen bzw. Atomen wird in Molekulardynamik Simulationen über Kraftfelder bestimmt. Diese Interaktionen lassen sich unterteilen in gebundene und nicht-gebundene Wechselwirkungen. Bindungslängen und Bindungswinkel werden über harmonische Potentiale und Diederwinkel über sinusförmige Funktionen beschrieben. Nicht-gebundene Wechselwirkungen wie das vdW-Wechselwirkungen werden mittels Lennard-Jones Potentials und elektrostatische Wechselwirkungen über Coulomb-Wechselwirkungen beschrieben (Quelle: modifiziert nach (158)).

K_r = Kraftkonstante für harmonische Potentiale auf allen Bindungslängen

r_{eq} = Gleichgewichtskonstante der entsprechenden Bindungslänge

K_θ = Kraftkonstante für harmonische Potentiale auf allen Bindungswinkel

θ_{eq} = Gleichgewichtskonstante des entsprechenden Bindungswinkels

n = Multiplizitäten des Diederwinkel ϕ

γ = Phasen des Diederwinkel ϕ

A, B = van der Waals Parameter

$q_i q_j$ = Partiaalladung von Atom i bzw. j

Die Beschreibungen der Bindungslängen basieren auf den Hook'schen Gesetzen und werden mit Hilfe eines harmonischen Potentials abgeschätzt. Die Gleichung schätzt die Bindungsenergie zwischen zwei Atomen unter Berücksichtigung einer Fluktuation zu einer Gleichgewichtskonstante. Bei großen Fluktuationen ist eine Abschätzung durch ein harmonisches Potential ungeeignet und lässt sich durch ein Morse-Potential realistischer beschreiben. Die Berechnungen der Bindungswinkel basieren ebenfalls auf den Hook'schen Gesetzen und können auch durch ein harmonisches Potential beschrieben werden. Die Kraftkonstante der Bindungswinkel beschreibt die Steifigkeit eines Winkels. Der Diederwinkel beschreibt die Rotation des Moleküls um eine Bindung und ist durch vier Atome definiert. Die Rotationsbarrieren werden durch eine Kosinusfunktion abgeschätzt. Bei den bisher diskutierten Energiefunktionen handelt es sich um gebundene Wechselwirkungen. Nicht-gebundene Wechselwirkungen lassen sich mit Hilfe von van der Waals (vdW) Wechselwirkungen und elektrostatischen Wechselwirkungen beschreiben. Die vdW-Interaktionen setzen sich aus einem anziehenden und einem abstoßenden Teil zusammen. Ursache für die repulsiven Kräfte ist das Pauli-Verbot. Anziehende Wechselwirkungen sind vdW-Kräfte (London-Kräfte), Dipol-Dipol Wechselwirkungen und Wasserstoffbrücken. Die

elektrostatischen Wechselwirkungen werden mittels Coulomb-Potential beschrieben. Die Ladungsverteilung eines Atoms wird durch Punktladungen approximiert.

Die Parametrisierung der Konstanten erfolgt auf atomarer Ebene. Zudem besitzen gleiche Atomtypen mit unterschiedlichem Hybridisierungsgrad unterschiedliche Parameterwerte (159). Zur Bestimmung der Parameter werden unter anderem Ergebnisse aus quantenmechanischen Rechnungen und X-Ray-Strukturinformationen herangezogen.

In dieser Arbeit wurden für sämtliche DNA Simulationen das *parmbc0*-Kraftfeld verwendet und für RNA Simulationen das *parmbc0XOL* (152, 153, 160, 161). Das *amber-parmbc0*-Kraftfeld ist eine Erweiterung des *amber-parm99* Kraftfeldes. Das *parm99*-Kraftfeld hat die Tendenz bei langen Simulationen (≥ 50 ns) irreversible α/γ Übergänge der Torsionswinkel hervorzurufen, sodass artifizielle Trajektorien entstehen (115). Für die modifizierten DNA-Basen wurden in dieser Arbeit die entsprechenden Kraftfeldkorrekturen hinzugefügt.

Das *parmbc0XOL*-Kraftfeld für RNA Strukturen wurde verwendet um „*ladder-like*“ Strukturen zu vermeiden, die beim *amber-parm99* bzw. *amber-parm99bcs0* Kraftfeld für RNA Strukturen entstehen können (161, 162).

2.2 DNA-Konformationsanalyse

In Abschnitt 1.2 wurde der Aufbau und die Stabilität von Nucleinsäuren diskutiert. Um spezifische strukturelle Eigenschaften von Nucleinsäuren zu ermitteln, müssen generelle Eigenschaften und Betrachtungsweisen von Nucleinsäuren definiert werden. Im folgenden Kapitel wird daher auf die Diederwinkel von Nucleinsäuren, die Furchenbildung und auf helikale Parameter von Nucleinsäuren eingegangen. Jede Komponente der Nucleinsäure besitzt physikochemische Eigenschaften die zur Beschreibung einer Konformation herangezogen werden können. Mit Hilfe dieser Eigenschaften können Konformationen von Nucleinsäuren beschrieben und miteinander verglichen werden.

2.2.1 Konformationen des Zucker-Phosphat-Rückgrats

Der (Desoxy)-Ribose-Ring in Nukleinsäuren sind nicht planar und besitzen daher eine begrenzte Flexibilität. Die Konformationen, die sich aus der nicht-planaren Eigenschaft des Zuckers ergeben, werden als Zucker-Pucker bezeichnet. Die Konformation des Zuckers bzw. der Zucker-Pucker kann anhand ihrer inneren Diederwinkel beschrieben werden. Die Stabilität der Konformation ist dabei stark abhängig von der nicht-gebundenen Wechselwirkung ihrer Substituenten, die eine möglichst weite Distanz zueinander anstreben. Eine Zuordnung des Konformationsraums kann anhand der Phasenwinkel der Pseudorotation P beschrieben werden. Diese definiert die Art des Pucker-Zustandes, wobei P über die inneren Diederwinkel τ_0 , τ_4 des Ringes bestimmt ist (163–165):

$$\tan P = \frac{(\tau_4 + \tau_1) - (\tau_3 + \tau_0)}{2 * \tau_2 * (\sin 36^\circ + \sin 72^\circ)} \quad (2.4)$$

Zusätzlich gilt, dass der maximale Winkel des Puckers, also die Amplitude, limitiert ist durch τ_m :

$$\tau_m = \tau_2 / (\cos P) \quad (2.5)$$

Der Pseudorotation Phasen Winkel kann Werte zwischen 0° und 360° einnehmen. Ferner gilt, dass, falls $\tau_2 < 0$, wird 180° auf den aktuellen Wert von P addiert wird. Experimentelle Werte für τ_m liegen zwischen 25° und 45° . Der Pseudorotation Phasenwinkel kann mittels eines Rads dargestellt werden (siehe Abbildung 2.3).

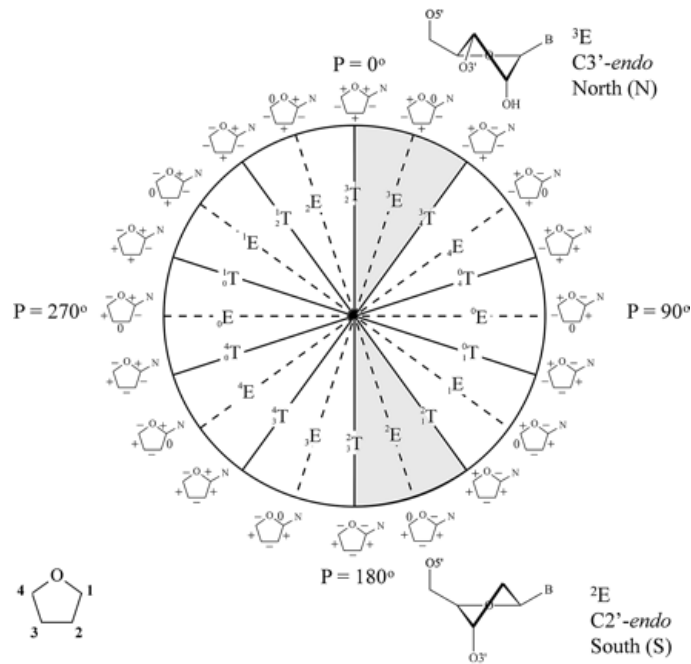


Abbildung 2.3 Klassifizierung der Konformationen anhand eines Pseudorotations-Rad. Dargestellt sind die bevorzugten Konformationen der Ribose (C3'-endo, A-RNA) und der Desoxyribose (C2'-endo, B-DNA) durch den grau markierten Bereich. Definiert werden die Konformationen über einen Pseudorotationswinkel P . Am Rand sind Werte der endozyklischen Torsionswinkel (v_0 bis v_4) beschrieben über ihr Vorzeichen (+, - oder 0) (Quelle: modifiziert nach (166)).

Der Phasenwinkel korreliert mit dem Diederwinkel δ des Phosphatrückgrats. Der Zucker-Pucker ist somit eine wichtige Determinante um Nukleinsäure Konformationen zu beschreiben (167). Neben dem Zucker-Pucker spielen die Diederwinkel des Phosphatrückgrats eine wichtige Rolle in der Beschreibung der Konformation von Nukleinsäuren.

Der Phosphatrückgrat erstreckt sich über sechs Diederwinkel: $\alpha, \beta, \gamma, \delta, \epsilon$ und ζ . In Abbildung 2.4 sind die bevorzugten Bereiche der jeweiligen Winkel angegeben.

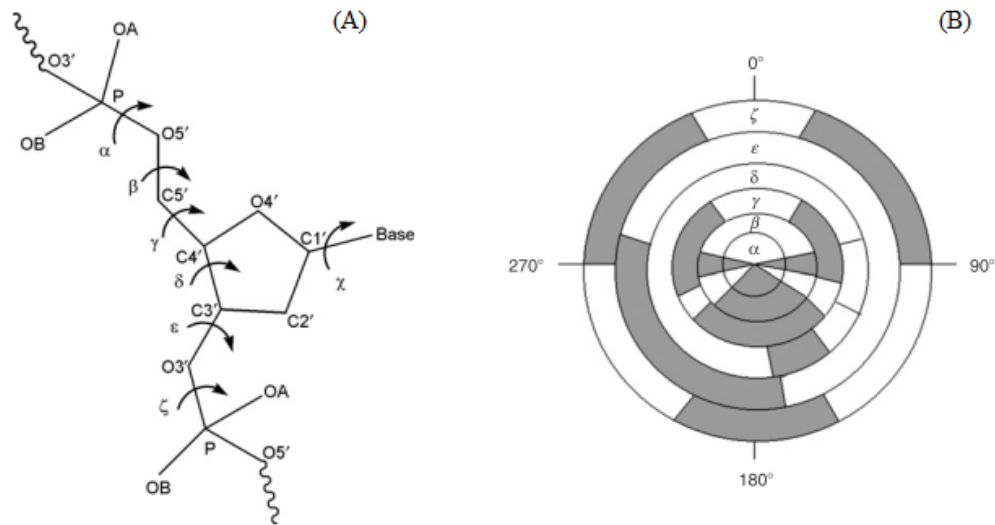


Abbildung 2.4 Diederwinkel des DNA-Phosphatrückgrats. (A) Ausgehend von 3'-Ende sind die Diederwinkel α , β , γ , δ , ϵ , ζ sowie der χ -Winkel zwischen der Base und dem Zucker dargestellt. (B) Bevorzugte Bereiche der Winkel in B-DNA Strukturen (Quelle: modifiziert nach (166)).

Aus sterischen Gründen kann nicht jeder Wert zwischen 0° und 360° eingenommen werden. Es ist jedoch wichtig darauf hinzuweisen, dass die α , β , γ und ζ drei zulässige Bereiche einnehmen können. Berücksichtigt man weiterhin den großen Bereich, den der ϵ -Winkel einnehmen kann, und die Flexibilität des Zucker-Ringes so lässt dich daraus schließen, dass der Phosphatrückgrat, trotz der Beschränkung durch Watson-Crick-Paarungen, eine Vielzahl von energetisch günstigen Konformationen einnehmen kann (3). Als generelle Konvention zur Beschreibung der Phosphatrückgrat-Winkel gilt: Werte im Bereich $\sim 60^\circ$ werden als *gauche+*(g+) bezeichnet. Liegen die Werte im Bereich -60° werden sie als *gauche-*(g-) und Werte im Bereich $\sim 180^\circ$ als *trans*(t) definiert. Die Winkel α und γ können in der g+, g- oder t Konformation vorliegen. Der β -Winkel bevorzugt eine t-Konformation. Für den γ -Winkel wurden g+, g- und t Konformation in Kristallstrukturen beobachtet, wobei überwiegend die g+ Konformation bei der kanonischen B-DNA auftritt.

Der δ -Diederwinkel ist stark mit dem Zucker-Pucker korreliert. Liegt der δ -Diederwinkel bei ungefähr 75° , so liegt ein C3'-endo Zucker-Pucker vor und bei 150° ein C2'-endo Zucker-Pucker. Es wurden eine ganze Reihe von Korrelationen zwischen den verschiedenen Winkel, sowie dem Winkel der glykosidischen Bindung und den Zucker-Puckern ermittelt (168–170, 114, 115) :

- Der Zucker-Pucker korreliert mit dem χ -Winkel der glykosidischen Bindung.
- Der γ -Winkel korreliert mit dem χ und α -Winkel.
- Der α -Winkel korreliert mit dem γ -Winkel.
- Der ε korreliert mit dem ζ -Winkel. Hierbei werden zwei Bereiche bevorzugt. Man unterscheidet zwischen BI und dem BII Zustand. Der BI-Zustand liegt vor, falls ε - ζ zwischen -160° und 20° liegen. Der BII-Zustand liegt vor, falls ε - ζ zwischen 20° und 200° liegen.

Neben den Winkeln des Phosphatrückgrats und dem Zucker-Pucker spielt die glykosidische Bindung zwischen Zucker und Base eine wichtige Rolle in der Beschreibung einer Konformation. Die Bindung ist definiert zwischen den Atomen C1' und N9 für Purine und C1' und N1 für Pyrimidine. Der Diederwinkel χ um diese Bindung kann eine Reihe von Werten einnehmen, jedoch existieren bevorzugte Bereiche für B-DNA bzw. A-RNA Strukturen.

Der Winkel ist über folgende Atome definiert: O4'-C1'-N9-C4 für Purine und O4'-C1'-N1-C2 für Pyrimidine. Man unterscheidet zwischen einer *syn* und einer *anti* Konformation. Die Mehrheit der Nukleotide nimmt eine *anti*-Konformation ein. Der Wert des χ Diederwinkel liegt zwischen 180° und 300° . Für *syn*-Konformationen liegt der Bereich zwischen 0° und 90° .

2.2.2 Helikale Parameter

Die Flexibilität einer Base liegt zum einen in den chemischen und physikalischen Eigenschaften der Base und des Basenpartners und zum anderen an den benachbarten Basen und der Umgebung. Da Basen durch ihr π -Elektronensysteme planare Strukturen ausbilden, ist die Orientierung und Stabilität zu ihrem gegenüberliegenden Nachbar abhängig von der Anzahl der möglichen Wasserstoffbrücken. Kanonische bzw. Watson-Crick Basenpaare (A:T, A:U, G:C) bieten die stabilste Kombination (171, 172). Durch diese Basenpaarung entsteht eine optimale geometrische Zuordnung der Basenpaare. So liegt die Distanz zwischen C1'-C1' Atomen zwischen A:T und G:C zwischen 10.1 und 10.9 Å. Neben der Wasserstoffbrücken-Interaktion spielen Stacking-Energien eine entscheidende Rolle in der Position und Orientierung der Basenpaare.

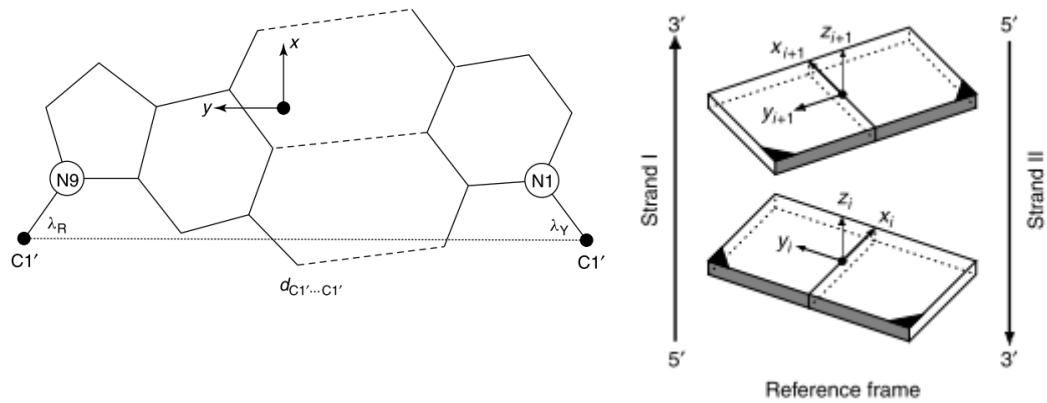


Abbildung 2.5 Darstellung eines idealen Basen-Paar Parameters. Beschreibung des Standardreferenz-Systems nach Olsen et. al. . $C1'-C1'$ Vektor λ definieren die optimale Geometrie. Die Orientierung der Basenpaare wird über die x , y , und z -Koordinaten bestimmt, wobei der Ursprung als \bullet gekennzeichnet ist (Quelle modifiziert nach (167)).

So entstehen sequenz-abhängige Eigenschaften und Konformationen der Nukleinsäure. Man unterscheidet hierbei zwischen lokalen und globalen Eigenschaften der Nukleinsäure-Helix. Lokale Eigenschaften lassen sich durch die Anordnung von zwei Basenpaaren zueinander beschreiben, wohingegen langreichweitige Effekte wie die Biegung (engl. *bending*) globale Eigenschaften Helix sind (173, 6). Um die dreidimensionale Ausrichtung einer Base bzw. eines Basenpaars zu beschreiben, mussten eine Reihe von helikalen Parametern definiert werden. Diese Parameter lassen sich zum einen über die Orientierung des Basenpaars und der Orientierung zum Nachbar-Basenpaar beschreiben.

Es ist jedoch an dieser Stelle anzumerken, dass die Berechnung dieser Parameter unterschiedlich erfolgen kann, abhängig von der Definition und Betrachtungsweise der Achsenzuordnung. Um Unklarheiten zu beseitigen schlugen Lu und Olsen 1999 ein Standardreferenz-System zur Berechnung dieser Parameter vor (Tsubaka Übereinkunft (174)).

Die x -Achse zeigt in dieser Repräsentation in Richtung der großen Furche (engl. *major groove*). Die y -Achse befindet sich entlang der Längsachse des Basenpaars und parallel zum $C1'-C1'$ -Vektor. Der Vorteil der Tsubaka Referenz-Betrachtung liegt in ihrer Eindeutigkeit, sodass unabhängig vom verwendeten Algorithmus die helikalen Parameter berechnet werden können.

Die Orientierung zweier Basen zueinander wird über folgende Parameter beschrieben: Propeller, Twist, Buckle, Inklination sowie X - und Y -displacement. Die Orientierung zweier Basenpaare zueinander wird über die Parameter Rise, Slide, Shift, Tilt, Roll und Twist beschrieben (174, 175):

- Propeller: Ist definiert über den Diederwinkel zwischen den Normalen zu den Basen ausgehend von der Längsachse des Basenpaares. Der Winkel besitzt in der kanonischen B-DNA ein negatives Vorzeichen.
- Twist: Der Winkel zwischen aufeinanderfolgenden Basenpaaren. Gemessen anhand der Änderung in der Richtung des C1'-C1' Vektors ausgehend von einem Basenpaar zum nächsten.
- Roll: Der Diederwinkel beschreibt die Drehung von einem Basenpaar bezüglich seines Nachbarn, um die Längsachse des Basenpaares. Ein positiver Roll-Winkel öffnet ein Basenpaar in Richtung der kleinen Furche der DNA.
- Tilt ist der Diederwinkel entlang der kurzen Achse der Basenpaare.
- Slide beschreibt die relative Verschiebung eines Basenpaares im Vergleich zum nächsten.

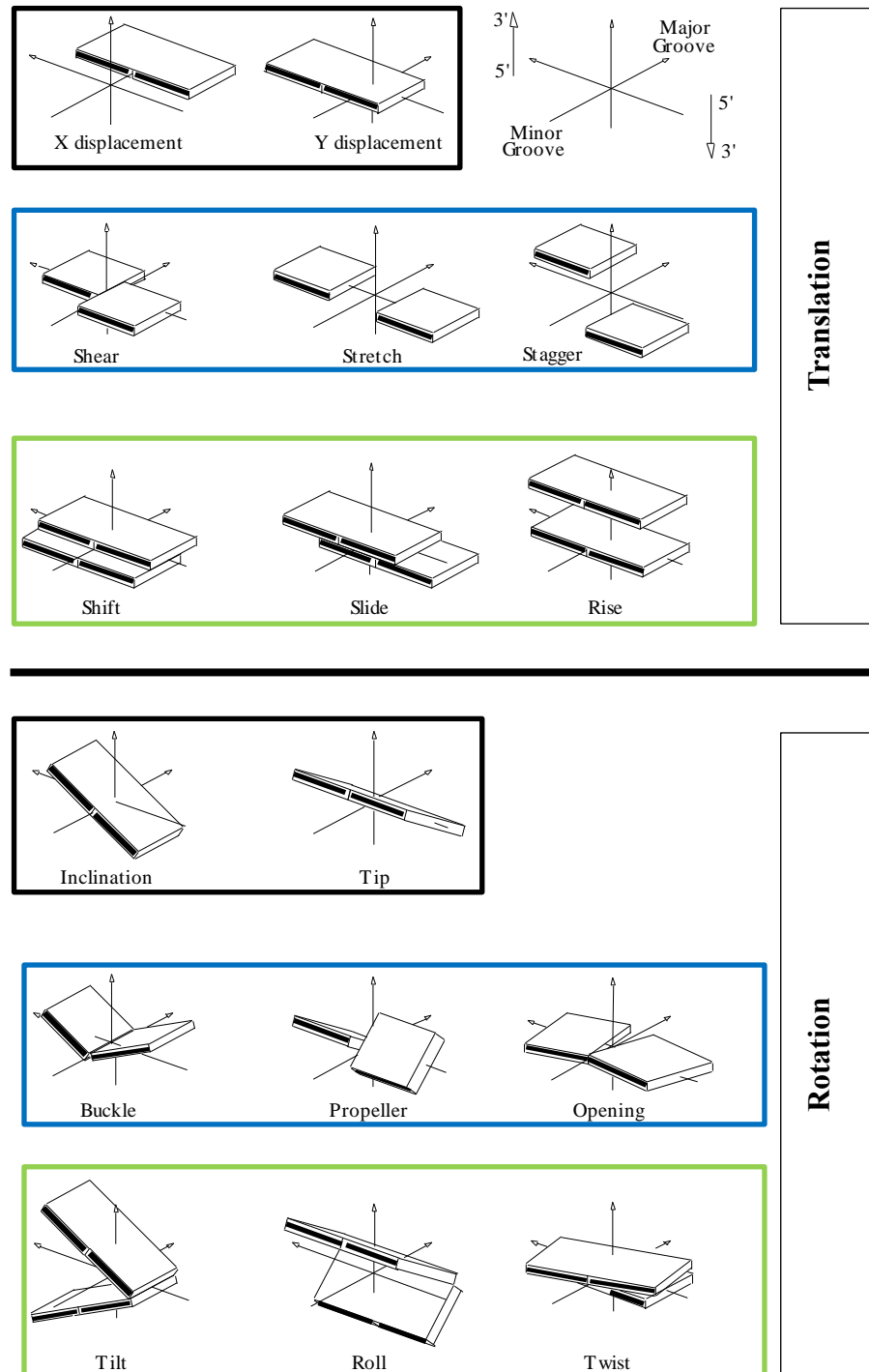


Abbildung 2.6 Helikale Parameter für Nucleinsäuren. Die Parameter zur Beschreibung helikaler Eigenschaften einer Helix lassen sich unterteilen in Translations- und Rotationsparameter. Diese wiederum werden unterteilt in Parameter intramolekulare Basenpaar Eigenschaften (blaue Box), intermolekulare Eigenschaften (grüne Box) und Geometrien des Basenpaares bezüglich der Achsen (schwarze Box). (Quelle: modifiziert nach (175))

2.3 Berechnung der Freien-Energie-Differenz

Wechselwirkungen zwischen Biomolekülen und deren Spezifität zueinander ist ein wichtiger Bestandteil für das Verständnis von biologischen Fragestellungen. Die Bestimmung von Energiebarrieren für Konformationsübergänge in Makromolekülen ist für molekulare Erkennung dieser Moleküle ausschlaggebend. Wie effektiv solch eine Konformationsübergang abläuft lässt sich mit Hilfe der Differenz der freien Energie beschreiben. Die Differenz der freien Energie ΔG beschreibt Beziehungen zwischen verschiedenen Energieformen und beantwortet die Frage, ob und unter welchen Bedingungen und in welchem Umfang solch ein Prozess abläuft. Die freie Energie ΔG wird über einen enthalpischen (ΔH) und einen entropischen Beitrag ($T\Delta S$) beschrieben (176):

$$\Delta G = \Delta H - T\Delta S \quad (2.6)$$

Die treibende Kraft einer Protein-Ligand-Wechselwirkung oder eines Konformationsüberganges kann eine Enthalpieabnahme, eine Entropiezunahme oder beides sein. Um ein besseres Bild der enthalpischen und entropischen Kräfte zu erhalten werden diese kurz zusammengefasst.

Der enthalpische Beitrag setzt sich zusammen aus:

- Elektrostatischen Wechselwirkungen (Salzbrücken, Wasserstoffbrücken)
- Dipol/Dipol Wechselwirkungen
- Solvatations / Desolvationseffekten
- VdW Wechselwirkungen

Die Coulomb und vdW-Wechselwirkungen beschreiben nicht-kovalente Wechselwirkungen. Enthalpische Beiträge lassen sich durch Kraftfeldtermen sehr gut beschreiben, wobei zu beachten ist, dass es sich nur um Approximationen der tatsächlichen Energie handelt (177).

Der entropische Beitrag einer Wechselwirkung wird durch den hydrophoben Effekt beschrieben. Unter dem hydrophoben Effekt versteht man das Verdrängen

von Wassermolekülen auf der Oberfläche eines Biomoleküls, was zur Folge hat, dass sich sowohl die Wassermoleküle, als auch der das Biomolekül reorganisieren müssen. So gehen z.B. bei der Bindung von einem Liganden an einen Rezeptor jeweils drei Translations- und Rotationsfreiheitsgrade verloren, die in sechs neuen Schwingungs-Freiheitsgraden wieder auftauchen. Entropische Beiträge bzw. Entropiedifferenzen lassen sich nach dem heutigen wissenschaftlichen Stand schwer abschätzen, da die Entropie eine globale Eigenschaft des biomolekularen Systems ist und als solche auch betrachtet werden muss, d.h. um die Entropie abschätzen zu können muss der gesamte Konformationsraum abgesucht werden (144).

2.3.1 Umbrella-Sampling Simulationen

Die Freie Energiedifferenz von zwei Zuständen kann mit Hilfe von Simulationsmethoden berechnet werden, indem die Wahrscheinlichkeitsverteilung der Zustände bestimmt wird (178):

$$\Delta F = F_B - F_A = k_B T \ln \frac{P_A}{P_B} \quad (2.7)$$

P_A und P_B sind die Wahrscheinlichkeiten, dass sich eine Konformation in Zustand A oder in Zustand B befindet. Diese Wahrscheinlichkeiten sind proportional zu der Zeit, in der sich ein System während einer MDS aufhält. Besteht zwischen den Zuständen A und B eine hohe Energiebarriere, so besteht die Gefahr, dass ausgehend von Zustand A , Zustand B nur selten oder gar nicht während der Simulation erreicht wird, sodass die Wahrscheinlichkeit des Zustands B falsch wiedergegeben wird. Um die Problematik der Energiebarriere zu lösen, besteht die Möglichkeit eine Umbrella-Sampling Simulation durchzuführen. Hierbei wird ein zusätzliches, fixes Potential (Umbrella oder Bias Potential) auf die Potentialfunktion des Systems addiert, sodass eine flache Energielandschaft zwischen den beiden Zuständen entsteht (179–181). Das Bias-Potential ω_i des Fensters i ist ein zusätzlicher Term, welcher nur von der Reaktionskoordinate ξ abhängig ist. Für die modifizierte Potentialfunktion E^b

$$E^b = E^0(r) + \omega_i(\xi) \quad (2.8)$$

Das zusätzliche Potential $\omega_i(\xi)$ wird häufig durch eine harmonische Funktion beschrieben. Für ein harmonisches Potential der Stärke K gilt:

$$\omega_i(\xi) = K_i(\xi - \xi_i^{ref})^2 \quad (2.9)$$

In jedem Fenster i wird somit durch das zusätzliche Potential dem System hinzugefügt um nahe der Referenz-Reaktionskoordinate des Fensters i (ξ_i^{ref}) zu sein. Die Stärke der Potentialfunktion K muss vor der Simulation definiert werden, und ist somit ein kritischer Faktor. Wird das System zu schnell aufgrund eines hohen Potentials von A nach B überführt, besteht die Gefahr, dass die Wahrscheinlichkeitsverteilung der Konformation entlang der Reaktionskoordinate nur schlecht repräsentiert ist. Das Ziel ist jedoch aus der Wahrscheinlichkeitsverteilung der Zustände entlang der Reaktionskoordinate ξ , die Freie Energie zu bestimmen. Die Wahrscheinlichkeitsverteilung eines Systems entlang der Reaktionskoordinate ξ ist über das Integral über alle Freiheitsgrade definiert (182):

$$Q(\xi) = \frac{\int \delta[\xi(r) - \xi] \exp[(-\beta E) d^N r]}{\int \exp[(-\beta E) d^N r]} \quad (2.10)$$

Mit $\beta=1/(k_B T)$. Für die Freie Energie (bzw. potential of mean force (PMF)), die nun abhängig ist von der Reaktionskoordinate und er Wahrscheinlichkeitsverteilung, gilt dann:

$$F(\xi) = -\frac{1}{\beta} \ln Q(\xi) \quad (2.11)$$

Problematisch ist jedoch, dass die Zustandssumme hier benötigt wird und dadurch das Integral über den Phasenraum nicht exakt berechenbar ist, sodass die Methode stark abhängig von der Güte der Verteilung des Zustands ist. Unter der Annahme, dass es sich um ein ergodisches System handelt und der Berücksichtigung des zusätzlichen Potential $\omega_i(\xi)$ gilt für die Wahrscheinlichkeitsverteilung entlang einer Reaktionskoordinate ξ :

$$P_i(\xi) = \frac{\int \exp\{-\beta[E(r) + \omega_i(\xi'(r))]\} \delta[\xi'(r) - \xi] d^N r}{\int \exp\{-\beta[E(r) + \omega_i(\xi'(r))]\} d^N r} \quad (2.12)$$

Für die Wahrscheinlichkeitsverteilung ohne zusätzliches Potential gilt:

$$P_i^0(\xi) = P_i(\xi) \exp[\beta\omega_i(\xi)] \langle \exp[-\beta\omega_i(\xi)] \rangle \quad (2.13)$$

Für die Freie Energie (PMF) des Fensters i gilt nun:

$$F_i(\xi) = -\left(\frac{1}{\beta}\right) \ln P_i(\xi) - \omega_i(\xi) + F_i \quad (2.14)$$

Aus der daraus resultierenden Freien Energie, kann der Bias wieder abgezogen werden und mit Hilfe der *Weighted histogram analysis Method* (WHAM) ermittelt werden (183).

Die Vorgehensweise bei einer Umbrella-Sampling Simulation ist folgende:

1. Definiere eine Reaktionskoordinate $\xi(\mathbf{r})$ die den Übergang von Zustand A zu Zustand B beschreibt.
2. Definiere die Kraftkonstante (Bias/Umbrella potential) mit der Zustand A in Zustand B überführt werden soll.
3. Definiere die Anzahl der Schritte (Fenster) die benötigt werden um Zustand A in Zustand B zu überführen.
4. Bestimme mit Hilfe von MDS die Wahrscheinlichkeitsverteilung entlang der Reaktionskoordinate.
5. Bestimme Freie Energie ohne zusätzliches Potential mittels WHAM.

2.3.2 MM/PBSA Rechnungen

Die Methode von Kollman *et al.* wurde ursprünglich entwickelt, um die relative Stabilität von unterschiedlichen Konformationen zu ermitteln (184). Die Einsatzgebiete dieser Methode sind jedoch vielseitig. So werden MM/PBSA-Rechnungen häufig bei der Vorhersage von freien Bindungsenthalpien zwischen Biomolekülen eingesetzt (185). Zudem wurde die Methode bereits erfolgreich für die Untersuchung von Aminosäuremutationen auf Bindungsaffinitäten von Protein- Protein-Interaktionen, Protein-Ligand-Interaktionen, Nukleinsäure-Konformationen sowie Nukleinsäure-Ligand-Interaktionen eingesetzt (184, 186–189). Bei Interaktionen zwischen zwei Biomolekülen muss die relative freie

Enthalpie zwischen gebundenem und ungebundenem Zustand berechnet werden. Solch eine Berechnung ist jedoch nicht ohne weiteres möglich, da die zur Berechnung der freien Enthalpie herangezogenen Kraftfelder keine entropischen Beiträge berücksichtigen. Ein weiteres Problem stellt die Wechselwirkungsenergie der Wassermoleküle dar, welche einen großen Anteil der Gesamtenergie des Systems ausmachen. Um dieser Problematik zu entgehen, behilft man sich einem Kreisprozess zur Berechnung der freien Bindungsenthalpie. Zudem werden Solvatationseffekte bzw. deren elektrostatischer Anteil durch Kontinuumsmodelle approximiert.

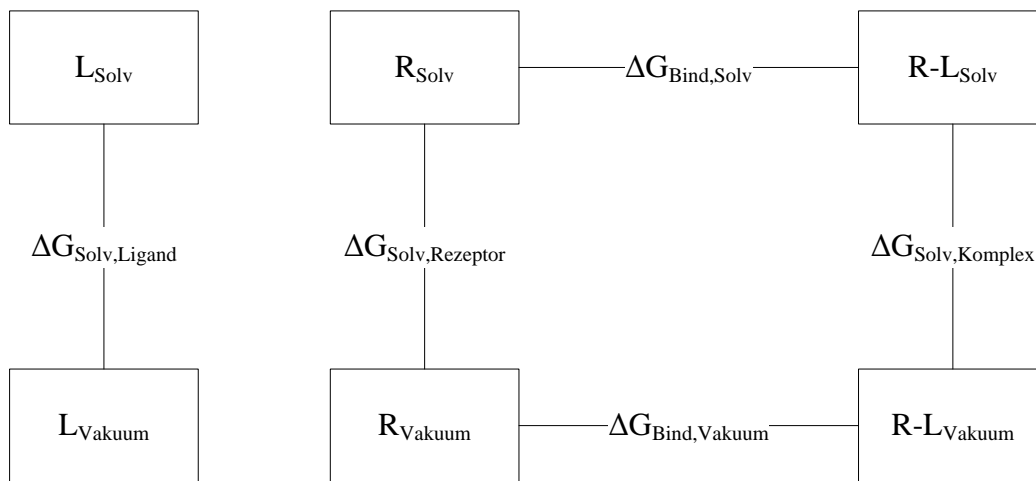


Abbildung 2.7 Thermodynamischer Kreisprozess bei einer MM/PBSA-Rechnung

Die freie Bindungsenthalpie lässt sich nun folgendermaßen beschreiben:

$$\Delta G_{Bind,Solv} = \Delta G_{Bind,Vakuum} + \Delta G_{Solv,Komplex} - (\Delta G_{Solv,Ligand} + \Delta G_{Solv,Rezeptor}) \quad (2.15)$$

Die Einzeltherme sind wiederum folgendermaßen definiert:

$$\Delta G_{Solv} = \Delta G_{elek,\varepsilon=80} + \Delta G_{elek,\varepsilon=1} + \Delta G_{Hydrophob} \quad (2.16)$$

$$\Delta G_{Vakuum} = \Delta E_{MM} - T\Delta S_{Normalmodenanalyse} \quad (2.17)$$

$$\Delta S_{Normalmodenanalyse} = S_{Komplex} - (S_{Rezeptor} - S_{Ligand}) \quad (2.18)$$

Für die nichtpolare Solvatationsenthalpie (F_{Np}) gilt:

$$F_{Np}(i) = \gamma SA(i) \quad (2.19)$$

γ = Proportionalitätsfaktor

Man geht davon aus, dass der nichtpolare Anteil der Solvatationsenthalpie sich proportional zur lösemittelzugänglichen Oberfläche (engl.: solvent accessible surface area, SASA bzw. SA) verhält (190). Für die Berechnung dieser Oberfläche existieren verschiedene Algorithmen. In dieser Arbeit wurde die LCPO-Methode (engl.: linear combination of pairwise overlaps) verwendet (191). Da es sich bei Nukleinsäuren um ein hochflexibles und dynamisches System handelt, muss der gesamte Konformationsraum des Systems berücksichtigt werden. Da dies jedoch zu rechenaufwendig ist, behilft man sich einer Molekulardynamik-Simulation, welches ein Ensemble von Konformationen generiert. Ziel hierbei ist den Konformationsraum möglichst gut abzudecken. Dieses Ensemble von Konformationen wird durch Koordinatenschnappschüsse generiert, was Momentaufnahmen während einer Simulation entspricht. Die freie Bindungsenthalpie setzt sich letztendlich aus dem Mittelwert der freien Bindungsenthalpien der einzelnen Konformationen zusammen.

2.4 Hamiltonian Replica Exchange Methode

Um die Eigenschaften und Funktionen eines Biomoleküls zu verstehen, muss die Anzahl der möglichen Zustände ermittelt werden. Form und Funktion eines Biomoleküls stehen in direkter Korrelation zueinander. Um den Konformationsraum mit Hilfe von Computersimulationen zu untersuchen besteht unter anderem die Möglichkeit der Verwendung von stochastischen Methoden, wie der Monte-Carlo Simulationstechnik oder die der Molekulardynamik-Simulation (MDS) Technik (192, 193). Die MDS hat den Vorteil, dass zeitabhängige Eigenschaften des Moleküls untersucht werden können. Die mit 2 fs kurzen Zeitschritte sind hierbei ein limitierender Faktor. Konformationsänderung, wie Faltungsprozesse von Molekülen, spielen sich jedoch im μ s, ms bis hin zu Sekunden-Bereich ab (194). Die Problematik liegt hierbei in der Energiehyperfläche des Faltungsprozesses von Molekülen, die aufgrund der hohen Anzahl von Freiheitsgraden viele energetische Minima aufweist und anisotrope Eigenschaften besitzen. Aufgrund des kleinen Zeitschritts

und der hohen Energiebarrieren kann die Energiefläche mit Hilfe der MDS nicht korrekt repräsentiert werden, da ein Zustand überrepräsentiert sein kann (kinetic trapping) (195–197).

Wie stark ein Zustand repräsentiert ist, ist abhängig von seiner Energie. Energetisch günstige Zustände besitzen hierbei eine höhere Auftretens-Wahrscheinlichkeit. Diese Wahrscheinlichkeit lässt sich durch die Boltzmann-Verteilung beschreiben. Hierbei gilt für Konformation x mit der Potentialfunktion $F(x)$:

$$P(x) \sim \exp\left[\frac{-F(x)}{k_B T}\right] \quad (2.20)$$

wobei, k_B die Boltzmann-Konstante und T die absolute Temperatur des Systems ist. Um den Gleichgewichtszustand eines Systems zu ermitteln müssen auch Zustände ermittelt werden, deren Potential gering ist. Um dieses Ziel zu realisieren wurden in dieser Arbeit Umbrella-Sampling Methoden und die Hamiltonian Replica Exchange Methode verwendet. Ziel war es eine Technik zu entwickeln und anzuwenden, die möglichst effizient die Energiehyperfläche nach Bereichen bzw. Konformationen niedriger Energie abgesucht und mit deren Hilfe sich die Energiebarrieren der verschiedenen Minima reduzieren lassen. Diese Reduktion kann unter anderem durch erhöhte Temperaturen oder unterschiedliche Kraftfeldern vorgenommen werden (198, 199). Ein System wird hierbei bei unterschiedlichen Temperaturen, jedoch mit der identischen Startstruktur simuliert. Es handelt sich also um Replikate eines Systems. Stochastische Übergänge zwischen den Replikaten mit größerer thermischer Fluktuation erlauben es neue Konformationen in den Replikaten zu generieren. Die stochastischen Übergänge sind abhängig von einem Metropolis-Kriterium. Für eine Replica-Exchange Simulation mit n Replikaten von einem System A gilt zum Startpunkt $A_1 = A_2 = A_3 \dots A_n$. Für die Gewichtung eines Zustandes C in einem verallgemeinerten Ensemble aller Replikate gilt:

$$W_{GE}(C) = \exp\left[\sum_{i=1}^n \frac{1}{k_B T} F(C_i)\right] \quad (2.21)$$

Für zwei Konformationen in Replika i und Replika j muss die detaillierte Balance Gleichung gelten, die ein kanonisches Ensemble gewährleistet:

$$W_{GE}(C_i)P_{ij} = W_{GE}(C_j)P_{ji} \quad (2.22)$$

Die Akzeptanzwahrscheinlichkeit eines Überganges P_{ij} ist abhängig von einem Metropolis-Kriterium:

$$P_{ij} = \min(1, e^{\Delta}), \quad (2.23)$$

$$\Delta = \beta_i (F_i(q_j) - F_i(q_i)) - \beta_j (F_j(q_j) - F_j(q_i))$$

F ist hierbei die potentielle Energie und β die inverse thermische Energie $1/k_B T$. Da die Akzeptanzwahrscheinlichkeit exponentiell mit dem Temperaturunterschied bzw. Potentialunterschied sinkt, muss, um eine hohe Akzeptanzwahrscheinlichkeit zwischen Replika i und Replika j zu ermöglichen, die Verteilung der potentiellen Energien beider Replikate eine Überlappung besitzen. In der hier verwendeten Methode werden alle Simulationen bei Raumtemperatur durchgeführt ($\beta_i = \beta_j$), jedoch mit einer unterschiedlichen Hamilton-Funktion (200–202). Die Replikate unterscheiden sich nicht mehr in ihrer potentiellen Energie F , sondern nur in dem unterschiedlichen Teil der Hamilton-Funktion. In der hier verwendeten Methode unterscheiden sich die Hamilton-Funktionen der Replikate in der Berechnung spezifisch definierter Diederwinkel. Für die Akzeptanzwahrscheinlichkeit eines Überganges mit dem Biasing-Potential v_i gilt nun:

$$P_{ij} = \min(1, e^{\Delta}), \Delta = \beta_i (v_i(q_j) - v_i(q_i)) - \beta_j (v_j(q_j) - v_j(q_i)) \quad (2.24)$$

Die Akzeptanzwahrscheinlichkeit hängt somit nur noch vom Biasing Potential der verschiedenen Replika ab. Die Grundlegende Idee dieser Methode beruht auf der Kopplung und Korrelation von Diederwinkel. Ziel hierbei ist es, die Energiebarriere für Übergänge von Zuständen der Diederwinkel zu reduzieren. Hierbei wird ein 2-dimensionales Potential verwendet um Regionen von zwei Diederwinkeln zu destabilisieren. Das Potential besitzt eine quasi-Gaussische Form mit einem flachen Plateau.

Diederwinkel wie zum Beispiel ε/ζ nehmen definierte Bereiche ein (BI/BII-Zustand). Diese Bereiche können mittels dieser Methode destabilisiert werden. Vorteil dieser Methode ist, dass die Destabilisierung der Winkel genau definiert

werden kann. So kann die Methode dazu genutzt werden, um spezifische Regionen eines Biomoleküls zu destabilisieren und es dadurch zu ermöglichen andere Konformationsbereiche zu explorieren. (201, 203).

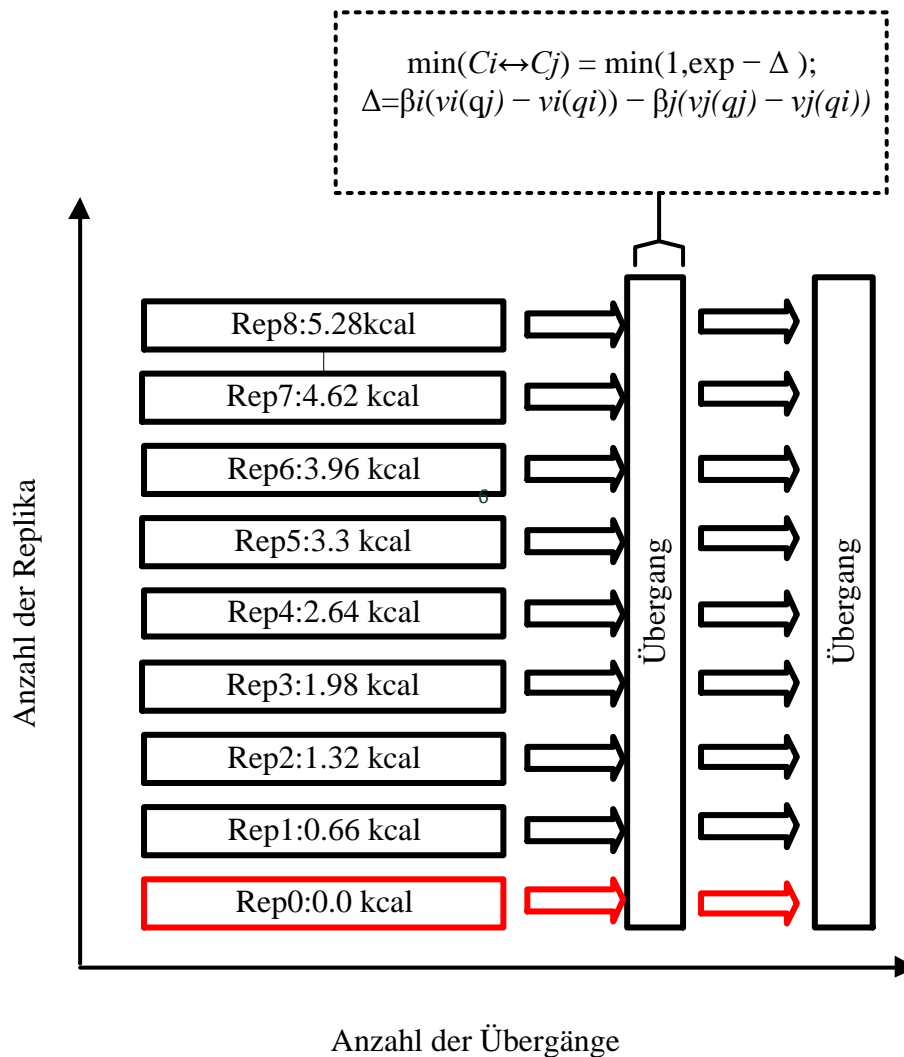


Abbildung 2.8 Schematische Darstellung der BP-REXMD. Replika 0 wird analysiert und kann direkt mit einer Trajektorie aus einer klassischen MD verglichen werden, da hier keine Kraftfeld Änderung vorliegt. Die Simulationszeit zwischen den Übergängen kann frei definiert werden. In der vorliegenden Arbeit betrug die Simulationszeit nach jedem Übergang 2ps. Die Übergänge werden nach einem Metropoliskriterium entweder akzeptiert oder abgelehnt.

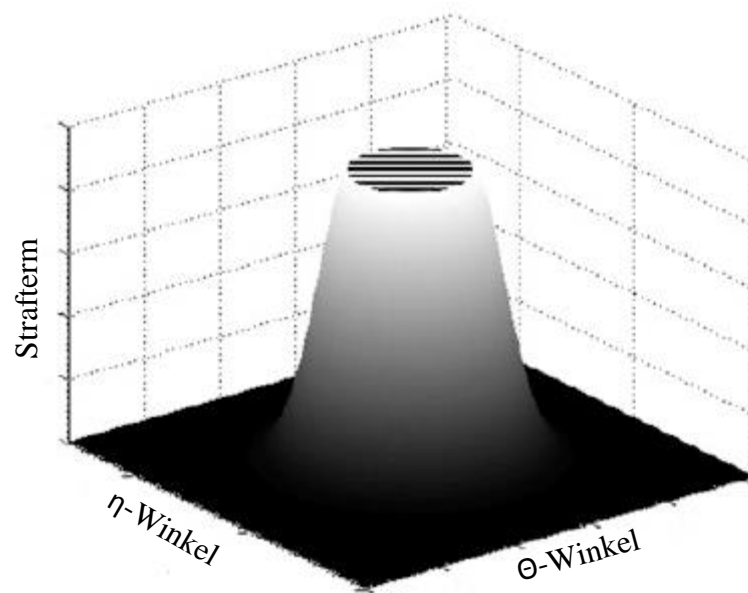


Abbildung 2.9 Das 2D-Biasing Potential. In dieser Methode wird ein 2D biasing Potential verwendet um bevorzugte η/θ -Diederwinkel zu destabilisieren. Die Potentialfunktion besitzt eine Plateau-Region (siehe Formel 2.26) und eine Übergangsregion (grau \rightarrow schwarz) in der sich das Potential langsam verringert (siehe Formel 2.25). In der schwarzen Region ist das Potential null. Die Potentialfunktion wird durch vier Parameter charakterisiert (Position des Zentrums, Radius der Plateau-Region, Radius der Übergangsregion und die Höhe des Strafterms).

Es handelt sich, im Gegensatz zur Temperatur, nicht um einen globalen Parameter des Systems. Dies bringt den Vorteil mit sich, dass die Überlappung der Potentiale zwischen den Replikaten hoch ist, sodass Übergänge von Konformationen zwischen den Replikaten mit einer höheren Wahrscheinlichkeit stattfinden. Abhängig von dem Bereich in dem sich die Winkel befinden, wird ein zusätzliches Potential (Bias) hinzugefügt um diesen Bereich zu stabilisieren. Die Methode wird daher als *Biasing-Potential Replica Exchange* bezeichnet (200, 201, 203). Das Biasing-Potential hängt von ihrer Region der Winkel-Paare und der Kraftkonstante k , die sich in jedem Replika unterscheidet und die Stärke der Destabilisierung bestimmt, ab und ist folgendermaßen definiert:

$$v_i(\eta, \theta) = k \times \begin{cases} ((x - r)^2 - (R - r)^2)^2, & \text{falls } r < x < R \\ (R - r)^4, & \text{falls } x < r \\ 0, & \text{ansonsten} \end{cases} \quad (2.25)$$

Für die Distanz x zwischen der aktuellen Position und dem Zentrum gilt der BP-Region:

$$x = \sqrt{(\eta - \eta_r)^2 + (\theta - \theta_r)^2} \quad (2.26)$$

Das zusätzliche Potential ist somit 0, falls keine dieser Bedingungen erfüllt ist. R und r sind Parameter die den Bereich des Plateaus und des abfallenden Region definieren. Die Effizienz dieser Methode hängt jedoch von der Anzahl der zu destabilisierenden Winkel ab. Um die Anzahl der zu destabilisierenden Winkel zu reduzieren wurden bei RNA Strukturen das vereinfachte RNA-Rückgrat Modell mit zwei Pseudowinkeln η ($C4'_{n-1}$, P_n , $C4'_n$, P_{n+1}) und θ (P_n , $C4'_n$, P_{n+1} , $C4'_{n+1}$) von Duarte und Pyle verwendet. Diese Repräsentation wurde bereits in einer Studien verwendet um strukturelle Eigenschaften von RNA Molekülen zu beschreiben (204). Für DNA Strukturen wurden zwei neue Pseudowinkel definiert, die es erlauben Übergänge vom ϵ/ζ und α/γ Konformationsraum gleichzeitig zu induzieren.

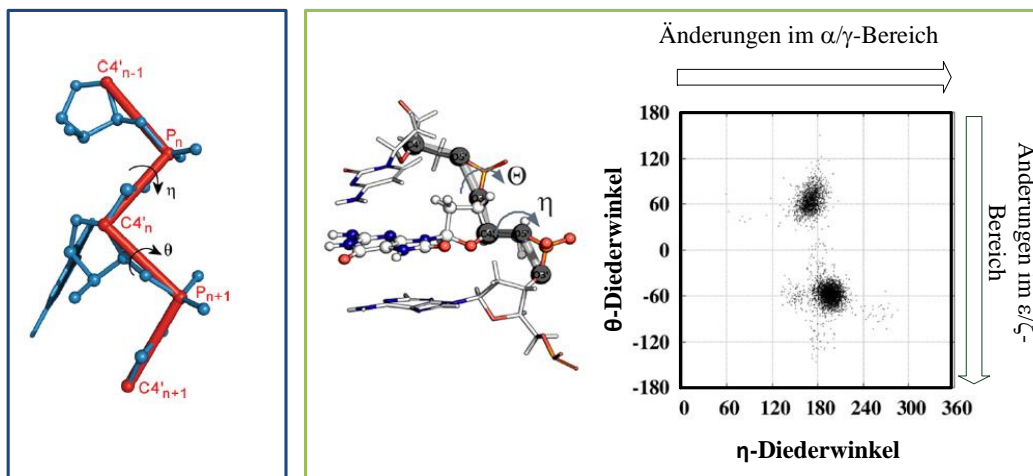


Abbildung 2.10 Veranschaulichung der verwendeten Pseudo-Diederwinkel η und θ . Für RNA-Bulge Simulationen wurde der Pseudo-Diederwinkel über die Atome P und $C4'$ definiert: η ($C4'_{n-1}$, P_n , $C4'_n$, P_{n+1}) und θ (P_n , $C4'_n$, P_{n+1} , $C4'_{n+1}$) (blaue Box). Die rote Linie beschreibt die Pseudo-Bindung zwischen den Atomen. Die Pseudo-Diederwinkel erstrecken sich über den vorherigen und den nächsten Nachbarn. Für RNA Strukturen konnte gezeigt werden, dass anhand dieses Pseudo-Diederwinkels verschiedene Konformationen beschrieben werden können (204). Diese Winkel-Definition ist geeignet um globale Konformationsänderungen zu induzieren. Die grüne Box beschreibt die Diederwinkel die für die DNA-Schaden Simulationen verwendet wurden: η ($O3'_{n-1}$, $O4_n$, $C4'_n$, $O3'_n$) und θ ($C4'_n$, $O3'_n$, $O5'_{n+1}$, $C4'_{n+1}$). Der Vorteil dieser Definition ist, dass simultan Übergänge im α/γ und ϵ/ζ -Bereich induziert werden. Beispielhaft wurde hier das 8oxoG Modell (balls and stick Repräsentation) und die nächsten Nachbarn (stick Repräsentation) dargestellt. Änderungen im η -Diederwinkel korrelieren mit Änderungen im α/γ Bereich, während Änderungen im θ -Diederwinkel mit dem ϵ/ζ -Bereich korrelieren.

2.5 Modellierung der Systeme

2.5.1 7,8-dihydro-8-oxoguanin

Molekulardynamik-Simulationen

Grundlage für alle Molekulardynamik Simulationen waren, mit Hilfe des Programms *nucgen* modellierte, reguläre B-DNA Strukturen. Für die Simulationen wurden selbst-komplementäre Strukturen mit folgenden Sequenzen verwendet: d(5'-CCAGCGCTGG)₂ (intakte DNA) und d(5'-CCA(8OG)CGCTGG)₂ (modifizierte DNA) (205). Der Vorteil solch einer Sequenz ist, dass die Struktur auf beiden Strängen identische Motive hat, sodass eine bessere Kontrolle der Konvergenz gewährleistet ist. Zusätzlicher Vorteil ist, dass für beide Sequenzen Kristallstrukturen zum Vergleich vorhanden sind. In dieser Arbeit wurde das *pmemd*-Programm des *amber12*-Pakets für die Durchführung der Molekulardynamik Simulationen verwendet (206). Um die Systeme miteinander vergleichen zu können, wurden alle Simulationen unter identischen Bedingungen durchgeführt. Für die Simulation wurde ein NPT-Ensemble verwendet. Das heißt, dass unter konstantem Druck, konstanter Temperatur und einer konstanten Anzahl an Teilchen, simuliert wurde.

Für die Erstellung der Eingabeparameter für die Simulation wurde das Programm *tLEaP* des *amber*-Pakets verwendet. Alle Systeme wurden unter periodischen Randbedingungen mit variabler TIP3P-Wasserbox simuliert. Hierbei ist eine Anpassung der Boxgröße während der MDS notwendig. Wenn der Druck größer als der Referenzdruck ist, wird in jedem Simulationsschritt der Druck auf die Seitenwände berechnet und das Volumen gegebenenfalls angepasst. Als Referenzdruck wurde in allen Simulationen 1 bar festgelegt. Die Größe des Systems wurde über einen cut-off von 10 Å festgelegt, d.h. dass alle Atome (Solut) einen Mindestabstand von 10 Å zur Seitenwand der Box besitzen. Zudem wurde in allen Systemen eine oktaedrische Box verwendet. Als Solvationsmodell wurde das TIP3P-Modell verwendet. In diesem Modell werden alle Wassermoleküle als starre Moleküle mit drei Interaktionspunkten angesehen. Die Particle-mesh-Ewald (PME)-Methode wurde für die Berechnung der weitreichenden elektrostatischen Wechselwirkungen verwendet (207, 208).

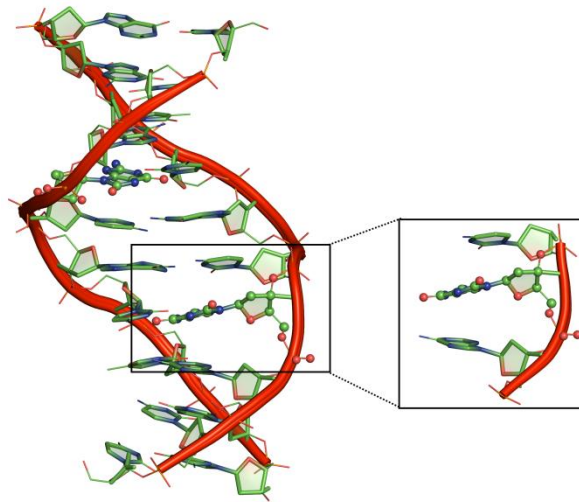


Abbildung 2.11 Selbstkomplementäre Struktur mit zwei 8oxoG-Basen. Die 8oxoG Base ist hier in der balls und stick Repräsentation dargestellt.

Alle System wurden mit Hilfe von Na^+ -Ionen neutralisiert. Als Kraftfeld wurde das amber-parmbsc0 Kraftfeld verwendet. Für die 8oxoG Base wurden die Parameter von Cheng et al. verwendet, wobei auch hier die Phosphatrückgrat Korrekturen des parmbsc0 Kraftfeld hinzugefügt wurden (74, 153). Als cut-off für weitreichende Wechselwirkungen wurde 9 \AA festgelegt. Für sämtliche Rechnungen wurde der SHAKE-Algorithmus (209) aktiviert. Mit Hilfe des SHAKE-Algorithmus wird das System in seinen Freiheitsgraden eingeschränkt. Die Verwendung von SHAKE ermöglicht eine Schrittweite von 2 fs. Hierbei werden alle Bindungslängen für alle Bindungen, die Wasserstoffatome eines beliebigen Typs beinhalten, konstant gehalten und die Energie aus den entsprechenden harmonischen Potenzialen nicht berechnet. Um möglichst natürliche Bedingungen zu erhalten wurden alle Simulationen bei einer Temperatur von 300 K durchgeführt. Um dies zu ermöglichen wird das System mit einem Thermostat gekoppelt. Hier wurde in allen Simulationen das Berendsen Thermostat für die Temperaturkontrolle verwendet. Zunächst wurden die Wassermoleküle des Systems mit der Methode des steilsten Abstiegs minimiert (5000 Schritte) um ungünstige Wechselwirkungen untereinander aufzuheben. Die DNA wurde mit Hilfe von harmonischen Rückstellkräften fixiert ($100 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$). Die Koordinaten aus der ersten Minimierung wurden übernommen. Anschließend wurde die DNA-Struktur mit der Methode des konjugierten Gradienten (2500 Schritte), gefolgt von 2500 Schritten mit der Methode des steilsten Abstiegs minimiert. Die Koordinaten der minimierten Struktur wurden für die Aufheizphase übernommen. Das System wurde in neun Schritten (jeder

Schritt 100ps) auf 300K aufgeheizt. Hierbei wurden die Rückstellkräfte während der Aufheizphase graduell von 25 kcal mol^{-1} auf $0.5 \text{ kcal mol}^{-1}$ reduziert. Der Aufheizphase folgte die Equilibrierungsphase. Die Equilibrierungsdauer betrug in allen Simulationen 1ns und wurde bei 300K ohne Rückstellkräfte durchgeführt. Die äquilibrierte Struktur wurde als Startstruktur für die kontinuierliche MD und die BP-REMD Simulation verwendet.

Biasing potential Replica-Exchange: Anwendung von Pseudo-Diederwinkel

Bei der Anwendung der Biasing-Potential Replica Exchange Molecular Dynamics (BP-REMD) Methode wurden die Pseudo-Diederwinkel η und θ destabilisiert. Es handelt sich hierbei um Pseudowinkel, die sich über mehrere Atome des DNA Phosphatrückgrats erstrecken. Der η -Winkel ist über die Atome $\text{O3}'(n-1)\text{-O5}'(n)\text{-C4}'(n)\text{-O3}'(n)$ und der θ -Winkel über die Atome $\text{C4}'(n)\text{-O3}'(n)\text{-O5}'(n+1)\text{-C4}'(n+1)$ definiert. Änderung dieser Winkel korrelieren mit Änderungen im α/γ und ϵ/ζ Konformationsbereich. Diese Repräsentation ermöglicht es, simultan α/γ und ϵ/ζ Winkel zu destabilisieren.

Die Position des Biasing-Potentials wurde anhand der Pseudo-Diederwinkel von 96 experimentell ermittelten 8oxoG bestimmt. Die Breite (Durchmesser) des Potentials wurde auf 60° festgelegt. Insgesamt wurden beide 8oxoG (bei intakter DNA die entsprechende Guanin) Nukleotide und die jeweils gegenüberliegenden Nukleotide destabilisiert.

Für die BP-REMD Simulation wurden neun Replika verwendet. Die Differenz der Biasing Potentiale zwischen den Replika betrug $0.66 \text{ kcal mol}^{-1}$ wodurch eine Akzeptanzrate für Übergänge von etwa $\sim 50\%$ erreicht wurde. Alle 1000 Schritte (2ps) wurde anhand des Metropolis Kriteriums ein Übergang akzeptiert bzw. abgelehnt.

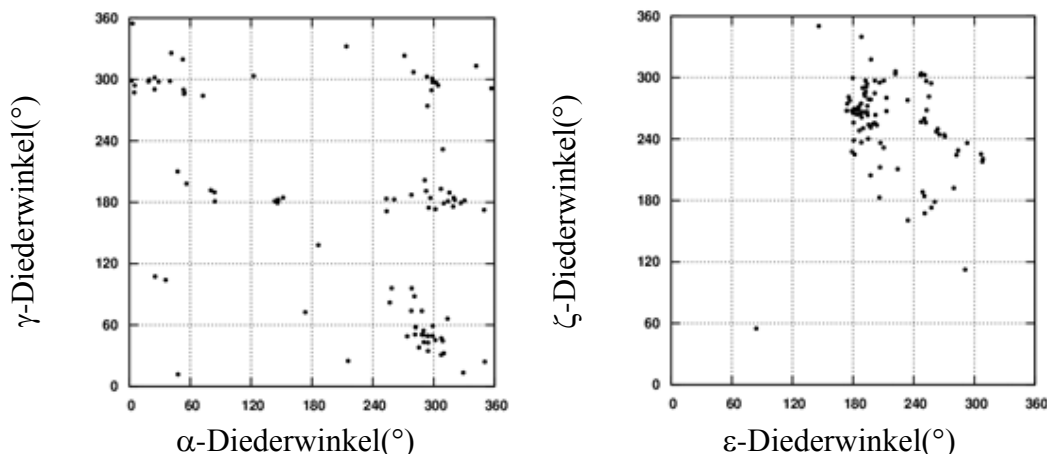


Abbildung 2.12 α/γ and ϵ/ζ Konformationsbereiche von 96 experimentellen 8oxoG-Basen. Folgende PDB Strukturen wurden für die Analyse verwendet. 1FYI; 1N2W; 183D; 178D; 3I4M; 3I4N; 3I0W; 3I0X; 3IN5; 2W7O; 2W7P; 3KNT; 3LZI; 3LZJ; 3LDS; 3OHA; 3OHB; 2XCA; 2XCP; 2XHI; 3Q8P; 3Q8Q; 3Q8R; 3Q8S; 3RJE; 3RJE; 3RJE; 3RJE; 3RJE; 3RJI; 3RJJ; 3RJK; 1EBM; 1LPQ; 1MQ2; 1MQ3; 1N3C; 1Q9Y; 1R2Y; 1RRQ; 1RRS; 1TK0; 1TK5; 1TK8; 1TKD; 1U45; 1U47; 1U48; 1U49; 1U4B; 1YQR; 1VRL; 1ZYQ; 2ASD; 2ASJ; 2ASL; 2F5Q; 2F5S; 2I5W; 2NOB; 2NOE; 2NOF; 2NOH; 2NOL; 2NOZ; 2C22; 2C28; 2C2D; 2C2E; 2C2R; 2UVR; 2UVU; 2UVV; 2UVW; 3CVS; 3CVT; 3CW7; 3FSQ; 3G0Q; 3GII; 3GIJ; 3GIK; 3GIL; 3GIM; 3GPP; 3GPU; 3GPY; 3GQ3; 3GQ4; 3GP1; 3GO8

Zur Analyse der Unterschiede zwischen intakter und modifizierter DNA wurde der RMSD (engl. root mean square deviation) berechnet. Der RMSD ist eine Methode zur Beurteilung von Konformationsänderungen. Die Konformation eines Moleküls ist über seine $3N-6$ Koordinaten im Raum definiert, wobei N die Anzahl der Atome ist. Der RMSD gibt die Ähnlichkeit von zwei Molekülen anhand ihrer Differenz in ihren internen Koordinaten wieder, wobei die mittlere quadratische Abweichung der Koordinaten als Beurteilungskriterium für die Differenz herangezogen wird. Um den RMSD-Wert zu ermitteln müssen die Moleküle gegeneinander aligniert werden. Der RMSD-Wert ist über die Konformationen A und B sowie deren kartesischen Koordinaten x,y und z definiert:

$$rmsd = \sqrt{\frac{1}{N} \left(\sum_{i=1}^N (x_{i,A} - x_{i,B})^2 + \sum_{i=1}^N (y_{i,A} - y_{i,B})^2 + \sum_{i=1}^N (z_{i,A} - z_{i,B})^2 \right)} \quad (2.27)$$

Tabelle 2.1 BP-REMD Parameter. Zur Beschreibung des BP-Potentials werden vier Parameter benötigt. Neben der Position der BP-Zentrum (η/θ -Winkel) muss der Radius der Plateau-Region (ausgehend vom Zentrum) und Radius der Tail Region (ausgehend vom Zentrum) definiert werden. Zusätzlich wird die maximale Höhe des Strafterms (E_{\max}) definiert.

	η -Winkel	Θ -Winkel	Plateau-Region	Tail-Region	E_{\max}
Zentrum 1	180°	60°	30°	60°	5.28
Zentrum 2	180°	-60°	30°	60°	5.28

Die RMSD-Werte wurden mit dem Programm *ptraj* des amber12-Paketes ermittelt. Zudem wurden die Phosphatrückgrat Winkel und die Wasserverteilung mit *ptraj* berechnet. Helikale Parameter der DNA wurden mittels Curves+ berechnet (175).

Berechnung der *potential of mean force* entlang der Diederwinkel

Um die Konformationsverteilung der Simulationen genauer zu untersuchen wurden PMF Rechnungen für den ε -Diederwinkel durchgeführt. Die Übergänge der Winkel in verschiedene Zwischenzustände werden mittels eines Umbrella Potentials induziert. Eine Rotation um den ε -Winkel beschreibt den Übergang vom BI zum BII Zustand. Alle PMF Rechnungen wurden am Rest 4 (8oxoG bzw. G) mit einer Kraftkonstante von 0.02 kcal/mol.deg² durchgeführt. Der Referenzwinkel wurde für den ε -Winkel in 18 Schritten jeweils um 10° von 0° zu -180° verändert. Jedes Fenster wurde 10ns simuliert. Die Freie Energie (potential of mean force) entlang dieser Reaktionskoordinate wurde mit der WHAM-Methode berechnet.

2.5.2 O⁶-Methylguanin

Für die Simulation wurde die DNA Sequenz d(5'-GCCATG(O⁶-MeG)CTAGTA)₂ der Struktur mit dem PDB – ID 1T38 herangezogen (210). Es handelt sich hierbei um ein Protein-DNA Komplex mit dem Reparaturenzym MGMT und einer DNA Struktur die eine O⁶-Methylguanin Base gepaart mit einem Cytosin beinhaltet. Die O⁶-MeG Base befindet sich in der aktiven Tasche des Enzyms. Die Sequenz diente als Grundlage zur Generierung der regulären B-DNA Struktur, welche

mittels *nucgen* generiert wurde. Zusätzlich wurde eine Struktur mit entsprechendem G:C Basenpaar erzeugt d(5'-GCCATGGCTAGTA)2.

Tabelle 2.2 Parameter der klassischen und der BP-REMD Simulation. Parameter der BP-REMD Simulation sind aus Tabelle 2.1 zu entnehmen.

	6MG:C	6MG:T	G:C
Kraftfeld	Parmbsc0	Parmbsc0	Parmbsc0
Wassermodell	TIP3P	TIP3P	TIP3P
Temperatur	300K	300K	300K
Druck	1bar	1bar	1bar
PME	Ja	Ja	Ja
Neutralisierung	Na+ Ionen	Na+ Ionen	Na+ Ionen
Simulationsdauer (kMDS) in ns	200	200	200
Simulationsdauer (BP-REMD) in ns	8	8	8
Anzahl Replika (BP-REMD)	9	9	9

Ferner wurde eine Struktur mit der Basenpaarung O⁶-MeG:T erzeugt, sodass ein direkter Vergleich zwischen G:C, O⁶-MeG:C und O⁶-MeG:T Basenpaarungen möglich war. Die Kraftfeldparameter für O⁶-MeG wurden aus der Modified Nucleic Acid Parameter Database von Pirro Saro und John SantaLucia entnommen (<http://ozone3.chem.wayne.edu/>). Die Kraftfeldparameter wurden entsprechend des parmbosc0 Kraftfeldes angepasst. Das Simulationsprotokoll entspricht der 8oxoG Simulationen (siehe Abschnitt 2.5.1). Für alle drei Strukturen wurden 200ns lange klassische MDS durchgeführt. Zusätzlich wurden für alle drei Strukturen BP-REMD Simulation durchgeführt. Für alle Simulationen galten identische Bedingungen:

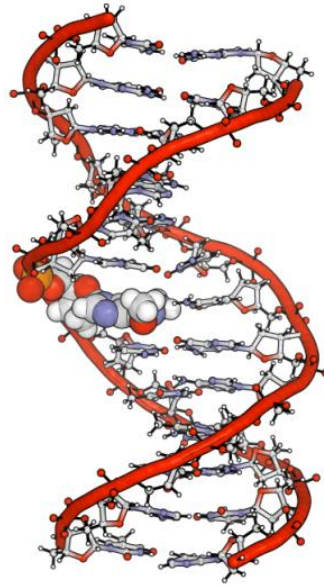


Abbildung 2.13 B-DNA Modell der O⁶-MeG:C Fehlpaarung. Der DNA-Schaden befindet sich im zentralen Bereich der DNA. Die O⁶-MeG-Base ist in der van der Waals Repräsentation dargestellt.

2.5.3 Dangling-End-Strukturen

Alle Molekulardynamik-Simulationen, sowie die Freien Energierechnungen wurden mit dem amber9 Packet durchgeführt. Die Startstrukturen (Doppelstrang RNA/DNA, Einzelstrang RNA/DNA) wurden mit dem *nucgen* Programm des amber9 Pakets generiert. Das Simulationsprotokoll entspricht dem Protokoll aus Abschnitt 2.5.1. Zur Generierung der Dangling-End-Struktur wurde die entsprechende Base am 3' bzw. 5' Ende entfernt. Allen Sequenzen ist gemein, dass sie einen GC-reichen Kern besitzen (siehe Tabelle 2.3). Um den Beitrag der Freien Energie eines Dangling-End Motivs zu bestimmen wurden Umbrella-Sampling Simulationen durchgeführt. Als Reaktionskoordinate wurde der Massenschwerpunkt der Schweratome der Dangling-End Base und der Massenschwerpunkt der Nachbarbase verwendet (siehe Abbildung 2.14). Die WHAM Methode wurde verwendet um das freie Energie Profil entlang der Reaktionskoordinate zu bestimmen (183). Wie aus der Tabelle 2.3 zu entnehmen ist, weisen die Strukturen zwei identische Motive auf. Dies ermöglicht es simultan an beiden Dangling-End Motiven die Freie Energie zu bestimmen, sodass eine bessere Kontrolle der Ergebnisse möglich ist.

Tabelle 2.3 *Dangling-End* Sequenzen. Die verwendeten Sequenzen bestehen nur aus Guanin und Cytosin Basen. Hierbei wurden alle Kombinationen der Motive untersucht. Alle Motive wurden sowohl für RNA als auch für DNA Strukturen untersucht.

	Sequenz	Motiv
Guanin-Überhang	5'-GGCGCGCGC -'3	5'-GG
	3'- CGCGCGCGG-'5	3'- C
	5'-GCGCGCGCG -'3	5'-GC
	3'- GCGCGCGCG-'5	3'- G
	5'- CGCGCGCGG-'3	5'- C
	3'-GGCGCGCGC -'5	3'- GG
	5'- GCGCGCGCG-'3	5'- G
	3'-GCGCGCGCG -'5	3'- GC
Cytosin-Überhang	5'-CCGCGCGCG -'3	5'-CC
	3'- GCGCGCGCC-'5	3'- G
	5'-CGCGCGCGC -'3	5'-CG
	3'- CGCGCGCGC-'5	3'- C
	5'- GCGCGCGCC-'3	5'- C
	3'-CCGCGCGCG -'5	3'- CC
	5'- CGCGCGCGC-'3	5'- G
	3'-CGCGCGCGC -'5	3'-CG

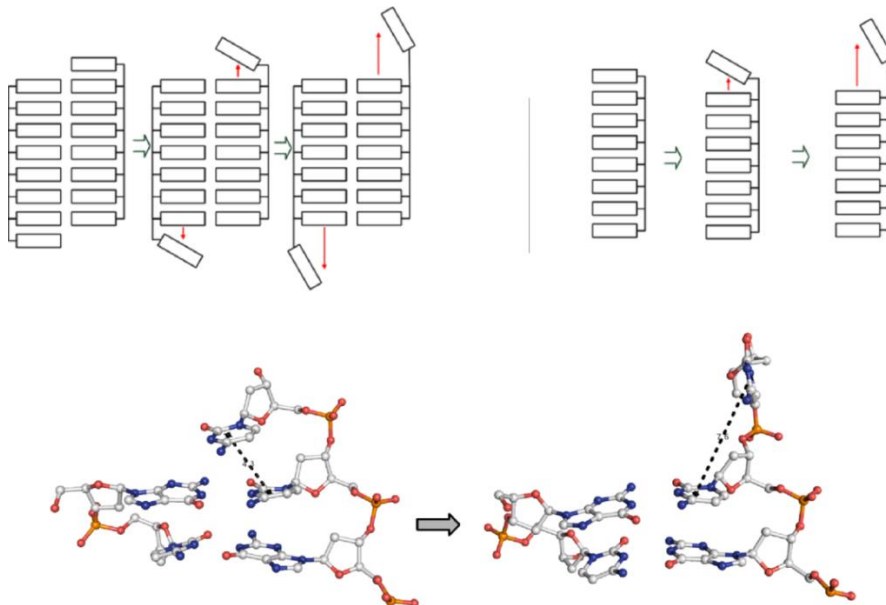


Abbildung 2.14 Reaktionskoordinate vom gestackten zum völlig solvatisierten Zustand. Als Distanz wurde der Massenschwerpunkt der Base verwendet und während der Simulation unter Verwendung von elf Fenstern von 5 Å → 10 Å erhöht.

Für die Umbrella-Sampling Simulationen wurden 11 Fenster verwendet, wobei die Distanz der Massenschwerpunkte von 5 Å auf 10 Å in 0.5 Å Schritten erhöht wurde. Jedes Fenster wurde 10 ns simuliert. Entsprechende Simulationen für jedes Motiv wurden auch für den Einzelstrang simuliert. Neben den Umbrella-Sampling Simulationen wurden Molecular Mechanics Poisson-Boltzmann Surface Area (MM/PBSA) Rechnungen durchgeführt. Für die MM/PBSA bzw. MM/GBSA-Rechnungen wurde das MMPBSA.py Skript von Holger Gohlke verwendet, welches standardmäßig im amber12-Programmpaket integriert ist (211). Für alle Berechnungen jeder Trajektorie wurden über eine Dauer von 10 ns in regelmäßigen Abständen von 2 ps insgesamt 5000 Koordinatenschnappschüsse erzeugt.

Wie bereits in Kapitel 2.3.2 erwähnt, greifen MM/PBSA bzw. MM/GBSA Rechnungen auf Kontinuum-Solvatationsmodelle (z.B. Poisson-Boltzmann-Modell oder Verallgemeinerte-Born-Rechnungen) zur Beschreibung von freien Solvatationsenthalpien zurück. Das elektrostatische Potential wird mit Hilfe der Poisson-Boltzmann-Gleichung berechnet.

$$-\varepsilon_0 \nabla(\varepsilon(r) \nabla \phi(r)) - \kappa \sinh(\phi(r)) = -4\pi \rho(r) \quad (2.28)$$

ε_0 = Dielektrizitätskonstante in einem homogenen Medium, $\varepsilon(r)$ = ortsabhängige Dielektrizitätskonstante, $\rho(r)$ = Ladungsdichte, κ wird über die inverse Debye-Hückel-Länge k berechnet.

Für die Poisson-Boltzmann-Gleichung gilt, dass \sinh alternativ als Taylor-Reihe dargestellt werden kann, sodass man folgenden Term erhält:

$$\nabla(\varepsilon(r) \nabla \phi(r)) - \kappa \phi(r) \left[1 + \frac{\phi(r)^2}{6} + \frac{\phi(r)^4}{120} + \dots \right] = -4\pi \rho(r) \quad (2.29)$$

Wird bei der Berechnung nur der erste Term betrachtet, d.h. nur das erste Glied der Taylor-Reihe, erhält man die linearisierte Poisson-Boltzmann-Gleichung:

$$\nabla(\varepsilon(r) \nabla \phi(r)) - \kappa \phi(r) = -4\pi \rho(r) \quad (2.30)$$

MM/PBSA Rechnungen basieren auf dem iterativen Lösen der linearisierten Poisson-Boltzmann-Gleichung. Da es sich hierbei um eine inhomogene Differentialgleichung zweiter Ordnung handelt, bedient man sich Methoden der

finiten Differenz um diese zu lösen. Hierbei werden Solut und Solvens in ein kubisches Gitter mit einer Gitterkonstante h gelegt.

Anschließend werden jedem Gitterpunkt elektrostatisches Potential F , Ladungsdichte r , Dielektrizitätskonstante ϵ und ionische Stärke I zugewiesen. Atomladungen werden aus den Kraftfeldparametern übernommen. Die Parameter α und β für den Freien Solvatationsbeitrag wurden aus Arbeiten von Srinivasan *et al.* 1998 übernommen. Die Dielektrizitätskonstante des Moleküls wurde auf 1, die des Solvens auf 80.0 und die Salzkonzentration auf 0.15M festgelegt.

2.5.4 RNA-Bulges

Zur Untersuchung von *RNA-Bulges* wurden zwei Strukturen verwendet deren Sequenz sich nur in der Bulge-Base unterscheidet (siehe Abbildung 2.16). Hierbei sollte das Verhalten von Purin (U) im Vergleich zu Pyrimidin (A) Basen untersucht werden. Alle klassischen MDS wurden mit dem pmemd.MPI Modul des Amber12-Packets durchgeführt. Als Startstrukturen wurden die NMR-Strukturen mit den PDBid's 2K41 und 2K3Z verwendet (212). Für die RNA Strukturen wurde das parmbsc0chiOL Kraftfeld verwendet (160, 161, 213). Das Simulationsprotokoll entspricht dem Simulationsprotokoll aus Abschnitt 2.5.1. Für die BP-REMD wurden die Pseudo-Diederwinkel $\eta(C4'_{n-1}, P_n, C4'_n, P_{n+1})$ und $\theta(P_n, C4'_n, P_{n+1}, C4'_{n+1})$ verwendet (siehe Abbildung 2.15).

Hierbei wurden die Bulge-Base und die beiden nächsten Nachbarn destabilisiert. Es wurden zwei Winkelbereiche destabilisiert. Weitere Simulationsdetails befinden sich in Tabelle 2.4. Da die Sequenz sehr kurz ist und ein fraying, d.h. die Auflösung der WC-Paarungen, an den terminalen Basen vermieden werden sollte, wurden Distanz-Einschränkungen ($0.35 \text{ kcal/mol}^{-1}$) zwischen den Wasserstoffbrücken-Partner des WC Paares festgelegt.

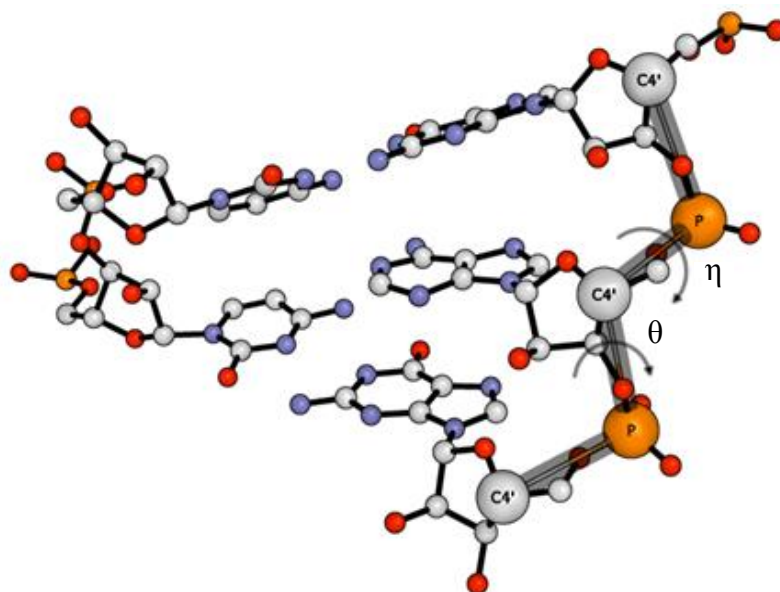


Abbildung 2.15 Pseudo-Diederwinkel des RNA-Bulge Modells. Die Pseudo-Bindungen zwischen den Atomen sind durch graue und transparente Linien markiert. Die Destabilisierung dieses Winkels ist geeignet um globale Konformationsänderungen zu induzieren.

Die Daten wurden einer Clusteranalysen unterzogen sowie RMSD Werte und Diederwinkel mittels ptraj berechnet. Die Klassifizierung der Strukturen erfolgte anhand der RMSD-Werte der Base und der nächsten Nachbarn, sowie deren WC-Partnern.

Tabelle 2.4 Parameter der klassischen und der BP-REMD Single RNA Bulge Simulation.

Parameter	kMDS	BP-REMD
Kraftfeld	parmbsc0chiOL	parmbsc0chiOL
Wassermodell	TIP3P	TIP3P
Temperatur	300K	300K
Druck	1bar	1bar
PME	Ja	Ja
Neutralisierung	Na+ Ionen	Na+ Ionen
Simulationsdauer in ns	350	16
Anzahl Replika (BP-REMD)		9

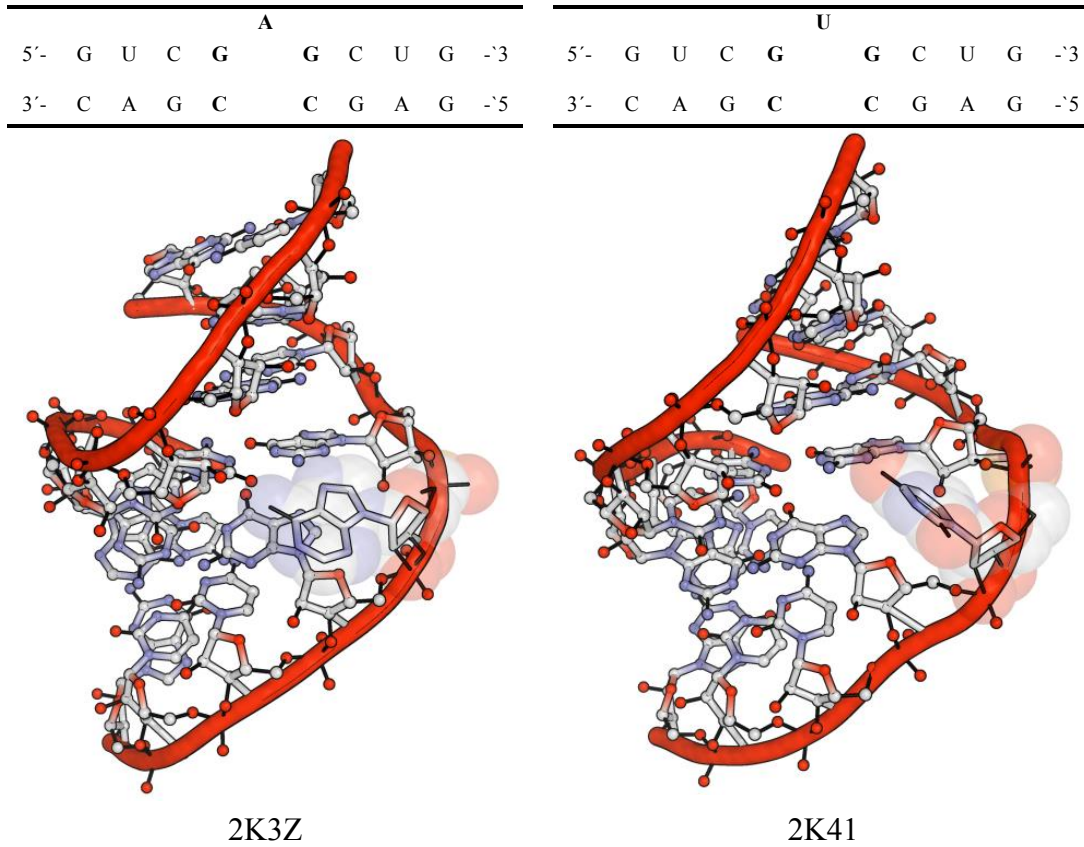


Abbildung 2.16 Primär und Tertiärstruktur der RNA-Bulge-Strukturen. Die Strukturen unterscheiden sich nur in ihrer Bulge-Base (U-Bulge vs. A-Bulge). Die Bulge Base ist jeweils in der vdW-Repräsentation transparent dargestellt. Abgebildet ist jeweils eine Struktur aus dem NMR-Ensemble.

Kapitel 3

3 Ergebnisse und Diskussion

In dieser Arbeit werden Methoden diskutiert, deren Ziel das effiziente Abtasten des Konformationsraumes von Nucleinsäuren ist. Das Verfahren beruht auf dem Prinzip der Destabilisierung von spezifisch definierten Pseudo-Diederwinkeln. Die Methode wird hinsichtlich ihrer Effizienz mit klassischen Molekulardynamik-Simulationen verglichen. Besonderes Augenmerk liegt hierbei im direkten Vergleich der beobachteten Konformationen in den jeweiligen Simulationen. Zum Vergleich der Methoden wurden *RNA-Bulge-Strukturen* und geschädigte DNA Strukturen untersucht. Die geschädigten DNA Strukturen wurden mit intakten DNA Strukturen verglichen. Annahme war hierbei, dass die geschädigten DNA Strukturen spezifische Zustände aufweisen die hilfreich für die Erkennung durch Reparaturenzyme sein können. Ziel der *RNA-Bulge* Simulationen war strukturelle Unterschiede zwischen Purin und Pyrimidin Basen im gleichen Sequenzkontext zu untersuchen. Zusätzlich wurden Umbrella-Sampling Simulationen und MM/PBSA Rechnungen durchgeführt um den Einfluss von *Dangling-End* Basen auf die Gesamtstruktur von Nucleinsäuren zu untersuchen.

3.1 Guanin und 8-oxoguanin

Im Zentrum dieses Abschnitts steht ein direkter Vergleich dem intakten Guanin Nucleotid mit dem modifizierten 8oxoG Nucleotid. Dabei wird die Interaktion zwischen der jeweiligen Base und ihren nächsten Nachbarn aus unterschiedlichen Blickwinkeln analysiert. Hierbei werden unter anderem Diederwinkel der DNA Nucleotide, die Zucker-Pucker Konformationen, Helikale Parameter und die Wasserverteilung miteinander verglichen. Annahme ist hierbei, dass die Erkennung der 8oxoG Base durch Reparaturenzyme anhand von spezifischen Konformationen erfolgt. Dabei werden keine Erkennungsmechanismen

diskutiert. Das Ziel ist vielmehr strukturelle Unterschiede zwischen beiden Modellen auszuarbeiten. Besonderes Augenmerk wird zudem auf die Stabilität der Struktur gelegt. Dabei wird insbesondere eine Korrelation zwischen den verschiedenen Diederwinkeln aufgezeigt. Um den Konformationsraum effizienter abzutasten und somit Unterschiede zwischen den Modellen auszuarbeiten wurde die BP-REMD Methode verwendet. Diese wurde wiederum mit kMDS verglichen.

Ziel ist es, zu verdeutlichen, welche Rolle und Auswirkung die 8oxoG-Base auf die Feinstruktur der DNA hat. Weiterhin soll verdeutlicht werden, inwiefern ein Unterschied zwischen den Ergebnissen der Simulationsmethoden vorliegt und welche Vorteile die BP-REMD Methode mit sich bringt.

3.1.1 Resultate der klassischen MD

In dieser Arbeit werden intakte und geschädigte DNA Strukturen miteinander verglichen. Es handelt sich bei den Modellen um selbstkomplementäre Sequenzen (siehe Abschnitt 2.5.1). Solch eine Sequenz bringt den Vorteil mit sich, dass die Struktur auf beiden Strängen im identischen Sequenzkontext eine 8oxoG Base aufweist (Position 4 und Position 14). Die intakte DNA besitzt an den entsprechenden Positionen ein Guanin. Annahme ist, dass alle äquivalente Nukleotide im äquilibrierten Zustand sehr ähnliche Konformationsverteilungen aufweisen. Ist dies nicht gegeben, kann davon ausgegangen werden, dass die Simulation nicht konvergiert ist. Die Konformationsbereiche der selbstkomplementären Sequenz dienen somit als Konvergenzkriterium für die Simulation. Im Folgenden wird auf strukturelle Unterschiede zwischen den Modellen eingegangen.

Als erstes Vergleichskriterium wurden RMSD-Werte der Schweratome des Phosphatrückgrats bestimmt. Als Referenzmolekül wurde die Startstruktur der entsprechenden Simulationen verwendet. Die RMSD-Werte der Struktur wurden mit Hilfe von *ptraj* ermittelt. Wie aus der Abbildung 3.1 zu entnehmen ist, weisen die RMSD-Werte der intakten DNA über die gesamte Trajektorie hinweg keine großen Schwankungen auf und befinden sich zwischen 1.5 und 2.5 Å (schwarze Linie). Die RMSD-Werte sprechen für eine stabile Struktur während der Simulation.

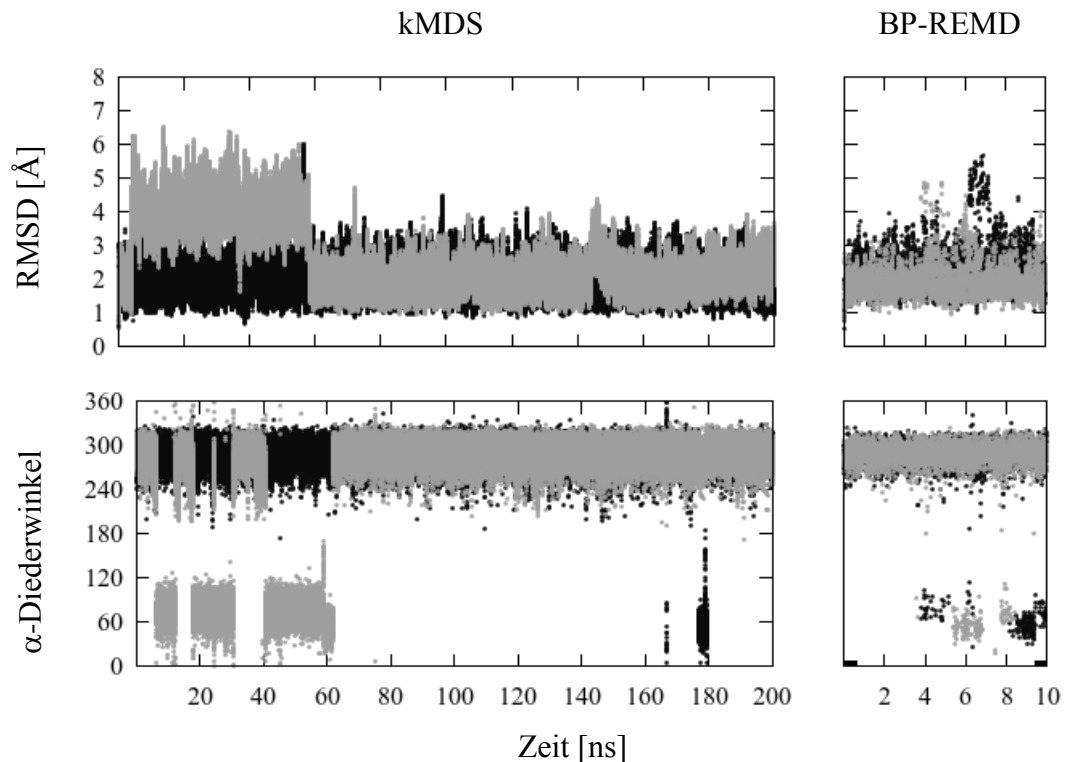


Abbildung 3.1 Zusammenhang zwischen RMSD und Diederwinkel. Im oberen Teil der Abbildung sind RMSD-Werte bezüglich der Startstruktur des Zucker-Phosphatrückgrats gegen die Zeit aufgetragen. Die graue Linie stellt den Verlauf der geschädigten DNA und die schwarze den Verlauf der intakten DNA dar. Die Werte der BP-REMD stammen aus dem Referenz-Replika (Replika 0). Der zeitliche Verlauf des α -Diederwinkel verdeutlicht den Zusammenhang zwischen dem RMSD Wert und dem Diederwinkel (gilt für 8oxoG-Simulation). Zudem wird deutlich, dass sich die Verteilungen des α -Diederwinkel des Restes 4 und des Restes 14 unterscheiden.

Die RMSD-Werte für die geschädigte DNA weisen auf stärkere Konformationsänderungen während der Simulation hin. So ist ein Anstieg von 2 auf etwa $\sim 4\text{--}5$ Å nach 4 ns zu beobachten. Dieser Zustand wird etwa 50 ns beibehalten. Nach 54 ns kehrt der RMSD-Wert auf einen ursprünglichen Zustand von ~ 2 Å zurück. Wichtig ist hierbei zu erwähnen, dass solch ein Verhalten für intakte DNA während der 200ns nicht beobachtet wird. Ursache für den Verlauf des RMSD-Wertes der geschädigten DNA sind Konformationsänderungen an der Base 8oxoG₁₄. Diese Konformationsänderungen sind hervorgerufen durch Übergänge im α/γ Konformationsbereich (crank shift Bewegungen). Unter Verwendung der kMDS über einen Simulationszeitraum von 200ns konnte keine Konvergenz für das Modell mit geschädigter DNA erreicht werden, da sich die Konformationsverteilung im α/γ Konformationsbereich zwischen 8oxoG₁₄ und 8oxoG₄ unterscheiden.

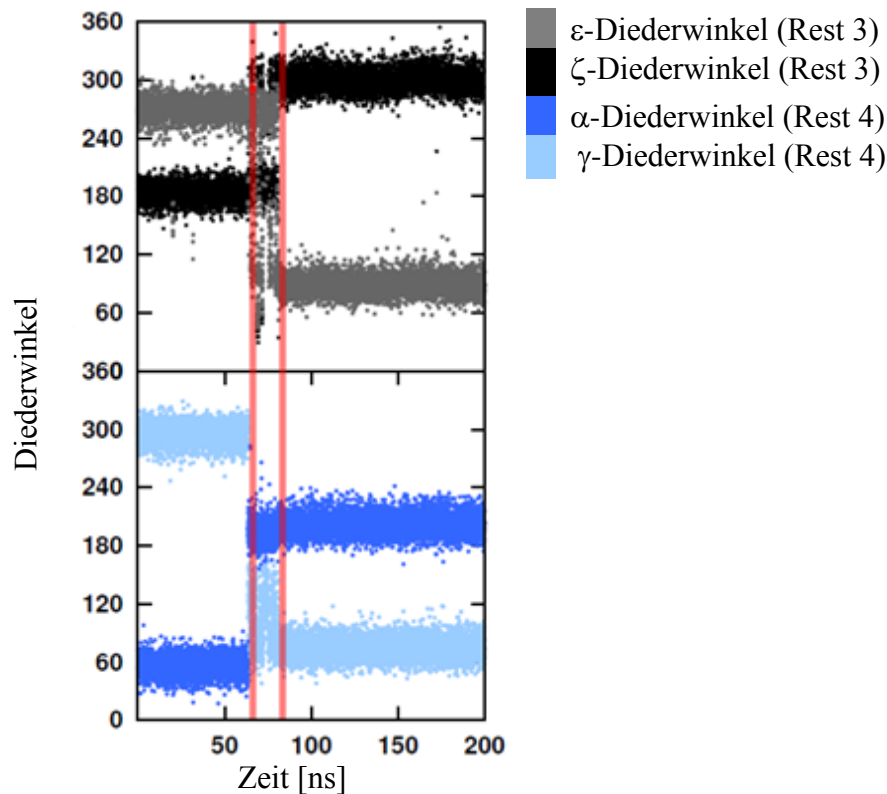


Abbildung 3.2 Zusammenhang zwischen Diederwinkel und die Auswirkungen des parm99 Kraftfeldes. Dargestellt sind die α und γ Diederwinkel der 8oxoG Base (Rest 4) und die ε und ζ Winkel des vorherigen Nachbarn. Ein irreversibler α/γ Übergang induziert einen irreversiblen ε/ζ Übergang. Aufgrund der Problematik ist das parm99 Kraftfeld nicht für lange MD Simulationen geeignet.

Diese Beobachtungen zeigen sich umso deutlicher, falls man statt dem parmbsc0-Kraftfeld, das parm99-Kraftfeld verwendet. So wurden 200ns lange Simulationen ohne die parmbsc0 Kraftfeld-Korrekturen mit dem identischen System durchgeführt. Die α/γ Übergänge sind hierbei irreversibel und beeinflussen die nächsten Basen (siehe Abbildung 3.2). Die Übergangsphase ist hierbei durch rote Balken markiert. Dieses Ergebnis bekräftigt die von der Literatur empfohlene Verwendung des parmbsc0 Kraftfeldes (118, 214, 215).

Neben den α/γ Diederwinkel sind auch ε/ζ miteinander gekoppelt und lassen sich unterteilen in BI und BII-Zustände (siehe Kapitel 2.2.1).

Die Verteilung dieser Zustände wurde mittels Curves+ ermittelt (175). Wie aus der Tabelle 2.1 zu entnehmen ist unterscheiden sich die Verteilungen zwischen 8oxoG und Guanin signifikant.

Tabelle 3.1 Relative Häufigkeiten des BI/BII Zustandes während der kMDS im Vergleich (in %). Die Verteilung der BI/BII der W (8oxoG bzw. G) und C (C) –Partner verdeutlicht in beiden die unterschiedliche Verteilung des BI/BII Zustandes.

8-oxoguanin				
	kMDS			
	BI		BII	
ResID	4	14	4	14
W	65	100	35	0
C	100	70	0	30
Guanin				
	kMDS			
	BI		BII	
ResID	4	14	4	14
W	92	100	8	0
C	100	91	0	9

An dieser Stelle ist jedoch anzumerken, dass die Unterschiede zwischen den Basen im jeweiligen System sich nur geringfügig unterscheiden. Sowohl für G_4 als auch für G_{14} konnte mit 92 % bzw. 91 % ein dominanter BI-Zustand bestimmt werden. Diese Dominanz ist nicht so stark ausgeprägt für die 8oxoG Basen. Der BI-Zustand ist mit 65% bzw. 70% zwar der dominantere, jedoch unterscheidet sich diese Verteilung deutlich zu den Guanin Basen. Um die unterschiedlichen Verteilungen genauer zu definieren wurden PMF Rechnungen am jeweiligen System an der Position 4 (G_4 bzw. $8oxoG_4$) durchgeführt. Als Reaktionskoordinate diente der ϵ -Diederwinkel (siehe Abschnitt 2.5.1). Die PMF-Rechnung sollte Aufschluss über die die Freien Energiedifferenzen von BI zum BII Zustand für die G_4 und $8oxoG_4$ Base zu liefern (siehe Abbildung 3.3). Wie aus Abbildung 3.3 zu entnehmen ist beträgt die Freie Energiedifferenz zwischen BI und BII für die Base G_4 $1.3 \text{ kcal}\cdot\text{mol}^{-1}$ und für $8oxoG_4$ $0.5 \text{ kcal}\cdot\text{mol}^{-1}$. Dieses Ergebnis ist im Konsens mit dem Ergebnis aus der kMDS (G_4 $1.4 \text{ kcal}\cdot\text{mol}^{-1}$ vs. $8oxoG_4$ $0.4 \text{ kcal}\cdot\text{mol}^{-1}$). Ursache für den Konsens sind die geringen Energiebarrieren vom BI zum BII Zustand (siehe Abbildung 3.3). Es kann aufgrund dieser Ergebnisse angenommen werden, dass die Verteilung der BI/BII-Zustände konvergiert ist.

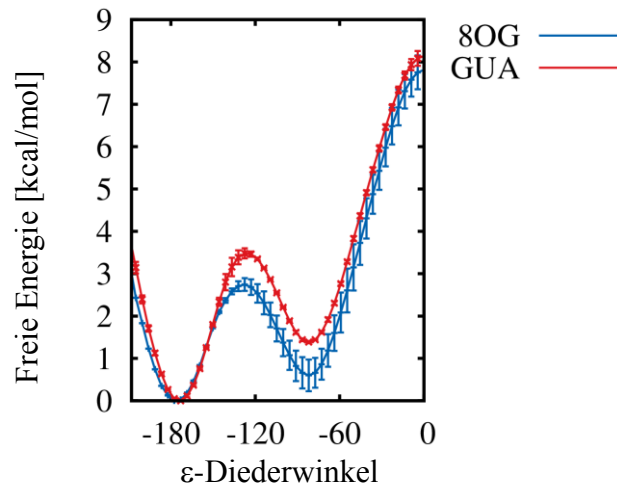


Abbildung 3.3 PMF-Kurve des ϵ -Diederwinkel. Berechnet wurden die Energieprofile entlang des ϵ -Diederwinkel für die 8oxoG-Base (Rest4, blaue Linie) und die Guanin-Base (Rest 4). Die Minima entsprechen dem BI ($\sim -180^\circ$) und dem BII (~ -80) Zustand. Aus dem Energieprofil wird deutlich, dass die Übergangsbarrriere vom BI \rightarrow BII Zustand der 8oxoG im Vergleich zur Guanin Base geringer ist. Zusätzlich unterscheiden sich die Minima des BII Zustanden.

Da jedoch die RMSD-Werte der Gesamtstruktur mit den α -Diederwinkel korrelieren, wurden der α/γ Konformationsbereich genauer untersucht. Eine detaillierte Analyse der α/γ Diederwinkel zeigt signifikante Unterschiede zwischen 8oxoG₄ und 8oxoG₁₄ in der kMDS, wohingegen für G₄ und G₁₄ keine α/γ Übergänge während der 200ns beobachtet wird (Abbildung 3.4). Insgesamt treten jedoch α/γ Übergänge im Vergleich zu ϵ/ζ seltener auf (Abbildung 3.5 und Tabelle 3.3). Ein direkter Vergleich zwischen 8oxoG₄ und 8oxoG₁₄ zeigt unterschiedliche Verteilungen der α/γ Zustände an der Position 4 und Position 14. Die Problematik besteht darin, dass durch den Übergang die Struktur in einem lokalen Minimum verweilt, sodass es zu einer Überrepräsentierung dieses lokalen Minimums in den Konformationsverteilungen kommt.

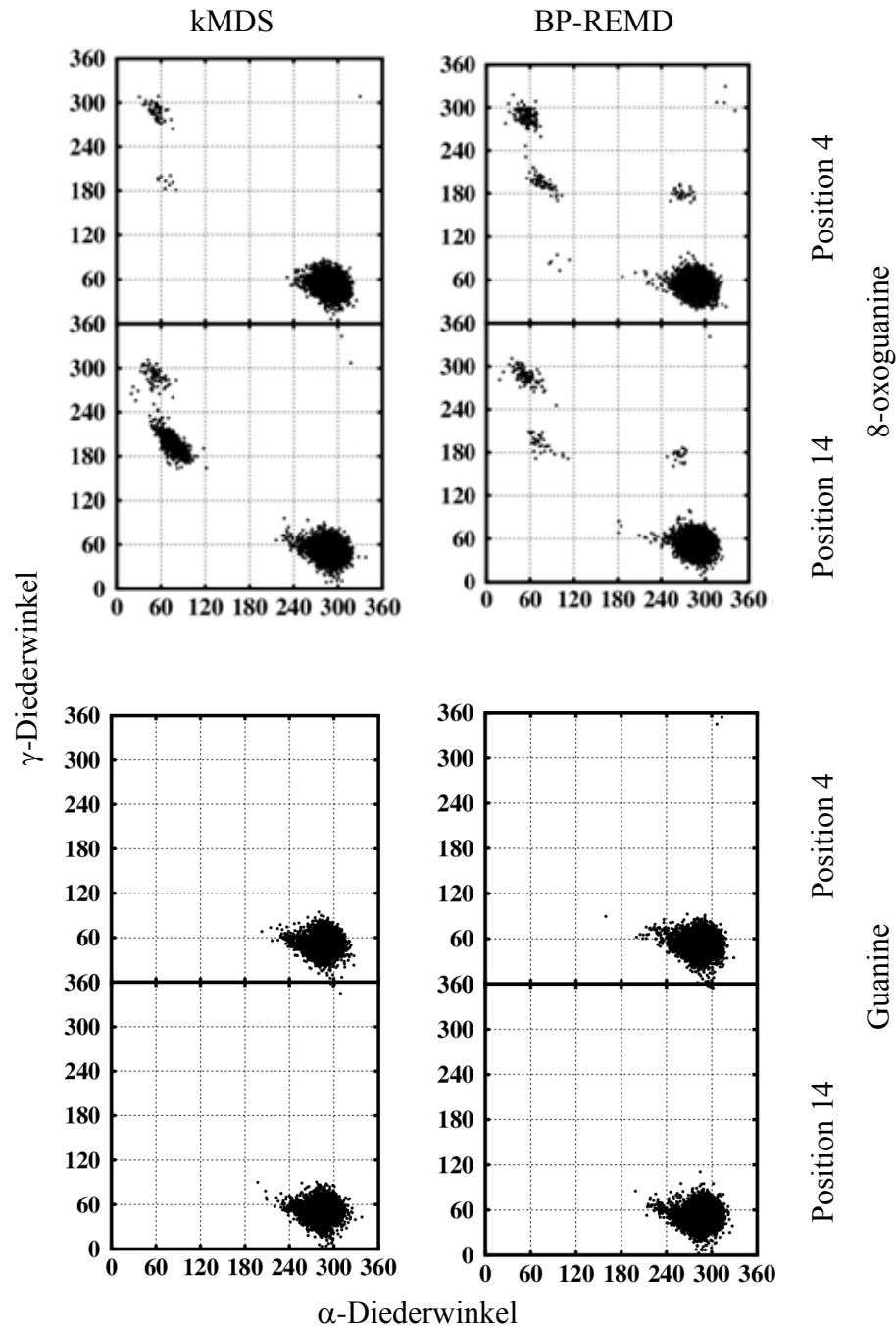


Abbildung 3.4 Konformationsverteilungen der α/γ -Diederwinkel im Vergleich. Ein Vergleich der Simulationen anhand des α/γ -Diederwinkel zeigt, dass mit der BP-REM Methode neue Konformationszustände ermittelt werden. Zusätzlich unterscheiden sich die Verteilungen zwischen Rest 4 und Rest 14 im für BP-REM Simulationen nur geringfügig. In kMDS Simulationen werden ebenfalls für die 8oxoG Base alternative Konformationszustände gefunden, jedoch unterscheiden sich die Verteilungen an den Positionen signifikant.

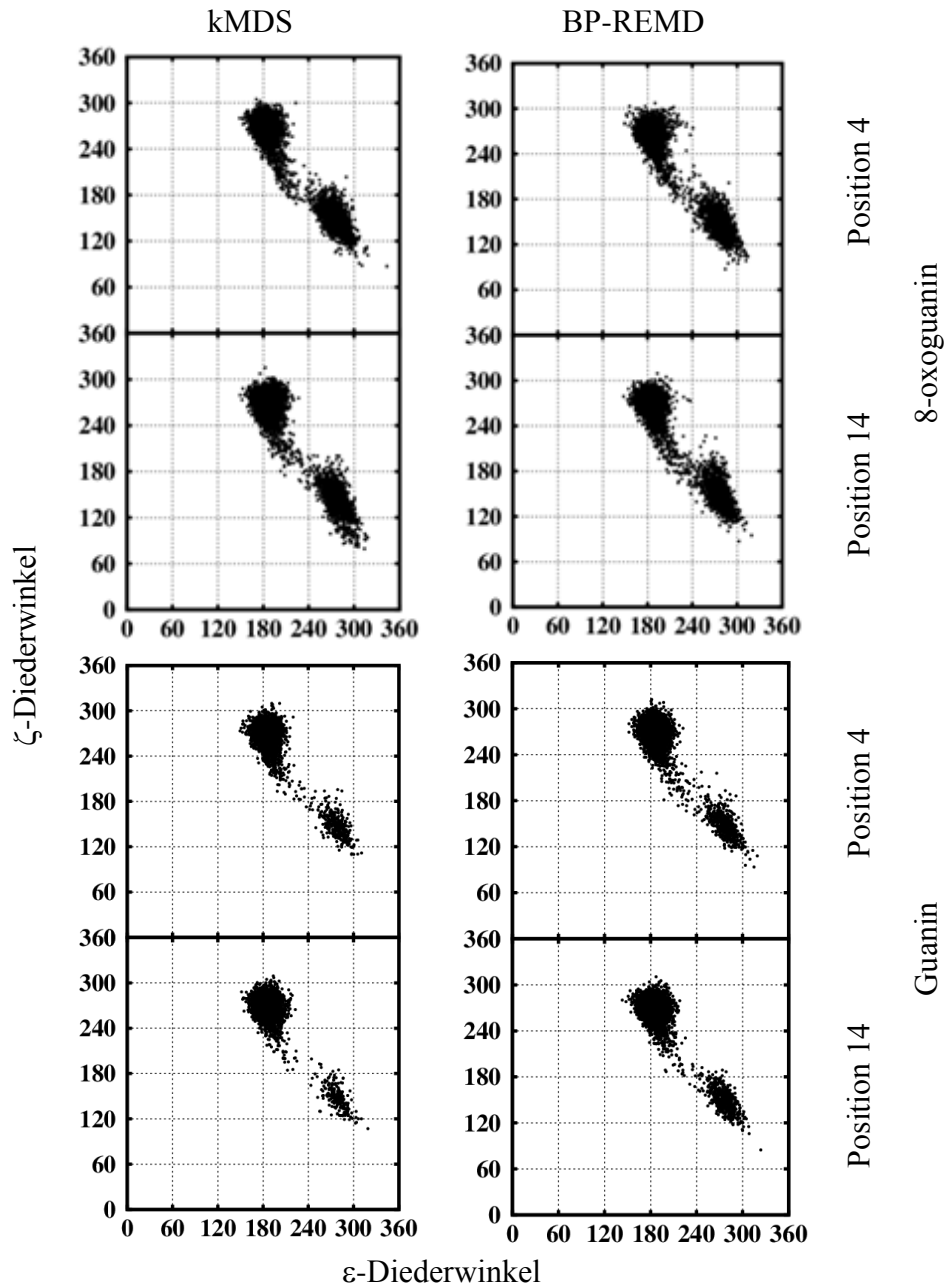


Abbildung 3.5 Konformationsverteilungen der ϵ/ζ -Diederwinkel im Vergleich. Beim direkten Vergleich der Methoden sind keine signifikanten Unterschiede zu beobachten. Beim Vergleich der 8oxoG Base mit der Guanin Base fällt auf, dass die BI/BII Verteilungen sich unterscheiden. Im Falle des 8oxoG werden mehr BII Zustände beobachtet.

3.1.2 Resultate der BP-REMD

In Abschnitt 3.1.1 wurde diskutiert, welche Folgen es haben kann, wenn in einer kMDS ein lokales Minimum in den Konformationsverteilungen überrepräsentiert wird. Um dieser Problematik gerecht zu werden, wurden BP-REMD-Simulationen durchgeführt. Test-Simulationen mit verschiedenen BP-Energien zeigten, dass eine BP-Energie von maximal $5.28 \text{ kcal}\cdot\text{mol}^{-1}$ ausreichend ist um Übergänge im α/γ und ε/ζ zu induzieren. Insgesamt wurden neun Replika verwendet, wobei die Energieunterschiede zwischen den Biasing Potentialen $0.66 \text{ kcal}\cdot\text{mol}^{-1}$ betragen. Dies hat den Vorteil, dass sich die Replika nur geringfügig unterscheiden, sodass eine Akzeptanzrate für die Übergänge der Replika etwa bei $\sim 50\%$ liegt. Die Gesamtsimulationsdauer pro BP-REMD-Simulation betrug 10ns. Analysiert und verglichen werden jeweils die kMDS Simulation mit der BP-REMD ohne Biasing Potential (Referenz-Replika).

Auch die Verwendung von Pseudo-Diederwinkeln zeigt, dass durch diese Repräsentation Übergänge im α/γ und ε/ζ erfolgreich induziert werden können. Die RMSD-Plots der BP-REMD-Simulation zeigen für die intakte und die geschädigte DNA Fluktuationen zwischen 2-4 Å (siehe Abbildung 3.1). Dies deutet auf eine erhöhte Flexibilität während der Replika-Simulation hin.

Tabelle 3.2 Relative Häufigkeiten des BI/BII Zustandes im Vergleich (in %). Die Verteilung der BI/BII der W (8oxoG bzw. G) und C (C) –Partner verdeutlicht in beiden die unterschiedliche Verteilung des BI/BII Zustandes.

8-oxoguanin								
	BP-REMD				kMDS			
	BI		BII		BI		BII	
ResID	4	14	4	14	4	14	4	14
W	65	100	35	0	65	100	35	0
C	100	67	0	33	100	70	0	30
Guanin								
	BP-REMD				kMDS			
	BI		BII		BI		BII	
ResID	4	14	4	14	4	14	4	14
W	89	100	11	0	92	100	8	0
C	100	90	0	10	100	91	0	9

Tabelle 3.3 Relative Häufigkeiten der α/γ Zustände im Vergleich (in %). Die Verteilung der α/γ Zustände Verteilungen der W (8oxoG bzw. G) und C (C) –Partner verdeutlicht den Vorteil der BP-REMD Methode.

8-oxoguanine									
ResID	BP-REMD				kMDS				
	4		14		4		14		
	W	C	W	C	W	C	W	C	
g-/t	1	0	0	1	0	0	0	1	
g-/g+	95	100	100	96	99	100	100	80	
g+/g-	3	0	0	2	1	0	0	2	
g+/t	1	0	0	1	0	0	0	18	

Guanine									
ResID	BP-REMD				kMDS				
	4		14		4		14		
	W	C	W	C	W	C	W	C	
g-/t	1	0	0	1	0	0	0	0	
g-/g+	99	100	100	99	99	100	100	99	
g+/g-	0	0	0	0	1	0	0	1	
g+/t	0	0	0	0	0	0	0	0	

Hinsichtlich der BI/BII Übergänge werden ähnliche Verteilungen wie bereits in der kMDS beobachtet. Wie bereits bei der kMDS, wurden die Energiedifferenzen anhand der Wahrscheinlichkeitsverteilung der Zustände bestimmt. Diese liegen bei $1.4 \text{ kcal}\cdot\text{mol}^{-1}$ für Guanin und $0.4 \text{ kcal}\cdot\text{mol}^{-1}$ für 8-oxoguanin, und somit im Konsens mit den PMF Rechnungen und den Ergebnissen aus der kMDS (siehe Abbildung 3.3 und Abbildung 3.4).

In den kMDS wurden für Guanin keine Übergänge im α/γ Bereich beobachtet (siehe Abbildung 3.4). Für die 8oxoG wurden Übergänge α/γ Bereich für beide Positionen (Position 4 und Position 14) beobachtet. Statistisch gesehen treten die alternativen α/γ Zustände zwar selten auf, jedoch ist wichtig zu erwähnen, dass an beiden Positionen Übergänge zu beobachten sind. Zudem wird durch den häufigen Austausch zwischen den Replika vermieden, dass ein Zustand überrepräsentiert ist (siehe Tabelle 3.3). Ein interessanter Nebeneffekt der Verwendung von Pseudo-Diederwinkeln ist, dass die Konformationsbereiche der Zucker-Pucker auch effizient untersucht werden können (siehe Tabelle 3.4). So konnten für alle Simulationen ein mit ca 50%, im Vergleich zu Guanin, sehr dominanter C1'-exo Zustand für 8oxoG bestimmt werden.

Tabelle 3.4 Relative Häufigkeit der Zucker-Pucker-Verteilung (in %). In allen Simulationen ist für die 8oxoG eine erhöhte C1'exo Präferenz zu beobachten.

	C3'n	C4'x	O1'n	C1'x	C2'n	C3'x
kMDS:8oxoG(ResID4)	0	3	30	49	17	1
kMDS:8oxoG(ResID14)	1	3	25	43	23	7
BP-REMD:8oxoG(Resid4)	0	3	28	47	19	3
BP-REMD:8oxoG(Resid14)	0	4	31	48	16	2
kMDS:dG(ResID4)	1	4	20	30	34	11
kMDS:dG(ResID14)	0	4	23	31	32	9
BP-REMD:dG(Resid4)	1	4	23	33	31	8
BP-REMD:dG(Resid14)	0	3	19	32	35	11

Neben dem Sampling Vorteil existiert auch ein Rechenvorteil von der BP-REMD im Vergleich zur kMDS Methode. So konnten in 10ns BP-REMD-Simulationen konvergierte Ergebnisse hinsichtlich mehrerer Diederwinkel erzielt werden. Die Resultate sprechen dafür, dass trotz der 200ns kMDS im α/γ Bereich keine Konvergenz erzielt werden konnte.

3.1.3 Stereochemische Ursachen der unterschiedlichen Konformationszustände

Für 8oxoG und GUA wurden identische Kraftfeldparameter für den Zucker-Phosphatrückgrat verwendet. Die 8oxoG Base wurde entsprechend der parmbse0 Parameter angepasst. Dies ist unproblematisch da die Basenmodifikationen einige Bindungen von den Diederwinkeln des Phosphatrückgrats entfernt sind. Die spezifischen Subzustände die beobachtet wurden für die 8oxoG-Base muss daher durch sterische Hinderungen durch die Modifikationen oder in Solvatationseffekten begründet sein. Um diese Annahme zu bestätigen wurden die Interaktionen der 8oxoG Atome detaillierter untersucht. Aus einem stereochemischen Gesichtspunkt lassen sich qualitativ die beobachteten Substände erklären. Um diese Aussage zu untermauern wurden Distanzen zwischen dem O8-Atom der Base und dem O4-Atom des Zuckers bestimmt. Wie aus Abbildung 3.9 zu entnehmen ist korreliert der BII-Zustand ($\epsilon \sim 280^\circ$) mit der O8/O4 Distanz. Für die 8oxoG Base gilt, dass diese Distanz im Vergleich zur H8/O4 Distanz im Falle des Guanin, um etwa $\sim 0.6 \text{ \AA}$ zunimmt. Es handelt sich hierbei zwar um eine kleine Distanz jedoch tritt der BII Zustand im Falle des 8oxoG signifikant häufiger auf.

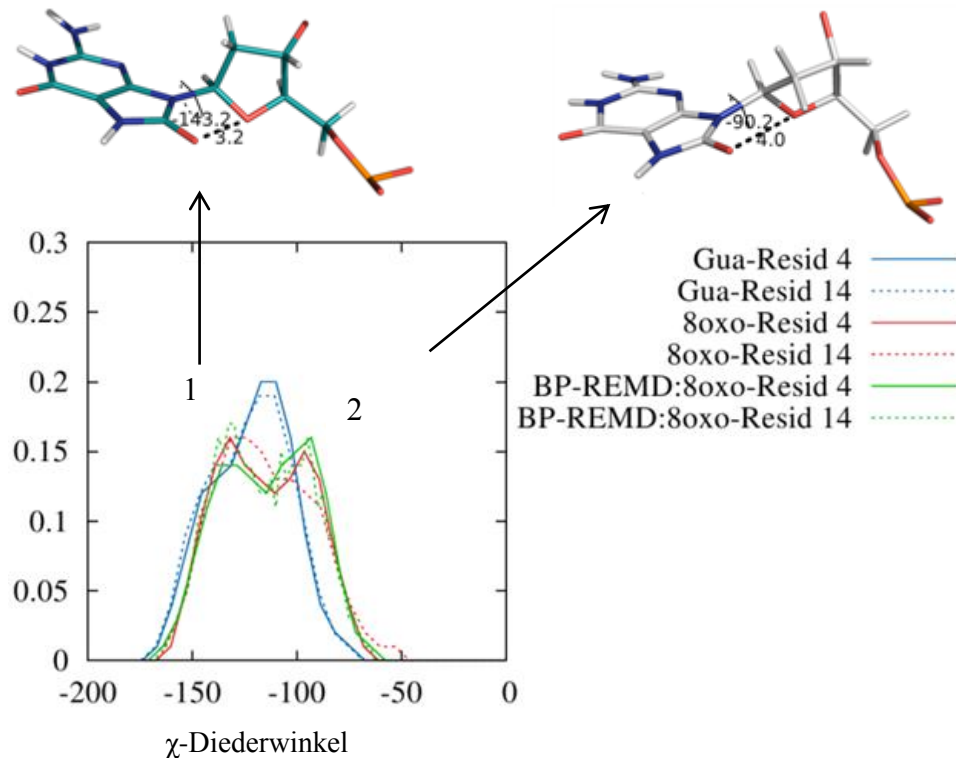


Abbildung 3.6 Verteilung der χ -Diederwinkel. Dargestellt sind die Verteilungen der Reste 4 und 14 für Guanin und 8oxoG aus der kMDS und der BP-REMD. Zudem sind zwei 8oxoG Strukturen aus den jeweiligen Verteilungen angegeben. Die Verteilungen zwischen kMDS und BP-REMD unterscheiden sich nicht signifikant. Die bimodale Verteilung des Diederwinkels wird durch elektrostatische Abstoßung der Sauerstoffatome hervorgerufen. Die Distanz zwischen den Sauerstoffatomen beträgt im Zustand 1 3,2 Å und in Zustand 2 4,0 Å.

Ursache hierbei sind elektrostatische Abstoßungen zwischen den beiden Sauerstoffen. Im Falle des Guanins liegt der gegensätzliche Fall vor. Hier wird der BI Zustand durch die H8/O4 Interaktion begünstigt.

Diese Eigenschaft des 8oxoG spiegelt sich auch im χ -Diederwinkel wieder, der mit dem BI/BII-Zustand gekoppelt ist (siehe Abbildung 3.8). Um die optimale Stacking Konformation und das Wasserstoffbrückennetzwerk aufrecht zu erhalten findet ein häufiger BI/BII Übergang statt. Diese häufigen Übergänge führen die einer bimodalen Verteilung des χ -Diederwinkel, sodass sich die Modelle auch in der Verteilung des χ -Diederwinkels unterscheiden (siehe Abbildung 3.6). Ein direkter Zusammenhang zwischen den α/γ Zuständen und den Basenmodifikationen konnte nicht ermittelt werden. Neben einigen stereochemischen Faktoren können diese Strukturen auch durch Stacking-Energien oder Solvatationsmuster stabilisiert werden.

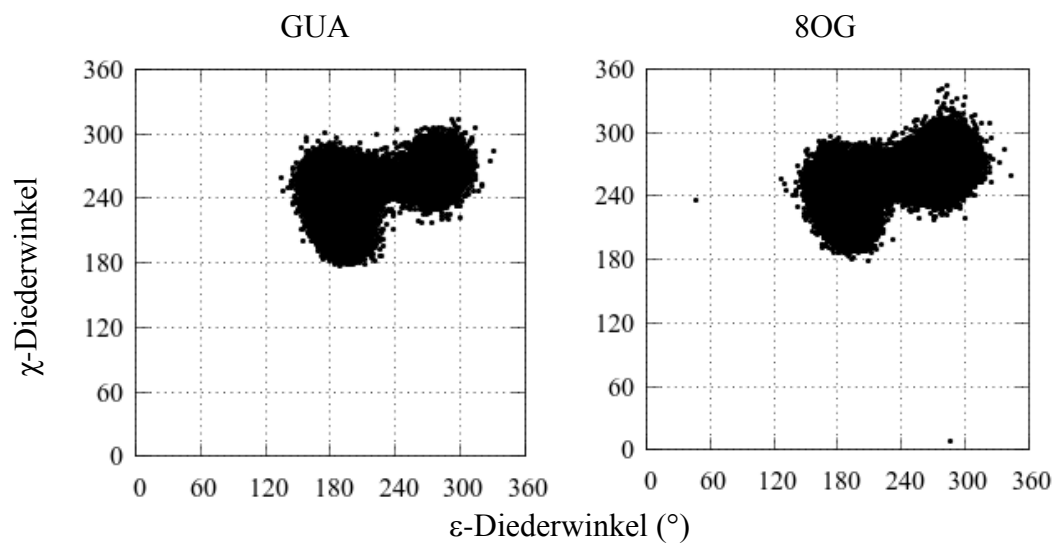


Abbildung 3.7 Korrelation zwischen ε und dem χ -Diederwinkel. Die Abbildung verdeutlicht den direkten Zusammenhang zwischen den beiden Diederwinkeln. Der BII Zustand ($\sim 270^\circ$) weist einen um 60° versetzten χ -Diederwinkel auf.

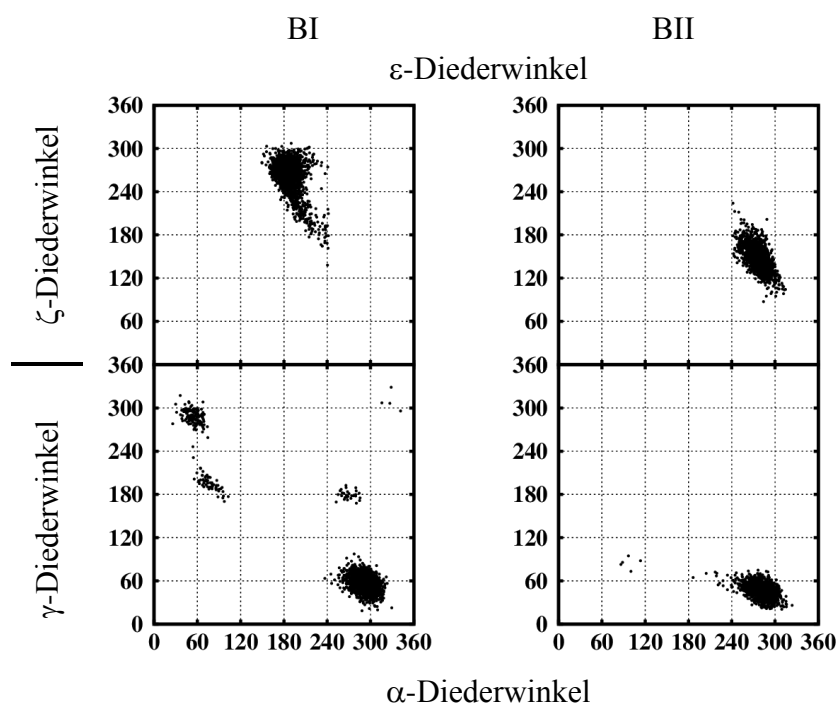


Abbildung 3.8 Zusammenhang zwischen BI/BII und den α/γ Zuständen. Unterteilt man die Trajektorien (kMDS: 8oxoG Modell) in BI und BII Konformationen stellt man fest, dass in BII Konformationen keine alternativen α/γ Zustände zu beobachten sind.

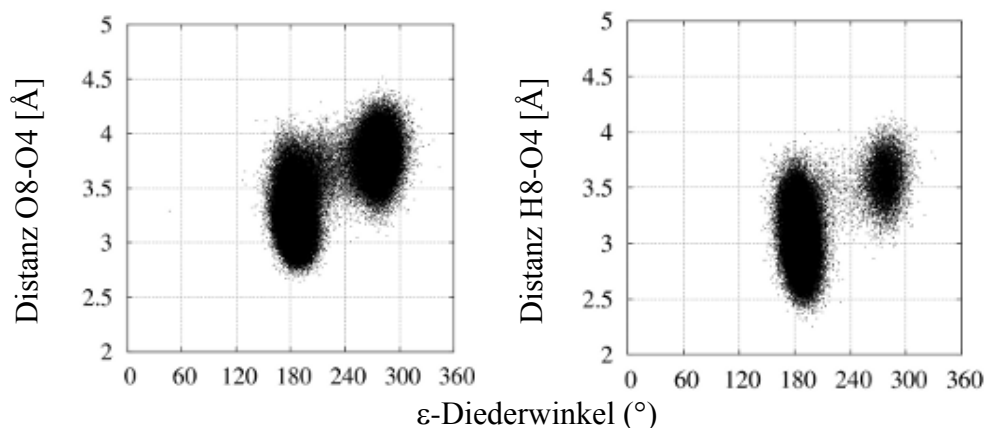


Abbildung 3.9 Korrelation der Winkel und der O8-O4 bzw. H8-O4 Distanzen. Diese Korrelation ist zudem in den Verteilungen der χ -Diederwinkel wiederzufinden.

3.1.4 Solvatisierung

Interaktionen zwischen DNA und dem umgebenden Medium (Wasser + Ionen) spielen eine wichtige Rolle in der Stabilisierung der Helix und können DNA Konformationen beeinflussen (128). Die 8oxoG Base ist durch das O8 Atom polarer als GUA mit dem H8 Atom an der entsprechenden Position. Um die Interaktion zwischen den Wasseratomen und den Basen zu bestimmen wurden die Radial-Verteilungen zwischen H8/Wasser und O8/Wasser bestimmt. Es konnte gezeigt werden, dass die Interaktion zwischen O8/8oxoG im Vergleich zu H8/GUA stärker ist, bzw. häufiger auftritt (siehe Abbildung 3.10). Dieses Verhalten wurde bereits in einer Studie gezeigt (81).

Das Wassermolekül, welches den Peak bei 1.8 Å in der Radial Verteilung darstellt, befindet sich in einem Wasserstoffbrücken-Netzwerk und hat eine durchschnittliche Aufenthaltsdauer von > 1 ns (gilt für kMDS). Eine Momentaufnahme dieses Wasserstoffbrücken-Netzwerkes findet sich in Abbildung 3.11. Ein gebundenes Wassermolekül an dieser Position kann auch einen Beitrag zur Erkennung der Base durch ein Reparaturenzym liefern.

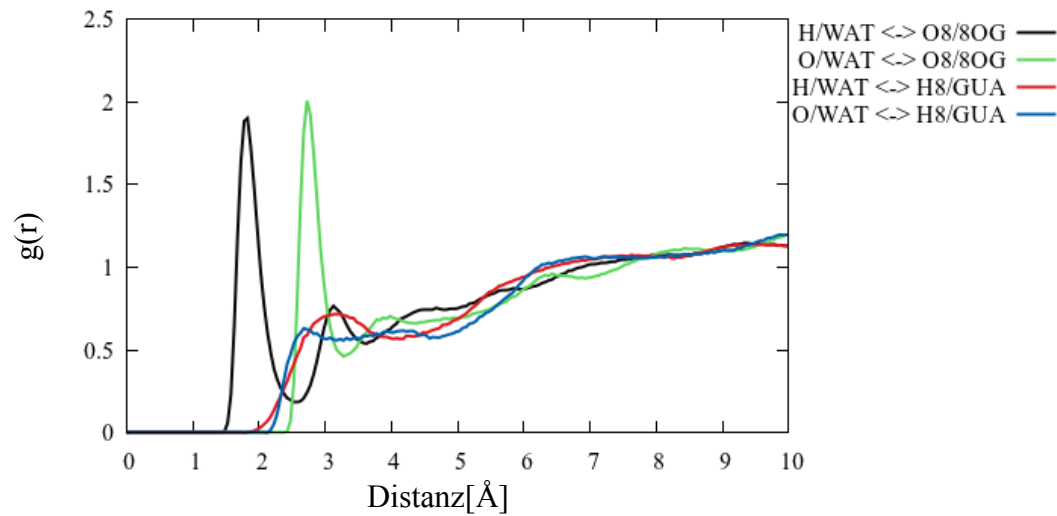


Abbildung 3.10 Vergleich der radialen Verteilungsfunktion von 8oxoG und Guanin. Folgende radiale Verteilungsfunktionen wurden berechnet: O8 Atom der 8oxoG Base \leftrightarrow Sauerstoff der Wasseratome (schwarze Linie); O8 Atom der 8oxoG Base \leftrightarrow Wasserstoff der Wasseratome (grüne Linie); H8 Atom der Guanin Base \leftrightarrow Sauerstoff der Wasseratome (blaue Linie); H8 Atom der 8oxoG Base \leftrightarrow Wasserstoffatom der Wasseratome (rote Linie).

So konnte im Encounter-Komplex mit 8oxoG:MutM keine direkte Interaktion zwischen dem Enzym und 8oxoG bestimmt werden, sodass Wassermoleküle oder spezifische Substände in der Erkennung eine wichtige Rolle spielen können (78, 66, 73, 216).

Neben den Radial-Verteilungen wurden weitere Analysen zur Beschreibung der Solvation durchgeführt. So wurden für die Basen 8oxoG₄, 8oxoG₁₄, G₄, G₁₄ die Anzahl der Wassermoleküle in einer Hülle von 3.5 Å um die Base mit dem Programm ptraj (*watershell*) bestimmt (siehe Abbildung 3.12). Die Ergebnisse bestätigen die Ergebnisse aus den radialen Verteilungsfunktionen. So ist die Wahrscheinlichkeit, dass zwei Wassermoleküle sich in der Hülle befinden im Falle des 8oxoG (blaue Balken) um etwa 25% höher.

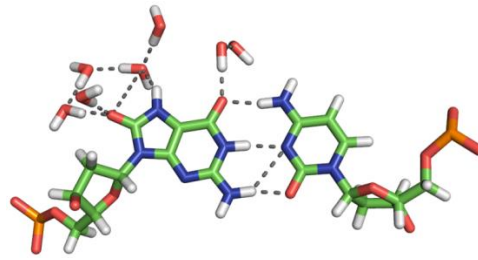


Abbildung 3.11 Wasserstoffbrücken-Netzwerk der 8oxoG Base. Eine Momentaufnahme aus der 8oxoG - BP-REMD Simulation zeigt die Interaktion zwischen 8oxoG:C und 8oxoG und den Wassermolekülen. Die möglichen Wasserstoffbrücken sind durch gestrichelte Linnien gekennzeichnet.

Weitere Analysen von BI und BII-Zuständen deuten auf kein spezifisches Muster hin. Die Wahrscheinlichkeit, dass sich zwei Wassermoleküle in der Hülle befinden, ist für BI und BII identisch. Es kann daher keine Aussage über einen spezifischen Solvatationseffekt hinsichtlich BI/BII getroffen werden.

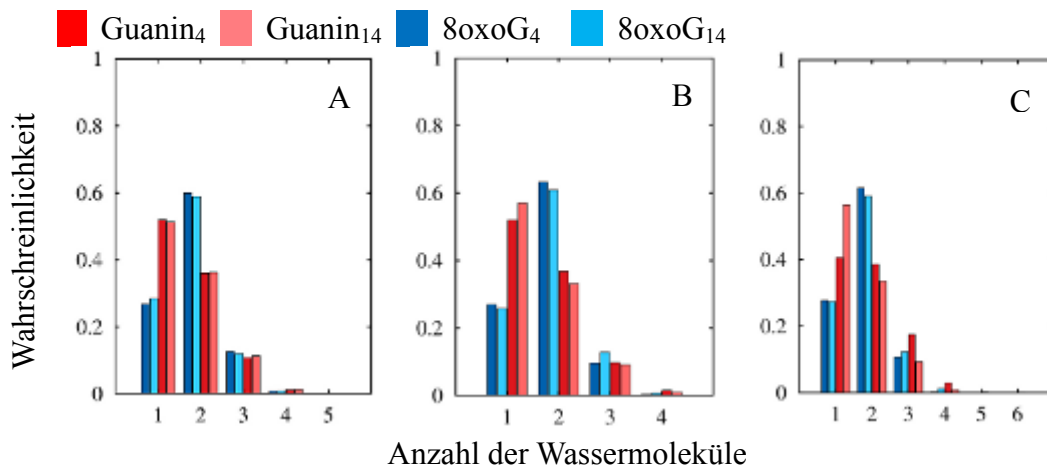


Abbildung 3.12 Verteilung von Wassermolekülen im BI bzw. BII Zustand. (A) Verteilung der Wassermoleküle über die gesamte kMDS Trajektorie. Bestimmt wurde die Anzahl der Wassermoleküle in einer Hülle mit einer Distanz von 3.5 Å um der Base herum. (B) Verteilung der Wassermoleküle in BI Zuständen. (C) Verteilung der Wassermoleküle in BII Zuständen. Es konnten keine Unterschiede in der Wasserverteilung zwischen BI und BII festgestellt werden.

3.1.5 Zusammenfassung und Schlussfolgerung

Ziel der vorliegenden Untersuchung war es strukturelle Eigenschaften von Guanin und einer modifizierten Variante, dem 8-oxoguanin, zu bestimmen. Hierfür wurden Molekulardynamik-Simulationen durchgeführt. Zusätzlich wurden Biasing-Potential Replika Exchange Simulationen durchgeführt. Es handelt sich hierbei um eine Methode, die speziell dafür entworfen wurde um effizienter DNA Phosphatrückgrat Konformation zu untersuchen (siehe Abschnitt 2.4). Die Methode benötigt weniger Replikate als die Temperatur-Replika-Exchange-Methode und basiert auf der Destabilisierung spezifischer Regionen von Winkel-Paaren entlang des Phosphatrückgrats. Durch die Destabilisierung entstehen in den höheren Replikaten neue Konformationen, die sich auch durch häufige Austausch zwischen den Replikaten im Referenz-Replika wiederfinden. In dieser Arbeit wurden Pseudo-Diederwinkel destabilisiert um neue Konformationen zu generieren. Der Vorteil durch die Pseudowinkel-Repräsentation liegt daran, dass simultan mehrere Diederwinkel destabilisiert werden, sodass die Anzahl der zu destabilisierenden Winkel reduziert wird und damit die Anzahl der benötigten Replika. Mit Hilfe der verwendeten Methoden und Simulationen konnte in dieser Studie gezeigt werden, dass durch die BP-REMD Methode eine effizientere Suche nach möglichen Subzustände bzw. DNA-Konformationen möglich ist. Vor allem im α/γ Bereich konnte gezeigt werden, dass durch die BP-REMD Methode die Überrepräsentation einer Konformation, im Vergleich zur kMDS, vermieden wird. Unter Berücksichtigung der α/γ Bereiche kann davon ausgegangen werden, dass die 200ns kMDS nicht konvergiert sind. Vorherige Molekulardynamik-Studien mit DNA Strukturen, die 8oxoG beinhalten, waren kürzer (< 200 ns) und verwendeten ein älteres Kraftfeld, von dem gezeigt wurde, dass die B-DNA Geometrie bei langen Simulation instabil wird. Für das Phosphatrückgrat konnten signifikante Unterschiede zwischen GUA und 8oxoG ermittelt werden. Eine geringere Energiebarriere von BI zum BII Zustand für 8oxoG konnte durch PMF Rechnungen bestimmt werden. Die geringere Energiebarriere hat stereochemische und elektrostatische Ursachen, da die gleichen Kraftfeldparameter für das Phosphatrückgrat verwendet wurden. Die BI/BII Übergänge beeinflussen auch den N-glykosidischen χ -Diederwinkel, welcher eine bimodale Verteilung im Falle des 8oxoG aufweist. Diese χ -Diederwinkel wurden auch in Kristallstrukturen die 8oxoG beinhalten an der entsprechenden Stelle gefunden. Zusätzliche Analysen an Kristallstrukturen zeigten zudem, dass auch dort alternative α/γ Bereiche zu finden sind. Die spezifischen Subzustände können bei der Erkennung der Base

durch ein Reparaturenzym eine Rolle spielen, da nur direkte Phosphatrückgrat-Protein Interaktionen im Encounter-Komplex mit beobachtet wurden. Neben den unterschiedlichen Verteilungen in den Diederwinkeln des Phosphatrückgrats wurden auch unterschiedliche Zucker-Pucker Verteilungen zwischen GUA und 8oxoG bestimmt. Experimentelle Studien mit DNA-Protein-Komplexen mit 8oxoG und Guanin an der entsprechenden Stelle weisen auch auf Unterschiede in den Zucker-Pucker hin (78). Im Komplex konnte experimentell ein C3'-endo Präferenz für 8oxoG und eine C2'-endo Präferenz für GUA gezeigt werden. Dieses Muster wurde konnte nicht für DNA Strukturen ohne Enzym bestätigt werden. Hier wurde eine Präferenz für C1'-exo beobachtet. Es kann angenommen werden, dass durch die Enzymbindung die Zucker-Pucker Konformation verändert wird. Zusätzliche Studien mit dem DNA-Enzym Komplex könnten einen Einblick in diese Fragestellung geben. Anhand der durchgeführten Simulationen konnte zudem ein unterschiedliches Wassermuster beobachtet werden. Da keine Interaktionen zwischen dem O8-Atom des 8oxoG und dem Reparaturenzym im Encounter-Komplex vorhanden ist, könnten diese Merkmale ausschlaggebend für die Erkennung der modifizierten Base sein. Aufgrund der diskutierten Daten wird angenommen, dass ein indirekter Erkennungsprozess durch spezifische Subzustände und Zucker-Pucker-Konformationen und/oder dem unterschiedlichen Wassermuster zwischen 8oxoG und dem Reparaturenzym vorhanden ist.

3.2 Vergleichende Modellierung von O⁶-MeG:C, O⁶-MeG:T und G:C

Dieser Abschnitt beschäftigt sich vor allem mit dem Einfluss von Alkylierungsschäden auf die DNA. Hierzu wurden drei verschiedene Modelle (siehe Abschnitt 2.5.2) mit kMDS und BP-REMD Simulationen untersucht. Der Fokus liegt auf den Konformationsraum der unterschiedlichen Basenpaare. Der Konformationsraum von Nukleinsäuren und ihre Abhängigkeit von der Sequenz spielt in der Erkennung und Funktion dieser Strukturen eine besonders wichtige Rolle. Im vorangehenden Kapitel wurde gezeigt, dass die BP-REMD-Methode geeignet ist um Unterschiede zwischen 8oxoG und Guanin auszuarbeiten. Wie bereits bei dem Vergleich zwischen 8oxoG und Guanin, ist auch hier das Ziel die Unterschiede zwischen verschiedenen Basenpaar-Konstellationen zu ermitteln.

Dabei sollen insbesondere die mit den jeweiligen Methoden ermittelten Unterschiede im Zucker-Phosphatrückgrat diskutiert werden. In dieser Arbeit angenommen wird, dass Unterschiede im Zucker-Phosphatrückgrat ausschlaggebend für die Erkennung der geschädigten DNA sind. Diese Annahme wurde bereits in mehreren Studien diskutiert. Marathe *et al.* konnten bei einem Vergleich zwischen freier DNA und DNA im Komplex mit Proteinen zeigen, dass eine Protein-Bindung den Übergang in ungünstige Zucker-Phosphatrückgrat Konformationen erfordern kann (217). Ferner weisen sie darauf hin, dass bereits kleine Unterschiede im Zucker-Phosphatrückgrat ausschlaggebend für die globale Geometrie der Helix sein können. Zudem zeigen sie, dass trotz größerer Konformationsänderungen, wie dem Übergang vom gestackten in den looped-out Zustand, die B-Form der Helix erhalten bleiben kann. Sowohl Marathe *et al.* als auch Svozil *et al.* weisen darauf hin, dass lokal unterschiedliche DNA-Formen eingenommen werden können (217, 218). Bestimmte Bereiche der DNA können einen Übergang von B-DNA in nahezu A-DNA aufweisen. Die Kombination aus A-DNA und B-DNA spiegelt sich auch in den Zucker-Phosphatrückgrat Konformationen wieder. Svozil *et al.* konnten eine Reihe von spezifischen α/γ -Zuständen für diese A+B-DNA Konformationen aufzeigen (218). Diese lokalen und sequenzspezifischen Konformationsänderungen können die DNA krümmen oder einen Knick verursachen, welches wiederum von Proteinen erkannt wird. Die Annahme ist hierbei, dass lokale Unterschiede in der DNA bereits wichtig sein können für die Erkennung durch DNA-bindende Proteinen. Mandel-Gutfreut *et al.* zeigten erste Hinweise für die Bestätigung dieser Annahme. Sie konnten nachweisen, dass das Zucker-Phosphat-Rückgrat in mehr als 50% aller DNA-Protein Interaktionen involviert ist (219). Zur Untersuchung der Fragestellung ob spezifische Konformationen für die jeweilige Basenpaar-Konstellation vorliegen wurden drei Modelle untersucht. Die ursprüngliche Sequenz der Modelle stammt aus einer DNA-Protein-Struktur (PDBid 1T38) und beinhaltet ein O⁶-MeG:C Basenpaar, wobei die O⁶-MeG Base sich im aktiven Zentrum des Proteins befindet. Es kann daher angenommen werden, dass diese Sequenz durch ein Reparaturenzym erkannt wird. Neben dem Modell, welches eine Sequenz mit den O⁶-MeG:C beinhaltet, wurden Modelle erstellt, welches an der entsprechenden Stelle ein O⁶-MeG:T bzw. G:C aufweisen. Zusammenfassend wurden drei Modelle mit folgenden Sequenzen: d(5'-GCCATG(O⁶-MeG:C)CTAGTA)₂, d(5'-GCCATG(O⁶-MeG:T)CTAGTA)₂, und d(5'-GCCATG(G:C)CTAGTA)₂ mit kMDS und BP-REMD untersucht.

3.2.1 Stabilität der Strukturen anhand der RMSD-Werte

Um die Konvergenz der Simulationen zu bestimmen, wurden - wie bereits in Abschnitt 3.1 - die RMSD Werte für die Simulationen der unterschiedlichen Strukturen, einmal mit der klassischen Variante (kMDS) und einmal mit der erweiterten Methode (BP-REMD), untersucht. Die RMSD-Werte aller Simulationen weisen über die gesamte Trajektorie Schwankungen auf und befinden sich zwischen 1.2 und 5 Å (siehe Abbildung 3.13).

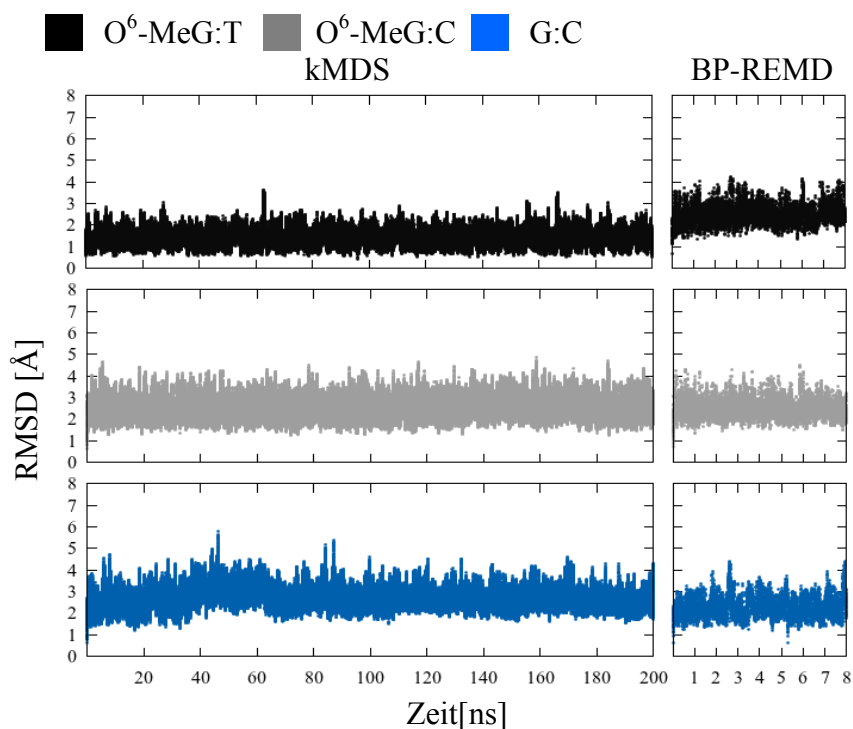


Abbildung 3.13 RMSD Werte der drei Modelle: O⁶-MeG:C, O⁶-MeG:T und G:C. Die RMSD-Werte sind bezüglich der Startstruktur des Zucker-Phosphat-Rückgrats gegen die Zeit aufgetragen. Die schwarze Linie stellt den Verlauf der Werte des O⁶-MeG:T-Modell, die graue Linie das O⁶-MeG:C Modell und die blaue Linie das G:C Modell. Die Werte der BP-REMD stammen aus dem Referenz-Replika (Replika 0).

Ursache für diese starken Schwankungen ist bei allen Simulationen das sogenannte „*fraying*“ an den terminalen Basen der DNA. Unter *fraying* versteht man den Verlust der WC-Paarung an den Enden der DNA Struktur. Die Ursachen für das *fraying* können in der Destabilisierung der WC-Partner durch Solventmoleküle sein.

Da die zu untersuchende Basenmodifikation im mittleren Abschnitt der Struktur liegt, kann davon ausgegangen werden, dass der Einfluss des *fraying* nur sehr gering ist auf diesen Bereich der Struktur. Um die Stabilität dieses Abschnitts zu bestimmen, wurden die RMSD-Werte der O⁶-MeG-Base bzw. Guanin Base und der nächsten Nachbarn bestimmt (Rest 6,7,8,19,20,21). Hieraus lässt sich ableiten, wie stark die Schwankungen dieses Teilbereichs der DNA sind. Sortiert nach den RMSD-Werten der Strukturen, lässt sich festhalten, dass G:C die geringsten und O⁶-MeG:C die größten Konformationsänderungen aufweisen (G:C mit 1.5 Å vs. O⁶-MeG:C mit 2.1 Å, siehe Tabelle 3.5). Dies bestätigt die experimentelle Annahme, dass O⁶-MeG:T stabiler ist als O⁶-MeG:C (56).

Tabelle 3.5 Vergleich der RMSD-Werte der Reste 6,7,8 und deren WC-Partner bezüglich der Startstruktur

Modell	Mittelwert		Std.		Minimum		Maximum	
	BP-REMD	kMDS	BP-REMD	kMDS	BP-REMD	kMDS	BP-REMD	kMDS
G:C	0.9	0.9	0.2	0.2	0.4	0.3	1.6	1.5
6MG:C	1.1	1.0	0.3	0.2	0.5	0.4	2.0	2.1
6MG:T	1.0	1.0	0.2	0.2	0.5	0.4	1.8	2.0

Neben dem Vergleich der modellierten Strukturen, wurde auch der RMSD-Wert der modellierten O⁶-MeG:C Struktur mit der Kristallstruktur bestimmt. Die Kristallstruktur diente hierbei als Referenz für die Trajektorie. Da die O⁶-MeG-Base in einer *looped-out* Konformation vorliegt, wurde sie aus der RMSD-Wert Berechnung ausgeschlossen. Als Referenz dienten die nächsten Nachbarn der O⁶-MeG-Base (Reste 6,8,19,21). Der niedrigste RMSD-Wert wurde für die BP-REMD Simulation ermittelt (~0.7 Å) (siehe Abbildung 3.14).

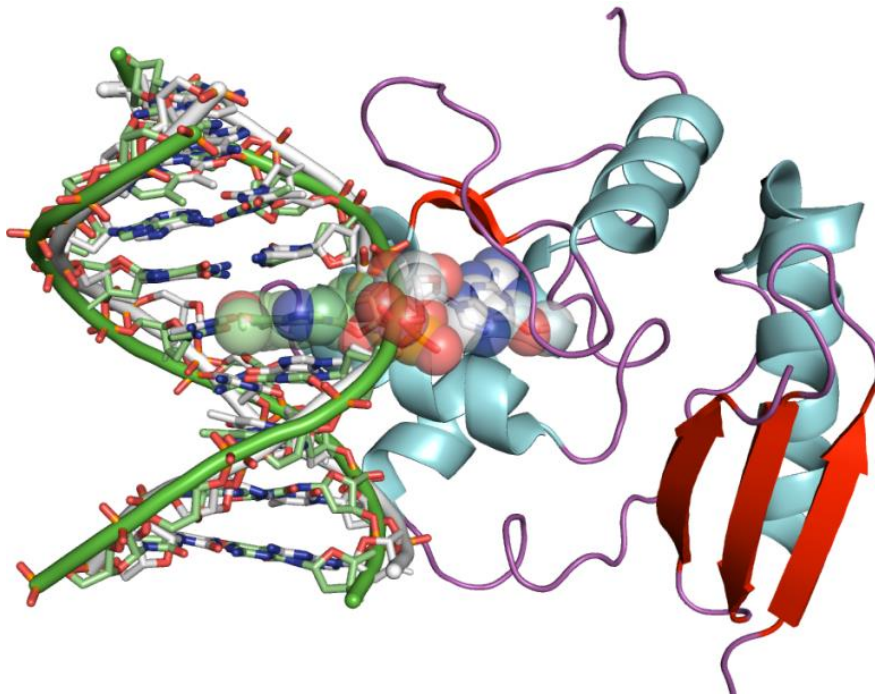


Abbildung 3.14 Superpositionierung der Kristall-Struktur 1T38 und der modellierten Struktur. Die Kristall-Struktur 1T38 ist ein DNA-Protein Komplex in dem sich die Base O⁶-MeG (vdW-Repräsentation, weiß) in einer extrahelikalen Konformation in der aktiven Bindetasche der Methylguanin-Transferase befindet (Cartoon-Repräsentation). Die Struktur mit dem geringsten RMSD-Wert bezüglich der nächsten Nachbarn der O⁶-MeG-Base (0.7 Å) ist in einer Cartoon-Repräsentation dargestellt (grün). Die O⁶-MeG befindet sich in einer intrahelikalen Konformation (vdW-Repräsentation, grün)

Der RMSD-Wert über die gesamte Struktur liegt bei ~ 2.9 Å. Ignoriert man jedoch die terminalen Basen so liegt der RMSD-Wert des O⁶-MeG:C Basenpaars bei ~ 2.0 Å für die BP-REMD Simulation und ~ 2.1 Å für die cMD Simulation (siehe Tabelle 3.5). In Abbildung 3.14 sind beide Strukturen abgebildet (ohne terminale Basen). Bei der grün gefärbten Struktur handelt es sich um die aus der Simulation. Wie aus der Abbildung zu entnehmen ist, wird die helikale Form der experimentellen Struktur in den Simulationen wiedergefunden.

3.2.2 Vergleich zwischen klassischen MD und BP-REMD

Dieser Abschnitt fokussiert sich auf einen Vergleich der BP-REMD Methode mit kMDS. Dabei werden die Diederwinkel des Zucker-Phosphatrückgrats der O⁶-

MeG bzw. Guanin Base und die der nächsten Nachbarn analysiert. Bei der Untersuchung des ϵ -Diederwinkel konnten keine Unterschiede zwischen den drei Modellen ermittelt werden. Der dominante Zustand ist in allen Simulationen mit >90% der BI Zustand (siehe Tabelle 3.6). Die sehr geringfügigen Abweichungen lassen sich durch statistische Fluktuationen begründen. Es kann aufgrund dieses Ergebnisses festgehalten werden, dass ein Einfluss der Methyl-Gruppe auf den ϵ -Diederwinkel der entsprechenden Base nicht sichtbar ist. Zusätzlich lässt sich anhand der Verteilungen festhalten, dass der ϵ -Diederwinkel des WC Partners auch nicht durch die Methylgruppe beeinflusst wird (siehe Tabelle 3.6).

Neben den ϵ -Diederwinkel wurden auch weitere Diederwinkel des Zucker-Phosphatrückgrats analysiert. Abbildung 3.15 zeigt den δ -Diederwinkel der O⁶-MeG Base (Position 7) und die des gegenüberliegenden Partners (gestrichelte Linie). Zusätzlich wurden die δ -Diederwinkel des vorigen Basenpaares untersucht um den Einfluss der Basenmodifikation auf die nächsten Nachbarn zu untermauern. Das G:C-Modell (G blaue durchgezogene Linie, C blaue gestrichelte Linie) zeigt für die Guanin-Base eine bimodale Verteilung mit den Maxima bei 85° und 140°. Die gegenüberliegende Base weist auch eine bimodale Verteilung auf, jedoch mit einem dominierenden Zustand bei 140°. Für die Position sechs, ebenfalls ein G:C-Basenpaar, ist im G:C Modell für die Guanin-Base nur ein Zustand bei 140° zu beobachten. Für die Cytosin Base ist wiederum eine bimodale Verteilung zu beobachten, jedoch mit einem dominierenden Zustand bei 85°.

Für die Verteilung aus der kMDS im O⁶-MeG:C Modell (O⁶-MeG grüne Linie, C gestrichelte Linie) sind signifikant unterschiedliche Verteilungen zum G:C Modell zu beobachten. Die Verteilungen beider Basen werden durch den Zustand bei 140° dominiert. Für die Position sind unterschiedliche Verteilungen für die Base 19 (Cytosin, grün gestrichelte Linie) zu beobachten. Im O⁶-MeG:T Modell sind wiederum bimodale Verteilungen für die Position sieben zu beobachten. Die Unterschiede zwischen dem O⁶-MeG:T und dem G:C Modell sind nicht so signifikant wie die Unterschiede zwischen dem O⁶-MeG:C/G:C Modell an der Position 7. Für die Position sechs sind jedoch an der Base 19 deutlichere Unterschiede in den Verteilungen zu beobachten. Für die Base 6 (Guanin) sind in allen Modellen ähnliche Verteilungen zu beobachten. Hinsichtlich der Position 7 lässt sich festhalten, dass die größten Unterschiede zwischen dem O⁶-MeG:C und dem G:C Modell bestehen, jedoch auch Unterschiede zwischen den O⁶-MeG:T und dem G:C Modell zu beobachten sind.

Tabelle 3.6 Relative Häufigkeiten der BI/BII Zustände im Vergleich (in %). Die Verteilung der BI/BII der W (O^6 -MeG bzw. G) und C (C bzw. T) –Partner verdeutlicht das keine signifikanten Unterschiede zwischen den Modellen und den Simulationsmethoden vorliegen.

O^6 -MeG(W):C(C)				
	BP-REMD		kMDS	
	BI	BII	BI	BII
W	96	4	95	5
C	100	0	100	0
O^6 -MeG(W):T(C)				
	BP-REMD		kMDS	
	BI	BII	BI	BII
W	97	3	96	4
C	99	1	99	1
G(W):C(C)				
	BP-REMD		kMDS	
	BI	BII	BI	BII
W	97	3	98	2
C	99	1	99	1

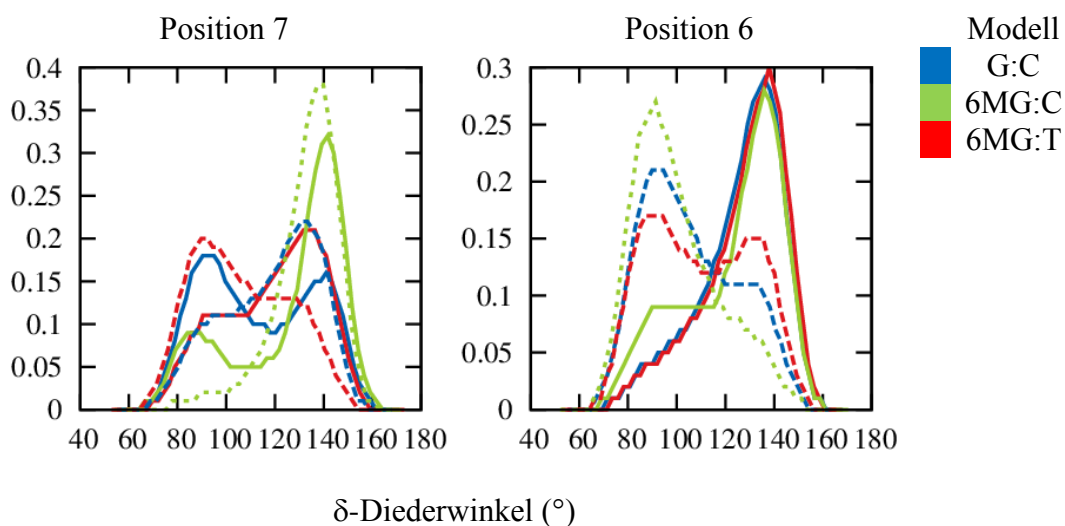


Abbildung 3.15 Verteilung des δ -Diederwinkel im O^6 -MeG:C, O^6 -MeG:C und G:C Modell. Dargestellt sind die Verteilung des δ -Diederwinkel der O^6 -MeG bzw. G Base (Position 7, durchgezogene Linie) und die des WC-Partners (gestrichelte Linie). Zusätzlich wurden die Verteilungen der WC-Partner an der Position 6 (durchgezogene Linie) miteinander verglichen.

Der Vorteil der BP-REMD Methode wird am Beispiel der α -Diederwinkel deutlich. Im G:C-Modell ist für beide Simulationsmethoden der g-/g+ Zustand zu beobachten. Dieser Zustand entspricht dem der kanonischen B-DNA. Im Vergleich zum G:C-Modell weist das O⁶-MeG:C Modell zwei weitere Zustände (g-/t und g+/t) auf. Das Auftreten dieser beiden Zustände wird nur in der BP-REMD Simulation beobachtet, wohingegen in der kMDS nur der Zustand g-/t repräsentiert wird. An dieser Stelle ist anzumerken, dass im Falle der kMDS der g-/t Zustand auch ermittelt wird. Für das O⁶-MeG:T Modell wurden nur in der BP-REMD Simulation zusätzliche Konformationsbereiche ermittelt (siehe Abbildung 3.16). Im Vergleich der unterschiedlichen Modelle zeigte sich, dass die Konformationsvielfalt für das O⁶-MeG:C Modell am größten ist. Dies lässt darauf schließen, dass die Energiebarrieren im O⁶-MeG:C Modell am geringsten sind. Fernerhin ist darauf hinzuweisen, dass die beobachteten Strukturen bzw. Konformationsräume auch experimentell nachgewiesen werden konnten (siehe Abbildung 3.17). Jedoch nehmen die alternativen Konformationen nur einen statistisch geringen Anteil ein (Ergebnisse aus der BP-REMD: O⁶-MeG:C: g-/g+ (90%), g-/t(7%), g+/t(3%); O⁶-MeG:T: g-/g+ (96%), g-/t(4%).

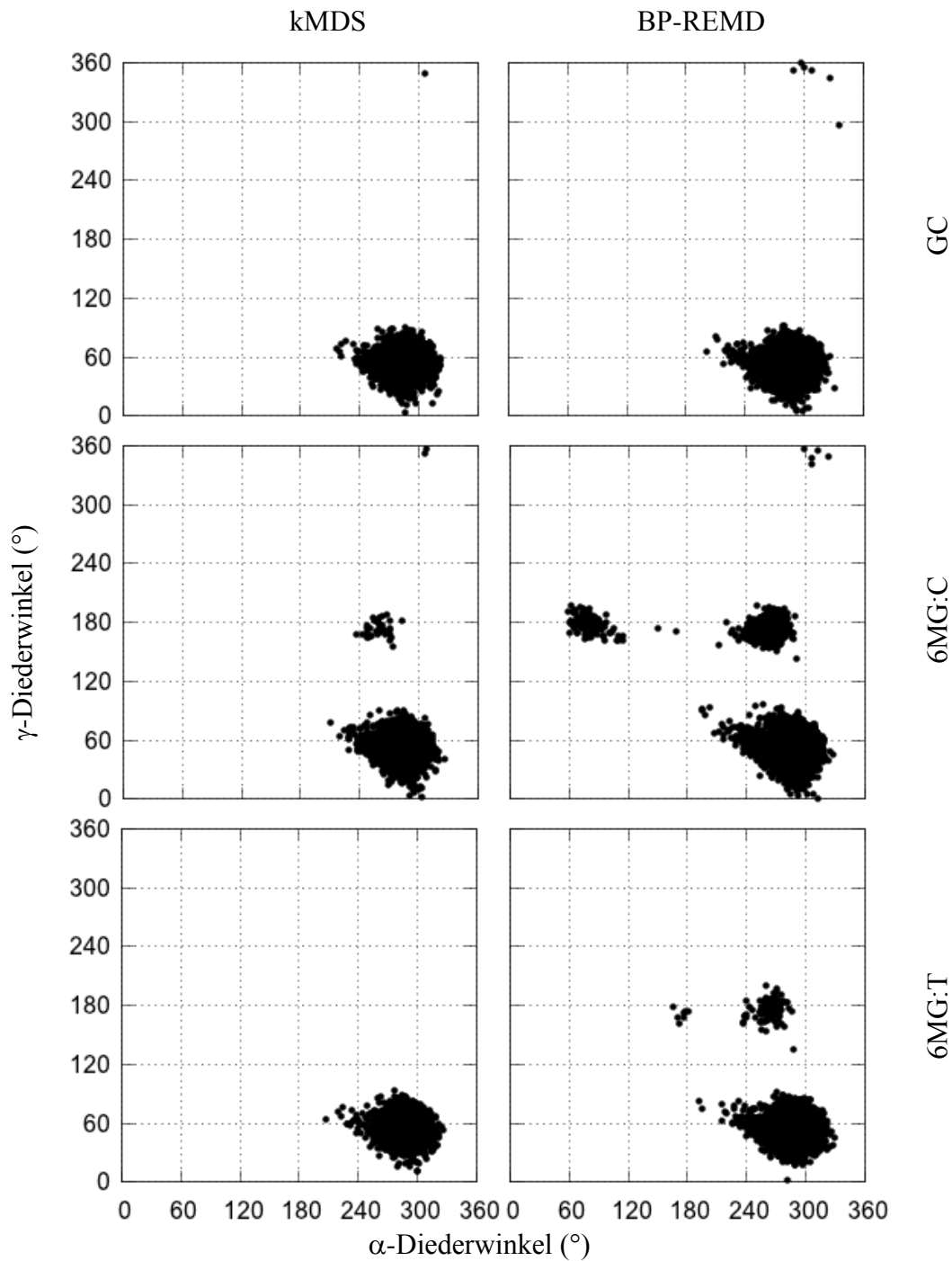


Abbildung 3.16 Konformationsverteilungen der α/γ -Diederwinkel im Vergleich. Ein Vergleich der Simulationen anhand des α/γ -Diederwinkel der O^6 -MeG:C, O^6 -MeG:T und G:C Modelle zeigt, dass mit der BP-REMD Methode neue Konformationszustände ermittelt werden.

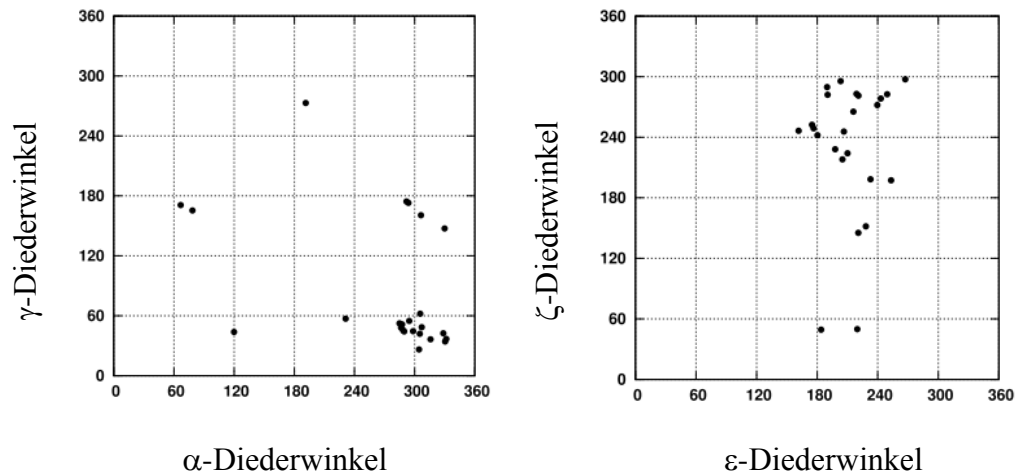


Abbildung 3.17 α/γ and ϵ/ζ Konformationsbereiche von experimentellen O⁶-MeG-Basen. Folgende PDB Strukturen wurden für die Analyse 1D24, 153D, 1D27, 1T38, 218D, 2HHQ, 2HHS, 2HHT, 2HHU, 2HHV, 2HHW, 2HHX, 2HVH, 2HW3, 2J6S, 2J6T, 2J6U, 2O8C, 3GX4, 3NGD, 3OSN verwendet.

Eine direkte Korrelation des δ - und des χ -Diederwinkel spiegelt sich auch in den Verteilungen der Konformationszustände dieser Winkel wieder (siehe Abschnitt 2.2). Interessanterweise ist für das O⁶-MeG:C-Modell eine bimodale Verteilung mit einer dominanten high-anti-Konformation an beiden Positionen zu beobachten. An der Position 7 (O⁶-MeG-Base) ist die *high anti*-Konformation ($\sim 100^\circ$) besonders stark ausgeprägt.

Im O⁶-MeG:T-Modell sind an den entsprechenden Positionen/Basen *anti*-Konformationen zu beobachten ($\sim 150^\circ$). Die Verteilung des χ -Diederwinkel im O⁶-MeG:T-Modell spricht für eine stabile Konformation hinsichtlich des Winkels. Die χ -Diederwinkel-Verteilung der O⁶-MeG Base im O⁶-MeG:T-Modell ($\sim 130^\circ$) befindet sich zwischen der Verteilung der Guanin-Base im G:C Modell ($\sim 150^\circ$) und der O⁶-MeG Base im O⁶-MeG:C-Modell ($\sim 100^\circ$). Dieses Ergebnis bestätigt die Annahme, dass die strukturellen Differenzen zwischen dem G:C Modell und dem O⁶-MeG:C-Modell größer sind im Vergleich zu dem G:C/ O⁶-MeG:T Modell. Die Verteilung im O⁶-MeG:C-Modell spricht dafür, dass während der Simulation eine Reihe von verschiedenen Konformationen eingenommen werden. In Abschnitt 2.2 wurde eine Reihe von Parametern diskutiert und auf ihre Korrelation untereinander hingewiesen. Neben der Korrelation der verschiedenen Diederwinkel besteht auch ein Zusammenhang zwischen dem δ -Diederwinkel und den Zucker-Puckern.

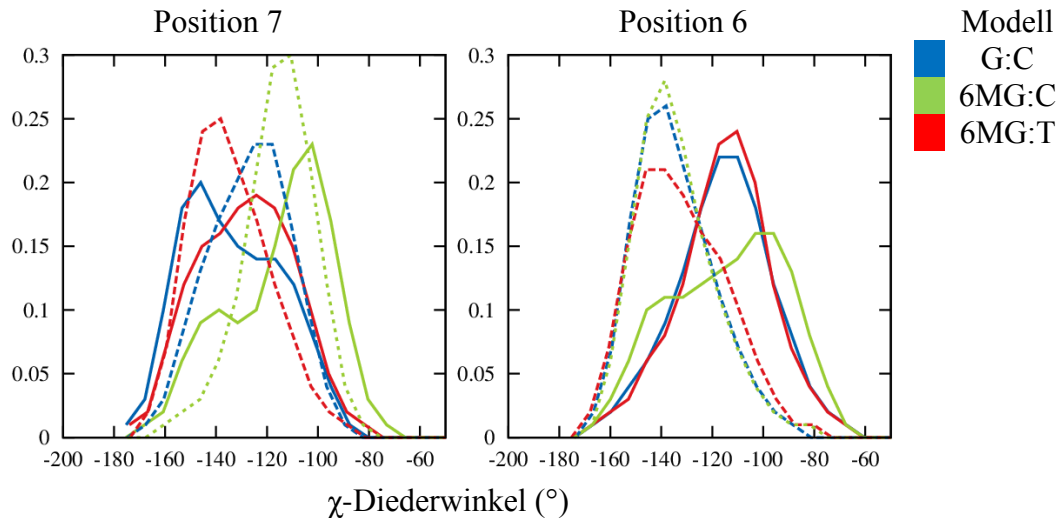


Abbildung 3.18 Verteilung des χ -Diederwinkel im O^6 -MeG:C, O^6 -MeG:C und G:C Modell. Dargestellt sind die Verteilung des χ -Diederwinkel der O^6 -MeG bzw. G Base (Position 7, durchgezogene Linie) und die des WC-Partners (gestrichelte Linie). Zusätzlich wurden die Verteilungen der WC-Partner an der Position 6 (durchgezogene Linie) miteinander verglichen.

Um zusätzlich die strukturellen Differenzen der verschiedenen Systeme zu untermauern und den Einfluss der O^6 -MeG Base auf die nächsten Nachbarn aufzuzeigen, wurden die relativen Häufigkeiten der Zucker-Pucker der Position 7 (O^6 -MeG bzw. G) und des gegenüberliegenden Nachbarn bestimmt. Ein Vergleich der Verteilungen in den verschiedenen Modellen und Simulationen ist in Tabelle 3.7 wiederzufinden. Wie aus der Tabelle 3.7 zu entnehmen ist, unterscheiden sich die Verteilungen aus der kMDS und der BP-REMD nur geringfügig. In den Verteilungen der Zucker-Pucker sind signifikante Unterschiede zwischen den verschiedenen Modellen zu beobachten. Das O^6 -MeG:C Modell zeichnet sich unter anderem durch die starke Präferenz des C2'-endo Zustand für die Cytosin Base aus. Interessanterweise wird der O1'-endo Zustand für die Cytosin Base nur selten beobachtet (2-3%). Die O^6 -MeG Base im O^6 -MeG:C Modell unterscheidet sich deutlich hinsichtlich der C3'-exo Verteilung zum O^6 -MeG:T bzw. G:C Modell. Die C3'-exo Konformation ist im O^6 -MeG:T Modell, mit 5% in beiden Simulationen, im Vergleich zum O^6 -MeG:C Modell, nur selten vertreten. Da in beiden Simulationen identische Kraftfelder eingesetzt wurden, lässt sich daraus schließen, dass die unterschiedlichen Verteilungen durch unterschiedliche Interaktionen mit dem gegenüberliegenden Nachbarn zu begründen sind. Das O^6 -MeG:T Modell zeichnet sich durch die C1'-exo Konformation für die O^6 -MeG Base und die C4'-exo Konformation der Thymin Base aus. Beide Konformation bzw. deren relative Häufigkeit unterscheiden sich signifikant von dem O^6 -MeG:C

bzw. G:C Modell. Im G:C Modell sind in beiden Simulationen für beide Nukleotide keine spezifischen Eigenschaften zu beobachten.

Neben den Simulationen wurden für experimentelle Strukturen die Zucker-Pucker bestimmt. Hierbei wurde darauf geachtet, dass nur DNA-Strukturen (ohne Protein) zur Analyse herangezogen wurden. Da nur sehr wenige Strukturen ohne im Komplex mit anderen Biomolekülen vorhanden sind, können nur Einzelfälle diskutiert werden. Für beide O⁶-MeG:T Modelle (PDBid 218D und 1D27) wurden für die O⁶-MeG die C1'-exo bzw. C2'-endo Konformation beobachtet.

Eine von beiden Konformationen tritt mit einer Wahrscheinlichkeit von etwa 70% in den O⁶-MeG:T Simulationen auf. Interessanterweise wurde zudem in der 218D Struktur die C4'-exo Konformationen beobachtet.

Tabelle 3.7 Relative Häufigkeit der Zucker-Pucker-Verteilung (in %) im Vergleich zu Kristall-Strukturen.

	Rest	Methode	C3'n	C4'x	O1'n	C1'x	C2'n	C3'x	C4'n	C2'x
6MG:C	6MG	kMDS	3	8	12	16	37	23	1	0
	6MG	BP-REMD	4	10	12	16	33	23	1	1
	C	kMDS	2	1	2	34	56	5	0	0
	C	BP-REMD	1	1	3	31	59	5	0	0
6MG:T	6MG	kMDS	1	4	21	40	29	5	0	0
	6MG	BP-REMD	1	5	24	39	26	5	0	0
	T	kMDS	2	15	36	32	15	1	0	0
	T	BP-REMD	2	18	37	31	12	0	0	0
G:C	G	kMDS	1	7	36	27	22	7	0	0
	G	BP-REMD	4	6	33	28	21	8	0	0
	C	kMDS	2	6	21	38	32	3	0	0
	C	BP-REMD	1	7	19	40	31	4	0	0
153D	6MG	X-Ray	0	0	0	0	0	2	0	0
	A	X-Ray	0	0	0	1	1	0	0	0
1D24	6MG	X-Ray	0	2	0	0	0	0	0	0
	C	X-Ray	0	0	0	0	2	0	0	0
1D27	6MG	X-Ray	0	0	0	0	2	0	0	0
	T	X-Ray	0	0	0	1	1	0	0	0
218D	6MG	X-Ray	0	0	0	2	0	0	0	0
	T	X-Ray	0	1	1	0	0	0	0	0

Diese Konformation tritt in den Simulationen mit einer Wahrscheinlichkeit von etwa 20% auf und unterscheidet sich deutlich zu den anderen Modellen.

Für die 1D24 Struktur (O^6 -MeG:C Basenpaar) wurde an der O^6 -MeG Position die C4'-exo Konformation bestimmt. Diese Konformation tritt in den Simulationen mit einer Wahrscheinlichkeit von etwa 10% nur vergleichsweise selten auf. Zur Kristallstruktur ist jedoch anzumerken, dass es sich hierbei um eine linkshändige Z-DNA Struktur handelt und diese somit nicht zum direkten Vergleich geeignet ist. Nichtsdestoweniger wurde für die Cytosin Base die C2'-endo Konformation beobachtet, welches zugleich mit 60% Wahrscheinlichkeit die dominante Konformation in den entsprechenden Simulationen ist.

Insgesamt lässt sich festhalten, dass die drei untersuchten Modelle sich in vielen strukturellen Eigenschaften unterscheiden. Die signifikantesten Unterschiede wurden zwischen dem G:C und dem O^6 -MeG:C Modell beobachtet. Ferner ist, ist darauf hinzuweisen, dass die O^6 -MeG Einfluss auf die Konformationen der umliegenden Nachbarn hat.

3.2.3 Bestimmung helikaler Parameter

Um den Einfluss der O^6 -MeG Base auf den DNA-Doppelstrang zu untersuchen wurden neben den Diederwinkeln auch die helikalen Parameter bestimmt. Neben der eigentlichen Base an Position 7 wurden auch die Nachbarbasen hinsichtlich der helikalen Parameter untersucht, da dies diese durch die geschädigte Base beeinflusst sein könnten. Zusätzlich ist anzumerken, dass die ermittelten Verteilungen und die hier vorgestellten Ergebnisse aus der kMDS stammen.

Beim Vergleich der unterschiedlichen Parameter fällt zunächst auf, dass, wie bereits bei der Verteilung der Diederwinkel, die größten Differenzen zwischen dem G:C und dem O^6 -MeG:C Modell bestehen. Besonders deutlich ist dies anhand der opening, slide und twist Parameter zu beobachten. Sowohl der opening, als auch der xdisp Parameter der O^6 -MeG Basen sprechen dafür, dass bereits in der ungebundenen Form die Basen verschoben, bzw. versetzt sind, sodass in der helikalen Struktur Signale zur Erkennung durch Proteine vorhanden sein können.

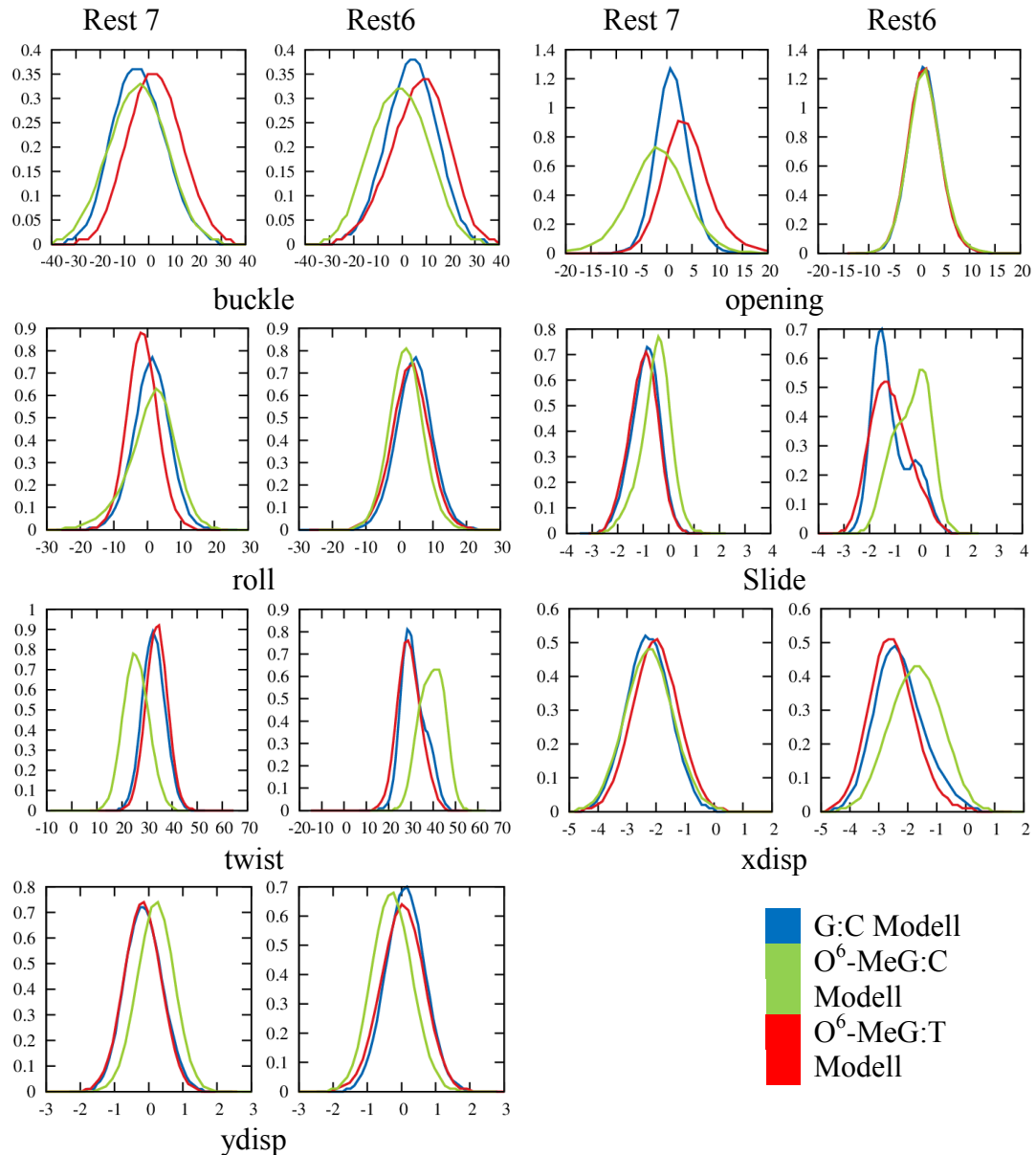


Abbildung 3.19 Helikale Parameter der der O⁶-MeG:C, O⁶-MeG:T und G:C Modelle im Vergleich

Dies spricht zudem dafür, dass die Energiebarrieren zum Herausdrehen der Base aus seinem gestackten Zustand geringer ist als im G:C Modell. Dies hat wiederum zur Folge, dass die hydrophoben Wechselwirkungen zu den Nachbar-Basen geringer sind, da kein optimales Stacking mehr gewährleistet ist.

Der versetzte twist, im O⁶-MeG:C Modell, an der Position 7 wird an der Position 6 wieder ausgeglichen, sodass die helikale Form aufrechterhalten bleibt. Von besonderem Interesse ist der slide Parameter, welcher für die Beschreibung von Sequenzeigenschaften herangezogen wird (218). Hierbei zeigen vor allem die Nachbar-Basen im O⁶-MeG:C Modell einen signifikanten Unterschied zu dem

G:C bzw. O⁶-MeG:T Modell. Es ist darauf hinzuweisen, dass die O⁶-MeG:T als Produkt des ersten Replikationszyklus entsteht und von dem MUTS α oder dem MGMT Enzym erkannt wird. Das O⁶-MeG:T Produkt entsteht also nur, wenn es nicht bereits das O⁶-MeG:C Paar durch das MGMT Enzym erkannt wird. Die deutlichsten Unterschiede zwischen dem O⁶-MeG:T und dem G:C Modell sind im opening Parameter wiederzufinden. Insgesamt, lässt sich jedoch festhalten, dass hinsichtlich der helikalen Parameter viele Gemeinsamkeiten zwischen dem O⁶-MeG:T und dem G:C Modell vorhanden sind. Dies lässt darauf schließen, dass andere Merkmale wie z.B. spezifische Zucker-Phosphatrückgrat Konformationen für die Erkennung des O⁶-MeG:T Basenpaares entscheidend sein können.

3.2.4 Clusteranalyse

Aus den vorigen Abschnitten ging hervor, dass eine Reihe von unterschiedlichen Konformationen von den unterschiedlichen Modellen eingenommen werden können. Um die Konformationen eindeutiger zu klassifizieren, wurden Clusteranalysen durchgeführt. Die Trajektorien wurden hierbei anhand der RMSD-Werte der Basen 6-8 und 19-21 klassifiziert.

In Abbildung 3.20 sind Ausschnitte (Basen 6-8 und 19-21) aus repräsentativen Strukturen jedes Cluster dargestellt. Für jeden Ausschnitt wurde mit Hilfe von PyMOL alle möglichen polaren Kontakte (Wasserstoffbrücken) dargestellt. Wie aus Abbildung 3.20 hervorgeht sind in allen Clustern im G:C Modell die Wasserstoffbrücken zwischen den WC-Partnern gewährleistet.

Im Falle des O⁶-MeG:C Modells wird deutlich, dass die repräsentative Struktur des ersten Clusters, zwischen dem O⁶-MeG:C Paar, nur eine mögliche Wasserstoffbrücke aufweist. Im O⁶-MeG:T Modell, werden an entsprechender Stelle zwei Wasserstoffbrücken beobachtet. Wie aus der Abbildung zu entnehmen ist, besteht jedoch in beidem Modellen eine Vielzahl von Möglichkeiten mit den nächsten Nachbarn eine Wasserstoffbrücke zu bilden. Dadurch entsteht eine Vielzahl von verschiedenen energetisch günstigen Konformationen.

In Abbildung 3.21 sind alle Cluster des O⁶-MeG:C Systems dargestellt. Hierbei soll vor allem darauf hingewiesen werden, dass die B-DNA Struktur trotz erhöhter Flexibilität an der Position 7 erhalten bleibt.

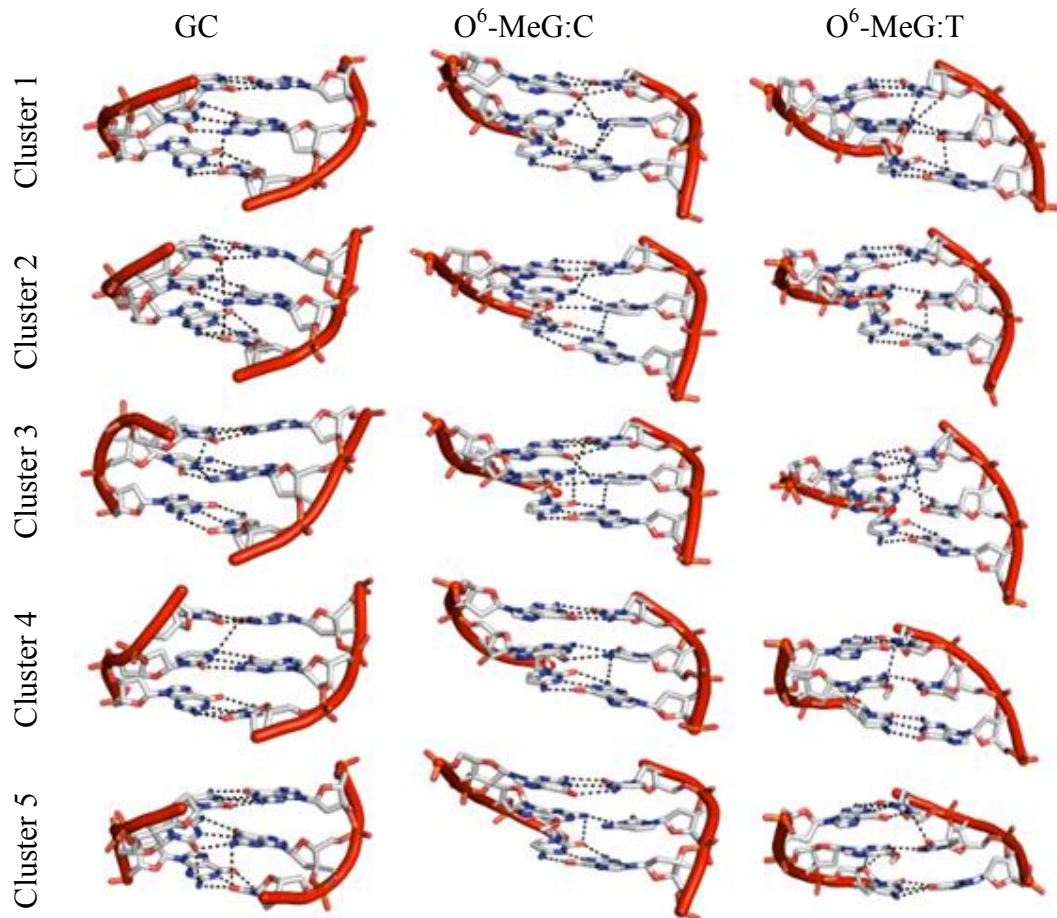


Abbildung 3.20 Repräsentative Strukturen aus der Clusteranalyse der kMDS-Trajektorie. Dargestellt sind jeweils die Reste 6-8 und deren WC-Partner. In gestrichelten Linien sind alle potentiellen Wasserstoffbrücken zwischen den Basen dargestellt.

Neben der Wasserstoffbrücken-Charakterisierung wurden die helikalen Parameter mit den Parametern der Kristallstruktur verglichen (siehe Tabelle 3.8).

Hierbei ist besonders Interessant, dass die Ax-bend Parameter der ersten Cluster des O⁶-MeG:C und des O⁶-MeG:T Clusters Gemeinsamkeiten aufzeigen. So beträgt die Differenz des Ax-bend Parameters in beiden Modellen an der O⁶-MeG Position 0.2 zur Kristallstruktur. Die Annahme ist hierbei, dass bereits in der Feinstruktur der DNA Informationen bzw. Erkennungsmerkmale für die Bindung von Enzymen vorhanden sind.

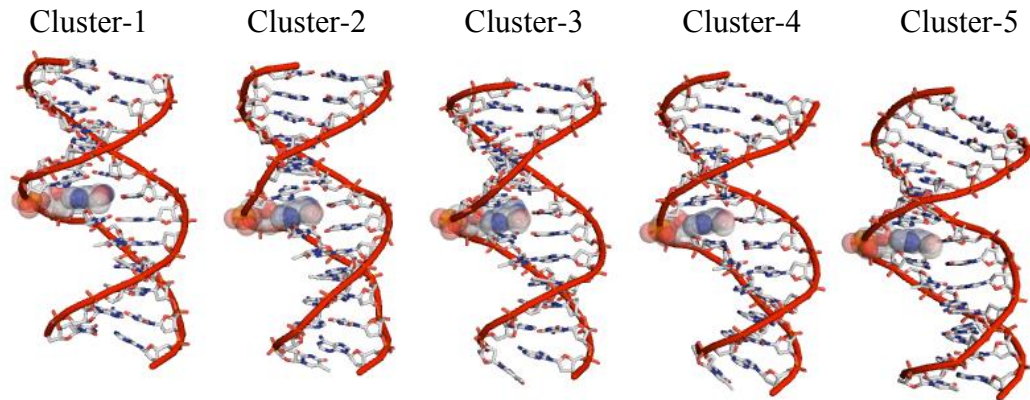


Abbildung 3.21 Repräsentative Strukturen aus der Clusteranalyse des O^6 -MeG:T Modells (kMDS). Die O^6 -MeG, dargestellt in der vdW-Repräsentation, befindet sich im Zentrum der Struktur. In allen Clustern bleibt die B-DNA Form erhalten.

Tabelle 3.8 Vergleich des Opening und Ax-Bending Parameters aus der kMDS mit der DNA der Kristall-Struktur 1T38

		O^6 -MeG:T		O^6 -MeG:C		G:C		1T38
		Cluster		Cluster		Cluster		
	ResID	1	2	1	2	1	2	
Opening	6	-1.1	-10.2	2,0	-1.4	0.1	0.2	-3.4
	7	3.3	-0.5	-4.7	-9.1	1.8	-1	-57
	8	3.4	4.5	2.0	0.3	1.7	1.7	0.6
Ax-bend	6	3.5	1.4	3.5	1.4	0.8	3.2	2.6
	7	2.9	1.1	2.9	1.1	0.3	1.4	3.1
	8	2.3	3.3	2.3	3.3	2,4	1.9	3.5

3.2.5 Zusammenfassung und Schlussfolgerung

Der Fokus dieses Abschnitts lag vor allem in der Bestimmung der Unterschiede zwischen drei unterschiedlichen Basenpaar-Konstellationen (G:C, O^6 -MeG:C, O^6 -MeG:T) und deren Einfluss auf die Feinstruktur der DNA Helix. Hierzu wurden neben klassischen MD Simulationen auch BP-REMD Simulationen durchgeführt. Ziel war es, strukturelle Unterschiede zwischen den verschiedenen Modellen herauszuarbeiten. Annahme ist hierbei, dass die strukturellen Unterschiede als Erkennungsmerkmal für Reparaturenzyme dienen. Zu diesem Zweck wurden neben den Diederwinkel des Zucker-Phosphat-Rückgrats auch helikale Parameter

der verschiedenen Modelle bestimmt. Hierbei stellte sich heraus, dass die größten strukturellen Unterschiede zwischen dem G:C und O⁶-MeG:C bestehen. Dieses Ergebnis wird durch experimentelle Daten bestätigt.

Mittels der BP-REMD Simulation konnten alternative α/γ Konformationszustände ermittelt werden, welche auch in experimentell bestimmten Strukturen vorkommen. Hinsichtlich der BI/BII Verteilungen konnten keine Unterschiede zwischen den verschiedenen Modellen bestimmt werden.

Bezüglich des opening und des Ax-bend Parameters konnten interessante Gemeinsamkeiten zwischen den modellierten System (O⁶-MeG:C und O⁶-MeG:T-Modell) und der Kristallstruktur (PDBid 1T18) bestimmt werden. Zusätzlich wurden mit Hilfe der BP-REMD Methode Strukturen simuliert die der experimentellen Struktur ähneln. Darüber hinaus konnten für alle drei Modelle spezifische Zucker-Pucker Verteilungen ermittelt werden, welche wiederum als Erkennungsmerkmal dienen können.

3.3 Einfluss von Dangling-Ends auf die Stabilität von Nukleinsäuren

Der Einfluss von Dangling-End Basen ist abhängig von einer ganzen Reihe von Faktoren. Der Fokus dieses Abschnitts liegt in der Analyse der Unterschiede zwischen einem RNA-Dangling-End und einem DNA-Dangling-End. Untersucht wurden nur GC-haltige Sequenzen.

Experimentell wurde bereits für DNA Strukturen nachgewiesen, dass 5'-Dangling-End Basen einen stärkeren Beitrag zur Gesamtstabilität leisten als und 3'-Dangling-End Basen (39). Für RNA Strukturen gilt ein gegensätzlicher Effekt. Es wird angenommen, dass der Ursprung der Stabilisierung in der Stacking-Interaktion zwischen dem Dangling-End und den Terminalen Basen liegt (47, 39).

Zu erwähnen ist zudem, dass Dangling-End Basen bereits im Einzelstrang einen stabilisierenden Effekt besitzen können. Entgegensetzt der Annahme einiger Vorhersage-Methoden kann ein Einzelstrang bereits eine geordnete Struktur aufweisen und liegt nicht als *random coil* vor (220). Die Stabilisierung der DNA-Doppelhelix durch eine Dangling-End Base wird daher durch zusätzliche

Stacking-Interaktionen und Cross-Stacking Interaktionen mit den terminalen Basen begründet.

Ziel dieser Abschnitts ist es die Doppelstrang-Stabilisierung anhand der Unterschiede in den Freien Energien von gestackten zum solvatisierten Zustand im Doppelstrang vs. Einzelstrang zu bestimmen. Zu diesem Zweck wurden kMDS und Freie-Energie Simulationen (Umbrella-Sampling Simulationen) an selbstkomplementären Doppelstrang-Strukturen und den entsprechenden Einzelsträngen durchgeführt. Das Modell der selbstkomplementären Strukturen hat den Vorteil, dass eine bessere Konvergenz der ermittelten Energien möglich ist, da auf beiden Seiten identische Dangling-End Motive vorhanden sind und die gleiche Reaktionskoordinate verwendet wird, sodass erwartet wird, dass die Energien sich nur geringfügig unterscheiden.

3.3.1 Optimierung von Dangling-End Konformationen in freien Simulationen

B-DNA und A-RNA Strukturen wurden als Startgeometrie für die Simulationen verwendet. Durch das Entfernen einer Base an beiden Strängen wurden zwei identische Dangling-End Motive modelliert. Für die freien Simulationen (10 ns, 300K) wurde eine mittlere Struktur bestimmt die keine starken Abweichungen bezüglich der Startstruktur aufwies. Nichtsdestoweniger ist zu erwähnen, dass für die 3'-Dangling-End Motive von DNA Strukturen größere Abweichungen im Vergleich zum 5'-Dangling-End Motiv beobachtet wurden. Für RNA Strukturen wurden stärkere Abweichungen im Falle des 5'-Dangling-End beobachtet. Diese Beobachtung spricht für eine höhere Flexibilität des 5'-Dangling-End im Falle von A-RNA Strukturen und 3'-Dangling-End für B-DNA Strukturen.

Zwar sind die Abweichungen nur gering, jedoch können bereits kleine Unterschiede einen Einfluss auf die Gesamtstabilität der Helix haben. Interessanterweise korreliert die beobachtete Flexibilität der Dangling-End-Strukturen mit den experimentell bestimmten thermodynamischen Parametern. Zusätzlich wurde gezeigt, dass eine gute Übereinstimmung der experimentellen Strukturen und der mittleren Struktur aus den freien Simulationen vorhanden ist (sofern experimentelle Daten fuer die einzelnen Motive vorhanden waren) (siehe Abbildung 3.22)

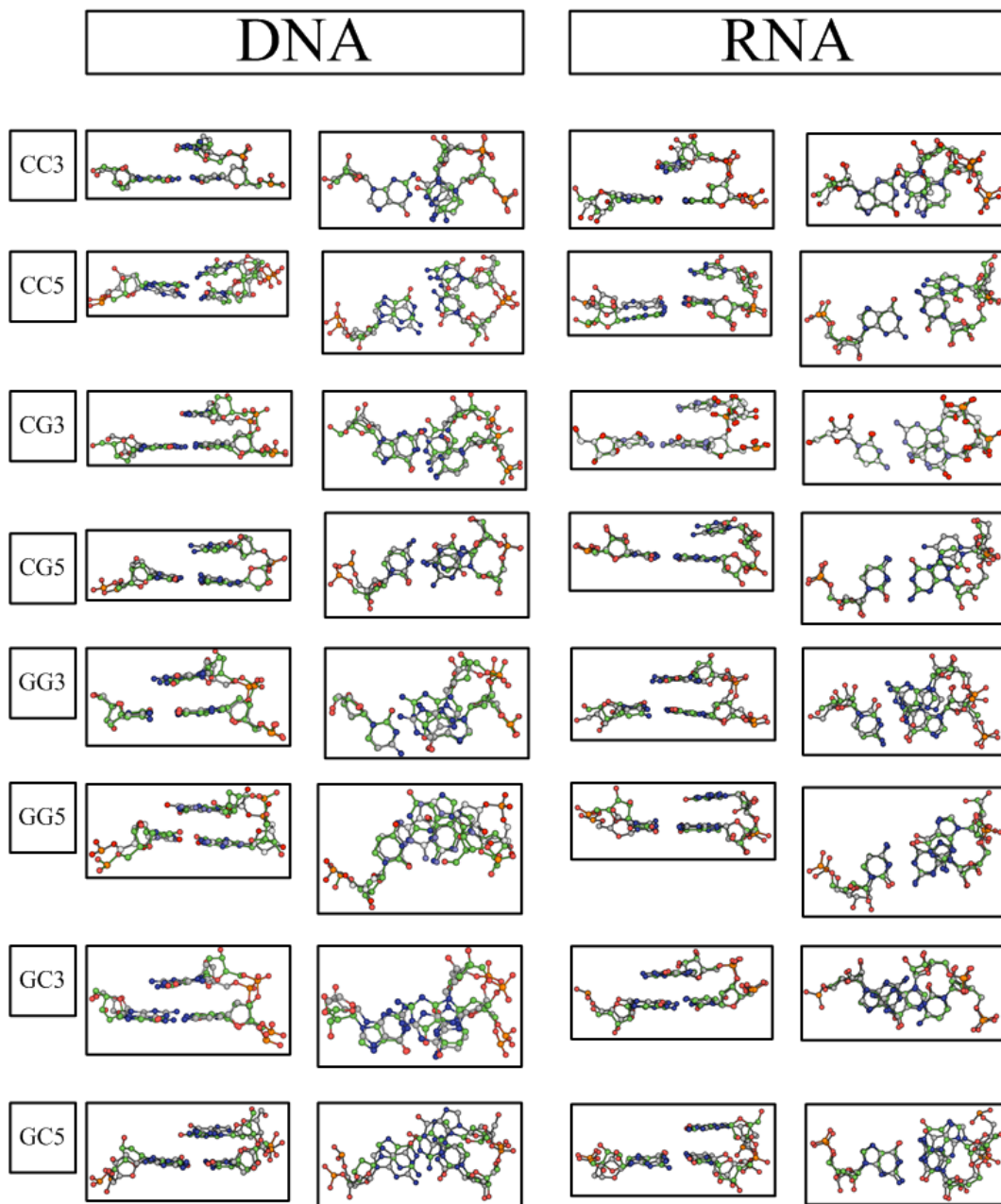


Abbildung 3.22 Superpositionierung von Average-Strukturen und Kristallstrukturen. In Fällen in denen keine Kristall-Struktur zur Verfügung stand wurden die Average-Strukturen mit den Startstrukturen superpositioniert. Folgende PDB-Strukturen wurden berücksichtigt: 1BGB (GC5-Motiv; DNA), 208D (CG5-Motiv; DNA), 476D (CG3-Motiv; DNA), 1DUL (GG5-Motiv; RNA), 1JBR (CG3-Motiv; RNA), 1L3D (CG5-Motiv; RNA), 1MJI (CC3-Motiv; RNA), 5MSF (GG3-Motiv; RNA), 429D (GC5-Motiv; RNA)

3.3.2 Berechnung der Freien Energie mittels Umbrella-Sampling Simulationen

Dangling-End-Strukturen haben einen stabilisierenden Einfluss auf Nukleinsäuren (39). Um den Beitrag dieses Effekts zu bestimmen, wurden Freie Energie-Simulationen (Umbrella-Sampling-Simulationen) durchgeführt. Als Reaktionspfad wurden Massenschwerpunkte der Dangling-End Base und des nächsten Nachbarn gewählt. Hierbei wurden mit einer Kraftkonstante von $1 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$ die Distanz zwischen den Basen pro Fenster um 0.5 \AA erhöht. So konnte ein langsamer Übergang vom gestackten zum völlig solvatisierten Dangling-End simuliert werden. Die Simulationen wurden für dsRNA, dsDNA und die entsprechenden Einzelstrang-Strukturen durchgeführt.

Die so bestimmten Freien-Energiekurven (PMF-Kurven) zeigen, dass das freie Energieminimum dem gestackten Zustand entspricht. So befindet sich das Minimum zwischen $4\text{-}5\text{\AA}$ und steigt für Distanzen über 6 \AA an. In einigen Simulationen wurden Sub-Minima beobachtet. Die Diskussion dieser semistabilen Intermediat-Zustände findet im nächsten Abschnitt statt.

Ein Vergleich der beiden PMF-Kurven aus den Doppelstrang-Strukturen zeigt, dass die Abweichungen zwischen den Kurven kleiner ist als $0.3 \text{ kcal}\cdot\text{mol}^{-1}$ ist. (siehe Abbildung 3.24 und Tabelle 3.9). Es kann somit davon ausgegangen werden, dass der Fehler der berechneten Freien Energiedifferenz etwa bei $\pm 0.5 \text{ kcal}\cdot\text{mol}^{-1}$ liegt. Neben den Fehlerbestimmungen aus den unterschiedlichen PMF Kurven wurden zusätzlich Monte Carlo Bootstrap Analysen durchgeführt. Die Methode ist standardmäßig in WHAM implementiert (183). Die Analyse ergab in allen Fällen einen Fehler $<0.1 \text{ kcal}\cdot\text{mol}^{-1}$. Der Unterschied (± 0.5 vs. <0.1) zeigt deutlich, dass durch die Monte Carlo Bootstrap Methode den Fehler zu gering abgeschätzt.

Obwohl in allen Simulationen der gestackte Zustand bevorzugt wurde, können dennoch Unterschiede in der Stabilisierung anhand der Position (3'-vs.-5'-Ende) und der Typ der Base (Purin vs. Pyrimidin) bestimmt werden. (siehe Tabelle 3.9).

Tabelle 3.9 Vergleich der experimentellen und berechneten Freien Energie Beiträge

		3'-Dangling-End				5'-Dangling-End				
		DNA		RNA		DNA		RNA		
		X				X				
		C	G	C	G	C	G	C	G	
→		-0.3±0.04 ^e	-0.0±0.04 ^e	-0.8 ^e	-1.7 ^e	→	-0.5±0.06 ^e	-0.7±0.05 ^e	-0.2 ^e	-0.2 ^e
CX						XC				
G		-0.3±0.17 ^c	-0.60±0.14 ^c	-1.0±0.11 ^c	-1.9±0.13 ^c	G	-0.8±0.11 ^c	-1.4±0.42 ^c	0.4±0.20 ^c	-0.5±0.08 ^c
←						←				
→		-0.2±0.05 ^e	-0.4±0.05 ^e	-0.4 ^e	-1.3 ^e	→	-0.3±0.04 ^e	-0.6±0.07 ^e	-0.3 ^e	-0.0 ^e
GX						XG				
C		-0.4±0.09 ^c	-1.2±0.46 ^c	-0.2±0.16 ^c	-0.8±0.18 ^c	C	-0.9±0.24 ^c	-0.6±0.21 ^c	-0.4±0.09 ^c	0.1±0.13 ^c
←						←				

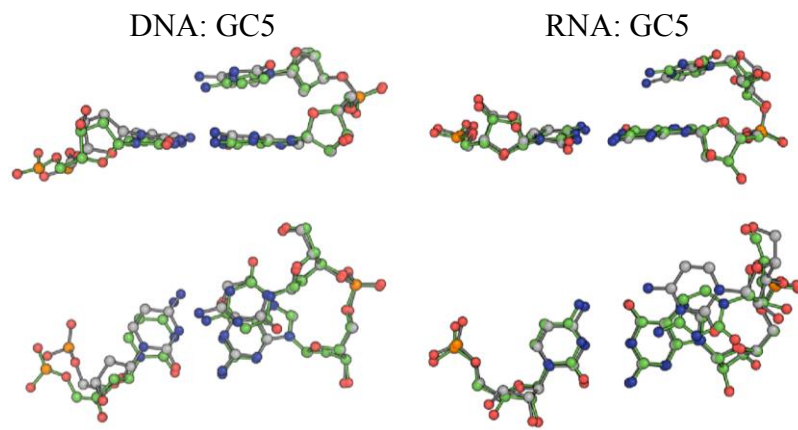
^e=Experimentelle Werte aus (51)^c=Berechnete freie Energie Beträge nach dem Schema in Abbildung 3.25

Abbildung 3.23 Superpositionierung der 5'-Cystein Dangling-End Base. Ein Vergleich der 5'-Cystein Dangling-End Average-Struktur (grün, balls und stick Repräsentation) während der kMDS Simulation (ohne restraints) und der experimentellen Struktur (PDBid: 1B8I, 1LNG) verdeutlicht die Stabilität der 5'-Dangling-End-Struktur in DNA Strukturen im Vergleich zu RNA Strukturen. Dargestellt sind hierbei die Dangling-End Base und das benachbarte Basenpaar in einer Seiten-Ansicht(oben) und eine um 90° und die helikale Achse rotierte Oben-Ansicht(unten)

Im Folgenden wird angenommen, dass der völlig solvatisierte Zustand einer Dangling-End Base keinen spezifischen Beitrag zur Gesamtstabilität der Helix leistet. Zu erwähnen hierbei ist jedoch, dass auch der solvatisierte Zustand einen Beitrag leisten kann.

Tabelle 3.10 RMSD (der Schweratome in Å) zwischen der Average-Struktur (Dangling-End Base + benachbartes Basenpaar) bezüglich der experimentellen Struktur (falls vorhanden, grau) oder der Startstruktur

	CC3	CC5	CG3	CG5	GG3	GG5	GC3	GC5
RNA	0.5	0.2	0.3	1.0	0.8	0.4	0.2	0.3
DNA	0.2	1.1	1.1	1.3	0.4	1.1	1.0	0.2

So kann die zusätzliche Phosphat-Gruppe der Dangling Base einen elektrostatischen Effekt auf die Wechselwirkung mit dem gegenüberliegenden Strang haben und zudem den Einfluss der Salzkonzentration auf die Stabilität der Gesamthelix mitbestimmen. Nichtsdestotrotz zeigen experimentelle Daten, dass eine zusätzliche Phosphatgruppe nur einen geringen Einfluss auf die Gesamtstabilität hat ($\sim 0.1 \text{ kcal}\cdot\text{mol}^{-1}$, (37)).

Für die Bestimmung des Dangling-End Beitrags zur Gesamtstabilität muss berücksichtigt werden, dass bereits im Einzelstrang der Nukleinsäuren der gestackte Zustand vorliegt, sodass der zusätzliche Beitrag nur durch ein Cross-Stacking der Dangling-End Base und der terminalen Base des gegenüberliegenden Stranges bestimmt wird.

Der Beitrag der Stabilisierung kann daher aus der Differenz der Freien-Energiedifferenzen des Doppelstranges und der Freien-Energiedifferenzen des Einzelstranges berechnet werden (siehe Abbildung 3.25).

Die so ermittelten Beiträge zeigen für alle Simulationen einen negativen Freien-Energie Beitrag ($\Delta\Delta G$) und sind mit akzeptablen Abweichungen in Übereinkunft mit den experimentellen Daten (siehe Tabelle 3.9). In einigen Fällen konnte eine sehr gute Übereinstimmung zwischen berechneten und experimentell bestimmten Beiträgen gezeigt werden (siehe Tabelle 3.9). Es ist weiterhin darauf hinzuweisen, dass die Vernachlässigung des Stacking-Beitrages im Einzelstrang zu einer Überschätzung des Dangling-End Beitrages zur Gesamthelix führen würde. Betrachtet man zum Beispiel die Freie-Energiedifferenz des 5'-Dangling-End Base (Cytosin) auf einem Terminalen C:G Paar (siehe Abbildung 3.24) so stellt man fest, dass die schwarze und die rote Kurve eine Freie-Energiedifferenz von $> 2.5 \text{ kcal}\cdot\text{mol}^{-1}$ aufweisen.

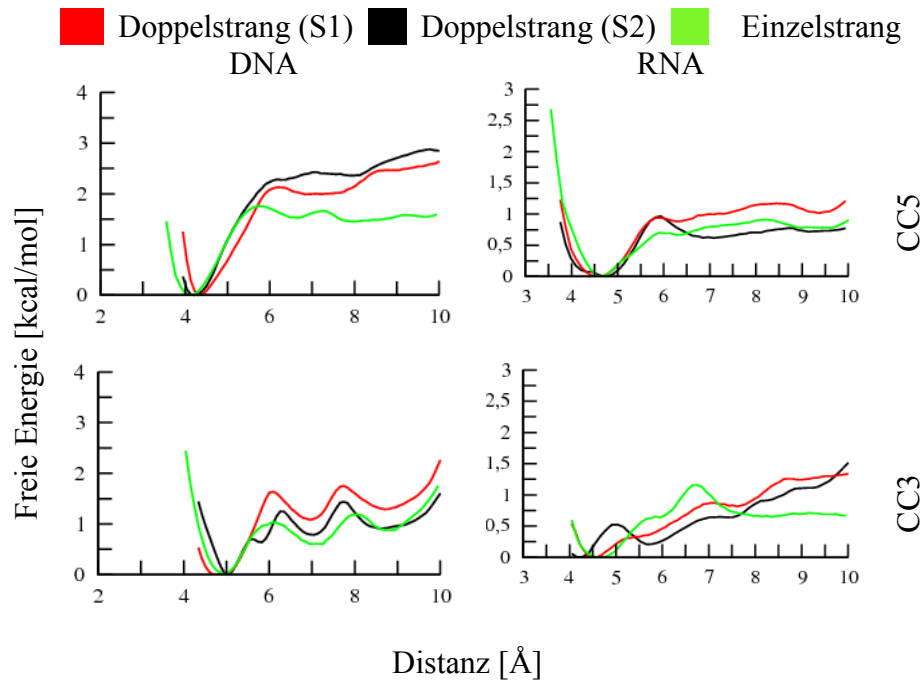


Abbildung 3.24 Vergleich zwischen den RNA und DNA Energieprofilen. Berechnet wurde die das Freie Energieprofil entlang der Distanz der Massenschwerpunkt der Dangling-End Base und der Nachbarbase. Hierbei wurde die Distanz von 5 Å \rightarrow 10 Å (0.5 Å/Simulationsfenster) erhöht. Das Potentials-of-mean force wurde für beide Dangling-End Motive des Doppelstranges (rot/schwarz) und dem Einzelstrang (grün) berechnet.

Angenommen der Freie-Energiebeitrag der Dangling-End Base hätte nur im Doppelstrang einen Einfluss auf die Stabilität (Vernachlässigung der Stacking-Stabilität im Einzelstrang (grüne Kurve, siehe Abbildung 3.24), erhält man eine signifikante Abweichung zu den experimentellen Daten (siehe Tabelle 3.9).

Dieses Beispiel bestätigt, dass es wichtig ist den Einfluss im Einzelstrang zu berücksichtigen. Obwohl keine quantitative Übereinstimmung der experimentellen Daten und der berechneten Daten vorhanden ist, kann dennoch ein qualitativer Trend zwischen berechneten und experimentell bestimmten Beiträgen beobachtet werden. Zusätzlich ist darauf hinzuweisen, dass experimentelle Messungen einen Restfehler aufweisen und sich die Ergebnisse der Messungen je nach Protokoll unterscheiden können.

Der Trend, dass 5'-Dangling-End Basen im Falle von DNA Strukturen einen stärkeren Effekt auf die Gesamtstabilität haben, konnte durch die Simulationen bestätigt werden (siehe Tabelle 3.9).

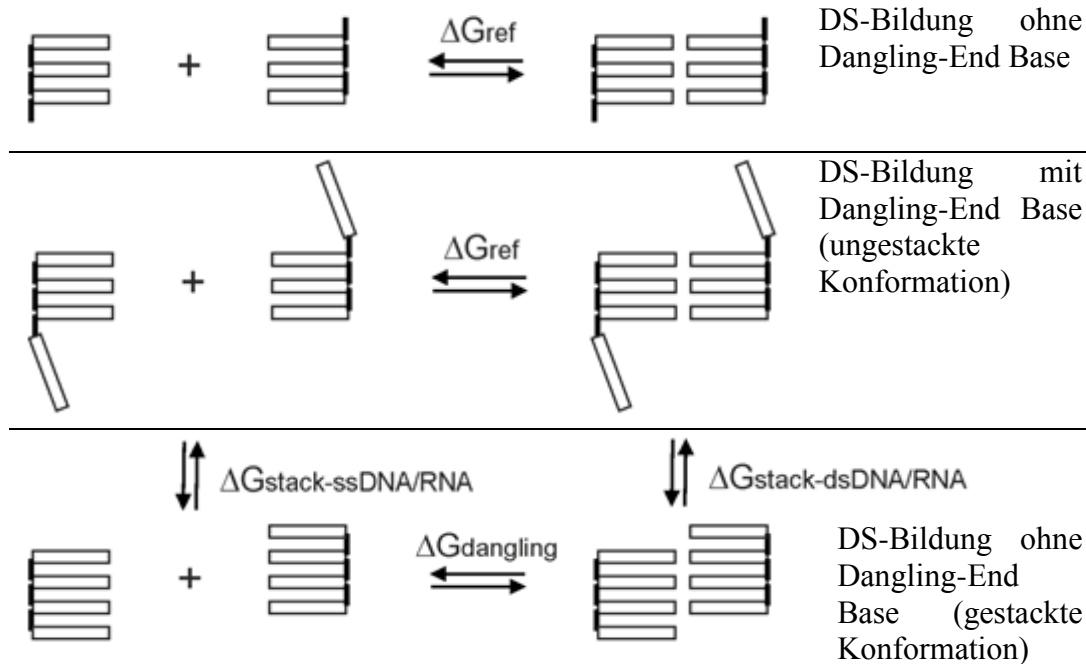


Abbildung 3.25 Thermodynamischer Kreis zur Berechnung des Freien Energiebeitrages zur Doppelstrang Stabilität. Annahme ist, dass die ungestackten Konformationen keinen Beitrag zur Duplex-Stabilität leistet. Daraus folgt, dass die Doppelstrang-Bildung (DS-Bildung) mit dieser Konformation und der Doppelstrang-Bildung ohne Dangling-End Base (nur Phosphat-Gruppe) Äquivalent sind (zu beachten ist jedoch, dass die Phosphat-Gruppe durch elektrostatische Interaktionen einen Beitrag zur Doppelstrang-Bildung leisten kann). Die Freie Energie der Doppelstrang-Bildung und der Dangling-End Base in der ungestackten Konformation entspricht der Freien Energie der Referenz (ΔG_{ref}). Der zusätzliche Beitrag durch die Dangling-End Base zur Doppelstrang-Bildung kann aus der Differenz zwischen der Stacking-Freien Energy im Duplex vs. Einzelstrang bestimmt werden. (Vertikal Pfad im thermodynamischen Kreis). $\Delta G_{\text{dangling}}$ beschreibt die Summe aus dem Beitrag des Dangling-End und ΔG_{ref} .

Zudem konnte auch der gegensätzliche Effekt bei RNA Strukturen durch die Simulationen bestätigt werden (siehe Tabelle 3.9). Aus struktureller Sichtweise, sind die Ursachen für diesen Effekt durch das Cross-Stacking der Dangling-End Base mit dem gegenüberliegenden Strang zu erklären. Ein weiterer Aspekt der mit den experimentellen Beobachtungen übereinstimmt ist, dass der Einfluss von Purin Basen stärker ist als der Einfluss von Pyrimidine Basen.

3.3.3 Analyse der Zusammensetzung der Energiebeiträge

Um die Energiebeiträge die zur Stabilisierung durch Dangling-End-Strukturen genauer zu untersuchen, wurden MM/PBSA Rechnungen durchgeführt. Hierbei wurden die Trajektorien in gestackte und völlig solvatisierte Konformationen eingeteilt. Zudem wurden die verschiedenen Energiekomponenten für beide Zustände miteinander verglichen.

Zum Vergleich wurden die MM/PBSA Ergebnisse mit den experimentell bestimmten Enthalpie-Werten (ΔH_{exp}) verglichen, sofern sie vorhanden waren. Wie bereits bei den Freien-Energie Rechnungen wurden auch hier die zum Experiment vergleichenden Beiträge mittels der Energiedifferenzen zwischen Doppelstrang und Einzelstrang bestimmt. Zur Berechnungen wurden jeweils die Mittelwerte über alle gestackten Konformationen und alle völlig solvatisierten Konformationen herangezogen (siehe Kapitel 2.3.2).

Da für RNA Strukturen keine vergleichbaren experimentellen Daten (ΔH_{exp}) gefunden werden konnten, ist kein direkter Vergleich möglich. Es konnte jedoch gezeigt werden, dass der Enthalpie-Beitrag des Dangling-End Purin Base im Vergleich zu Pyrimidin Base stärker ist (45)

Bei dem Vergleich der experimentellen Daten mit berechneten Daten ist zusätzlich zu erwähnen, dass die experimentell bestimmten Daten mit einem starken Fehler, als die Freie-Energiedifferenzen (ΔG_{exp}), behaftet sind. Grund hierfür ist, dass ΔH_{exp} aus relativ kleinen Energiedifferenzen aus unterschiedlichen Studien und Temperaturen berechnet werden.

Mit Ausnahme des GG3/GG5 Falles kann gezeigt werden, dass die 5'-Dangling-End Basen (DNA) energetisch bevorzugt werden im Vergleich zu 3'-Dangling-End (DNA). Dieser Trend ist auch in experimentellen Daten wiederzufinden (siehe Tabelle 3.11). Für den GG3/GG5 Fall weisen die experimentellen ΔH_{exp} Beiträge darauf hin, dass der Stabilisierungseffekt für beide Dangling-End Positionen in der gleichen Größenordnung liegen (siehe Tabelle 3.11). Diese Beobachtung stimmt qualitativ mit den berechneten Daten überein, die nur einen geringen Unterschied in den MM/PBSA Energien aufweisen (siehe Tabelle 3.11).

Nicht-gebundene Wechselwirkungen wie die vdW-Wechselwirkungen und elektrostatische Interaktionen leisten im Doppelstrang im Vergleich zum

Einzelstrang den größten Beitrag zur Gesamtstabilität. VdW-Interaktionen haben hierbei einen positiven Beitrag zur Dangling-End Stabilisierung, jedoch wird kein Sequenz-spezifischer Effekt beobachtet. Elektrostatische Interaktionen destabilisieren ein Stacking der Dangling-End Base, jedoch haben der Coulomb-Interaktionen im Falle von 5'-Dangling-End Basen einen stabilisierenden und für 3'-Dangling-End Basen einen destabilisierenden Effekt auf das Stacking. Diese Beobachtung gilt jedoch nur für die DNA-Strukturen (siehe Tabelle 3.11).

Tabelle 3.11 Energiebeiträge der Dangling-End Base zur Doppelstrang-Bildung (alle Beiträge in kcal·mol⁻¹)

DNA								
STACK-UNSTACK	CC3	CC5	CG3	CG5	GG3	GG5	GC3	GC5
GEBUNDENE WW	-0.1	-1.1	-0.0	-2.4	0.2	-0.5	0.4	1.5
VDWAALS-DIFF	-8.9	-11.1	-9.9	-8.5	-15.8	-13.4	-19.0	-14.2
EEL-DIFF	7.1	-16.4	9.4	-33.2	7.2	-10.5	8.4	18.8
EPB-DIFF	-2.1	22.6	-3.7	38.1	-0.4	16.4	4.8	-17.3
ENPOLAR-DIFF	-0.8	-0.9	-0.4	-0.7	-1.0	-0.7	-1.0	-0.8
Estatic	5.0	6.2	5.7	4.9	6.8	6.0	13.1	1.5
Etot	-4.1	-6.0	-4.3	-5.9	-8.8	-7.9	-5.5	-11.1
Exp. Wert (51)*	-2.6	-4.4	-0.2	-4.0	-3.9	-3.9	-3.2	-5.1
RNA								
STACK-UNSTACK	CC3	CC5	CG3	CG5	GG3	GG5	GC3	GC5
GEBUNDENE WW	0.4	-0.7	0.6	-0.4	0.3	0.5	0.5	-0.3
VDWAALS-DIFF	-3.4	-1.5	-4.3	-1.1	-7.8	-1.4	-11.7	-2.5
EEL-DIFF	12.9	8.0	-8.1	-16.5	7.5	34.4	17.6	12.0
EPB-DIFF	-12.2	-8.4	8.9	18.2	-6.1	-29.4	-13.6	-8.4
ENPOLAR-DIFF	-0.3	-0.2	-0.4	-0.2	-0.4	-0.2	-0.5	0.3
Estatic	0.7	-0.4	0.8	1.8	1.4	5.0	4.0	3.6
Etot	-2.3	-2.6	-3.0	0.2	-6.1	4.1	-7.2	0.8

*entspricht den ΔH_{exp} (in kcal·mol⁻¹)

Gebundene WW bezieht sich auf Differenzen in den Termen der gebunden Wechselwirkungen (Bindungslängen, Bindungswinkel und Diederwinkel). VDWAALS-DIFF bezieht sich auf Differenzen in den Lennard-Jones van der Waals Term; EEL-DIFF und EPB-DIFF beziehen sich auf Differenzen im Coulomb Term und Dipol-Dipol Interaktionen und ENPOLAR-DIFF bezieht sich auf nicht-polare Oberflächen Wechselwirkungen abhängig von der Solvatationsenergie. (Estatic: Differenz der aller Elektrostatik-Terme)

Für RNA-Strukturen wurde ein geringerer Beitrag für die Stabilität durch Dangling-End-Motive bestimmt (siehe Tabelle 3.11, RNA). In Übereinkunft mit den experimentellen Daten konnte gezeigt werden, dass die 3'-Dangling-End Motive einen stärkeren Effekt auf die Gesamtstabilität haben als 5'-Dangling-End Motive (mit Ausnahme von CC3 vs. CC5). Der geringere Beitrag im Falle der

RNA Strukturen kann in dem geringeren Beitrag der vdW-Wechselwirkung begründet sein, da hier die cross-stacking Energien geringer ausfallen. Bestätigt wird dies zudem durch den Beitrag der nicht-polaren Wechselwirkungen. Ein direkter Vergleich der RNA und DNA Beiträge lässt darauf schließen, dass die Unterschiede in den Beiträgen in der wasserzugänglichen Oberfläche beim Prozess vom gestackten in den völlig solvatisierten Zustand liegen (Doppelstrang vs. Einzelstrang RNA).

Die MM/PBSA Energien sind also in einer vergleichbaren Größenordnung wie die experimentellen Daten. Ein qualitativer Trend zwischen experimentellen Daten und berechneten Daten kann nachgewiesen werden (siehe Tabelle 3.11). Es ist jedoch darauf hinzuweisen, dass keine quantitative Übereinstimmung erzielt werden konnte (siehe Tabelle 3.11).

3.3.4 Konformationsanalyse der Umbrella-Sampling Simulationen

Die vorhergesagten Freien-Energiekurven zeigen in einigen Fällen kleine Energiebarrieren um vom gestackten in den ungestackten Zustand überführt zu werden (siehe Abbildung 3.24). Es handelt sich hierbei um Übergangszustände die in Lösung auftreten können. Da die Reaktionskoordinate so gewählt wurde, dass Distanz zwischen den Massenschwerpunkten der Dangling-End Base und der terminalen Nachbarbase erhöht wird, sind die meisten Intermediate-Zustände nicht in spezifischer Interaktion mit dem gegenüberliegenden Strang (siehe Abbildung 3.26 und Abbildung 3.27).

Exemplarisch für einige DNA-Strukturen sind typische Motive der Dangling-End Base, die während der Umbrella-Sampling-Simulation beobachtet wurden, und sich nicht in einem gestackten Zustand befinden, auf Abbildung 3.26 dargestellt.

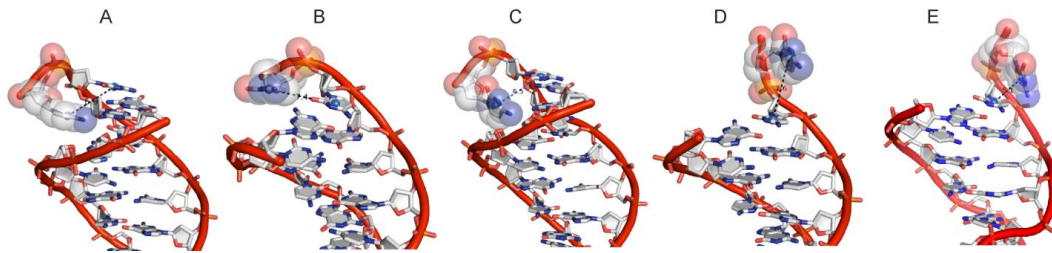


Abbildung 3.26 Vergleich der Zwischenstände während der Umbrella-Sampling Simulation. (A-E) Dargestellt sind die Zwischenzustände der Cytosin Dangling-End-Base (in der vdW-Repräsentation, restliche Basen des Doppelstrangs in der stick Repräsentation). Konformationen A-C entsprechen ungestackten Zwischenzuständen. Sie interagieren mit den Nachbarbasen oder der Zucker-Phosphat-Gruppe (in der kleinen Furche). Die Konformationen D und E befinden sich in einer looped-out Konformationen, interagieren jedoch mit dem Phosphat-Rückgrat.

Üblicherweise wurden die kurzlebigen Intermediate-Zustände zwischen 6.5-8 Å der Massenschwerpunkt-Distanz beobachtet. Hierbei positionierten sich die Dangling-End Base in der kleinen Furche des Doppelstrangs (siehe Abbildung 3.26 (A-C)) oder interagieren mit dem Phosphatrückgrat der Nachbarbase des Doppelstrangs (siehe Abbildung 3.26 (D, E)).

Für RNA Strukturen wurde ein charakteristischer Intermediate-Zustand beobachtet. So wurde in allen Sequenzen (RNA-Umbrella-Sampling-Simulationen) ein Zustand beobachtet in dem die Dangling-End Base senkrecht mit dem terminalen Basenpaar in der kleinen Furche interagiert (siehe Abbildung 3.27).

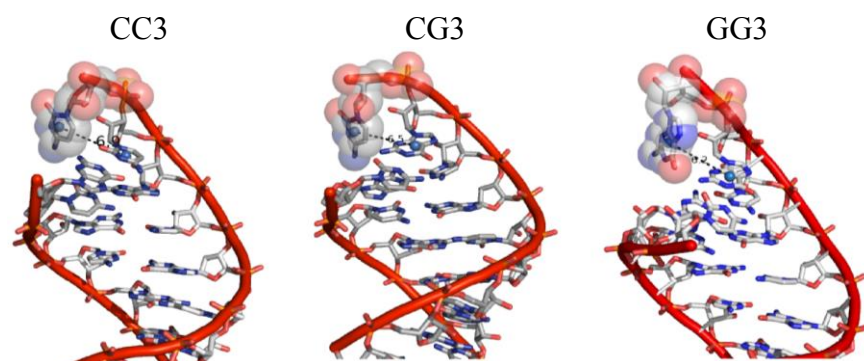


Abbildung 3.27 Charakteristische Zwischenzustände von RNA Dangling-End-Strukturen während der Umbrella-Sampling Simulation. Eine spezifische Konformation wurde für in allen US-Sampling Simulationen beobachtet. Sie ist charakterisiert durch die senkrechte Orientierung zu dem benachbarten Basenpaar und befindet sich in der kleinen Furche des Doppelstrangs. (Dangling-End Base in der vdW Repräsentation, Doppelstrang in der Stick-Repräsentation).

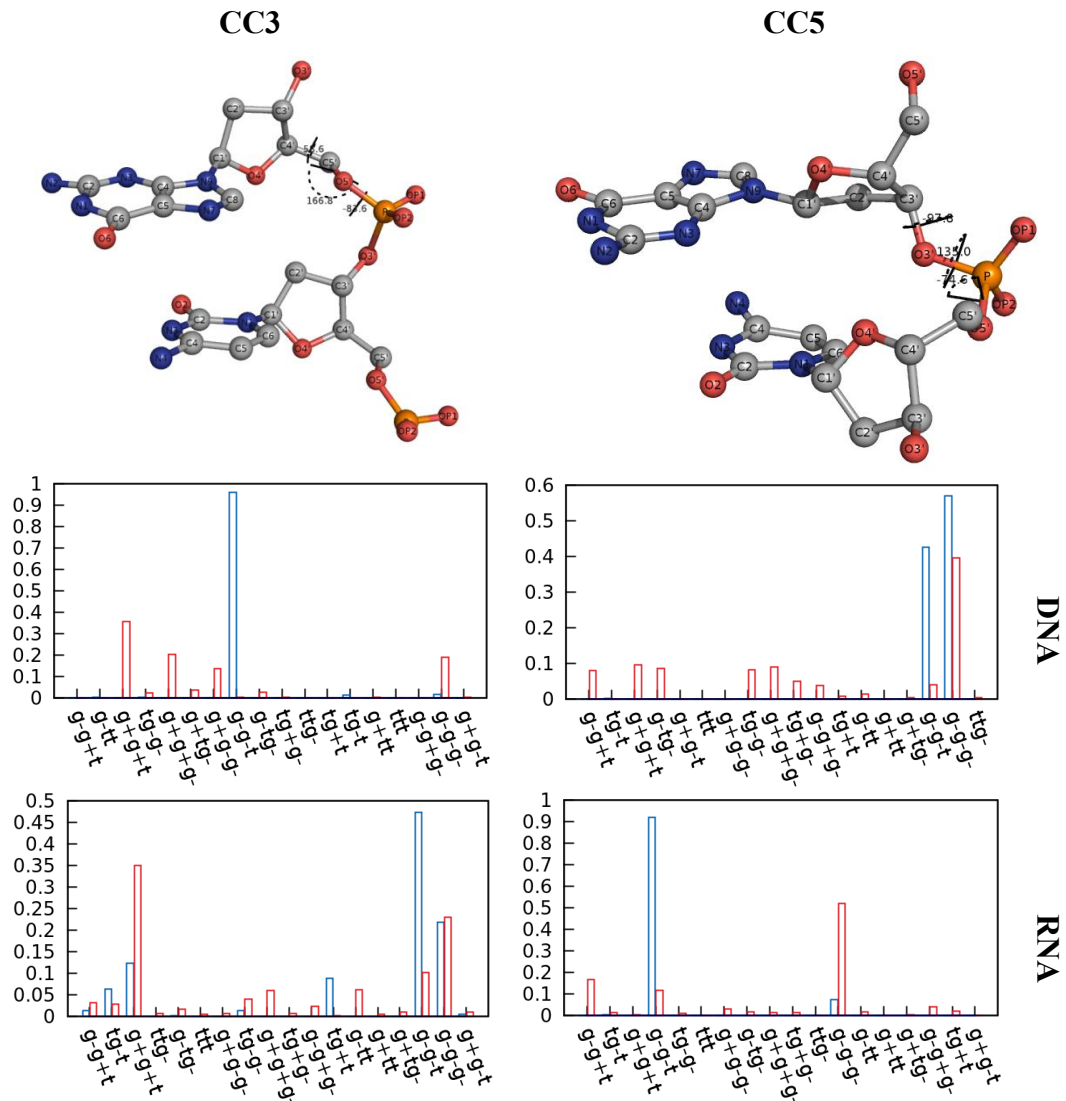


Abbildung 3.28 Verteilung der Phosphatrückgrat Diederwinkel. Verteilung der Phosphat-Rückgrat Diederwinkel im gestackten (blau) und ungestackten Zustand (rot). Die Strukturen wurden anhand ihrer γ, β, α - Diederwinkel für 3'-Dangling-End-Strukturen und ϵ, ζ und $\alpha+1$ für 5'-Dangling-Ends Strukturen unterteilt.

Dieser Intermediat-Zustand tritt in den Simulationen mit einer Distanz über 8 Å nicht mehr auf, da sich die Dangling-End Base dann in einem völlig solvatisierten Zustand befindet. Beschreibt man die verschiedenen Zustände anhand der Phosphatrückgrat Diederwinkel, so stellt man fest, dass die ungestackten Konformationen einen größeren Konformationsraum einnehmen (siehe Abbildung 3.28). Zur Klassifizierung der 3'-Dangling-End Base wurden die α, β und γ Diederwinkel des Phosphatrückgrats verwendet.

Für die 5'-Dangling-End Basen wurden die ε , ζ und α zur Klassifizierung herangezogen (siehe Abbildung 3.28). Interessanterweise wurden für die 3'-Dangling-End Basen im gestackten Zustand im Vergleich zum entsprechenden 5'-Dangling-End mehr mögliche Konformationszustände beobachtet.

Die unterschiedlichen Verteilungen der Diederwinkel weisen darauf hin, dass bereits im gestackten Zustand die Base mehrere Konformationen einnehmen kann. Für DNA gilt der gegensätzliche Effekt. Hierbei ist das 5'-Dangling-End flexibler als das 3'-Dangling-End. Die erhöhte Mobilität erleichtert die Interaktion mit dem gegenüberliegenden Strang, sodass die wasserzugängliche Fläche geringer wird. Dies ermöglicht ein besseres Stacking zwischen der Dangling-End Base und dem terminalen Basenpaar.

3.3.5 Zusammenfassung und Schlussfolgerung

Dangling-End-Strukturen sind häufig auftretende Motive in Nukleinsäuren. Speziell bei RNA-Strukturen nehmen sie eine wichtige funktionale Bedeutung ein. Die Mechanismen der Doppelstrang-Stabilisierung durch Dangling-End-Strukturen sind weiterhin nicht vollständig aufgeklärt. Häufig wird ein Zusammenhang zwischen der Stacking Interaktion zwischen dem Dangling-End und dem terminalen Basenpaar im Doppelstrang hergestellt. Nichtsdestoweniger können bereits im Einzelstrang die terminalen Basen im gestackten Zustand vorliegen (im Äquilibrium mit dem ungestackten Zustand). Die zusätzliche Base kann während der Doppelstrang Bildungen einen Beitrag zur Stabilisierung des gestackten Zustandes leisten. Durch Cross-Stacking mit dem gegenüberliegenden Strang wird ein weiterer Beitrag zur Gesamtstabilität der Helix geleistet.

In der vorliegenden Studie wird angenommen, dass die Unterschiede in den Stacking-Energien zwischen Doppelstrang und Einzelstrang die dominanten Faktoren sind, die zu einer Verbesserung der Gesamtstabilität führen. Die berechneten Freie-Energiedifferenzen sind im akzeptablen Einklang mit den experimentellen Daten, sodass von einer realistischen Kraftfeldbeschreibung des Stacking-Effekts ausgegangen werden kann.

Weiterhin deuten die berechneten Daten darauf hin, dass der Stacking-Effekt ausreichend ist, um den Stabilisierungseffekt durch Dangling-End Motive zu begründen. Ferner ist zu erwähnen, dass die experimentellen Daten bzw. die Trends, die sich hieraus ergeben, qualitativ anhand der Simulationen reproduziert werden können.

Die Analyse der Energiebeiträge ergab, dass vdw-Interaktionen einen starken Beitrag zum Stabilitätseffekt durch die Dangling-End Base leisten. Die berechneten Daten weisen darauf hin, dass dieser Effekt in DNA Strukturen stärker ausgeprägt ist, als in RNA Strukturen. Neben vdw-Interaktionen spielen auch elektrostatische Interaktionen eine besondere Rolle. Es konnte gezeigt werden, dass elektrostatische Interaktionen maßgeblich zur Erklärung der unterschiedlichen Stabilitätseffekte zwischen 5'-Dangling-End Basen und 3'-Dangling-End Basen sowie für sequenzspezifische Effekte sind.

Konformationsanalysen der Simulationen zeigten, dass eine Reihe von Intermediatzuständen der Dangling-End Base existieren. Die energetisch günstige Konformation ist jedoch der gestackte Zustand.

Die vorliegenden Freien-Energieprofile könnten hilfreich sein um spezifischer Dangling-End Motive zu entwerfen, mit dem Ziel die Gesamtstabilität der Helix zu verbessern. Insbesondere in der Sekundärstrukturvorhersage und dem Sequenz-Entwurf für Hybridisierungs-Experimente (DNA-Microarray Experiment) sind solche Informationen hilfreich.

3.4 RNA-Bulge

RNA Bulge Strukturen sind strukturelle RNA Motive, die in gefalteten Strukturen auftreten, in RNA-Ligand Bindungen involviert sind und wichtig sind für die tertiär Struktur der RNA. Es handelt sich bei RNA-Bulge-Strukturen um Basen die keinen WC-Partner aufweisen (siehe Abschnitt 1.3.1). Die RNA-Bulge-Struktur kann hierbei eine Reihe von unterschiedlichen Konformationen eingehen, welche die Gesamtstruktur der Helix beeinflussen. Die Konformation der RNA ist hierbei von einer Reihe von Faktoren abhängig (z.B. Salzkonzentration und Sequenz) ((21). In dieser Arbeit werden Single-RNA Bulge Strukturen diskutiert. Hierbei sollen vor allem unterschiedliche Konformationsbereiche einer U-Bulge

bzw. A-Bulge RNA Struktur ermittelt werden. Zu diesem Zweck wurden kMDS und BP-REMD Simulationen durchgeführt. Zusätzlich ist anzumerken, dass beide Systeme sich nur in ihrer Bulge-Base unterscheiden.

3.4.1 Gesamtstabilität der Strukturen während der Simulation

Um die Stabilität der Gesamt-Struktur zu bestimmen, wurden die RMSD Werte des Zucker-Phosphatrückgrats bestimmt. Da die terminalen Basen mittels Distanz-Einschränkungen zwischen den Wasserstoffbrücken der WC-Partner aufrecht gehalten wurden, kann ein *fraying* am Terminus ausgeschlossen werden. Somit kann eine Korrelation zwischen den RMSD-Werten der Gesamthelix und den Konformationsänderungen der Bulge-Base erstellt werden. Wie aus der Abbildung 3.29 zu entnehmen ist, bleibt die Struktur des 2K3Z-Modells während der kMDS stabil. Der RMSD-Wert fluktuiert hierbei zwischen 1-2 Å. Zwischen ~190ns und 200ns befindet sich der RMSD Wert zwischen 2-3 Å. Dieser Zustand ist zwar kurzlebig, tritt jedoch während der Simulation immer wieder auf. Die RMSD-Werte der entsprechenden BP-REMD weisen deutlich größere Fluktuationen auf. So befindet sich der RMSD-Wert, wie bei der kMDS zunächst zwischen 1-2 Å. Zwischen 6-10 ns fluktuiert der RMSD-Wert hingegen zwischen 2-5 Å. Im weiteren Verlauf der Simulation sind immer wieder Konformationen mit höheren RMSD Werten zu beobachten. Diese Schwankungen befinden sich zwischen 5-7 Å (siehe Abbildung 3.29).

Die 2K41-kMDS wird von zwei Zuständen dominiert. Nach etwa ~45ns ist ein Übergang der Konformation zu beobachten. Der RMSD-Wert ändert sich hierbei von 1-2 Å auf 2-3 Å. Dieser Zustand bleibt etwa 15 ns erhalten und tritt während der Simulation bei etwa 150ns und zwischen 280 und 320 ns wieder auf. Bei 150 ns ist ein dritter Zustand mit einem RMSD-Wert von 4 Å zu beobachten. Dieser Zustand ist jedoch nur von kurzer Dauer (~10ns).

In Abbildung 3.30 sind die dominanten Konformationen der kMDS dargestellt. Es handelt sich hierbei um einen Übergang des U-Bulges in eine intrahelikale Konformation. Wie aus der Abbildung zu entnehmen ist, wird diese unter anderem durch polare Interaktionen zwischen der Ribose des U-Bulges und dem Zucker-Phosphatrückgrat der nächsten Nachbarn stabilisiert. Der Verlauf der BP-REMD ist mit dem Verlauf der BP-REMD des 2K3Z Modells vergleichbar.

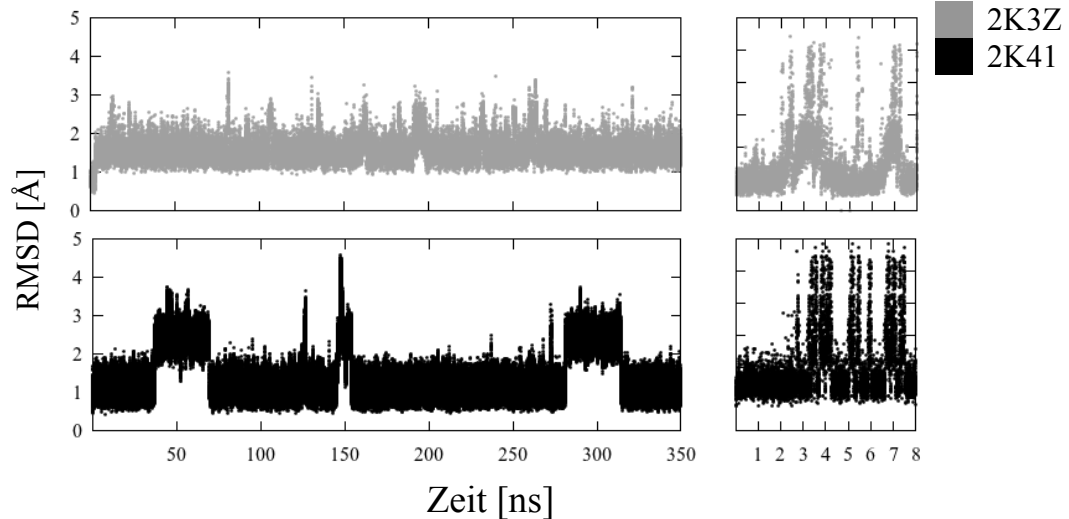


Abbildung 3.29 RMSD Werte der A-Bulge und U-Bulge Modelle im Vergleich. Die RMSD-Werte sind bezüglich der Startstruktur des Zucker-Phosphat-Rückgrats gegen die Zeit aufgetragen. Die schwarze Linie stellt den Verlauf der Werte des 2K41-Modell, die graue Linie das 2K3Z Modell. Die Werte der BP-REMD stammen aus dem Referenz-Replika (Replika 0). Für das 2K41-Modell werden in der kMDS zwei dominante Zustände beobachtet (siehe Abbildung 3.30)

Erste Übergänge der RMSD-Werte lassen sich nach etwa 6ns beobachten. Anschließend findet in regelmäßigen Abständen ein Übergang zwischen 2-5 Å statt. Insgesamt sind in beiden BP-REMD Simulationen mehr Übergänge zwischen den Konformationen zu beobachten. Weiterhin lässt sich an dieser Stelle auf den Vorteil der BP-REMD hinweisen. So sind die Übergänge nur von kurzer Dauer, während in der kMDS die Gefahr der Überpräsentierung einer Konformation vorhanden ist.

In der 2K41-kMDS finden nur wenige Übergänge statt, die über mehrere nsec in einem Zustand verweilen. Anhand der Simulationen dieser Systeme wird der Vorteil der BP-REMD Simulation besonders deutlich. Neben dem Vorteil der Simulationsdauer, ist es mit dieser Methode, aufgrund der häufigen Übergänge zwischen den Replika, auch möglich schneller konvergierte Ergebnisse zu erhalten und Konformationsräume abzusuchen die mit Hilfe der kMDS während der 350ns nicht bestimmt werden konnten jedoch experimentell nachgewiesen wurden für dieses Sequenzmotiv.

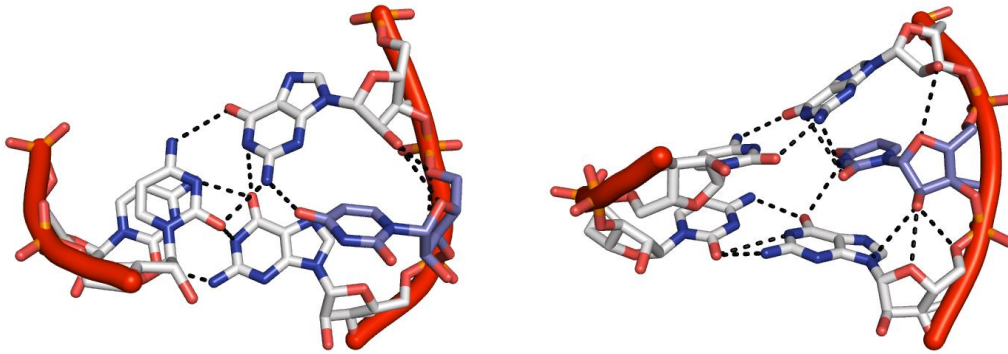


Abbildung 3.30 Charakteristische Zustände des U-Bulges während der kMDS. Zwei charakteristische Zustände während der kMDS werden beobachtet. Die U-Bulge Base (violett, stick-Repräsentation) befindet sich entweder extrahelikal (Startstruktur, links) oder intrahelikal (rechts)

3.4.2 Vergleich zwischen BP-REMD und kMDS Konformationen

Experimentell verfügbare Strukturen mit dem gleichen Sequenzkontext (GUG bzw. GAG) wurden mittels der RNA-Strukturdatenbank *RNA FRABASE 2.0* (<http://rnafrabase.ibch.poznan.pl/>) ermittelt und untersucht. Dabei konnten eine Reihe von verschiedenen Konformationen für das jeweilige Motiv gefunden werden (siehe Abbildung 3.31). Interessanterweise wurden für beide Strukturen sowohl intra- als auch extrahelikale Strukturen gefunden. Die Strukturen durften dabei nicht mit weiteren Biomolekülen wie Proteinen oder anderen Nukleinsäuren interagieren. Dieses Strukturvergleiche verdeutlichen, dass RNA-Bulges dynamische Strukturen sind, die eine Vielzahl von Konformationen einnehmen können. Um den Konformationsraum genauer zu untersuchen und die Verteilung der verschiedenen Konformationen zu bestimmen wurden kMDS und BP-REMD Simulationen durchgeführt. In Abschnitt 3.4.1 wurde bereits auf die Problematik der Überrepräsentation einer Konformation in der kMDS hingewiesen.

Die Vielzahl neu beobachteter Konformationen in der BP-REMD hat ihren Ursprung in den extrahelikalen Konformationen. Da die Strukturen völlig solvatisiert sind existiert kein spezifisches Minimum.

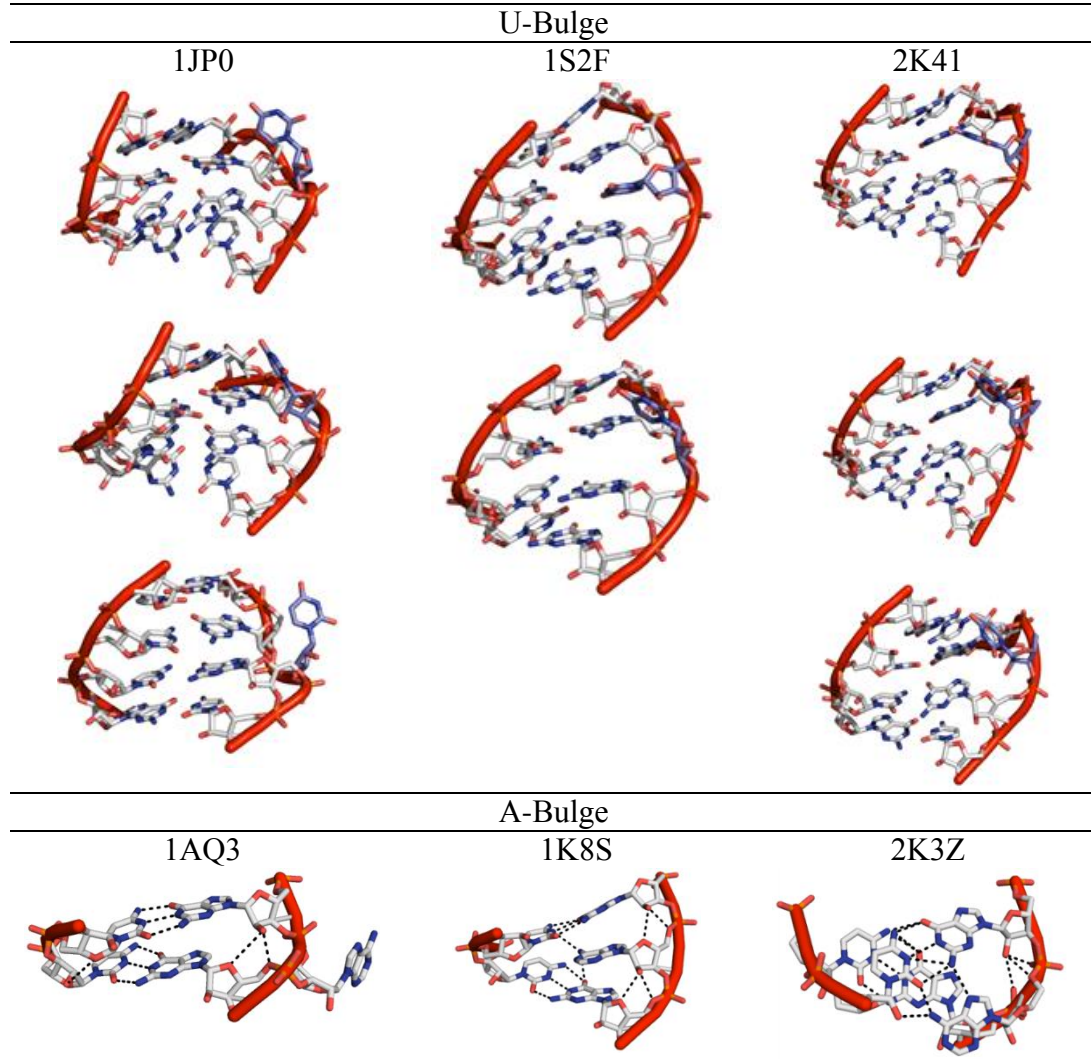


Abbildung 3.31 Experimentell ermittelte Konformationszustände des A- und U-Bulges. Die experimentell ermittelten Strukturen verdeutlichen die Konformationsvielfalt des A- und U-Bulges.

Im direkten Vergleich der Diederwinkel wird der Vorteil der BP-REMD Methode deutlich. Sowohl im ϵ/ζ als auch im α/γ Konformationsbereich werden mehr Konformationen ermittelt. In Abbildung 3.32 wird auch der Vorteil der Verwendung von Pseudo-Diederwinkeln deutlich. Intrahelikale Strukturen befinden sich in der η/θ -Darstellung bei etwa $150^\circ/240$ und nehmen nur einen spezifischen Bereich ein. Im Falle des U-Bulges ist dieser Bereich genauer definiert. So beobachtet man zwei Konformationszustände in der kMDS des 2K41-Modells. Die Destabilisierung des η/θ -Konformationsbereiches erlaubt es somit parallel mehrere Diederwinkel zu destabilisieren, sodass die Anzahl der BP-Potentiale reduziert wird. Der hier verwendete Pseudo-Diederwinkel ist daher

geeignet, globale Konformationsänderungen zu induzieren. Die Konformationszustände unterscheiden sich in ihren η/θ -Diederwinkel nur geringfügig. Die beiden Konformationen entsprechen den Konformationen in Abbildung 3.30. Die Konformationsbereiche der 2K3Z kMDS weisen zwei kleine Minima bei $130^\circ/250^\circ$ und $180^\circ/300^\circ$ auf. Wie aus Abbildung 3.33 hervorgeht, spiegeln sich die Konformationsänderungen im η/θ -Konformationsbereiches auch sehr gut sowohl im ε/ζ als auch im α/γ Konformationsbereich wieder. Besonders im α/γ -Konformationsbereich sind Übergänge und neue lokale Minima zu beobachten.

Die Analyse der experimentellen Strukturen bestätigen, dass die beobachteten Konformationen auch experimentell nachgewiesen werden können. Weiterhin ist bekannt, dass RNA-Kraftfelder intrahelikale Konformationen über-stabilisieren. Dieses Problem wird mit der BP-REMD umgangen, indem man diese Konformationen destabilisiert.

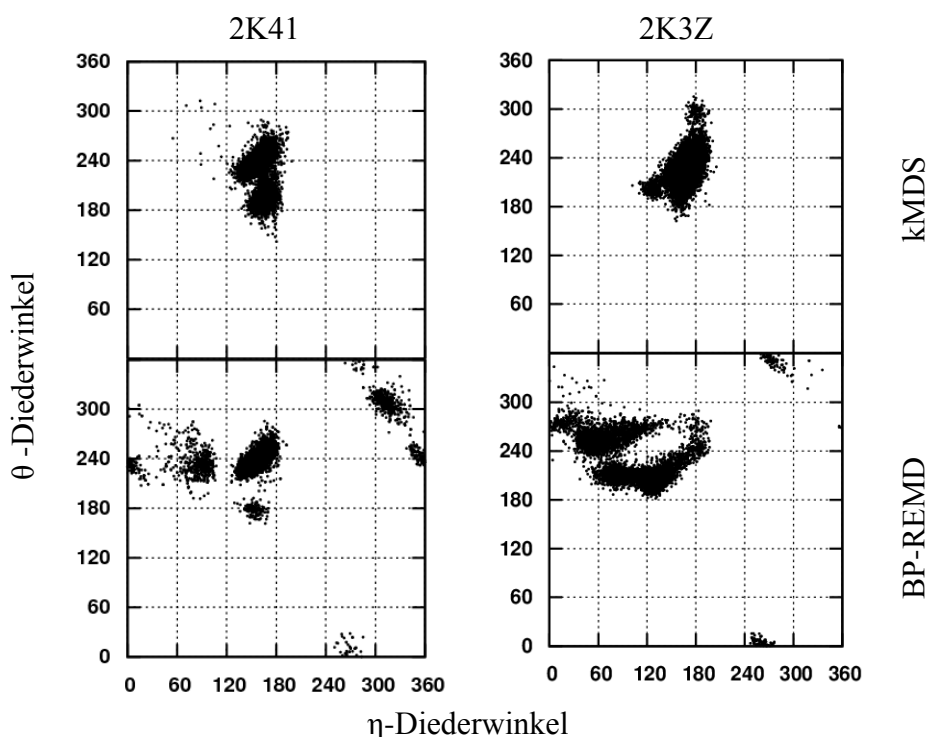


Abbildung 3.32 η/θ -Konformationszustände im Vergleich. Beim direkten Vergleich der Simulationmethoden fällt auf, dass in der kMDS ein dominanter Zustand bei $180^\circ/240^\circ$ in beiden Modellen vorhanden ist. Dieser Zustand wird in der BP-REMD destabilisiert, sodass neue Konformationen generiert werden und nach dem Metropolis-Kriterium austauschen können. Verglichen wird hierbei die kMDS mit dem Replika 0 der BP-REMD.

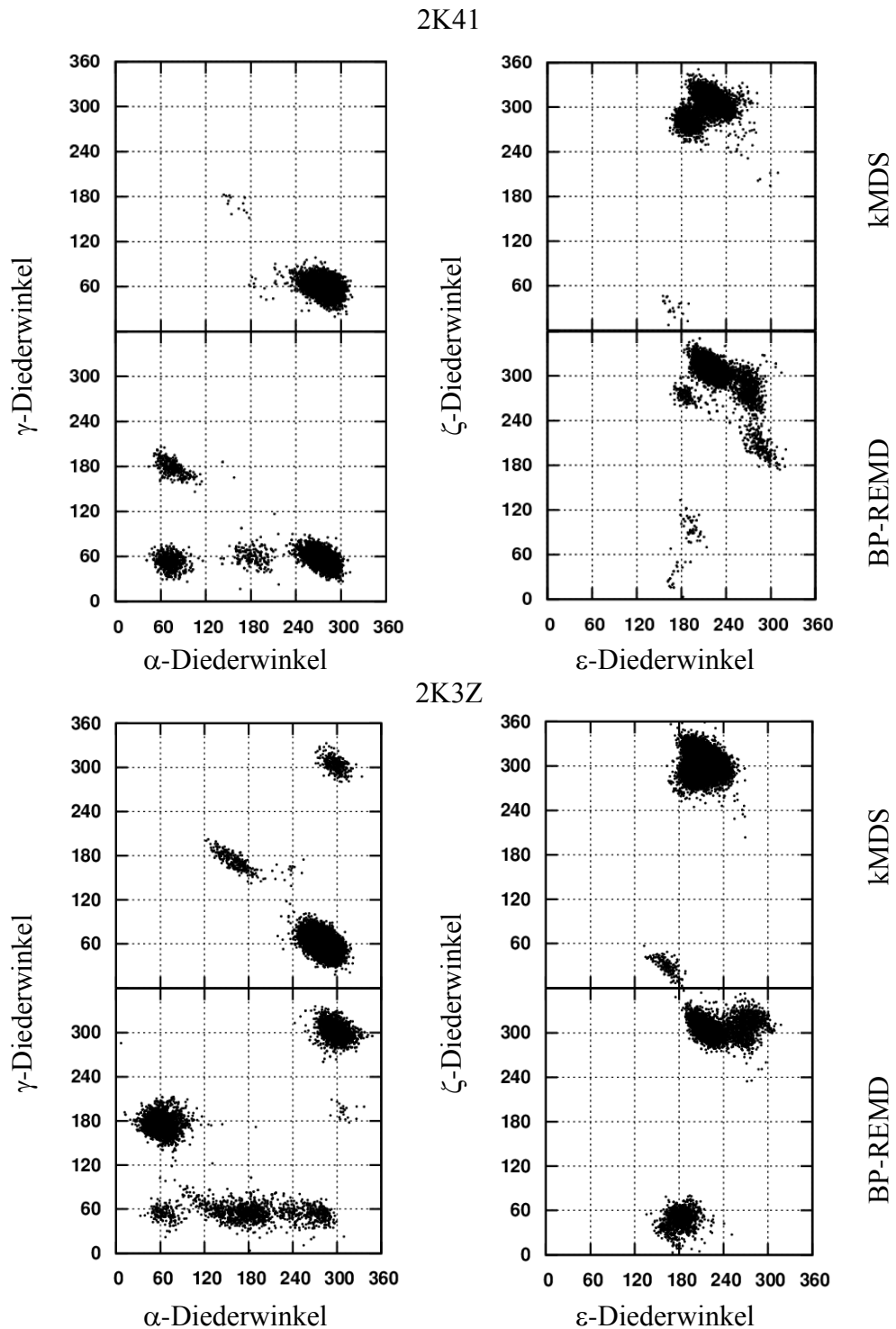
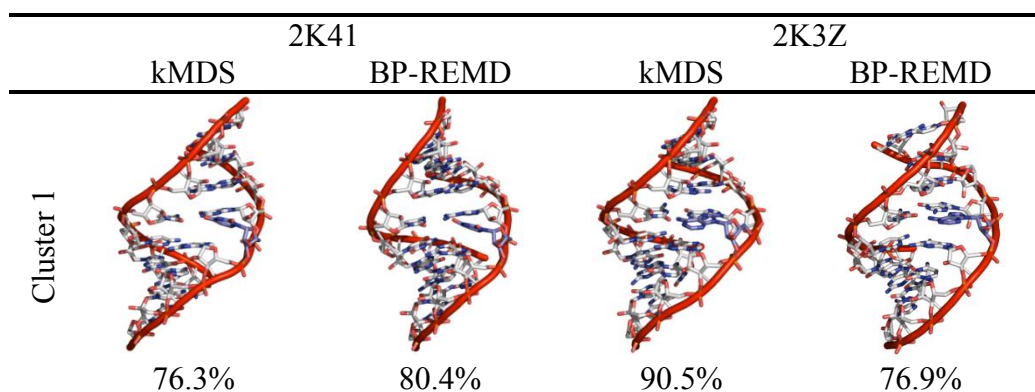


Abbildung 3.33 α/γ und ϵ/ζ Konformationszustände im Vergleich. Ein direkter Vergleich der Methoden zeigt den Vorteil der BP-REMD Methode. Sowohl im α/γ als auch ϵ/ζ Konformationsbereich werden neue Konformationen mit Hilfe der BP-REMD Methode ermittelt. Zusätzlich wird der Vorteil der Verwendung von Pseudo-Diederwinkel deutlich. Die Anzahl der zu destabilisierenden Zustände wird durch die Pseudo-Diederwinkel reduziert.

Eine Clusteranalyse der Trajektorien bestätigt das Ergebnis, dass durch die Verwendung der BP-REMD Methode neue Konformationen ermittelt werden. Zudem erhält man die relative Häufigkeit einer Konformation während der Simulation (siehe Abbildung 3.34). Hierbei wird besonders deutlich, dass Konformationen in kMDS überrepräsentiert sein können. In Abbildung 3.34 werden die relativen Häufigkeiten der kMDS und der BP-REMD miteinander verglichen. Die Strukturen sind hierbei nach ihrem RMSD-Wert sortiert, bzw. nach der Distanz der RMSD-Werte zur Startstruktur. Die Strukturen aus dem Cluster 2 der kMDS des 2K41 Modells entsprechen den Strukturen aus dem Cluster 3 der BP-REMD Simulation. Hierbei handelt es sich in beiden Simulationen um die zweit-häufigste Konformation. Interessanterweise unterscheiden sich jedoch die relativen Häufigkeiten der Cluster signifikant (kMDS 22.6% vs. BP-REMD 7.5%). Die weiteren Konformationen des 2K41-Modells werden in der kMDS entweder nicht ermittelt oder sind mit <1% vertreten. Dieses Ergebnis weist darauf hin, dass die kMDS nicht konvergiert ist. Die Beobachtungen der 2K41 Modelle lassen sich auch in den 2K3Z-Simulationen wiederfinden. Zudem wird an der 2K3Z Simulation deutlich, dass die jeweiligen Cluster der kMDS sich nur geringfügig unterscheiden.

Im Vergleich dazu werden in den BP-REMD sowohl intra- als auch extrahelikale Strukturen gefunden. Die unterschiedlichen Verteilungen und Konformationen des U- und A-Bulges lassen sich durch die Anzahl möglicher polarer Interaktionen zu den nächsten Nachbarn erklären (siehe Abbildung 3.35). Wie aus Abbildung 3.35 zu entnehmen ist, wird die Konformation Cluster 1 des A-Bulges durch drei Wasserstoffbrücken zwischen der Bulge Base und den WC-Paaren stabilisiert. Die U-Bulge Base hat im Vergleich hier zu nur eine mögliche Wasserstoffbrücke. Da es sich bei dem A-Bulge um eine sehr stabile Struktur handelt, werden keine Konformationsübergänge während der 350 ns kMDS beobachtet.



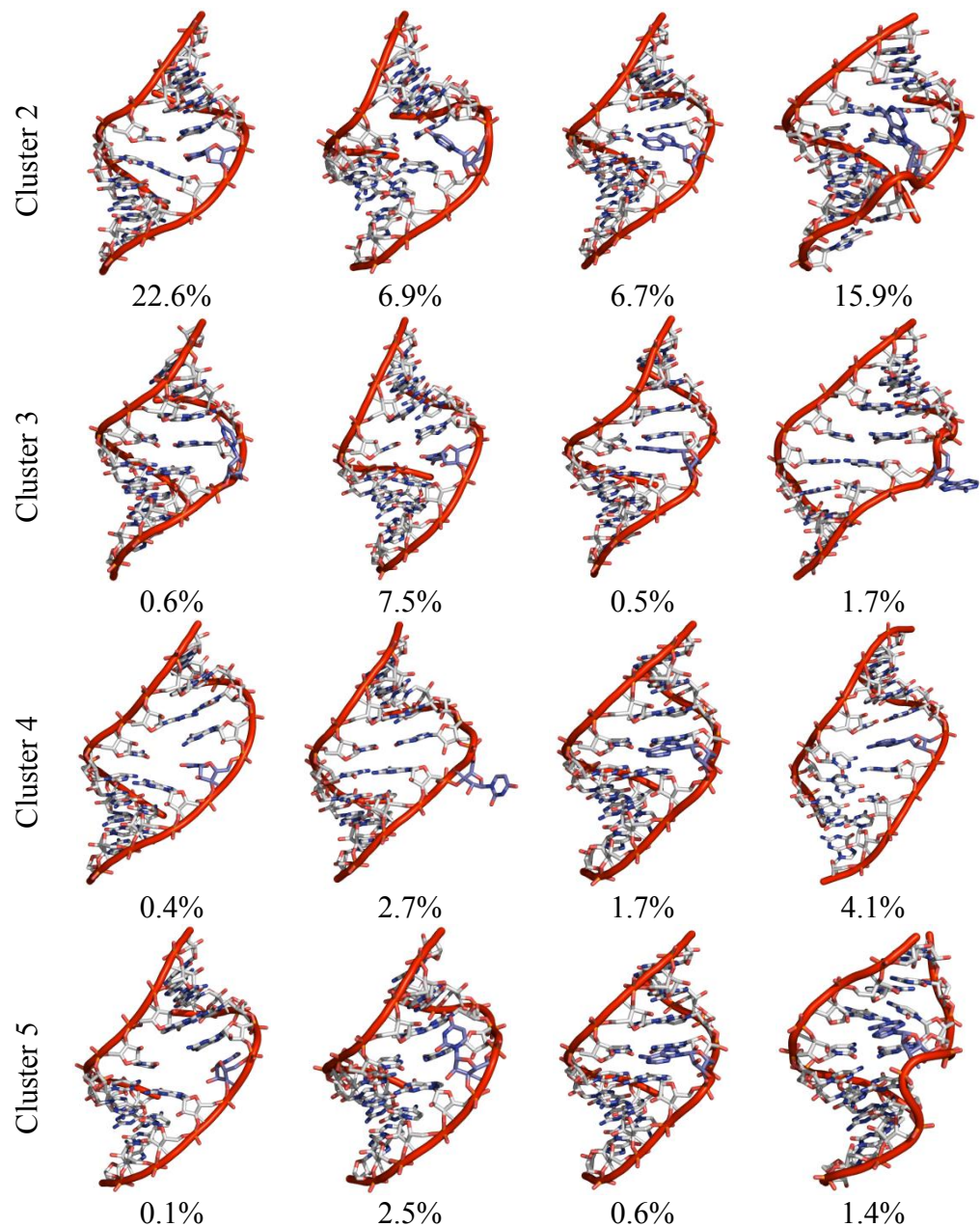


Abbildung 3.34 Clusteranalyse der kMDS und BP-REMD Trajektorie. Die Cluster der kMDS A-Bulge Simulation unterscheiden sich nur sehr geringfügig. In der BP-REMD Simulation hingegen werden unterschiedliche Konformationszustände ermittelt. Interessanterweise, können die in der BP-REMD Simulation ermittelten Zustände auch experimentell nachgewiesen werden (siehe Abbildung 3.31) Beim Vergleich des U-Bulges ist darauf hinzuweisen, dass in der kMDS keine extrahelikale Konformation ermittelt werden konnte. Zusätzlich ist die intrahelikale Konformation überrepräsentiert.

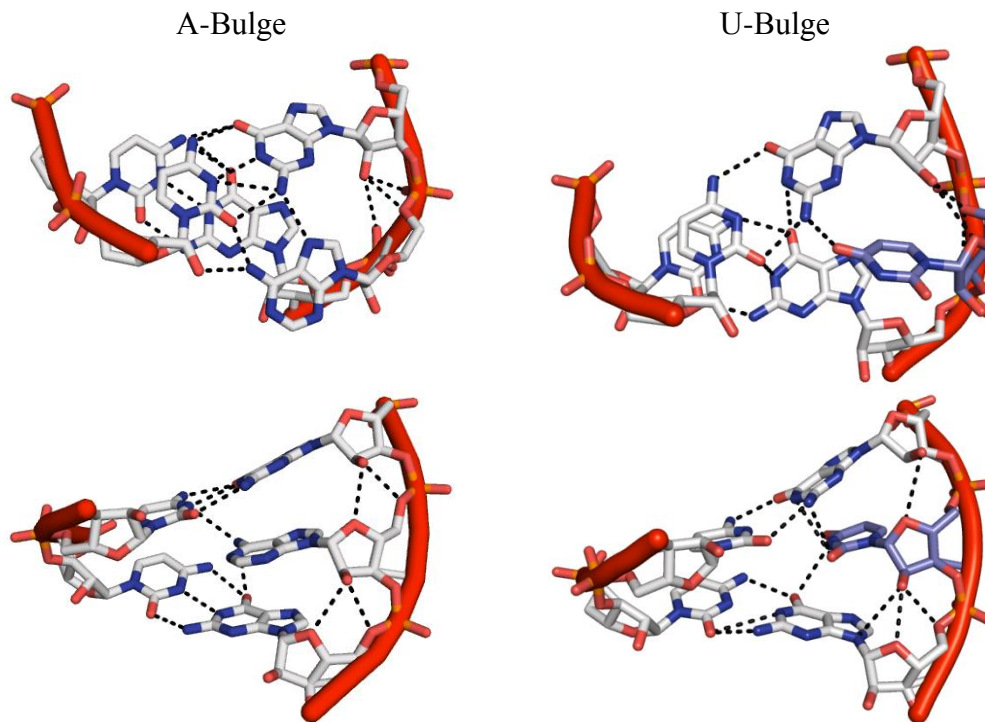


Abbildung 3.35 Vergleich des U- und A-Bulges. Die Startkonformationen (1. Zeile) zeigen, dass die Konformation des A-Bulges mehr mögliche Wasserstoffbrückenpartner hat in Vergleich zum U-Bulge. Die intrahelikale Konformation des U-Bulges weist hingegen mehr mögliche Wasserstoffbrückenbindungen auf, als der U-Bulge. Die intrahelikalen Konformationen werden vorwiegend durch Wasserstoffbrücken des Zucker-Phosphat-Rückgrats stabilisiert.

3.4.3 Zusammenfassung und Diskussion

Ziel dieser Arbeit war es anhand von MD-Simulationen Konformationszustände von RNA-Single-Bulge Strukturen und deren Verteilungen zu ermitteln. Zu diesem Zweck wurden kMDS und BP-REMD Simulationen durchgeführt. Untersucht wurde eine U und ein A-Bulge Struktur. Dabei wurde bei dem System darauf geachtet, dass der Sequenzkontext identisch ist. Zusammenfassend lassen sich folgende Ergebnisse festhalten:

- Mit Hilfe der BP-REMD Methode lassen sich RNA-Single Bulge Strukturen effizienter untersuchen als kMDS (siehe Abbildung 3.33)
- Für den verwendeten Sequenzkontext ist die Wahrscheinlichkeit einer intrahelikalen Konformation für einen U-Bulge höher als für einen A-Bulge (siehe Abbildung 3.34)
- Klassische MD Simulationen haben den Nachteil, dass sie über einen langen Zeitraum in einem lokalen Minimum verweilen (siehe Abbildung 3.29)
- Intrahelikale Konformationen werden über Wasserstoffbrücken des Zucker-Phosphat-Rückgrats stabilisiert. Die Zucker-Zucker Interaktion der Bulge-Base mit den nächsten Nachbarn spielt hierbei eine wichtige Rolle (siehe Abbildung 3.30)

Kapitel 4

4 Abschließende Diskussion und Ausblick

Der Fokus dieser Arbeit liegt zum einen in der erstmaligen Anwendung der BP-REMD Methode anhand von Pseudo-Diederwinkeln und zum anderen in der Charakterisierung von Nukleinsäure-Sekundärstrukturen mit Hilfe von etablierten Computersimulations-Methoden. Im Vordergrund stand neben der Analyse der Konformationsbereiche der untersuchten Nukleinsäuren auch eine vergleichende Bewertung der klassischen Molekulardynamik-Simulation und der BP-REMD Methode. Ziel war es hierbei die Vor- und Nachteile der BP-REMD Methode aufzuzeigen.

Die BP-REMD Methode basiert auf dem Prinzip der Destabilisierung von spezifischen Pseudo-Diederwinkeln entlang des Zucker-Phosphatrückgrats und wird in Abschnitt 2.4 vorgestellt. Anzumerken ist zudem, dass für die RNA-Simulationen unterschiedliche Pseudo-Diederwinkel gewählt wurden als für die DNA Simulationen. Grund für die Verwendung unterschiedlicher Pseudo-Diederwinkel ist die Fragestellung bzw. das Ziel der jeweiligen Simulationen. In den DNA Simulationen wurde ein Vergleich zwischen modifizierten Basen und der ursprünglichen Base untersucht. Es handelt sich um Unterschiede in der Feinstruktur der DNA und somit um lokale Konformationsänderungen. In den RNA-Bulge-Simulationen hingegen wurden globale Konformationsänderungen untersucht. Die in den DNA Simulationen verwendeten Pseudo-Diederwinkel sind nicht geeignet um globale Konformationsänderungen zu induzieren. Daher wurde für die RNA-Simulationen ein Pseudo-Diederwinkel gewählt von dem bereits bekannt ist, dass er geeignet ist, um unterschiedliche RNA Konformationen zu charakterisieren (204). Die Ergebnisse aus Abschnitt 3.1 und Abschnitt 3.2 zeigen deutlich, dass die Verwendung der BP-REMD Methode gute Resultate in der Beschreibung und Suche von Konformationen liefert. Zudem konnte gezeigt werden, dass mit der Pseudo-Diederwinkel-Darstellung parallel Konformationsübergänge unterschiedlicher Diederwinkel induziert werden können. Zusammenfassend lässt sich festhalten, dass mit Hilfe der BP-REMD

Methode sowohl für die 8oxoG Base, als auch für die O⁶-MeG Base alternative Konformation-Zustände im Vergleich zu Guanin ermittelt wurden, die in der Erkennung dieser Base eine ausschlaggebende Rolle spielen können.

Neben der Untersuchung von DNA-Schäden wurde auch erfolgreich gezeigt, dass die Verwendung der BP-REMD Methode geeignet ist um RNA-Bulge-Strukturen zu untersuchen (siehe Abschnitt 3.4). Besonderes Augenmerk gilt hierbei dem Vergleich zu experimentell ermittelten Strukturen. So konnten sowohl für A-Bulge, als auch für U-Bulge alle experimentell bestimmten Konformationen im gleichen Sequenzkontext nur mit Hilfe der BP-REMD Methode ermittelt werden. Diese Konformationen konnten in zeitaufwendigen kMDS nicht bestimmt werden (Simulationsdauer: 350 ns).

Neben der Anwendung und dem Vergleich der kMDS und der BP-REMD wurden Dangling-End Motive untersucht. Hierbei konnten die experimentell bestimmten thermodynamischen Parameter tendenziell reproduziert werden. Die Ermittlung der Zusammensetzung der Energiebeiträge lieferte zudem zusätzliche Hinweise für die Erklärung der experimentellen Ergebnisse.

In dieser Arbeit wurde gezeigt, dass die BP-REMD-Methode geeignet ist, effizienter als kMDS, den Konformationsraum von Nukleinsäuren zu untersuchen. Hierbei wurden die Nukleinsäuren jedoch ohne Interaktion mit anderen Biomolekülen oder chemischen Wirkstoffen untersucht. Die Anwendung der BP-REMD Methode für die Untersuchung von Nukleinsäure-Protein bzw. Nukleinsäure-Ligand-Komplexen kann sinnvolle Hinweise zur Interaktion dieser Biomoleküle geben. So könnte die Methode angewandt werden, um die Ergebnisse aus RNA-Protein-Docking Experimenten zu verbessern. Hierbei könnten vor allem die Interaktion der Bindestellen zwischen RNA und Protein optimiert und mit experimentellen Ergebnissen verglichen werden. Ein weitere interessantes Anwendung wäre der Vergleich zwischen der ungebundenen HIV-1 TAR RNA Struktur und der an verschiedene Liganden gebundenen Form. Hierbei könnte vor allem der Einfluss des Liganden auf die Flexibilität der Struktur bestimmt werden.

Zudem beschränkten sich die DNA-Schäden nur auf zwei unterschiedliche Basenmodifikationen. Auch hier kann die BP-REMD-Methode genutzt werden um systematisch unterschiedliche Basenmodifikationen, sowohl für DNA-Strukturen (z.B. 3-Ethyladenine, 5-Methylcytosin, O⁶-Ethylguanin) als auch für

RNA-Strukturen (z.B. N⁶-Methyladenin, N²-Methylguanin, 3-Methyluridin) zu untersuchen. Zudem kann die Untersuchung spezifischer auf Basenmodifikationen ausgelegt sein. In dieser Arbeit wurden die Nukleinsäuren aus unterschiedlichen Blickwinkeln analysiert (z.B. helikale Parameter, Diederwinkel, Zucker-Pucker und Wasserverteilung). Die Idee eines sequenzabhängigen Deskriptors scheint jedoch gerade bei Basenmodifikationen sinnvoll. Dieser Deskriptor könnte als Klassifizierer zwischen intakter und geschädigter DNA dienen. Ein neuronales Netz müsste hierzu mit Parametern für jeden Di- bzw. Tri-Nukleotid-Schritt trainiert werden. Diese Parameter können aus Simulationen gewonnen werden. Annahme wäre hierbei, dass im gleichen Sequenzkontext zur intakten DNA, jeder DNA-Schaden in einem spezifisch definierten Parameter Unterschiede zur intakten DNA aufweist.

Neben der Anwendung für Basenmodifikationen könnte die BP-REMD auch dazu genutzt werden um Faltungsprozesse von Nukleinsäuren genauer zu untersuchen oder Strukturverfeinerungen an schlecht aufgelösten experimentellen Strukturen vorzunehmen. Die Anwendungen in diesen Bereichen könnten wie folgt gestaltet werden:

- Ausgehend von einer linearen, modellierten Struktur (RNA oder DNA) könnten entlang des Zucker-Phosphatrückgrats Pseudo-Diederwinkel destabilisiert werden, die in der Lage sind globale Konformationsänderungen zu induzieren. Als Test-Systeme können RNA-Hairpin, DNA-Hairpin oder G-Quadruplexe dienen.
- Ausgehend von einer experimentell schlecht aufgelösten Struktur könnte die Methode genutzt werden, um Strukturverfeinerungen in Teilbereiche der Struktur (z.B. Loop-Strukturen) vorzunehmen. Hierzu wären lokale Konformationsänderungen in diesen Bereichen notwendig. Anwendungsgebiete könnten hierbei Strukturverfeinerungen von Ribosom-Strukturen sein.

Danksagung

Mein aufrichtiger Dank gilt allen Personen, die mir die Anfertigung und Fertigstellung dieser Arbeit ermöglicht haben. Ein besonderer Dank gilt Herrn Prof. Zacharias für die Überlassung des interessanten Themas und das Vertrauen in meine Person und meine Arbeit. Ein besonderer Dank gilt ihm an dieser Stelle noch einmal für die ständige Diskussionsbereitschaft und Ideen-Vielfalt, während meiner Dissertation. Die sehr gute zwischenmenschliche Beziehung zu ihm und dem gesamten Arbeitskreis erfüllten meinen Alltag mit Freude und Spaß.

Dr. Sebastian Schneider, Sina Kazemi und insbesondere Mehmet Kara möchte ich herzlich für das Korrekturlesen und für konstruktive Kritik danken.

Ein Dankeschön geht auch an die Prüfungskommission, Frau Prof. Antes für die Begutachtung meiner Arbeit und Herrn Prof. Torsten Hugel für den Vorsitz in meiner Prüfung.

Ich möchte diese Arbeit meinen Eltern und meiner Familie widmen, denen ich sehr dankbar bin für ihre Geduld und Unterstützung während meines Studiums und meiner Dissertation.

Publikationsliste

Die folgenden Publikationen sind im Lauf dieser Arbeit entstanden:

Kara M. and Zacharias M.

Stabilization of duplex DNA and RNA by dangling ends studied by free energy simulations. *Biopolymers.*, 2013 (doi:10.1002/bip.22398)

Drsata, Tomas; Kara, Mahmut; Zacharias, Martin; Lankas, Filip

Effect of 8-Oxoguanine on DNA Structure and Deformability, *J. Phys. Chem. B*, 2013, 117(39), 11617–11622

Kara M. and Zacharias M.

Influence of 8-oxoguanosine on the fine structure of DNA studied with biasing-potential replica exchange simulations. *Biophysical Journal*, 2013, 104(5):1089-97

Kara M. and Zacharias M.

Theoretical Studies of Nucleic Acids Folding. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2013 (doi:10.1002/wcms.1146)

Sushil Kumar Mishra, Mahmut Kara , Martin Zacharias, and Jaroslav Koča

Enhanced Conformational Sampling of Carbohydrates by Hamiltonian Replica Exchange Simulation, akzeptiert in *Glycobiology*

Kara M. and Zacharias M.

Conformation stability of RNA Single Base Bulges depends on nearest-neighbor effects: A computational study of RNA-Bulges, in Vorbereitung

Kara, Mahmut; Drsata, Tomas; Lankas, Filip; Zacharias, Martin

Comparative study of O6-MeG:C, O6-MeG:T and G:C Basepairs, in Vorbereitung

Reinhard Zelm, Mahmut Kara, Martin Zacharias and Dietmar Blohm

Influence of internal structure and sequence of oligonucleotides on the hybridisation efficiency to surface bound DNA, in Vorbereitung

Anhang

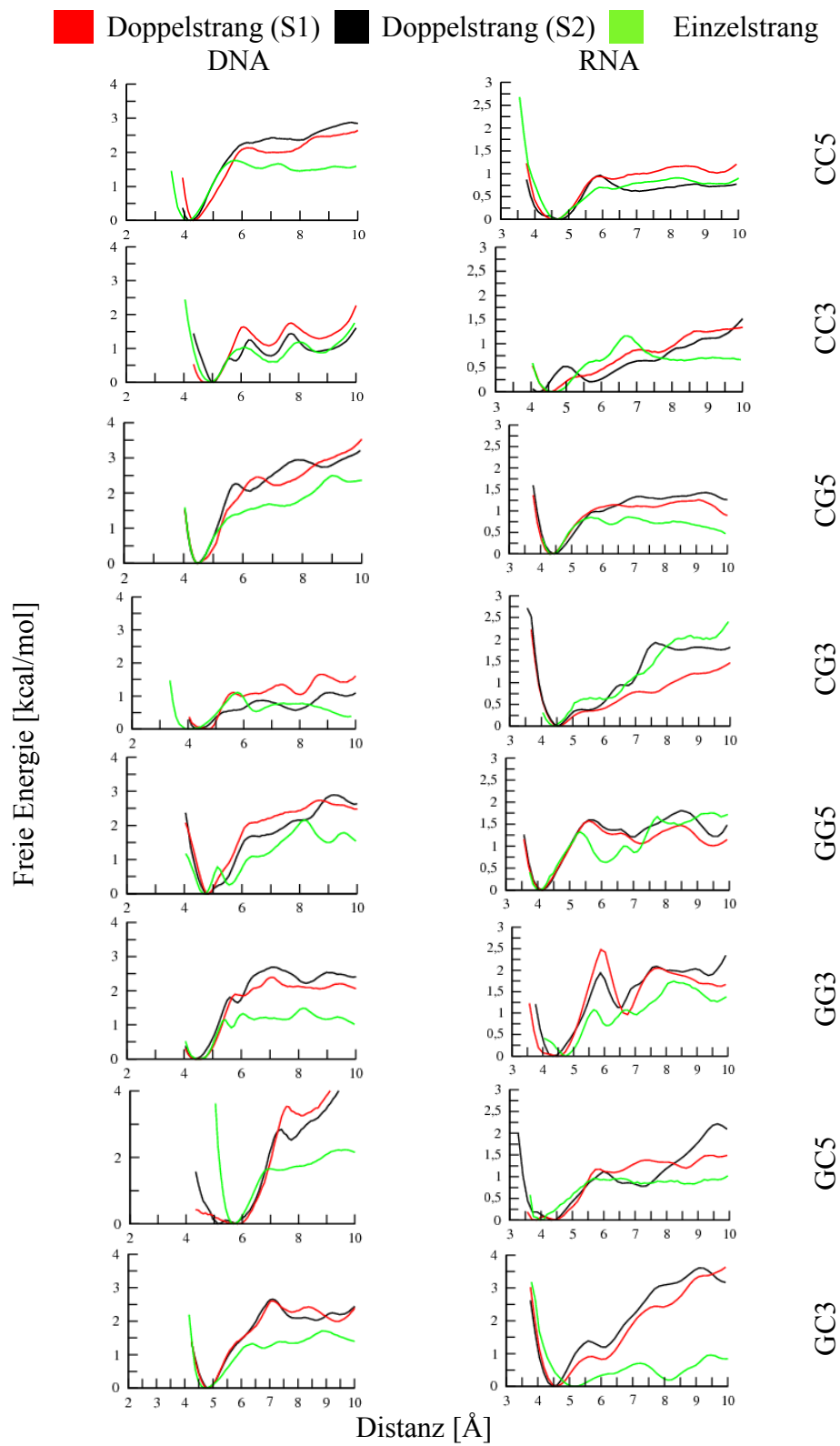


Abbildung A. 1 Vergleich zwischen allen RNA und DNA Energieprofilen.

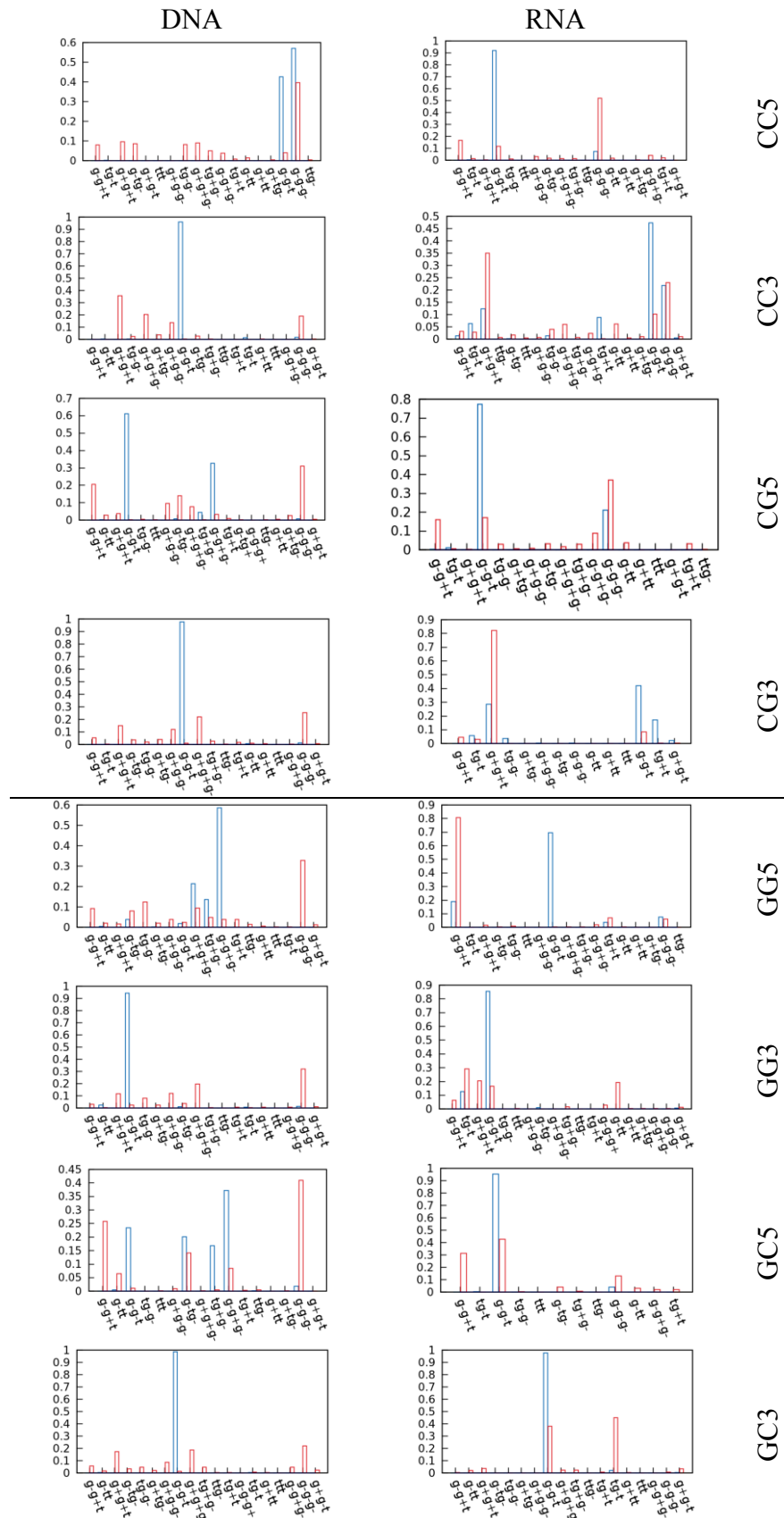


Abbildung A. 2 Verteilung aller Phosphatrückgrat Diederwinkel

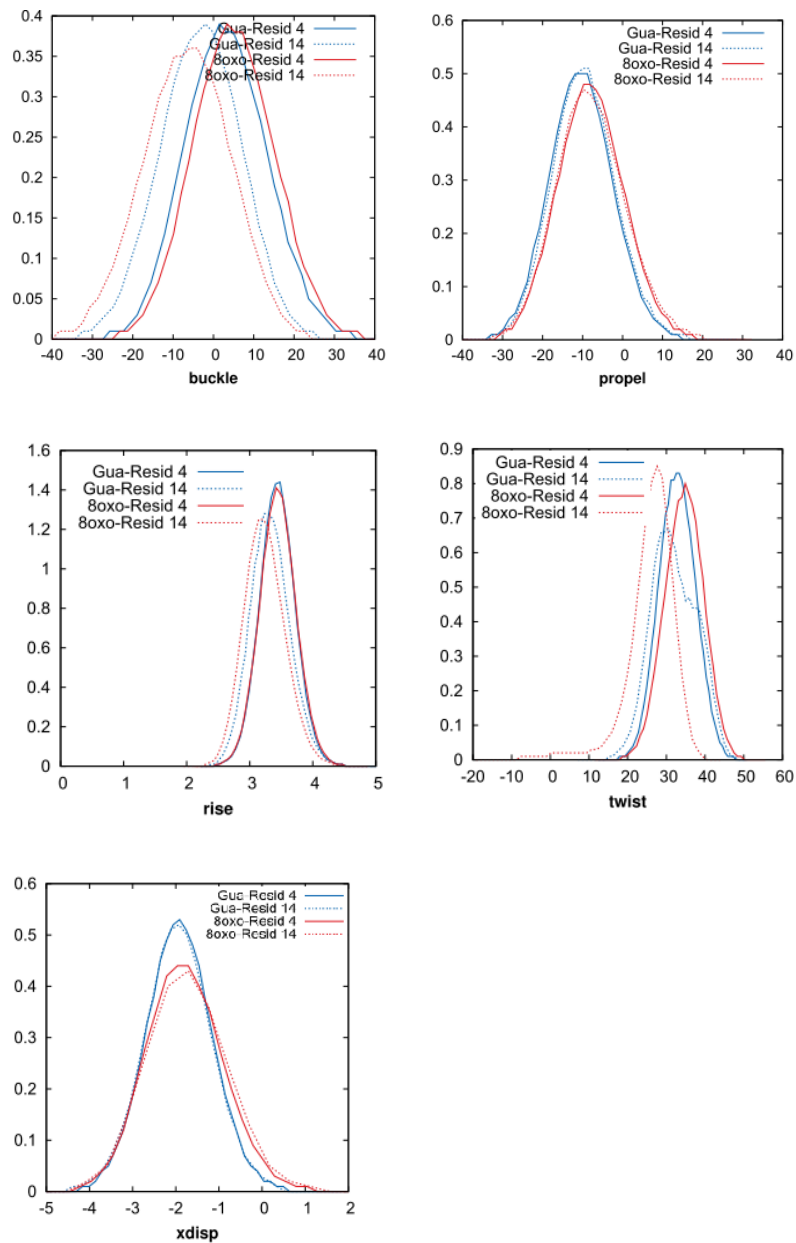


Abbildung A. 3 Vergleich der helikalen Parameter zwischen Guanin und 8-oxoguanin

Literaturverzeichnis

1. Watson,J.D. and Crick,F.H.C. (1953) Molecular structure of nucleic acids. *Nature*, **171**, 737–738.
2. Crick,F. (1970) Central Dogma of Molecular Biology. *Nature*, **227**, 561–563.
3. Schneider,B. and Berman,H.M. (2006) Basics of Nucleic Acid Structure. In Šponer,J., Lankas,F. (eds), *Computational Studies of RNA and DNA*. Springer-Verlag, pp. 1–44.
4. Donohue,J. and Trueblood,K.N. (1960) Base pairing in DNA. *J Mol Biol*, **2**, 363–371.
5. Moran,L.A., Horton,H.R., Scrimgeour,K.G., Perry,M.D. and Marc D. Perry (2012) Principles of Biochemistry Fifth Edition Pearson Education, Inc.
6. Dickerson,R.E., Drew,H.R., Conner,B.N., Wing,R.M., Fratini,A. V and Kopka,M.L. (1982) The anatomy of A-, B-, and Z-DNA. *Science*, **216**, 475–85.
7. Finley,J.B. and Luo,M. (1998) X-ray crystal structures of half the human papilloma virus E2 binding site: d(GACCGCGGTC). *Nucleic Acids Res.*, **26**, 5719–27.
8. Rozenberg,H., Rabinovich,D., Frolow,F., Hegde,R.S. and Shakked,Z. (1998) Structural code for DNA recognition revealed in crystal structures of papillomavirus E2-DNA targets. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 15194–9.
9. Ban,C., Ramakrishnan,B. and Sundaralingam,M. (1996) Crystal structure of the self-complementary 5'-purine start decamer d(GCGCGCGCGC) in the Z-DNA conformation. I. *Biophys. J.*, **71**, 1215–21.
10. Brion,P. and Westhof,E. (1997) Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct.*, **26**, 113–37.
11. Macke,T.J. (2001) RNAMotif, an RNA secondary structure definition and search algorithm. *Nucleic Acids Res.*, **29**, 4724–4735.
12. Zuker,M. and Sankoff,D. (1984) RNA secondary structures and their prediction. *Bull Math Biol*, **46**, 591–621.
13. Williamson,J.R. (2000) Induced fit in RNA-protein recognition. *Nat. Struct. Biol.*, **7**, 834–7.

14. Xayaphoummine, a, Bucher, T. and Isambert, H. (2005) Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Res.*, **33**, W605–10.
15. Markham, N.R. and Zuker, M. (2008) UNAFold: software for nucleic acid folding and hybridization. *Methods Mol. Biol.*, **453**, 3–31.
16. Jonikas, M. a, Radmer, R.J., Laederach, A., Das, R., Pearlman, S., Herschlag, D. and Altman, R.B. (2009) Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters. *RNA*, **15**, 189–99.
17. Moore, P.B. (1999) Structural motifs in RNA. *Annu. Rev. Biochem.*, **68**, 287–300.
18. Barthel, A. and Zacharias, M. (2006) Conformational transitions in RNA single uridine and adenosine bulge structures: a molecular dynamics free energy simulation study. *Biophys. J.*, **90**, 2450–62.
19. Leontis, N.B. and Westhof, E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA*, **7**, 499–512.
20. Hermann, T. and Patel, D.J. (2000) RNA bulges as architectural and recognition motifs. *Structure*, **8**, R47–54.
21. Hermann, T. and Patel, D.J. (2000) RNA bulges as architectural and recognition motifs. *Structure*, **8**, 47–54.
22. Sandbrink, J., Ossipov, D., Aström, H. and Strömberg, R. (2005) Investigation of potential RNA bulge stabilizing elements. *J. Mol. Recognit.*, **18**, 318–26.
23. Kumar, P., Lehmann, J. and Libchaber, A. (2012) Kinetics of Bulge Bases in Small RNAs and the Effect of Pressure on It. *PLoS One*, **7**, 1–8.
24. Barthel, A. and Zacharias, M. (2006) Conformational transitions in RNA single uridine and adenosine bulge structures: a molecular dynamics free energy simulation study. *Biophys. J.*, **90**, 2450–62.
25. Zacharias, M. and Sklenar, H. (1997) Analysis of the Stability of Looped-Out and Stacked-In Conformations of an Adenine Bulge in DNA Using a Continuum Model for Solvent and Ions. *Biophys J*, **73**, 2990–3003.
26. Sudarsanakumar, C., Xiong, Y. and Sundaralingam, M. (2000) Crystal structure of an adenine bulge in the RNA chain of a DNA.RNA hybrid, d(CTCCTCTTC).r(gaagagagag). *J. Mol. Biol.*, **299**, 103–12.
27. Ye, X., Kumar, R.A. and Patel, D.J. (1995) Molecular recognition in the bovine immunodeficiency virus Tat peptide-TAR RNA complex. *Chem. Biol.*, **2**, 827–40.

28. Harrich,D., Mavankal,G., Mette-Snider,A. and Gaynor,R.B. (1995) Human immunodeficiency virus type 1 TAR element revertant viruses define RNA structures required for efficient viral gene expression and replication. *J. Virol.*, **69**, 4906–13.
29. Watson,K. and Edwards,R.J. (1999) HIV-1-trans-activating (Tat) protein: both a target and a tool in therapeutic approaches. *Biochem. Pharmacol.*, **58**, 1521–8.
30. Altman,S. (2007) A view of RNase P. *Mol. Biosyst.*, **3**, 604–7.
31. Kazantsev,A. V and Pace,N.R. (2006) Bacterial RNase P: a new view of an ancient enzyme. *Nat. Rev. Microbiol.*, **4**, 729–40.
32. Schmitz,M. (2004) Change of RNase P RNA function by single base mutation correlates with perturbation of metal ion binding in P4 as determined by NMR spectroscopy. *Nucleic Acids Res.*, **32**, 6358–66.
33. White,S. a and Draper,D.E. (1989) Effects of single-base bulges on intercalator binding to small RNA and DNA hairpins and a ribosomal RNA fragment. *Biochemistry*, **28**, 1892–7.
34. Peabody,D.S. (1993) The RNA binding site of bacteriophage MS2 coat protein. *EMBO J.*, **12**, 595–600.
35. Valegård,K., Murray,J.B., Stonehouse,N.J., van den Worm,S., Stockley,P.G. and Liljas,L. (1997) The three-dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence-specific protein-RNA interactions. *J. Mol. Biol.*, **270**, 724–38.
36. Lesnik,E. a and Freier,S.M. (1995) Relative thermodynamic stability of DNA, RNA, and DNA:RNA hybrid duplexes: relationship with base composition and structure. *Biochemistry*, **34**, 10807–15.
37. Freier,S.M., Alkema,D., Sinclair,A., Neilson,T. and Turner,D.H. (1985) Contributions of dangling end stacking and terminal base-pair formation to the stabilities of XGGCCp, XCCGGp, XGGCCYp, and XCCGGYp helices. *Biochemistry*, **24**, 4533–9.
38. Freier,S.M., Kierzek,R., Jaeger,J. a, Sugimoto,N., Caruthers,M.H., Neilson,T. and Turner,D.H. (1986) Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci. U. S. A.*, **83**, 9373–7.
39. SantaLucia,J. and Hicks,D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, **33**, 415–40.

40. Mathews,D.H., Burkard,M.E., Freier,S.M., Wyatt,J.R. and Turner,D.H. (1999) Predicting oligonucleotide affinity to nucleic acid targets. *RNA*, **5**, 1458–1469.
41. Williams,J.C., Case-Green,S.C., Mir,K.U. and Southern,E.M. (1994) Studies of oligonucleotide interactions by hybridisation to arrays: the influence of dangling ends on duplex yield. *Nucleic Acids Res.*, **22**, 1365–1367.
42. Ma,J.-B., Ye,K. and Patel,D.J. (2004) Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature*, **429**, 318–22.
43. O’Toole,A.S., Miller,S. and Serra,M.J. (2005) Stability of 3’ double nucleotide overhangs that model the 3’ ends of siRNA. *RNA*, **11**, 512–516.
44. Miller,S., Jones,L.E., Giovannitti,K., Piper,D. and Serra,M.J. (2008) Thermodynamic analysis of 5’ and 3’ single- and 3’ double-nucleotide overhangs neighboring wobble terminal base pairs. *Nucleic Acids Res.*, **36**, 5652–9.
45. Bommarito,S., Peyret,N. and SantaLucia,J. (2000) Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res.*, **28**, 1929–34.
46. Isaksson,J. and Chattopadhyaya,J. (2005) A uniform mechanism correlating dangling-end stabilization and stacking geometry. *Biochemistry*, **44**, 5390–401.
47. Liu,J.D., Zhao,L. and Xia,T. (2008) The dynamic structural basis of differential enhancement of conformational stability by 5’- and 3’-dangling ends in RNA. *Biochemistry*, **47**, 5962–75.
48. Petersheim,M. and Turner,D.H. (1983) Proton magnetic resonance melting studies of CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry*, **22**, 269–77.
49. Isaksson,J., Acharya,S., Barman,J., Cheruku,P. and Chattopadhyaya,J. (2004) Single-stranded adenine-rich DNA and RNA retain structural characteristics of their respective double-stranded conformations and show directional differences in stacking pattern. *Biochemistry*, **43**, 15996–6010.
50. Quartin,R.S. and Wetmur,J.G. (1989) Effect of ionic strength on the hybridization of oligodeoxynucleotides with reduced charge due to methylphosphonate linkages to unmodified oligodeoxynucleotides containing the complementary sequence. *Biochemistry*, **28**, 1040–7.
51. SantaLucia,J. and Hicks,D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, **33**, 415–40.

-
52. Guckian, K.M., Schweitzer, B. a, Ren, R.X.-F., Sheils, C.J., Tahmassebi, D.C. and Kool, E.T. (2000) Factors Contributing to Aromatic Stacking in Water: Evaluation in the Context of DNA. *J. Am. Chem. Soc.*, **122**, 2213–2222.
53. Guckian, K.M., Schweitzer, B. a, Ren, R.X.-F., Sheils, C.J., Paris, P.L., Tahmassebi, D.C. and Kool, E.T. (1996) Experimental Measurement of Aromatic Stacking Affinities in the Context of Duplex DNA. *J. Am. Chem. Soc.*, **118**, 8182–8183.
54. Lindahl, T. (1999) Quality Control by DNA Repair. *Science (80-.)*, **286**, 1897–1905.
55. David, S.S., O’Shea, V.L. and Kundu, S. (2007) Base-excision repair of oxidative DNA damage. *Nature*, **447**, 941–50.
56. Fu, D., Calvo, J. a and Samson, L.D. (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat. Rev. Cancer*, **12**, 104–20.
57. Seeberg, E., Eide, L. and Bjørås, M. (1995) The base excision repair pathway. *Trends Biochem. Sci.*, **20**, 391–7.
58. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–15.
59. Bruner, S.D., Norman, D.P. and Verdine, G.L. (2000) Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature*, **403**, 859–66.
60. Tubbs, J.L., Pegg, A.E. and Tainer, J.A. (2008) DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O6 -alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA Repair (Amst.)*, **6**, 1100–1115.
61. Roberts, R.J. and Cheng, X. (1998) Base flipping. *Annu. Rev. Biochem.*, **67**, 181–98.
62. Lindahl, T. and Wood, R.D. (1999) Quality control by DNA repair. *Science*, **286**, 1897–905.
63. Carell, T., Burgdorf, L.T., Kundu, L.M. and Cichon, M. (2001) The mechanism of action of DNA photolyases. *Curr. Opin. Chem. Biol.*, **5**, 491–8.
64. Friedman, J.I. and Stivers, J.T. (2010) Detection of damaged DNA bases by DNA glycosylase enzymes. *Biochemistry*, **49**, 4957–67.
65. Miller, J.H., Straatsma, T.P., Kennedy, M.A. and Uni, S. (2003) 8-Oxoguanine Enhances Bending of DNA that Favors Binding to Glycosylases. *J Am Chem Soc*, **115**, 6331–6.

-
66. Qi, Y., Spong, M.C., Nam, K., Banerjee, A., Jiralerspong, S., Karplus, M. and Verdine, G.L. (2009) Encounter and extrusion of an intrahelical lesion by a DNA repair enzyme. *Nature*, **462**, 762–766.
67. Batty, D.P. and Wood, R.D. (2000) Damage recognition in nucleotide excision repair of DNA. *Gene*, **241**, 193–204.
68. Sancar, A. (1996) DNA excision repair. *Annu. Rev. Biochem.*, **65**, 43–81.
69. De Laat, W.L., Jaspers, N.G. and Hoeijmakers, J.H. (1999) Molecular mechanism of nucleotide excision repair. *Genes Dev.*, **13**, 768–85.
70. Sugawara, K., Okamoto, T., Shimizu, Y., Masutani, C., Iwai, S. and Hanaoka, F. (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev.*, **15**, 507–21.
71. Le Page, F., Kwok, E.E., Avrutskaya, A., Gentil, A., Leadon, S.A., Sarasin, A. and Cooper, P.K. (2000) Transcription-coupled repair of 8-oxoguanine: requirement for XPG, TFIIH, and CSB and implications for Cockayne syndrome. *Cell*, **101**, 159–71.
72. Delaney, S., Jarem, D. a, Volle, C.B. and Yennie, C.J. (2012) Chemical and biological consequences of oxidatively damaged guanine in DNA. *Free Radic. Res.*, **46**, 420–41.
73. Singh, S.K., Szulik, M.W., Ganguly, M., Khutsishvili, I., Stone, M.P., Marky, L. a and Gold, B. (2011) Characterization of DNA with an 8-oxoguanine modification. *Nucleic Acids Res.*, **39**, 6789–801.
74. Cheng, X., Kelso, C., Hornak, V., de los Santos, C., Grollman, A.P. and Simmerling, C. (2005) Dynamic behavior of DNA base pairs containing 8-oxoguanine. *J. Am. Chem. Soc.*, **127**, 13906–18.
75. Cheng, K.C., Cahill, D.S., Kasai, H., Nishimura, S. and Loeb, L. a (1992) 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions. *J. Biol. Chem.*, **267**, 166–72.
76. Shibutani, S., Takeshita, M. and Grollman, A.P. (1991) Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature*, **349**, 431–4.
77. Kirpota, O.O., Endutkin, A. V, Ponomarenko, M.P., Ponomarenko, P.M., Zharkov, D.O. and Nevinsky, G. a (2011) Thermodynamic and kinetic basis for recognition and repair of 8-oxoguanine in DNA by human 8-oxoguanine-DNA glycosylase. *Nucleic Acids Res.*, **39**, 4836–50.

-
78. Sung,R.-J., Zhang,M., Qi,Y. and Verdine,G.L. (2012) Sequence-dependent structural variation in DNA undergoing intrahelical inspection by the DNA glycosylase MutM. *J. Biol. Chem.*, **287**, 18044–54.
79. Pinak,M. (2003) 8-oxoguanine lesioned B-DNA molecule complexed with repair enzyme hOGG1: a molecular dynamics study. *J. Comput. Chem.*, **24**, 898–907.
80. Barone,F., Lankas,F., Spackova,N., Sponer,J., Karran,P., Bignami,M. and Mazzei,F. (2005) Structural and dynamic effects of single 7-hydro-8-oxoguanine bases located in a frameshift target DNA sequence. *Biophys. Chem.*, **118**, 31–41.
81. Naômé,A., Schyman,P., Laaksonen,A. and Vercauteren,D.P. (2010) Molecular dynamics simulation of 8-oxoguanine containing DNA fragments reveals altered hydration and ion binding patterns. *J Phys Chem B*, **114**, 4789–801.
82. Plum,G.E., Grollman,A.P., Johnson,F. and Breslauer,K.J. (1995) Influence of the oxidatively damaged adduct 8-oxodeoxyguanosine on the conformation, energetics, and thermodynamic stability of a DNA duplex. *Biochemistry*, **34**, 16148–60.
83. Mackay,W.J., Han,S. and Samson,L.D. (1994) DNA alkylation repair limits spontaneous base substitution mutations in Escherichia coli. *J. Bacteriol.*, **176**, 3224–30.
84. Casorelli,I., Russo,M.T. and Bignami,M. (2008) Role of mismatch repair and MGMT in response to anticancer therapies. *Anticancer. Agents Med. Chem.*, **8**, 368–80.
85. Sedgwick,B. (2004) Repairing DNA-methylation damage. *Nat. Rev. Mol. Cell Biol.*, **5**, 148–57.
86. Murray,V. (1987) O6-Methylguanine specifically induces AT→GC transition mutations. *Mut Res Lett*, **190**, 267–70.
87. Pegg,A.E., Swenn,K., Chae,M.Y., Dolan,M.E. and Moschel,R.C. (1995) Increased killing of prostate, breast, colon, and lung tumor cells by the combination of inactivators of O6-alkylguanine-DNA alkyltransferase and N,N'-bis(2-chloroethyl)-N-nitrosourea. *Biochem. Pharmacol.*, **50**, 1141–8.
88. Hickman,M.J. and Samson,L.D. (2004) Apoptotic signaling in response to a single type of DNA lesion, O(6)-methylguanine. *Mol. Cell*, **14**, 105–16.
89. Ochs,K. and Kaina,B. (2000) Apoptosis induced by DNA damage O6-methylguanine is Bcl-2 and caspase-9/3 regulated and Fas/caspase-8 independent. *Cancer Res.*, **60**, 5815–24.

-
90. Kaina,B., Christmann,M., Naumann,S. and Roos,W.P. (2007) MGMT: key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents. *DNA Repair (Amst)*, **6**, 1079–99.
 91. Kalnik,M.W., Li,B.F., Swann,P.F. and Patel,D.J. (1989) O6-ethylguanine carcinogenic lesions in DNA: an NMR study of O6etG.C pairing in dodecanucleotide duplexes. *Biochemistry*, **28**, 6182–92.
 92. Ginell,S.L., Kuzmich,S., Jones,R.A. and Berman,H.M. (1990) Crystal and molecular structure of a DNA duplex containing the carcinogenic lesion O6-methylguanine. *Biochemistry*, **29**, 10461–5.
 93. Vojtechovsky,J., Eaton,M.D., Gaffney,B., Jones,R. and Berman,H.M. (1995) Structure of a new crystal form of a DNA dodecamer containing T.(O6Me)G base pairs. *Biochemistry*, **34**, 16632–40.
 94. Leonard,G.A., Thomson,J., Watson,W.P. and Brown,T. (1990) High-resolution structure of a mutagenic lesion in DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **87**, 9573–6.
 95. Patel,D.J., Shapiro,L., Kozlowski,S.A., Gaffney,B.L. and Jones,R.A. (1986) Structural studies of the O6meG.T interaction in the d(C-G-T-G-A-A-T-T-C-O6meG-C-G) duplex. *Biochemistry*, **25**, 1036–42.
 96. Yu,B., Edstrom,W.C., Benach,J., Hamuro,Y., Weber,P.C., Gibney,B.R. and Hunt,J.F. (2006) Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. *Nature*, **439**, 879–84.
 97. Eichman,B.F., O'Rourke,E.J., Radicella,J.P. and Ellenberger,T. (2003) Crystal structures of 3-methyladenine DNA glycosylase MagIII and the recognition of alkylated bases. *EMBO J.*, **22**, 4898–909.
 98. Sedgwick,B., Bates,P.A., Paik,J., Jacobs,S.C. and Lindahl,T. (2007) Repair of alkylated DNA: recent advances. *DNA Repair (Amst)*, **6**, 429–42.
 99. Duguid,E.M., Rice,P. a and He,C. (2005) The structure of the human AGT protein bound to DNA and its implications for damage detection. *J. Mol. Biol.*, **350**, 657–66.
 100. Nishikori,S. (2004) Mutational Effects on O6-Methylguanine-DNA Methyltransferase from Hyperthermophile: Contribution of Ion-Pair Network to Protein Thermostability. *J Biochem*, **135**, 525–532.
 101. Lopez,S., Margison,G.P., McElhinney,R.S., Cordeiro,A., McMurry,T.B.H. and Rozas,I. (2011) Towards more specific O6-methylguanine-DNA methyltransferase (MGMT) inactivators. *Bioorg. Med. Chem.*, **19**, 1658–65.

-
102. Pegg, A.E. (2000) Repair of O(6)-alkylguanine by alkyltransferases. *Mutat. Res.*, **462**, 83–100.
103. Ginell, S.L., Vojtechovsky, J., Gaffney, B., Jones, R. and Berman, H.M. (1994) Crystal structure of a mispaired dodecamer, d(CGAGAATTC(O6Me)GCG)₂, containing a carcinogenic O6-methylguanine. *Biochemistry*, **33**, 3487–93.
104. Gao, Y.G., Sriram, M., Denny, W.A. and Wang, A.H. (1993) Minor groove binding of SN6999 to an alkylated DNA: molecular structure of d(CGC[e6G]AATTCGCG)-SN6999 complex. *Biochemistry*, **32**, 9639–48.
105. Calladine, C.R., Drew, H.R., Luisi, B.F. and Travers, A.A. (2004) *Understanding DNA Third Edition* Elsevier Academic Press.
106. Neidle, S. (2009) The structures of quadruplex nucleic acids and their drug complexes. *Curr. Opin. Struct. Biol.*, **19**, 239–50.
107. Zacharias M. (2003) Perspectives of Drug Design that Targets RNA. *Curr. Med. Chem. - Anti-Infective Agents*, **2**, 161–172.
108. Cheatham, T.E. (2004) Simulation and modeling of nucleic acid structure, dynamics and interactions. *Curr. Opin. Struct. Biol.*, **14**, 360–7.
109. Auffinger, P. and Hashem, Y. (2007) Nucleic acid solvation: from outside to insight. *Curr. Opin. Struct. Biol.*, **17**, 325–33.
110. Draper, D.E. (2008) RNA Folding : Thermodynamic and Molecular Descriptions of the Roles of Ions. *Biophys. J.*, **95**, 5489–5495.
111. Bowman, G.R., Huang, X., Yao, Y., Sun, J., Carlsson, G., Guibas, L.J. and Pande, V.S. (2008) Structural insight into RNA hairpin folding intermediates. *J. Am. Chem. Soc.*, **130**, 9676–8.
112. Kannan, S. and Zacharias, M. (2007) Folding of a DNA hairpin loop structure in explicit solvent using replica-exchange molecular dynamics simulations. *Biophys. J.*, **93**, 3218–28.
113. Kannan, S. and Zacharias, M. (2011) Role of the closing base pair for d(GCA) hairpin stability: free energy analysis and folding simulations. *Nucleic Acids Res.*, **39**, 8271–80.
114. Madhumalar, a and Bansal, M. (2005) Sequence preference for BI/BII conformations in DNA: MD and crystal structure data analysis. *J. Biomol. Struct. Dyn.*, **23**, 13–27.
115. Várnai, P., Djuranovic, D., Lavery, R. and Hartmann, B. (2002) Alpha/gamma transitions in the B-DNA backbone. *Nucleic Acids Res.*, **30**, 5398–406.

116. Djuranovic,D. and Hartmann,B. (2004) DNA fine structure and dynamics in crystals and in solution: the impact of BI/BII backbone conformations. *Biopolymers*, **73**, 356–68.
117. Beveridge,D.L., Barreiro,G., Byun,K.S., Case,D. a, Cheatham,T.E., Dixit,S.B., Giudice,E., Lankas,F., Lavery,R., Maddocks,J.H., et al. (2004) Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. I. Research design and results on d(CpG) steps. *Biophys. J.*, **87**, 3799–813.
118. Lavery,R., Zakrzewska,K., Beveridge,D., Bishop,T.C., Case,D. a, Cheatham,T., Dixit,S., Jayaram,B., Lankas,F., Laughton,C., et al. (2010) A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic Acids Res.*, **38**, 299–313.
119. Zhu,J. and Wartell,R.M. (1999) The effect of base sequence on the stability of RNA and DNA single base bulges. *Biochemistry*, **38**, 15986–93.
120. Perez,A. and Orozco,M. (2010) Real-time atomistic description of DNA unfolding. *Angew. Chemie Int. Ed.*, **49**, 4805–8.
121. Garcia,A.E. and Paschek,D. (2008) Simulation of the pressure and temperature folding/unfolding equilibrium of a small RNA hairpin. *J. Am. Chem. Soc.*, **130**, 815–7.
122. Kannan,S. and Zacharias,M. (2007) Folding of a DNA hairpin loop structure in explicit solvent using replica-exchange molecular dynamics simulations. *Biophys. J.*, **93**, 3218–28.
123. Balkwill,G.D., Williams,H.E.L. and Searle,M.S. (2007) Structure and folding dynamics of a DNA hairpin with a stabilising d(GNA) trinucleotide loop: influence of base pair mis-matches and point mutations on conformational equilibria. *Org. Biomol. Chem.*, **5**, 832–9.
124. Yoo,J. and Aksimentiev,A. (2012) Improved Parametrization of Li + , Na + , K + , and Mg 2+ Ions for All-Atom Molecular Dynamics Simulations of Nucleic Acid Systems. *J Phys Chem Lett*, **3**, 45–50.
125. Draper,D.E., Grilley,D. and Soto,A.M. (2005) Ions and RNA folding. *Annu. Rev. Biophys. Biomol. Struct.*, **34**, 221–43.
126. Bowman,J.C., Lenz,T.K., Hud,N. V and Williams,L.D. (2012) Cations in charge: magnesium ions in RNA folding and catalysis. *Curr. Opin. Struct. Biol.*, **22**, 262–72.

-
127. Wagoner, J. and Baker, N. A. (2004) Solvation forces on biomolecular structures: a comparison of explicit solvent and Poisson-Boltzmann models. *J. Comput. Chem.*, **25**, 1623–9.
128. Macke Thomas J. and Case David A. (1997) Molecular Modeling of Nucleic Acids Leontis, N.B., SantaLucia, J. (eds) American Chemical Society, Washington, DC.
129. Zacharias, M. (2000) Simulation of the structure and dynamics of nonhelical RNA motifs. *Curr. Opin. Struct. Biol.*, **10**, 311–7.
130. Pérez, A., Luque, F.J. and Orozco, M. (2012) Frontiers in molecular dynamics simulations of DNA. *Acc. Chem. Res.*, **45**, 196–205.
131. Access, O. (2009) Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters. 10.1261/rna.1270809.
132. Savelyev, A. and Papoian, G.A. (2010) Chemically accurate coarse graining of. *Proc Natl Acad Sci*, **2010**, 20340–20345.
133. Jonikas, M. A., Radmer, R.J., Laederach, A., Das, R., Pearlman, S., Herschlag, D. and Altman, R.B. (2009) Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters. *RNA*, **15**, 189–99.
134. Frellsen, J., Moltke, I., Thiim, M., Mardia, K. V, Ferkinghoff-Borg, J. and Hamelryck, T. (2009) A probabilistic model of RNA conformational space. *PLoS Comput. Biol.*, **5**, e1000406.
135. Rázga, F., Spackova, N., Réblova, K., Koca, J., Leontis, N.B. and Sponer, J. (2004) Ribosomal RNA kink-turn motif—a flexible molecular hinge. *J. Biomol. Struct. Dyn.*, **22**, 183–94.
136. Lilley, D.M. (1995) Kinking of DNA and RNA by base bulges. *Proc. Natl. Acad. Sci. U. S. A.*, **92**, 7140–2.
137. Rázga, F., Zacharias, M., Réblová, K., Koca, J. and Sponer, J. (2006) RNA kink-turns as molecular elbows: hydration, cation binding, and large-scale dynamics. *Structure*, **14**, 825–35.
138. Curuksu, J., Sponer, J. and Zacharias, M. (2009) Elbow flexibility of the kt38 RNA kink-turn motif investigated by free-energy molecular dynamics simulations. *Biophys. J.*, **97**, 2004–13.
139. Kannan, S. and Zacharias, M. (2009) Simulation of DNA double-strand dissociation and formation during replica-exchange molecular dynamics simulations. *Phys. Chem. Chem. Phys.*, **11**, 10541–2.

140. Maleki,A., Alavi,S. and Najafi,B. (2011) Molecular Dynamics Simulation Study of Adsorption and Patterning of DNA Bases on the Au(111) Surface. *J Phys Chem C*, **115**, 22484–22494.
141. Lee,O. and Schatz,G.C. (2009) Interaction between DNAs on a Gold Surface. *J Phys Chem C*.
142. Leach,A.R. (2001) Molecular modelling : principles and applications 2nd ed. Pearson Prentice Hall.
143. Snow,C.D., Sorin,E.J., Rhee,Y.M. and Pande,V.S. (2005) How well can simulation predict protein folding kinetics and thermodynamics? *Ann Rev Biophys Biomol Struc*, **34**, 43–69.
144. Peter,C., Oostenbrink,C., van Dorp,A. and van Gunsteren,W.F. (2004) Estimating entropies from molecular dynamics simulations. *J. Chem. Phys.*, **120**, 2652–61.
145. Levitt,M. (1983) Computer simulation of DNA double-helix dynamics. *Cold Spring Harb. Symp. Quant. Biol.*, **47 Pt 1**, 251–62.
146. McCammon,J.A. and Harvey,S.C. (1988) Dynamics of Proteins and Nucleic Acids Cambridge University Press.
147. De Groot,B.L. and Grubmüller,H. (2001) Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF. *Science*, **294**, 2353–7.
148. Pérez,A., Luque,F.J. and Orozco,M. (2007) Dynamics of B-DNA on the microsecond time scale. *J. Am. Chem. Soc.*, **129**, 14739–45.
149. Pérez,A., Luque,F.J. and Orozco,M. (2012) Frontiers in molecular dynamics simulations of DNA. *Acc. Chem. Res.*, **45**, 196–205.
150. Verlet,L. (1967) Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys Rev*, **159**, 98–103.
151. Hockney,R.W. (1970) The potential calculation and some applications. *Methods Compu t. Phys.*, **9**, 136 – 211.
152. Banáš,P., Hollas,D., Zgarbová,M., Jurečka,P., Orozco,M., Cheatham,T.E., Šponer,J. and Otyepka,M. (2010) Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins. *J. Chem. Theory Comput.*, **6**, 3836–3849.
153. Pérez,A., Marchán,I., Svozil,D., Sponer,J., Cheatham,T.E., Loughton,C. a and Orozco,M. (2007) Refinement of the AMBER force field for nucleic

- acids: improving the description of alpha/gamma conformers. *Biophys. J.*, **92**, 3817–29.
154. Reddy,S.Y., Leclerc,F. and Karplus,M. (2003) DNA polymorphism: a comparison of force fields for nucleic acids. *Biophys. J.*, **84**, 1421–49.
155. Brooks,B.R., Bruccoleri,R.E., Olafson,B.D., States,D.J., Swaminathan,S. and Karplus,M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, **4**, 187–217.
156. Cornell,W.D., Cieplak,P., Bayly,C.I., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W., Kollman,P. a., et al. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.
157. Soares,T.A., Hünenberger,P.H., Kastenholz,M.A., Kräutler,V., Lenz,T., Lins,R.D., Oostenbrink,C. and van Gunsteren,W.F. (2005) An improved nucleic acid parameter set for the GROMOS force field. *J. Comput. Chem.*, **26**, 725–37.
158. Durrant,J.D. and McCammon,J.A. (2011) Molecular dynamics simulations and drug discovery. *BMC Biol.*, **9**, 71.
159. Wang,J., Wolf,R.M., Caldwell,J.W., Kollman,P.A. and Case,D.A. (2004) Development and testing of a general amber force field. *J. Comput. Chem.*, **25**, 1157–74.
160. Mlýnský,V., Banás,P., Hollas,D., Réblová,K., Walter,N.G., Sponer,J. and Otyepka,M. (2010) Extensive molecular dynamics simulations showing that canonical G8 and protonated A38H⁺ forms are most consistent with crystal structures of hairpin ribozyme. *J Phys Chem B*, **114**, 6642–52.
161. Zgarbová,M., Otyepka,M., Sponer,J., Mládek,A., Banáš,P., Cheatham,T.E. and Jurečka,P. (2011) Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.*, **7**, 2886–2902.
162. Banaš,P., Hollas,D., Zgarbova,M., Jurecka,P., Orozco,M., III,T.E.C., Sponer,J. and Otyepka,M. (2010) Performance of Molecular Mechanics Force Fields for RNA Simulations : Stability of UUCG and GNRA Hairpins. *J. Chem. Theory Comput.*, **6**, 3836–3849.
163. Altona,C. and Sundaralingam,M. (1972) Conformational analysis of the sugar ring in nucleosides and nucleotides. New description using the concept of pseudorotation. *J. Am. Chem. Soc.*, **94**, 8205–8212.

164. Murray-Rust,P. and MOTHERWELL,S. (1978) Computer Retrieval and Analysis of Molecular Geometry. III. Geometry of the t- l¹-Aminofuranoside Fragment. *Acta Cryst*, **B34**, 2534–2546.
165. Dunitz,J.D. (1972) Approximate relationships between conformational parameters in 5- and 6-membered rings. *Tetrahedron*, **28**, 5459–5467.
166. Lescrinier,E. (2003) Difference in conformational diversity between nucleic acids with a six-membered “sugar” unit and natural “furanose” nucleic acids. *Nucleic Acids Res.*, **31**, 2975–2989.
167. Neidle,S. (2008) Principles of Nucleic Acid Structure Oxford University Press.
168. Schneider,B., Neidle,S. and Berman,H.M. (1997) Conformations of the sugar-phosphate backbone in helical DNA crystal structures. *Biopolymers*, **42**, 113–24.
169. Sundaralingam,M. (1969) Stereochemistry of nucleic acids and their constituents. IV. Allowed and preferred conformations of nucleosides, nucleoside mono-, di-, tri-, tetraphosphates, nucleic acids and polynucleotides. *Biopolymers*, **7**, 821–860.
170. Altona,C. and Sundaralingam,M. (1972) Conformational analysis of the sugar ring in nucleosides and nucleotides. A new description using the concept of pseudorotation. *J. Am. Chem. Soc.*, **94**, 8205–12.
171. Dong,F. and Miller,R.E. (2002) Vibrational transition moment angles in isolated biomolecules: a structural tool. *Science*, **298**, 1227–30.
172. Luisi,B., Orozco,M., Sponer,J., Luque,F.J. and Shakked,Z. (1998) On the potential role of the amino nitrogen atom as a hydrogen bond acceptor in macromolecules. *J. Mol. Biol.*, **279**, 1123–36.
173. Dickerson,R.E. (1989) Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Res.*, **17**, 1797–803.
174. Olson,W.K., Bansal,M., Burley,S.K., Dickerson,R.E., Gerstein,M., Harvey,S.C., Heinemann,U., Lu,X.J., Neidle,S., Shakked,Z., et al. (2001) A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.*, **313**, 229–37.
175. Lavery,R., Moakher,M., Maddocks,J.H., Petkeviciute,D. and Zakrzewska,K. (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.*, **37**, 5917–29.
176. Gilson,M.K. and Zhou,H.-X. (2007) Calculation of protein-ligand binding affinities. *Annu. Rev. Biophys. Biomol. Struct.*, **36**, 21–42.

-
177. Leach,A.R. (2001) *Molecular Modelling: Principles and Applications*.
178. Wereszczynski,J. and Mccammon,J.A. (2011) Statistical mechanics and molecular dynamics in evaluating thermodynamic properties of biomolecular recognition. *Quart Rev Biophys*.
179. Roux,B. (1995) The calculation of the potential of mean force using computer simulations. *Comp Phys Comm*, **91**, 275–282.
180. Torrie,G.M. and Valleau,J.P. (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella-Sampling. *J Comp Phys*, **23**, 187–199.
181. Torrie,G.M. and Valleau,J.P. (1974) Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chem Phys Lett*, **28**, 578–581.
182. Kästner,J. (2011) Umbrella-Sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **1**, 932–942.
183. Kumar,S., Rosenberg,J.M., Bouzida,D., Swendsen,R.H. and Kollman,P.A. (1995) Multidimensional free-energy calculations using the weighted histogram analysis method. *J. Comput. Chem.*, **16**, 1339–1350.
184. Kollman,P.A., Massova,I., Reyes,C., Kuhn,B., Huo,S., Chong,L., Lee,M., Lee,T., Duan,Y., Wang,W., et al. (2000) Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc Chem Res*, **33**, 889–897.
185. Gouda,H., ID,K., DA,C. and PA.,K. (2003) Free energy calculations for the phylline binding to an RNA aptamer: comparison of MM-PBSA and thermodynamic integration methods. *Biopolymers*, **68**, 16–34.
186. Wang,J., Morin,P., Wang,W. and Kollman,P.A. (2001) Use of MMPBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. *J Am Chem Soc*, **123**, 5221–5230.
187. Spomer,J., Cang,X. and Cheatham,T.E. (2012) Molecular dynamics simulations of G-DNA and perspectives on the simulation of nucleic acid structures. *Methods*, **57**, 25–39.
188. McDowell,S.E., Špačková,N., Šponer,J. and Walter,N.G. (2007) Molecular Dynamics Simulations of RNA: An In Silico Single Molecule Approach. *Biopolymers*, **85**, 169–184.
189. Kara,M. and Zacharias,M. (2013) Theoretical studies of nucleic acids folding. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 10.1002/wcms.1146.

-
190. Richards,F.M. (1977) Areas, volumes, packing, and protein structure. *Ann Rev Biophys Bioeng*, **6**, 151–176.
191. Weiser,J., Shenkin,P.S. and Still,W.C. (1999) Approximate atomic surfaces from linear combinations of pairwise overlaps(LCPO). *J. Comput. Chem.*, **20**, 217–230.
192. Wang,F. and Landau,D.P. (2001) Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.*, **86**, 2050–3.
193. McCammon,J.A., Gelin,B.R. and Karplus,M. (1977) Dynamics of folded proteins. *Nature*, **267**, 585–590.
194. Piana,S., Lindorff-Larsen,K. and Shaw,D.E. (2012) Protein folding kinetics and thermodynamics from atomistic simulation. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 17845–50.
195. Ansari, a, Kuznetsov,S. V and Shen,Y. (2001) Configurational diffusion down a folding funnel describes the dynamics of DNA hairpins. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 7771–6.
196. Treiber,D.K. and Williamson,J.R. (2001) Beyond kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.*, **11**, 309–14.
197. Mitsutake,A., Sugita,Y. and Okamoto,Y. (2001) Generalized-ensemble algorithms for molecular simulations of biopolymers. *Biopolymers*, **60**, 96–123.
198. Hansmann,U.H. (1997) Parallel Tempering Algorithm for Conformational Studies of Biological Molecules. *Chem Phys Lett*, **281**, 140–50.
199. Fukunishi,H., Watanabe,O. and Takada,S. (2002) On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *J. Chem. Phys.*, **116**, 9058.
200. Kannan,S. and Zacharias,M. (2010) Application of biasing-potential replica-exchange simulations for loop modeling and refinement of proteins in explicit solvent. *Proteins*, **78**, 2809–19.
201. Curuksu,J. and Zacharias,M. (2009) Enhanced conformational sampling of nucleic acids by a new Hamiltonian replica exchange molecular dynamics approach. *J. Chem. Phys.*, **130**, 104110.
202. Curuksu,J. (2011) Adaptive conformational sampling based on replicas. *J. Math. Biol.*, 10.1007/s00285-011-0432-6.

-
203. Kara, M. and Zacharias, M. (2013) Influence of 8-Oxoguanosine on the Fine Structure of DNA Studied with Biasing-Potential Replica Exchange Simulations. *Biophys J*, **104**, 1089–1097.
204. Wadley, L.M., Keating, K.S., Duarte, C.M. and Pyle, A.M. (2007) Evaluating and learning from RNA pseudotorsional space: quantitative validation of a reduced representation for RNA structure. *J. Mol. Biol.*, **372**, 942–57.
205. Elizabeth, M., Marshall, L., Richs, A., John, M., Dean, L., Lipscomb, L.A., Elizabeth Peek, M., Morningstern, M.L., Verghis, S.M., Miller, E.M., et al. (1995) X-ray structure of a DNA decamer containing 7,8-dihydro-8-oxoguanine. *Biochemistry*, **92**, 719–723.
206. Case, D.A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A., Simmerling, C., Wang, B. and Woods, R.J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem.*, **26**, 1668–88.
207. Darden, T., York, D. and Pedersen, L. (1993) Particle mesh Ewald: An $N^2 \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089–10092.
208. Cheatham, T.E.I., Miller, J.L., Fox, T., Darden, T.A. and Kollman, P.A. (1995) Molecular Dynamics Simulations on Solvated Biomolecular Systems: The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Proteins. *J. Am. Chem. Soc.*, **117**, 4193–4194.
209. Ryckaert, J.-P., Ciccotti, G. and Berendsen, H.J.. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comp Phys*, **23**, 327–341.
210. Daniels, D.S., Woo, T.T., Luu, K.X., Noll, D.M., Clarke, N.D., Pegg, A.E. and Tainer, J.A. (2004) DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nat. Struct. Mol. Biol.*, **11**, 714–20.
211. Miller, B.R., McGee, T.D., Swails, J.M., Homeyer, N., Gohlke, H. and Roitberg, A.E. (2012) MMPBSA.py : An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.*, **8**, 3314–3321.
212. Popena, L., Bielecki, L., Gdaniec, Z. and Adamiak, R.W. (2009) Structure and dynamics of adenosine bulged RNA duplex reveals formation of the dinucleotide platform in the C:G-A triple. *ARKIVOC*, **3**, 130–144.
213. Bana, P., Hollas, D., Zgarbova, M., Jurec, P., Iii, T.E.C., Orozco, M. and Otyepka, M. (2010) Performance of Molecular Mechanics Force Fields for RNA Simulations : Stability of UUCG and GNRA Hairpins. *J. Chem. Theory Comput.*, **6**, 3836–3849.

214. Pérez,A., Marchán,I., Svozil,D., Sponer,J., Cheatham,T.E., Laughton,C. a and Orozco,M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys. J.*, **92**, 3817–29.
215. Henriksen,N. and Cheatham,T.E. (2011) RNA structure refinement and force field evaluation using molecular dynamics simulations. *Proc. 2011 TeraGrid Conf. Extrem. Digit. Discov. - TG '11*, 10.1145/2016741.2016745.
216. Miller,J.H., Fan-Chiang,C.-C.P., Straatsma,T.P. and Kennedy,M.A. (2003) 8-Oxoguanine enhances bending of DNA that favors binding to glycosylases. *J. Am. Chem. Soc.*, **125**, 6331–6.
217. Marathe,A., Karandur,D. and Bansal,M. (2009) Small local variations in B-form DNA lead to a large variety of global geometries which can accommodate most DNA-binding protein motifs. *BMC Struct. Biol.*, **9**, 24.
218. Svozil,D., Kalina,J., Omelka,M. and Schneider,B. (2008) DNA conformations and their sequence preferences. *Nucleic Acids Res.*, **36**, 3690–706.
219. Mandel-Gutfreund,Y., Schueler,O. and Margalit,H. (1995) Comprehensive analysis of hydrogen bonds in regulatory protein DNA-complexes: in search of common principles. *J. Mol. Biol.*, **253**, 370–82.
220. Vesnaver,G. (1991) The Contribution of DNA Single-Stranded Order to the Thermodynamics of Duplex Formation. *Proc Natl Acad Sci*, **88**, 3569–3573.