

## 2. Struktur und Funktion der Nukleinsäuren

Die dreidimensionale Struktur der Deoxyribonukleinsäure (DNA) wurde erstmals 1953 von Watson und Crick als Doppelhelix zweier DNA-Stränge beschrieben [1]. Diese Entdeckung der DNA-Doppelhelix stellte einen Durchbruch und den Beginn der modernen Strukturbiologie dar, deren wesentliches Interesse bis heute in der Aufklärung der Struktur von biologisch relevanten Makromolekülen und der damit verbundenen Struktur-Funktions-Beziehungen besteht.

### 2.1. Chemischer Aufbau und Modelle der Raumstruktur

Die monomeren Bausteine der Nukleinsäuren sind die Nukleotide, die aus jeweils einer Base, einer Pentose und einer Phosphatgruppe gebildet werden, wobei die Variabilität der unterschiedlichen Nukleotide durch die Basen bestimmt wird. Die Base ist dabei kovalent über eine N-glykosidische Bindung an das C1'-Atom der Pentose gebunden. Im Fall von DNA ist die Pentose eine  $\beta$ -D-2'-Desoxyribose, wogegen die Pentose der Ribonukleinsäure (RNA) eine  $\beta$ -D-Ribose ist. Die Phosphatgruppe verbindet kovalent zwei benachbarte Zuckereinheiten, wobei eine Phosphorsäure mit dem C5'-Atom der Pentose verestert ist. Diese kovalente Verknüpfung über eine Phosphodiesterbindung zwischen der 3'-Hydroxylgruppe einer Pentose und der 5'-Hydroxylgruppe der Pentose des darauffolgenden Nukleotids führt zur Ausbildung eines kovalent verbundenen Rückgrats der DNA mit alternierenden Phosphat- und Zuckergruppen. Die Nukleotide bilden als monomere Bausteine der DNA, wie in Abb. 2.1 dargestellt, ein lineares Polymer mit den Basen als vom Rückgrat ausgehenden Verzweigungen. Aufgrund der negativen Ladungen der Phosphatgruppen und der gleichen Orientierung aller Phosphodiesterbindungen ist dieses polyanionische Polymer hochpolar. Die Richtung eines Nukleinsäure-Stranges wird nach einer Konvention immer nach den C5'- und C3'-Atomen der Zuckereinheit bezeichnet, so daß die Enden des Stranges als 5'- und 3'-Ende bezeichnet werden und die Sequenz eines DNA-Stranges stets in 5'-3'-Richtung angegeben wird.

Die Basen der DNA sind Derivate der Purine (R) bzw. der Pyrimidine (Y), die mit ihren konjugierten  $\pi$ -Elektronensystemen planare heterozyklische Systeme bilden. Die chemischen Strukturen der beiden Purinbasen Adenin (A) und Guanin (G) sowie der beiden Pyrimidinbasen Cytosin (C) und Thymin (T) sind in Abb. 2.1 dargestellt. Die RNA unterscheidet sich auch dadurch von der DNA, daß sie im wesentlichen statt Thymin die Pyrimidinbase Uracil enthält. Neben den erwähnten häufig vorkommenden Basen gibt es einige weitere, aber seltene Derivate dieser Basen.

Watson und Crick konnten 1953 mit Hilfe der Analyse von Röntgenbeugungsbildern von DNA-Fasern ein Postulat für die räumliche Struktur der DNA formulieren [1], das darauf beruht, daß zwei zueinander antiparallel verlaufende Polynukleotid-Stränge um eine gemeinsame Achse gewunden sind, wodurch eine Doppelhelix entsteht, deren Achse als

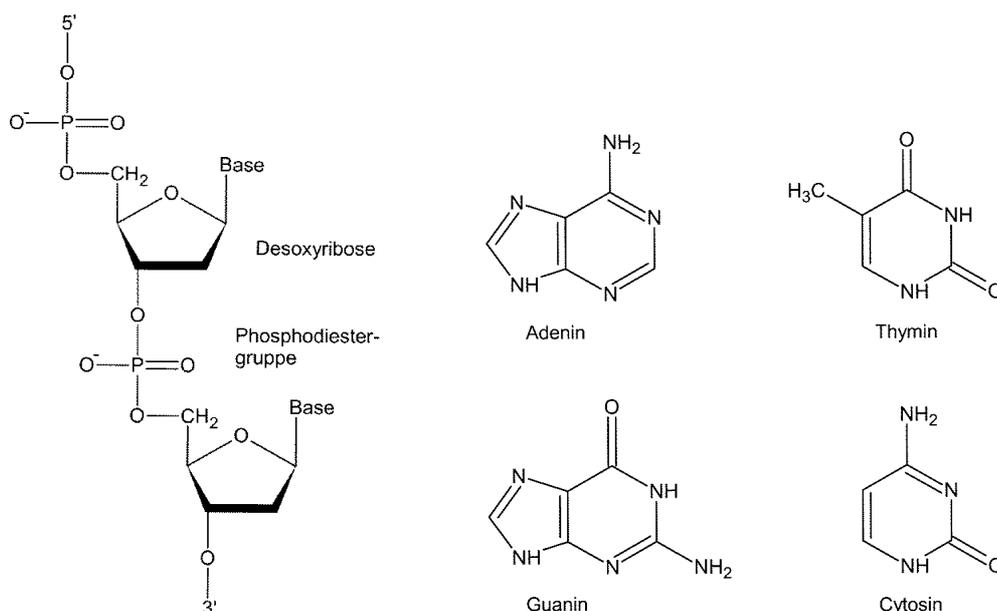


Abb. 2.1.: Chemischer Aufbau des Desoxyribose-Phosphodiester-Rückgrats und der vier Basen der Desoxyribonukleinsäure

helikale Achse bezeichnet wird. Das Zucker-Phosphat-Rückgrat ist dabei nach außen, die Basen sind hingegen nach innen orientiert, wobei die planaren Basen einen rechten Winkel mit der helikalen Achse bilden. Der Durchmesser der Doppelhelix beträgt dabei ca. 20 Å, benachbarte Basenpaare folgen auf der Helixachse im Abstand von ca. 3,4 Å, und die Basenpaare sind um die Helixachse um ca. 36° gegeneinander verdreht. Die beiden antiparallelen Polynukleotid-Einzelstränge sind durch Wasserstoffbrückenbindungen miteinander verknüpft. Aus sterischen Gründen bilden jeweils eine Purin- und eine Pyrimidinbase ein Basenpaar im Inneren der Doppelhelix, wobei ein Basenpaar durch Wasserstoffbrückenbindungen zwischen Akzeptor- und Donatoratomen der beiden beteiligten Basen stabilisiert wird. Die Basenpaarung ist aufgrund der geometrischen Lage der Donator- und Akzeptorbindungsstellen basenspezifisch, so daß Adenin und Thymin ein Basenpaar mit zwei sowie Guanin und Cytosin ein Basenpaar mit drei Wasserstoffbrücken bilden [2]. Darüberhinaus zeigt die DNA keine Einschränkung in der Basensequenz. Chargaff et al. berichteten über ein statistisches Sequenzverhältnis der Basen von  $A:T \cong G:C \cong 1,0$  [3]. Neben der Basenpaarung durch Wasserstoffbrücken wird die DNA durch eine hydrophobe Wechselwirkung der Basen stabilisiert, die zur Stapelung benachbarter Basenpaare führt und deshalb auch Stapelwechselwirkung genannt wird. Die Basenpaarung und die Stapelwechselwirkung gemeinsam ermöglichen die Abschirmung der hydrophoben, heterozyklischen Basen vom Lösungsmittel, das als Wasser eine hohe Polarität besitzt und dem lediglich das hochpolare Rückgrat der DNA ausgesetzt ist.

Die auf diese Weise stabilisierte DNA-Doppelhelix kann allerdings in Abhängigkeit von der Polarität und dem Salzgehalt des Lösungsmittels aber auch der Basensequenz eine unterschiedliche Anatomie aufweisen. Die von Watson und Crick beschriebene doppelhelikale DNA-Struktur wird als B-DNA bezeichnet [1, 4]. Die B-DNA liegt meist in wäßriger Lösung vor und ist eine rechtsgängige Doppelhelix, die einen mittleren Helixdurchmesser von 23,7 Å, einen Abstand zwischen den Basenpaaren auf der helikalen

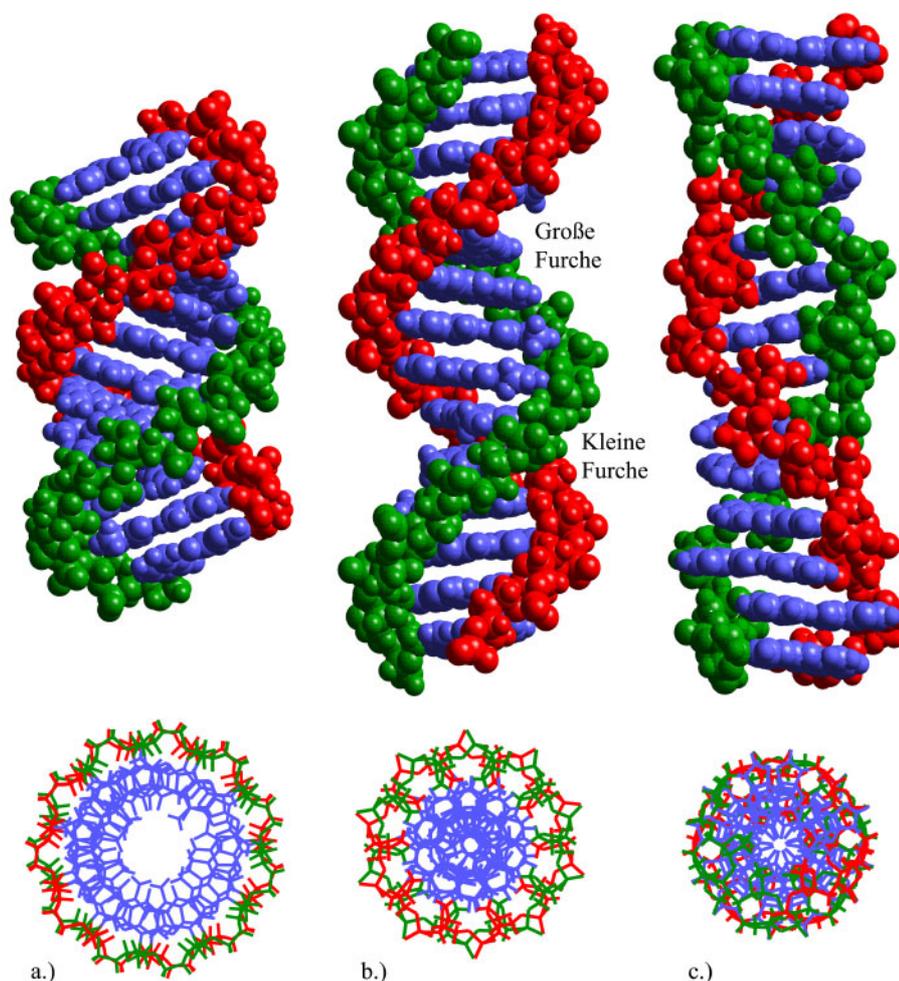


Abb. 2.2.: Konformationen der a.) A-DNA, b.) B-DNA und c.) Z-DNA. Ansichten oben senkrecht zur Helixachse und unten entlang der Helixachse.

Achse von  $3,4 \text{ \AA}$  und eine mittlere Neigung der Basenpaare zur Helixachse von nur  $1^\circ$  aufweist. Eine Helixwindung der B-DNA benötigt  $10,4$  Basenpaare und führt zu einer Ganghöhe von  $35,4 \text{ \AA}$ . Die dagegen bei geringerer Polarität des Lösungsmittels zu beobachtende A-DNA [4] ist ebenfalls eine rechtsgängige Doppelhelix, besitzt allerdings einen größeren mittleren Helixdurchmesser von  $25,5 \text{ \AA}$ , jedoch einen kleineren Abstand der Basenpaare auf der helikalen Achse von nur  $2,6 \text{ \AA}$  und eine mittlere Neigung der Basenpaare zu dieser Achse von  $19^\circ$ . Für eine Helixwindung der A-DNA werden  $11$  Basenpaare benötigt, die im Gegensatz zur B-DNA nur eine Ganghöhe von  $25,3 \text{ \AA}$  erreichen. Eine dritte, seltener und nur bei einem sehr hohen Salzgehalt des Lösungsmittels vorkommende Form ist die Z-DNA [5], deren wesentlichster Unterschied zur A- und B-DNA darin besteht, daß sie eine linksgängige Helix ist, deren Rückgrat einen zickzackförmigen Verlauf zeigt und deren mittlerer Helixdurchmesser lediglich  $18,4 \text{ \AA}$  beträgt, wobei der Abstand zwischen benachbarten Basenpaaren auf der Helixachse  $3,8 \text{ \AA}$  und die mittlere Neigung der Basenpaare zur helikalen Achse  $9^\circ$  beträgt. Bei der Z-DNA tragen  $12$  Basenpaare zu einer Helixwindung und einer Ganghöhe von  $45,6 \text{ \AA}$  bei. In Kap. 2.2 werden die Konformationen der A-, B- und Z-Form der DNA-Doppelhelix mit Hilfe dort eingeführter helikaler Parameter genauer beschrieben. Aus Abb. 2.2

werden neben den bereits beschriebenen Eigenschaften der verschiedenen Topologien Unterschiede in den Furchengeometrien deutlich. Die große Furche der B-DNA ist weit und tief, wobei diese Furche bei der A-Form verengt und weiter vertieft ist und bei der Z-Form sehr flach ist. Die kleine Furche der B-DNA ist dagegen eng und ebenfalls recht tief, die der A-Form hingegen ist flach und weit, die kleine Furche der Z-Form ist wiederum eng und tief. Neben der beschriebenen A-, B- und Z-Form der DNA-Doppelhelix [6] gibt es eine Reihe weiterer, selten vorkommender Konformationen wie die D-Form als extreme B-DNA mit einer sehr engen und tiefen kleinen Furche sowie einer starken gegenseitigen Drehung zweier benachbarter Basenpaare um die helikale Achse [6, 7]. Im Vergleich zu den diskutierten DNA-Konformationen sind RNA-Strukturen viel komplexer, obgleich doppelsträngige Bereiche in RNA-Molekülen zu einer A- oder A-ähnliche Konformation tendieren [6, 8]. Weitere charakteristische Formen der DNA und der RNA werden wie die vorangegangenen Formen von Saenger diskutiert [6]. So können auch drei oder vier Nukleinsäurestränge Tripelhelices bzw. Quadruplexe bilden [6, 9]. Voraussetzung für diese komplexeren Formen ist zumeist die Bildung von nicht-kanonischen Basenpaarungen wie beispielsweise der Hoogsteen-Basenpaarung [6], die nicht der Watson-Crick-Basenpaarung entsprechen [10, 11, 12].

## 2.2. Strukturelle Analyse von Nukleinsäuren

### 2.2.1. Beschreibung durch helikale Parameter

Die Basen der Nukleinsäuren bilden durch ihre konjugierten  $\pi$ -Elektronensysteme jeweils heterozyklische, planare Systeme, die als starre Körper ohne innere Freiheitsgrade approximiert werden können. Die Lage und Orientierung der Basen relativ zueinander und relativ zu einer definierten Helixachse werden durch helikale Parameter beschrieben. Aus Abb. 2.2 der A-, B- und Z-Form der DNA-Doppelhelix wird ersichtlich, daß eine Beschreibung der Konformation der Helices unter Ausnutzung der helikalen Symmetrie geeignet ist, um die Konformationen verschiedener helikaler Formen eindeutig analysieren und vergleichen zu können. Grundlage dafür ist die eindeutige Festlegung einer helikalen Achse. Im Rahmen dieser Arbeit wird der dem Programm Curves zugrunde liegende Algorithmus für die Definition der helikalen Achse und Berechnung der helikalen Parameter verwendet [13, 14], dessen Konventionen Teil eines internationalen Konsenses sind [15, 16].

Dabei wird für jede individuelle Base  $i$  ein körperfestes Koordinatensystem  $(\vec{j}_i, \vec{k}_i, \vec{l}_i)$  in einem basenbezogenen Referenzpunkt  $\vec{E}_i$  definiert [13, 14]. Die  $\vec{j}$ -Achse dieses Koordinatensystems zeigt in Richtung der großen Furche und liegt in der Standard-B-DNA auf der Dyadenachse. Die  $\vec{l}$ -Achse ist die Normale der Basenebene, wobei diese Achse in die 5'-3'-Richtung des Nukleinsäurestranges zeigt. Die mit diesen beiden Achsen ein rechtwinkliges Koordinatensystem bildende  $\vec{k}$ -Achse zeigt somit in Richtung des an die Base gebundenen Zuckerringes. Neben dem körperfesten Koordinatensystem  $(\vec{j}_i, \vec{k}_i, \vec{l}_i)$  der Base  $i$  wird ein auf die helikale Achse bezogenes Koordinatensystem  $(\vec{v}_i, \vec{w}_i, \vec{u}_i)$  in einem auf der helikalen Achse positionierten Referenzpunkt  $\vec{P}_i$  definiert, dessen  $\vec{v}$ -Achse in die 5'-3'-Richtung der helikalen Achse zeigt, dessen  $\vec{v}$ -Achse auf der lokalen Dyadenachse in die große Furche gerichtet ist, und dessen  $\vec{w}$ -Achse die zu den beiden Achsen Senkrechte bildet. Dieses achsenbezogene Koordinatensystem  $(\vec{v}_i, \vec{w}_i, \vec{u}_i)$  stimmt in einer idealen B-Form mit dem körperfesten Achsensystem der jeweiligen Base überein.

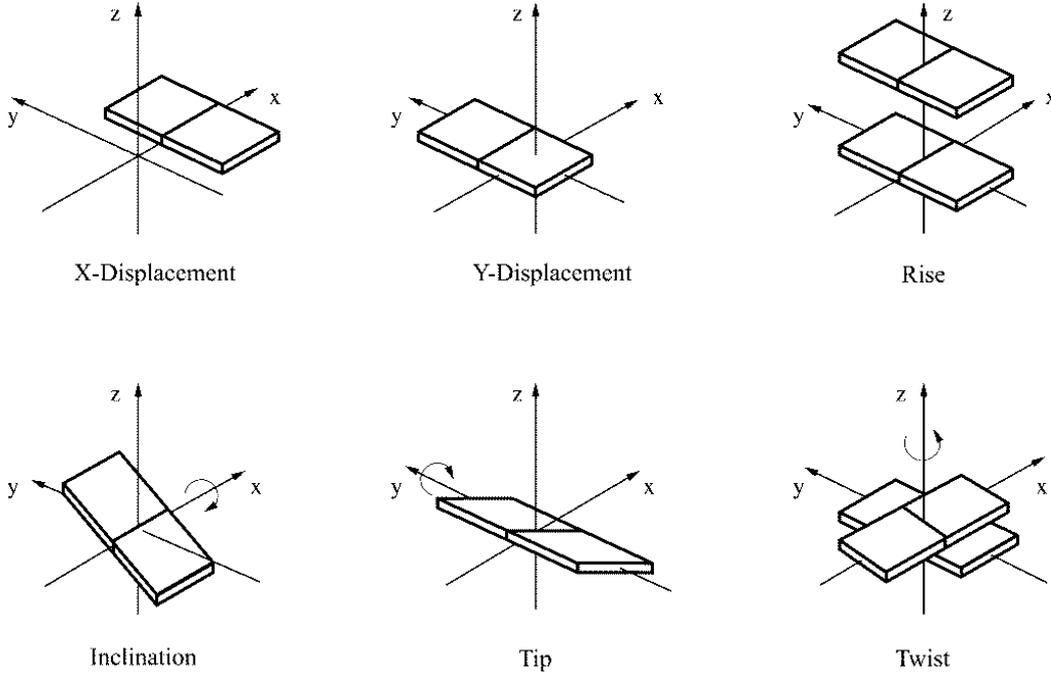


Abb. 2.3.: Schematische Darstellung globaler Basen-Achsen-Parameter [13, 14]

Die Lage des körperfesten Koordinatensystems  $(\vec{j}_i, \vec{k}_i, \vec{l}_i)$  in bezug auf das achsenfeste Koordinatensystem  $(\vec{u}_i, \vec{v}_i, \vec{w}_i)$  kann durch vier helikale Parameter der Base  $i$  beschrieben werden. Dieses sind die beiden Translationen in Richtung der  $\vec{v}$ -Achse (X-Displacement  $dx$ ) und der  $\vec{w}$ -Achse (Y-Displacement  $dy$ ) und die beiden rechtsgängigen Rotationen um eine durch  $\vec{E}_i$  verlaufende zur  $\vec{v}$ -Achse parallele Achse (Inclination  $\eta$ ) und um die  $\vec{k}$ -Achse (Tip  $\vartheta$ ) [14]. Die Lage einer Base relativ zu einer benachbarten Base kann durch den Abstand der Ursprünge (Rise  $Dz$ ) sowie den eingeschlossenen Winkel der  $\vec{w}$ -Achsen (Twist  $\Omega$ ) zweier aufeinanderfolgender achsenbezogener Koordinatensysteme beschrieben werden [14]. Diese sechs am häufigsten zur Beschreibung helikaler Strukturen verwendeten helikalen Parameter sind in Abb. 2.3 schematisch dargestellt und werden wie in 2.1-2.6 angegeben berechnet. Dabei werden die drei Rotationen  $\eta$ ,  $\vartheta$  und  $\Omega$  positiv bei rechtsgängigen Drehungen und entsprechend negativ bei linksgängigen Drehungen [14]. Bei der Berechnung der sich im Nukleinsäurestrang auf die vorhergehende Base  $i - 1$  beziehenden helikalen Parameter  $dz$  und  $\Omega$  wird in 2.1-2.6 eine lineare raumfeste helikale Achse angenommen, weshalb im Fall einer gekrümmten Achse zur Berechnung ein zwischen den beiden benachbarten Achsensystemen vermittelndes mittleres Achsensystem definiert wird [13].

$$\text{X-Displacement:} \quad dx(i) = \vec{v}_i \cdot (\vec{E}_i - \vec{P}_i) \quad (2.1)$$

$$\text{Y-Displacement:} \quad dy(i) = \vec{w}_i \cdot (\vec{E}_i - \vec{P}_i) \quad (2.2)$$

$$\text{Rise:} \quad Dz(i) = |\vec{P}_i - \vec{P}_{i-1}| \quad (2.3)$$

$$\text{Inclination:} \quad \eta(i) = \pm \arccos(\vec{k}_i \cdot \vec{w}_i) \quad (2.4)$$

$$\text{Tip:} \quad \vartheta(i) = \pm \arccos(\vec{j}_i \cdot \vec{v}_i) \quad (2.5)$$

$$\text{Twist:} \quad \Omega(i) = \pm \arccos(\vec{w}_{i-1} \cdot \vec{w}_i) \quad (2.6)$$

In Ergänzung dieser die Lage und Orientierung der Base relativ zur Helixachse ( $dx$ ,  $dy$ ,  $\eta$ ,  $\vartheta$ ) bzw. zweier benachbarter Basen zueinander ( $Dz$ ,  $\Omega$ ) definierender helikaler Parameter charakterisieren vier weitere Parameter die Lage benachbarter Basen zueinander (Shift  $Dx$ , Slide  $Dy$ , Tilt  $\tau$ , Roll  $\rho$ ), sechs Parameter die Orientierung zweier gepaarter Basen zueinander (Shear  $Sx$ , Stretch  $Sy$ , Stagger  $Sz$ , Buckle  $\kappa$ , Propeller  $\omega$ , Opening  $\sigma$ ) und letztlich vier Parameter den Verlauf der helikalen Achse ( $Ax$ ,  $Ay$ ,  $\eta_A$ ,  $\vartheta_A$ ). Dabei sind die für einzelne Basen definierten Parameter auch für Basenpaare berechenbar, wozu die helikalen Parameter der den beiden Nukleinsäure-Strängen  $I$  und  $II$  zugehörigen Basen durch Differenzbildung oder Summation auf das Basenpaar bezogene helikale Parameter definieren [13]. Im Rahmen dieser Arbeit werden neben den in Abb. 2.3 dargestellten helikalen Parametern auch die in 2.7-2.9 unter Annahme einer raumfesten helikalen Achse angegebenen Parameter Buckle  $\kappa$ , Propeller  $\omega$  und Roll  $\rho$  diskutiert werden [13, 14].

$$\text{Buckle:} \quad \kappa(i) = \eta(I) - \eta(II) \quad (2.7)$$

$$\text{Propeller:} \quad \omega(i) = \vartheta(I) + \vartheta(II) \quad (2.8)$$

$$\text{Roll:} \quad \rho(i) = \vartheta(i) - \vartheta(i-1) \quad (2.9)$$

Da die Definition der eingeführten helikalen Parameter eine globale Helixachse benutzt, sind diese Parameter globale helikale Parameter. Als Helixachse kann entweder als gute Approximation die beste durch die Struktur verlaufende lineare Achse oder eine auch möglicherweise gekrümmte globale Achse durch Optimierung mit Curves [13, 14] berechnet werden. Neben den auf eine globale Achse bezogenen globalen helikalen Parametern können lokale helikale Parameter berechnet werden, wenn sie in bezug auf die körperfesten lokalen Achsensysteme benachbarter Basen bzw. Basenpaare definiert werden [14]. Die Basennormalen  $\vec{l}_i$  der regulären B-DNA sind entlang der globalen helikalen Achse orientiert, wohingegen die Basennormalen  $\vec{l}_i$  der A-DNA sich um die globale Helixachse winden [17]. Ein Vorteil der Angabe globaler helikaler Parameter besteht darin, die A-, B- und andere reguläre Formen der DNA nur durch globale helikale Parameter eindeutig klassifizieren und interpretieren zu können. Die Krümmung einer Nukleinsäurestruktur wird ebenfalls besser durch die globale helikale Achse beschrieben. Die biologische Bedeutung der vielen, heute bekannten irregulären und gekrümmten Nukleinsäurestrukturen sowie deren Komplexe mit Liganden und Proteinen läßt einer konsequenten Analyse der Krümmungseigenschaften dieser Strukturen einen wachsenden Stellenwert zukommen [17].

### 2.2.2. Furchengeometrien

Neben den helikalen Parametern unterscheiden sich die unterschiedlichen Nukleinsäurekonformationen, wie in Abb. 2.2 ersichtlich, in der Geometrie ihrer beiden Furchen. Die Glykosidbindungen der an einem Basenpaar beteiligten Basen sind bei den am häufigsten beobachteten Basenpaarungen auf eine Seite der Paarebene gerichtet, so daß zwei unterschiedlich große Furchen zwischen den beiden Nukleinsäuresträngen entstehen, die große (M: major) und die kleine (m: minor) Furche [6]. Die Geometrie der beiden Furchen ist infolge der Flexibilität der Doppelhelix variabel, wobei allerdings Furchentiefen  $d^M$  und  $d^m$  sowie -weiten  $w^M$  und  $w^m$  angegeben werden können, die für die A-, B- und Z-Form der DNA charakteristisch sind. Aufgrund der Variabilität der Furchengeome-

trie sind die Furchen biologisch und pharmakologisch bedeutsame Bindungsstellen für Liganden und Proteine. Durch eine solche Furchenbindung kann sich die Geometrie der Furchen dramatisch ändern. Die Analyse der Furchengeometrie erfolgte im Rahmen dieser Arbeit durch einen in das Programm Curves [13, 14] implementierten Algorithmus, der die berechnete helikale Achse und den für jedes Basenpaar  $i$  definierten Pseudodyadenvektor  $\vec{v}_i$  benutzt. Außerdem definieren die Phosphoratome eines Nucleinsäurestranges jeweils eine kubische Raumkurve. Die Furchenweiten  $w_i^M$  und  $w_i^m$  werden über die Schnittpunkte der beiden Raumkurven mit einer um den Pseudodyadenvektor  $\vec{v}_i$  rotierenden Ebene definiert. Die Verbindungslinie dieser beiden Schnittpunkte schneidet ebenfalls in jeder der beiden Furchen die Pseudodyadenachse, so daß der Abstand dieser Punkte zu den Schnittflächen der Lennard-Jones-Oberfläche des Basenpaares mit dem Pseudodyadenvektor  $\vec{v}_i$  die Furchentiefen  $d_i^M$  und  $d_i^m$  definiert [18].

### 2.2.3. Torsionswinkel und Zuckerkaltung

Die helikalen Parameter beschreiben die Lage der als starre Körper betrachteten Basen in einem helikalen Achsensystem. Das die Basen verbindende flexible Desoxyribose-Phosphodiester-Rückgrat wird mit diesen Parametern nicht charakterisiert. Bei Annahme konstanter Bindungslängen und Valenzwinkel kann die innere Flexibilität des Rückgrats durch Torsionen um die chemischen Einfachbindungen beschrieben werden. Die Rotationsfreiheitsgrade eines Nucleotids sind in Abb. 2.4 schematisch dargestellt.

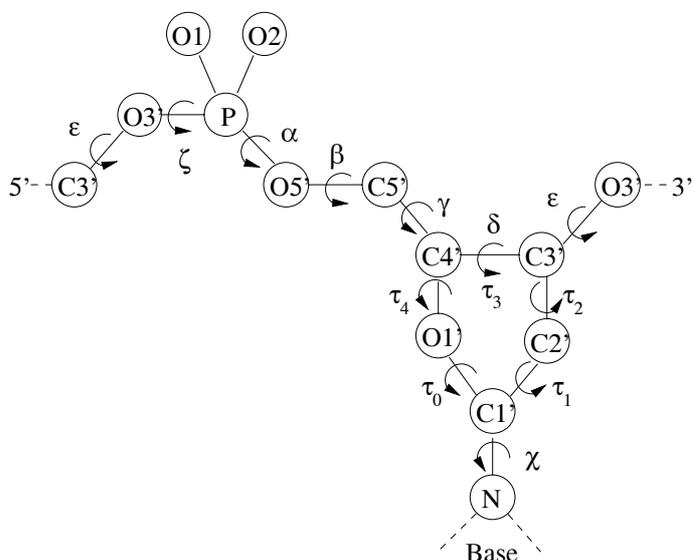


Abb. 2.4.: Torsionswinkel eines Nucleotids

um die glykosidische Bindung C1'-N und beschreibt damit die Orientierung der Base relativ zum Zuckerring. Der Glykosidwinkel  $\chi$  nimmt Werte in zwei Bereichen an, die sterisch verschiedenen Konformationen entsprechen, in denen die Base entweder in der anti-Konformation vom Zucker weg gerichtet ist oder in der syn-Konformation oberhalb des Zuckers liegt und zum Zucker hin gerichtet ist [6].

Der fünfgliedrige Furanosering bildet aufgrund der  $sp^3$ -Hybridisierung seiner Schweratome keine planare Ebene sondern ist gefaltet. Der Zuckerring kann so gefaltet sein,

Das Rückgrat einer monomeren Einheit besteht aus fünf aufeinanderfolgenden kovalenten Bindungen und einem Pentosering, der aufgrund eines beteiligten Sauerstoffatoms im Fall von Nucleinsäuren immer ein Furanosering ist. Die Konformation dieser sich wiederholenden Einheit des Rückgrats wird durch die sechs endozyklischen Torsionswinkel  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$  und  $\zeta$  außerhalb des Furanoserings und die fünf endozyklischen Torsionswinkel  $\tau_0$ ,  $\tau_1$ ,  $\tau_2$ ,  $\tau_3$  und  $\tau_4$  innerhalb des Furanoserings beschrieben. Außerdem charakterisiert der Torsionswinkel  $\chi$  die Rotation

daß entweder drei Atome (Twist-Form) oder vier Atome (Envelope-Form) in einer annähernd planaren Ebene liegen. Aus dieser Ebene herausgefaltete Atome befinden sich in einer endo- oder exo-Konformation in Abhängigkeit davon, ob sie sich auf derselben Seite der Zuckerebene befinden wie das C5'-Atom [6]. Der Zuckerring ist außerordentlich flexibel und kann eine Vielzahl verschiedener Konformationen annehmen, die durch eine Pseudorotation des Furanoserings analytisch charakterisiert werden können [19]. Dabei wird die Zuckerkonformation durch eine Phase  $P$  und eine Amplitude  $A$  beschrieben, wobei die Phase das Atom identifiziert, das aus der Zuckerebene am weitesten herausgefaltet ist, und die Amplitude als ein Torsionswinkel definiert ist, der den Betrag der Herausfaltung des am weitesten aus der Zuckerebene gefalteten Ringatoms angibt. Die analytischen Beziehungen zwischen den endozyklischen Torsionswinkeln  $\tau_0$ - $\tau_4$  und der Phase  $P$  und Amplitude  $A$  sind in den Formeln 2.10 und 2.11 angegeben [19].

$$\tan P = \frac{(\tau_4 + \tau_1) - (\tau_3 + \tau_0)}{2 \tau_2 (\sin \frac{\pi}{5} + \sin \frac{2\pi}{5})} \quad (2.10)$$

$$A = \frac{\tau_2}{\cos P} \quad \text{mit} \quad P = P + \pi \quad \text{für} \quad \tau_2 < 0 \quad (2.11)$$

Diese Gleichungen gelten approximativ, da ein Austausch der Torsionswinkel  $\tau_i$  gegeneinander aufgrund der unterschiedlichen Längen der C-C- und C-O-Bindungen zu geringen Abweichungen von  $\leq \pm 0,6^\circ$  in Phase und Amplitude führt [19]. Die Phase umschreibt eine Pseudorotation, denn ihr Wertebereich umfaßt das gesamte Intervall von  $P = 0^\circ$ - $360^\circ$ . Die Amplitude erreicht hingegen meist Werte im Bereich von  $A = 35^\circ$ - $45^\circ$ , kann jedoch auch deutlich darüber oder darunter liegen. Aus der Phase und Amplitude lassen sich die endozyklischen Torsionswinkel approximativ berechnen [19].

$$\tau_k = A \cos(P + k \frac{4\pi}{5}) \quad (2.12)$$

Die zwei am häufigsten auftretenden Furanosefaltungen besitzen mit der C2'-endo-Zuckerkonformation eine Phase von  $P = 144^\circ$ - $180^\circ$  und mit der C3'-endo-Zuckerkonformation eine Phase von  $P = 0^\circ$ - $36^\circ$  [19]. Die C2'-endo-Konformation ist der typische Zuckerfaltungsmodus der B-DNA, wohingegen die C3'-endo-Konformation für die A-DNA und die RNA charakteristisch ist [6]. In einer DNA sind diese beiden Zuckerfaltungskonformationen energetisch annähernd gleichwertig, wobei die C2'-endo- und C3'-endo-Zuckerkonformation durch eine relativ geringe Energiebarriere voneinander getrennt sind [20]. In der RNA ist aufgrund einer zusätzlich an der Furanose vorhandenen Hydroxylgruppe aus sterischen Gründen nur die C3'-endo-Zuckerkonformation möglich [21]. Diese Hydroxylgruppe führt jedoch zu einem zusätzlichen Freiheitsgrad, dem Torsionswinkel  $\rho$ , der die Rotation der Hydroxylgruppe um die C2'-O2'-Bindung beschreibt. Da die Zuckerfaltung auch den Abstand der Phosphatgruppen der beiden miteinander gepaarten Nucleinsäurestränge bestimmt, hat sie einen wesentlichen Einfluß auf die Konformation der Nucleinsäuren.

## 2.3. Ergebnisse experimenteller Methoden

### 2.3.1. Faserdiagramme

Eine strukturelle Beschreibung der unterschiedlichen Konformationen von Nucleinsäurehelices erfolgte erstmals meist durch die Interpretation von Faserdiagrammen, die durch Röntgenbeugung an quasikristallinen Nucleinsäurefasern gewonnen worden sind. Aufgrund der fehlenden Ordnung der Moleküle in der quasikristallinen Faser bleibt die Auflösung der Röntgenbeugungsbilder wegen Überlagerungen auf ca. 3 Å beschränkt. Obgleich deshalb eine Berechnung der Elektronendichte nicht möglich ist, kann die Topologie der Doppelhelix mit Faserdiagrammen experimentell aufgelöst werden. Arnott et al. konnten mit dieser Methode die helikalen Eigenschaften der A- und B-Form [4, 22, 23], der D-Form [7], der Z-Form [24] und der Tripelhelix [9] der DNA sowie der A-ähnlichen Formen der RNA [8] beschreiben. Aus Faserdiagrammen abgeleitete Struktur-Parameter der A-, B- und Z-Form der DNA-Doppelhelix sind in Tab. 2.1 angegeben.

Struktur-Parameter	A-DNA	B-DNA	Z-DNA	
			<sup>1</sup> C/ <sup>2</sup> CpG	<sup>1</sup> G/ <sup>2</sup> GpC
X-Displacement <sup>1</sup> [Å]	-5,3	0,0	-2,4	-1,4
Y-Displacement <sup>1</sup> [Å]	0,0	0,0	-2,4	2,5
Rise <sup>2</sup> [Å]	2,6	3,4	3,1	4,4
Inclination <sup>1</sup> [°]	20,7	1,5	9,6	3,5
Tip <sup>1</sup> [°]	-3,8	-6,6	-177,8	177,6
Twist <sup>2</sup> [°]	32,7	36,0	-58,8	-1,2
Weite der großen Furche [Å]	2,9	11,3	15,5...9,1	15,5...11,0
Tiefe der großen Furche [Å]	10,3	3,9	0,5...-4,2	0,4...-4,6
Weite der kleinen Furche [Å]	10,7	6,1	4,3...4,2	4,4...1,9
Tiefe der kleinen Furche [Å]	0,0	5,7	4,9...4,8	4,9...4,0
$\epsilon$ (C4'-C3'-O3'-P) <sup>1</sup> [°]	175	-133	-72	-103
$\zeta$ (C3'-O3'-P-O5') <sup>1</sup> [°]	-45	-157	102	-91
$\alpha$ (O3'-P-O5'-C5') <sup>1</sup> [°]	-90	-41	52	-110
$\beta$ (P-O5'-C5'-C4') <sup>1</sup> [°]	-149	136	-153	-168
$\gamma$ (O5'-C5'-C4'-C3') <sup>1</sup> [°]	47	38	178	54
$\delta$ (C5'-C4'-C3'-O3') <sup>1</sup> [°]	83	139	76	147
$\chi$ (O1'-C1'-N-C2/C4) <sup>1</sup> [°]	-154	-102	89	-159
Zuckerfaltungsmodus <sup>1</sup>	C3'-endo	C2'-endo	C2'-endo	C3'-endo

Tabelle 2.1.: Aus Faserdiagrammen von Arnott et al. [24] gewonnene Struktur-Parameter der A-, B- und Z-DNA: Helikale Parameter [13, 14], Furchengeometrien [18] und Torsionswinkel [24]

Bei der Bewertung der Parameter muß allerdings berücksichtigt werden, daß diese Daten historisch zwar von großer Bedeutung und beim Vergleich der A-, B-, und Z-Form auch durchaus instruktiv sind, daß aber heute die Strukturaufklärung viel größere Genauigkeiten erreicht, wodurch es gerade bei den Torsionswinkeln zu deutlichen Abweichungen von den in Tab. 2.1 angegebenen Werten kommen kann.

Ein Vorteil der Faserdiagramme besteht darin, daß Konformationsänderungen als Funk-

tion der Konzentration von Wasser und Salz untersucht werden können. Sequenzeffekte sind aber nur in begrenztem Maße zugänglich, da die Röntgenbeugungsbilder über die gesamte Sequenz sehr langer DNA-Fasern mitteln. Faserdiagramme ermöglichen somit keinen Zugang zu atomar aufgelösten Strukturdaten und deren Abhängigkeit von der Basensequenz.

### 2.3.2. Röntgenkristallographie

Im Gegensatz zu Faserdiagrammen quasikristalliner Polymere kann die Röntgenkristallographie die Strukturen von Makromolekülen mit einer weit höheren Auflösung von bis zu 1 Å oder teilweise darunter beschreiben, da die Moleküle auf einem kristallinen Gitter vollkommen regelmäßig angeordnet sind und somit zu Röntgenbeugungsbildern führen, die eine weitgehend exakte Berechnung der Elektronendichte für alle Atome des Moleküls zulassen [25]. Mit Hilfe der sehr leistungsfähigen Methode der Röntgenkristallographie konnten viele interessante Nukleinsäurestrukturen wie die Z-DNA [5], das EcoRI Dodecamer [26], die A-DNA [27, 28] und eine Reihe von RNA-Strukturen [29] bis hin zur Struktur des Ribosoms [30, 31] aufgeklärt werden, wobei die erreichten Auflösungen bei gleichzeitiger Limitierung durch immer größere Makromoleküle und Molekülverbände stets verbessert werden konnten. Neben den in Kap. 2.5 zu besprechenden DNA-Protein-Komplexen sollen die Fortschritte bei der Strukturaufklärung kurzlebiger, während chemischer Reaktionen auftretender Moleküle mittels der zeitaufgelösten Kristallographie [32] nicht unerwähnt bleiben, wobei diese Methodik allerdings keinen Zugang zur Analyse der Flexibilität von Makromolekülen bietet.

Die Röntgenkristallographie ermöglicht die Untersuchung auch kleiner Variationen von Nukleinsäurestrukturen auf der Ebene individueller Basenpaare. Dadurch wurde die Aufklärung und Beschreibung von Sequenzeffekten [33, 34, 35, 36] und von Einflüssen des umgebenden Mediums auf die Konformation [37] erst möglich, wodurch die Röntgenkristallographie eine zentrale Bedeutung gewinnt. Als ein wesentliches Ergebnis einer Vielzahl röntgenkristallographischer Untersuchungen von Nukleinsäurestrukturen kann der Nachweis dafür angesehen werden, daß die Konformationen von Nukleinsäuren stark von der Basensequenz abhängig sind [33, 34, 35, 36]. Die helikalen Parameter der 10 verschiedenen Dinukleotid-Schritte aller mit Hilfe der Röntgenkristallographie gelösten DNA-Strukturen sind statistisch ausgewertet worden und zeigen nochmals deutlich die Sequenzabhängigkeit der DNA-Konformationen [38, 39].

Neben den bedeutenden durch die Röntgenkristallographie gewonnenen Erkenntnissen über die 3-dimensionale Struktur von Nukleinsäuren und Proteinen muß allerdings erwähnt werden, daß die Methodik in der Regel auf statische, im kristallinen Zustand vorliegende Strukturen beschränkt ist. Bei der Auswahl der zu untersuchenden Strukturen kann eine Tendenz zu leicht kristallisierbaren Molekülen entstehen. Außerdem können Artefakte durch Kristallwechselwirkungen nicht ausgeschlossen werden.

### 2.3.3. Spektroskopische Methoden

Die Bedeutung der Spektroskopie besteht darin, nicht auf einen kristallinen Zustand beschränkt zu sein, sondern in vielen Fällen *in vivo*, in Lösung und im Kristall anwendbar zu sein und so einerseits einen Zugang zur Dynamik und Flexibilität der Strukturen und andererseits einen direkten Vergleich mit röntgenkristallographischen

Daten zu ermöglichen. Damit ergänzen die spektroskopischen Methoden zur Strukturaufklärung die Röntgenkristallographie komplementär.

### Optische Spektroskopie

Aufgrund ihrer konjugierten  $\pi$ -Elektronensysteme sind die Basen der Nukleinsäuren Chromophore, deren relative Lage zueinander mit Methoden der optischen Spektroskopie untersucht werden kann. Mit der Absorptions- und Fluoreszenzspektroskopie sind Effekte der Dimerisierung zweier oder mehrerer heterozyklischer  $\pi$ -Elektronensysteme zugänglich, so daß die Spindelwechselwirkungen der Basen und auch Interkalationen planarer heterozyklischer Liganden durch Farbverschiebungen dieser Spektren beobachtet werden können [40]. In enger Beziehung zu den direkten Messungen der elektronischen Übergänge  $S_0 \rightarrow S_1$  (Absorption) und  $S_1 \rightarrow S_0$  (Fluoreszenz) stehen der Lineare und Circulare Dichroismus, die die Richtungsabhängigkeit der Absorption heterozyklischer Moleküle ausnutzen. Die anisotrope Absorption linear polarisierten Lichtes wird als Linearer Dichroismus (LD) bezeichnet. Der LD ermöglicht die Messung der Inklination der Basen relativ zur Helixachse, wobei auch dynamische Fluktuationen meßbar sind [41]. Es können sowohl die A-, B- und Z-Formen der DNA voneinander unterschieden als auch Krümmungen und Längenänderungen von Nukleinsäurestrukturen gemessen werden [41].

Der Circulare Dichroismus (CD) entsteht durch die Wechselwirkung elektrischer Übergangsmomente benachbarter Chromophore, wodurch die Orientierung der Chromophore zueinander gemessen werden kann, oder durch die Beteiligung eines magnetischen Übergangsmomentes an einem elektronischen Übergang. Durch chirale Gruppen wie dem Zuckerring entstehen Wechselwirkungen mit den Übergangsmomenten der benachbarten Base, wodurch die Orientierung der Base relativ zum Zuckerring, also der Glykosidwinkel, meßbar wird [41]. Außerdem führt die Wechselwirkung der Übergangsmomente benachbarter Basen zu einem CD-Signal, wodurch eine rechtsgängige sehr deutlich von einer linksgängigen Helix unterschieden werden kann. Darüberhinaus kann das CD-Signal Aufschluß über die Orientierung der Übergangsmomente interkalierter Chromophore relativ zu den direkt benachbarten Basen geben [41].

### Schwingungs-Spektroskopie

Intramolekulare Schwingungen können mit Hilfe der Infrarot (IR)-Spektroskopie oder der Raman-Spektroskopie untersucht werden, wobei beiden Methoden unterschiedliche Mechanismen und Auswahlregeln zugrunde liegen. Die Infrarot-Spektroskopie beruht auf der Absorption schwingender Dipole. Die Raman-Spektroskopie beruht auf der Erzeugung und Vernichtung von Phononen infolge von Lichtstreuung. Zu den IR-Absorptionsspektren tragen Schwingungen asymmetrisch gebundener Atome bei, zur Raman-Spektroskopie hingegen die Schwingungen elektronenreicher Gruppen in den heterozyklischen Basen und den Phosphatgruppen des Rückgrats [42]. Der wesentliche Nachteil der IR-Spektroskopie besteht in der Absorption des Wassers im infraroten Spektralbereich, so daß Messungen nur in  $D_2O$  durchgeführt werden

können. Die Raman-Spektroskopie ist *in vivo*, in Lösung und im Kristall anwendbar [42], wobei große Moleküle oder Molekülverbände untersucht werden können und nur sehr geringe Probenmassen notwendig sind. Intermolekulare Wechselwirkungen sowie die Flexibilität und Dynamik molekularer Strukturen können hiermit untersucht werden [42].

Ein Nachteil der bisher besprochenen spektroskopischen Methoden besteht darin, daß die spektralen Peaks nicht immer eindeutig zugeordnet werden können, da die Spektren von kleinen Referenzmolekülen, aus denen das jeweilige Makromolekül aufgebaut ist, oft nicht zur Verfügung stehen.

### **NMR-Spektroskopie**

Die auf atomarem Niveau weitreichendsten Strukturdaten aller spektroskopischen Methoden liefert die Nuclear Magnetic Resonance (NMR)-Spektroskopie, die die Möglichkeit der Messung bestimmter Atomabstände bietet [43, 44]. Dabei können aus den chemischen Verschiebungen der Wasserstoffatome sehr genaue Angaben über die Geometrie der Basenpaarung und der Basenstapelung gewonnen werden [6, 43]. Außerdem sind Informationen über die Zuckerfaltung als auch über bestimmte Torsionswinkel zugänglich [6]. Eine Reihe von NMR-Untersuchungen haben die Sequenzabhängigkeit von Nukleinsäurestrukturen bestätigen können [43, 45]. Eine Analyse der sequenzspezifischen Geometrie von Basenpaaren und Basenpaarschritten ist mit NMR-Daten möglich [45]. Die NMR-Spektroskopie findet ihre Grenzen in der Nichtzugänglichkeit bestimmter Strukturparameter aufgrund z.B. fehlender Wasserstoffatome in der direkten Umgebung. Außerdem sind durch die NMR-Spektroskopie zwar auch nicht kristallisierbare Strukturen auflösbar, nur wird für die Messungen eine im Vergleich zu den anderen Methoden sehr hohe Probenmasse benötigt. NMR-Messungen *in vivo* sind nicht möglich.

## **2.4. Ergebnisse aus Theorie und Simulationen**

Die aus den Experimenten gewonnenen Erkenntnisse über makromolekulare Strukturen sind trotz ihrer Bedeutung meist begrenzt auf einen bestimmten Aggregatzustand oder nicht eindeutig interpretier- bzw. erklärbar. Außerdem sind Informationen über die Dynamik und Flexibilität der Strukturen auf atomarem Niveau in der Regel nicht verfügbar. Aus diesen Gründen bilden die theoretischen Methoden eine zweite Säule der Strukturbiologie. Mit diesen Methoden können einerseits Strukturen durch ihre energetische Analyse erklärt werden und andererseits Informationen über ihre Flexibilität auf atomarem Niveau gewonnen werden. Im folgenden werden einige für Nukleinsäurestrukturen wesentliche Ergebnisse der Theorie bzw. des Molecular Modeling vorgestellt, im Kap. 3 wird in die wesentlichen theoretischen Methoden eingeführt.

### **2.4.1. Die DNA als Polyelektrolyt**

Die DNA besitzt aufgrund ihrer elektrisch negativ geladenen Phosphatgruppen polyanionische Eigenschaften, die durch elektrisch positiv geladene Gegenionen in einer diese Ionen bereitstellenden Umgebung neutralisiert werden. Die DNA-Konformation reagiert empfindlich auf die Verteilung und Konzentration dieser Ionen wie auch des

übrigen Lösungsmittels.

Manning beschrieb die Kompensation der negativen Phosphatladungen durch Kondensation eines Teils der in Lösung befindlichen Natrium- oder Magnesium-Ionen an der DNA-Oberfläche. Dabei beschrieb er die DNA als unendlich langen, geladenen Draht, der sich in einem Lösungsmittelkontinuum mit konstanter Dielektrizitätskonstanten befindet [46]. Durch Minimierung der elektrostatischen Freien Energie gelangte Manning zu dem Resultat, daß 24% der Natrium- bzw. 12% der Magnesium-Ionen unabhängig von der Salzkonzentration frei beweglich bleiben. Somit kondensieren 76% der monovalenten bzw. 88% der divalenten Ionen an der DNA [47]. Der Radius des Zylinders um die DNA, in dem sich die kondensierten Ionen befinden, liegt ca. 7 Å außerhalb der DNA-Oberfläche und beträgt somit aufgrund der Annahme einer B-Konformation ca. 17 Å ausgehend von der Helixachse [47]. Durch Monte-Carlo-Simulationen unter Berücksichtigung der atomaren DNA-Struktur und bei Verwendung verschiedener Modelle einer adäquateren Beschreibung des Lösungsmittels konnten Jayaram et al. diese Resultate qualitativ bestätigen [48, 49].

### 2.4.2. Statische Modelle

#### Energieminimierungen

Molekül-Konformationen können durch in Kap. 3.2.1 beschriebene Energieminimierungen auf der Potentialhyperfläche des Moleküls energetisch untersucht werden. Die durch die Energieminimierungen gefundenen, lokalen Minima entsprechenden Konformationen ermöglichen eine detaillierte, in Kap. 2.2 beschriebene, strukturelle Analyse. Auch lassen sich durch das Studium dieser energetischen Minima entlang vorgegebener Pfade im Konformationsraum der Moleküle mögliche Konfigurationsübergänge untersuchen. Im Vorfeld dieser Arbeit wurden auf diese Weise Auffaltungswege und Aggregationswechselwirkungen von Proteinsekundärstrukturelementen untersucht [50].

Nucleinsäurestrukturen betreffend, wurden eine Reihe lokaler Konformationsänderungen durch Energieminimierungen energetisch und strukturell untersucht [51]. Beispielsweise berichteten Kollman et al., daß Nucleotide mit Pyrimidinbasen aus energetischen Gründen leichter ihre Zuckerkonformation wechseln können als Nucleotide mit Purinbasen [52].

Die Dynamik der Wechselwirkungen von Nucleinsäuren mit Liganden ist auch heute noch weitgehend ungeklärt, so daß die statischen Modelle für das Verständnis der Ligandenbindung eine nach wie vor große Bedeutung besitzen. Ein Beispiel ist die im Rahmen dieser Arbeit untersuchte Bindung von Methyleneblau an die DNA [53].

#### Berechnung Freier Energien

Konformationen können einer Analyse der Freien Energie 3.12 unterzogen und dadurch energetisch erklärt werden. Stabilisierende und destabilisierende Komponenten der Freien Energie werden berechnet und quantitativ miteinander verglichen. Dieser Vergleich kann zwischen verschiedenen Konformationen eines Moleküls oder auch zwischen verschiedenen, dieselbe Konformation umgebenden Lösungsmitteln erfolgen [54]. Beim Vergleich der A- und B-DNA in Wasser und Ethanol stellen sich die elektrostatischen Wechselwirkungen, und dabei sowohl die innere elektrostatische Energie der

DNA als auch die Wechselwirkung mit dem Lösungsmittel und den Ionen, als die dominierende Komponente der Freien Energie heraus [54, 55]. Bei der Untersuchung der Proteinfaltung [56] wie auch der Komplexbildung der DNA mit Proteinen [57, 58] hat sich die Analyse der Freien Energie als hilfreich beim Verständnis von Strukturbildungen erwiesen.

### 2.4.3. Flexible Elemente in Nukleinsäurestrukturen

Die Flexibilität von Nukleinsäurestrukturen entsteht durch thermische Schwingungen der Atome um ihre Gleichgewichtslagen oder die für Konformationsänderungen nötige kollektive Bewegung ganzer Atomgruppen. Zu diesen kollektiven Bewegungen gehören Translations- und Rotationsbewegungen individueller Basen, die Umfaltung des Zuckerrings oder Rotationen von Atomen um kovalente Bindungen innerhalb des Nukleinsäure-Rückgrats. Kollektive Bewegungen ereignen sich aus sterischen Gründen im Vergleich zu atomaren Schwingungen sehr selten. Außerdem gibt es eine Reihe von Bewegungen auf noch größerer Zeitskala. Beispiele dafür sind die Bildung und das Aufbrechen von Wasserstoffbrücken, das Supercoiling oder die Aufwindung rechtsgängiger und Bildung linksgängiger DNA. Für die Erklärung biologischer Prozesse ist allerdings die Beschreibung und das Verständnis der Dynamik und Flexibilität der Nukleinsäuren essentiell.

#### Flexibilität des Zucker-Phosphat-Rückgrats

Die möglichen Konformationsänderungen des Zuckerrings und der Torsionswinkel des Rückgrats einschließlich des Glykosidwinkels sind in Kap. 2.2.3 besprochen worden. Der Beitrag der Theorie besteht zunächst in der energetischen Analyse dieser Konformationsänderungen. Für den Übergang der Zuckerfaltungskonformation von der C2'-endo- zur C3'-endo-Konformation ist von Levitt und Warshel eine Energiebarriere von nur 0,6 kcal/mol angegeben worden [20], wohingegen Kollman et al. für diese Umfaltung des Zuckerrings sequenzabhängige Energiebarrieren von 2 bis 4 kcal/mol angeben [52]. In Untersuchungen der Dynamik dieses Übergangs erfolgten die Übergänge immer relativ schnell innerhalb 1 ns und über die O1'-endo-Konformation [59].

Neben der flexiblen Zuckereinheit fluktuieren die Werte der in Kap. 2.2.3 eingeführten Torsionswinkel des Nukleinsäure-Rückgrats. Die Fluktuationen der Torsionswinkel werden mit 6°-21° angegeben [59]. Bei der Flexibilität des DNA-Rückgrats infolge der Fluktuation seiner Torsionswinkel sind eine Reihe von Korrelationen bekannt [6].

#### Korrelationen der Torsionswinkel

Aus sterischen Gründen sind zunächst die Änderungen benachbarter Torsionswinkel miteinander korreliert, die als Wechselwirkung erster Ordnung bezeichnet wird [21]. Wesentlicher sind allerdings die Korrelationen zwischen Torsionswinkeln, die eine kovalente Bindung voneinander entfernt sind. Diese als Wechselwirkungen zweiter Ordnung bezeichneten Korrelationen sind im Modell als Kurbelwelle gut vorstellbar, denn es erfolgen beispielsweise Änderungen dieser miteinander korrelierten Torsionswinkelpaare von der (-gauche,+gauche)-Konformation in die (trans,trans)-Konformation [21, 59], die auch als Flip bezeichnet werden. Solche Korrelationen mit negativem Korrelationskoeffizienten sind für die Torsionswinkelpaare  $(\alpha,\gamma)$  und  $(\delta,\zeta)$  beobachtet worden. Im

Gegensatz zu diesen Wechselwirkungen wird die Korrelation der Torsionswinkel  $(\beta, \delta)$  durch einen positiven Korrelationskoeffizienten beschrieben [59]. Weitere betrachtete Korrelationen bestehen in den Torsionswinkelpaaren  $(\alpha, \beta)$ ,  $(\beta, \varepsilon)$  und  $(\varepsilon, \zeta)$  [6] sowie verschiedene Korrelationen mit  $\chi$  und andere Korrelationen weiter voneinander entfernter Torsionswinkel [60].

Charakteristisch für Nucleinsäurestrukturen ist der  $B_I \rightarrow B_{II}$ -Übergang, der die Korrelation zwischen  $\varepsilon$  und  $\zeta$  beschreibt. Bei diesem Übergang wechseln die Torsionswinkel  $(\varepsilon, \zeta)$  ihre Konformationen von (trans,-gauche) zu (-gauche,trans) [6].

Der Torsionswinkel  $\delta$  ist überdies mit der Zuckerkonformation derart korreliert, daß die mögliche Fluktuation von  $\delta$  in der C3'-endo-Konformation deutlich geringer als in der C2'-endo-Konformation ist. Somit führt auch umgekehrt eine Änderung in der Phase des Zuckers in der C3'-endo-Konformation zu einer geringeren Änderung von  $\delta$  als in der C2'-endo-Konformation [61]. Außerdem ist über den Zuckerring der Glykosidwinkel  $\chi$  mit  $\delta$  korreliert [6].

### 2.4.4. Dynamische Modelle

Die Kenntnis der Dynamik und Flexibilität makromolekularer Strukturen ist für das Verständnis biologischer Prozesse von großer Bedeutung. Lokalen Minima entsprechende Übergangszustände lassen sich durch Energieminimierung zwar studieren, die tatsächliche Dynamik eines Konformationsüberganges ist damit allerdings nicht aufgeklärt. Molekulardynamik(MD)-Simulationen ermöglichen hingegen das Studium von Trajektorien, die Aufschluß darüber geben, inwiefern lokale Konformationsänderungen zu Konformationsübergängen führen. Während Energieminimierungen die Energetik von Strukturen untersuchen, beruht die in Kap. 3.2.2 erläuterte Methodik der MD-Simulationen auf der Integration der Bewegungsgleichungen. Auch wenn in dieser Arbeit DNA-Strukturen im Vordergrund stehen, sei auf die Bedeutung dynamischer Modelle bei der Untersuchung der Proteinfaltung verwiesen [56, 62].

### Resultate aus Molekulardynamik-Simulationen

Levitt berichtete über die ersten MD-Simulationen von Nucleinsäurestrukturen, die vor der Einführung expliziter Wassermoleküle und Ionen noch im Vakuum durchgeführt wurden [63]. Die Dynamik von Wasserstoffbrückenbindungen, die Umfaltung der Zuckerringe und mit  $B_I$ - $B_{II}$ -Übergängen und  $\alpha/\gamma$ -Flips die Flexibilität des Rückgrats konnten in MD-Simulationen untersucht werden [64]. Aber erst seit Mitte der neunziger Jahre wurden aufgrund besserer Kraftfelder, wachsender Rechnerleistungen und akkuraterer Methoden zur Lösungsmittelbeschreibung stabile MD-Simulationen der Feinstrukturen von Nucleinsäuren wie Sequenzeffekten, der Abhängigkeit der Strukturen von den Umgebungseinflüssen und der Rolle spezifischer Wasser- und Ionen-Wechselwirkungen möglich [55, 64].

In einer MD-Simulation über  $> 5$  ns des EcoRI Dodecamers mit Ionen in Lösung blieb die Konformation erstmals nah an der kanonischen B-Form, und die strukturellen Parameter zeigten in der MD-Simulation eine Sequenzspezifität [65]. Abhängig von der Basensequenz konnten in MD-Simulationen spezifische Krümmungen der DNA beobachtet werden [66, 67]. Die Sequenzabhängigkeit der Konformation konnte ebenfalls

bei der MD-Simulation einer A-Form-Duplex reproduziert werden, wobei die MD-Simulationen eine geringere Flexibilität der A- im Vergleich zur B-Form zeigten [68]. Konformationsänderungen der Rückgrat-Torsionswinkel wie der  $B_I \rightarrow B_{II}$ -Übergang oder auch Zuckermalfaltungen konnten in MD-Simulationen über Zeitskalen von 5 bis  $> 600$  ps beobachtet werden [64, 69]. Cheatham und Kollman gelang es 1996 erstmals, einen A $\rightarrow$ B-Übergang in einer MD-Simulation über  $\approx 500$  ps nachzuvollziehen [69]. Zwar zeigten die MD-Simulationen verschiedener Sequenzen keinen konzertierten Ablauf der für den A $\rightarrow$ B- oder B $\rightarrow$ A-Übergang nötigen lokalen Konformationsänderungen, allerdings nahmen beim B $\rightarrow$ A-Übergang zunächst das X-Displacement sowie der Twist und erst später der Rise und die Inclination A-ähnliche Werte an, wenn der Zucker in die C3'-endo-Konformation gezwungen wurde [64].

Umgebungseinflüsse auf die MD-Simulationen wurden untersucht, indem die im Wasser erzielte MD-Simulation des spontanen A $\rightarrow$ B-Übergangs [69] in Ethanol wiederholt wurde und zur Stabilisierung der A-Form führte [70]. Zur Erklärung von Umgebungseinflüssen wurde die Lokalisation von Lösungsmittelmolekülen und Ionen an der DNA diskutiert [55, 64, 71], der Einfluß von Wassermolekülen und Ionen in den Furchen miteinander verglichen [72] oder die Wirkung verschiedener Lösungsmittel auf Wechselwirkungen wie die Basenstapelung untersucht [73]. Bei der MD-Simulation gekrümmter Nukleinsäurestrukturen wiesen diese Konformationen in ihren Krümmungszentren lokalisierte Ionen auf [72].

### Erfolge und Grenzen von MD-Simulationen

Molekulardynamik-Simulationen ermöglichten erstmals realistische Simulationen der Flexibilität von Makromolekülen auf der Basis aller Atome und unter Einschluß von Ionen und Lösungsmittelmolekülen, so daß MD-Simulationen bei immer besseren Beschreibungen von Lösungsmittel- und Salzeffekten den Übergang in eine Zeit prägten, die durch relevante Beiträge theoretischer Studien zur Strukturbiologie gekennzeichnet ist. MD-Simulationen bestätigen die Abhängigkeit der Konformationen von der Basensequenz [55, 64]. Umgebungseinflüsse werden ebenfalls in MD-Simulationen beobachtet. Die Resultate der MD-Simulationen stimmen mit Experimenten gut überein, obgleich die Strukturen im Vergleich zu Experimenten ein Unwinding aufweisen [55, 64]. Dieses ist auf einen mittleren Twist benachbarter Basenpaare zurückzuführen, der um  $3^\circ$ - $4^\circ$  gegenüber experimentellen Daten verringert ist [72].

Einschränkend muß auch erwähnt werden, daß heute zwar Simulationen über Zeitskalen von  $> 15$  ns Standard geworden sind, eine ausreichende Equilibrierung der Systeme in diesen Zeitintervallen allerdings nicht stattfindet. Beispielsweise sind Ionen im Lösungsmittel in ihrer Beweglichkeit stark eingeschränkt, und strukturelle Ereignisse finden nur vereinzelt statt [72]. Da selbst durch heutige Simulationen, die hunderte Stunden auf Supercomputern erfordern [72], eine Equilibrierung der Systeme nicht möglich scheint, findet die Entwicklung einer alternativen Simulationstechnik ihre Rechtfertigung.

## 2.5. Komplexbildung mit Liganden und Proteinen

Wechselwirkungen biologischer Makromoleküle sind die Grundlage physiologischer Prozesse. Dazu gehören Wechselwirkungen von Enzymen mit ihren Substraten, von Prote-

inen mit Liganden, von Nucleinsäuren mit Proteinen und kleineren Liganden und auch zwischen Makromolekülen und Ionen [74]. Da im Rahmen dieser Arbeit die Komplexbildung zwischen der DNA und einem Farbstoff-Molekül untersucht wurde, wird im folgenden auf die Komplexbildung von DNA-Molekülen mit Liganden näher eingegangen.

### 2.5.1. DNA-Liganden-Komplexe

Eine Reihe kleinerer Moleküle mit meist heterozyklischen  $\pi$ -Elektronensystemen kann mit DNA-Molekülen Komplexe bilden und dadurch zur Veränderung biologischer Eigenschaften der Nucleinsäuren führen. Als Liganden können Farbstoffe, Mutagene, Karzinogene, Antibiotika und andere Pharmaka wirksam werden [74]. Die Komplexbildung mit der DNA kann durch drei verschiedene Bindungsmodi erfolgen: (I) durch Interkalation des Liganden zwischen zwei benachbarte Basenpaare, (II) durch Bindung des Liganden in der kleinen Furche oder (III) in der großen Furche [6, 53]. Mögliche Kandidaten für die Interkalation sind dabei alle heterozyklischen planaren Moleküle wie beispielsweise Ethidiumbromid, Proflavin, Acridinorange, Aminoacridin, Actinomycin [6, 74, 75] oder das in dieser Arbeit untersuchte Methylenblau [53]. Interkalierende Liganden können in Abhängigkeit von der Basensequenz der DNA oder des umgebenden Mediums oft auch in den Furchen an die DNA binden [53, 76], wobei es allerdings Liganden wie beispielsweise Netropsin gibt, das bevorzugt in der kleinen Furche bindet und durch seine kationische Ladung die polyanionische Wirkung des DNA-Rückgrats teilweise kompensiert [75].

Die Interkalation eines Liganden zwischen zwei benachbarte Basenpaare führt zu einer Veränderung der helikalen DNA-Struktur: das DNA-Rückgrat ist an der Interkalationsstelle gestört, was sich durch einen  $\alpha/\gamma$ -Flip ausdrücken kann. Außerdem kommt es durch die Interkalation zu einem Unwinding der DNA [6, 53]. Der interkalierende Ligand führt zu einer Verringerung des Twists zwischen den benachbarten Basenpaaren an der Bindungsstelle um ca.  $26^\circ$ , so daß sich an der Interkalationsstelle ein Twist von ca.  $10^\circ$  einstellt [53, 77]. Außerdem führt die Interkalation zu einem Längenzuwachs des Targetmoleküls, da sich der Rise an der Bindungsstelle annähernd verdoppeln muß, um die Interkalation des Liganden zu ermöglichen [6, 53]. Röntgenkristallographische und NMR-Untersuchungen verschiedener Interkalationskomplexe bestätigten eine Umfaltung des Zuckers der Nucleotide mit Purinbasen an der Interkalationsstelle auf C3'-endo, während der Zucker der Nucleotide mit Pyrimidinbasen in der C2'-endo-Konformation verblieb [74, 77]. Durch diese Umfaltung veränderten sich auch die Werte der Torsionswinkel  $\beta$  und  $\chi$  charakteristisch [6, 77].

Liganden, die in den Furchen der DNA binden, führen oft zu einer Verengung der Furche, in der sie gebunden sind, und zur entsprechenden Weitung der anderen Furche. Analysiert worden ist dieses charakteristische Verhalten beispielsweise für die Liganden DAPI [78], Hoechst 33258 [79] und Methylenblau [53].

Die Struktur der DNA-Liganden-Komplexe ist in manchen Fällen röntgenkristallographisch [6, 77] oder durch NMR-Messungen [74, 80] untersucht worden. Aufgrund ihrer heterozyklischen  $\pi$ -Elektronensysteme können die Wechselwirkungen interkalierender Liganden mit der DNA durch Methoden der optischen Spektroskopie untersucht werden [40, 81]. Oft ermöglichen erst theoretische Methoden eine detaillierte strukturelle Analyse [53, 75] und energetische Erklärung von Konformationen der DNA-

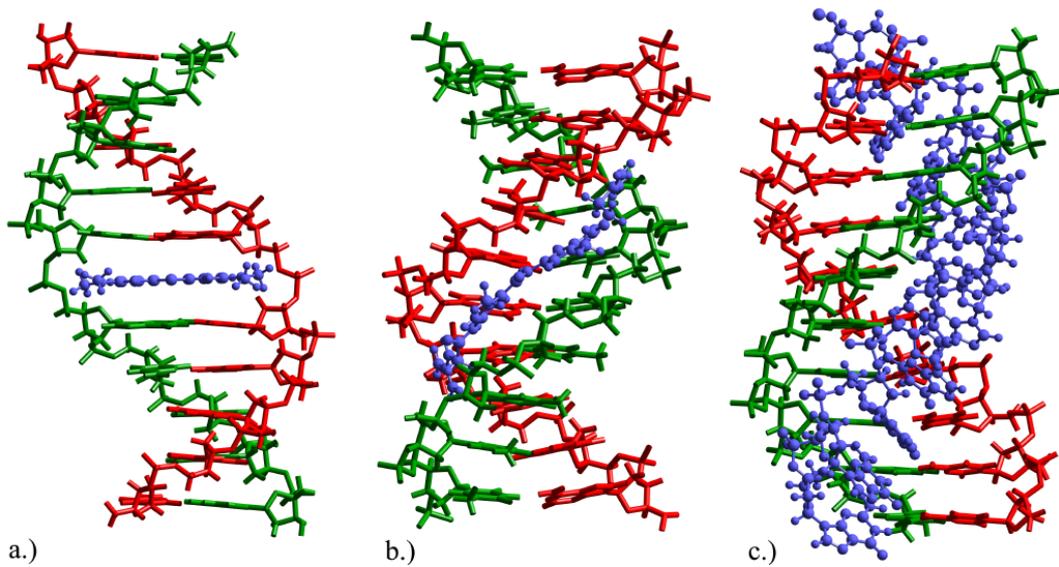


Abb. 2.5.: Konformationen a.) eines Methylenblau-DNA-Komplexes (Interkalation des Methylenblau), b.) eines Netropsin-DNA-Komplexes (Netropsin in der kleinen Furche) und c.) einer DNA-Tripelhelix (DNA-Einzelstrang in der großen Furche der Doppelhelix).

Liganden-Komplexe auf atomarem Niveau [53, 79, 82]. Die Komplexbildung mit Liganden wird bisher meist in statischen Modellen untersucht, durch die Berechnung freier Energien [83] oder im Rahmen dieser Arbeit zunächst durch Energieminimierungen [53, 76]. Dynamische Modelle beschränken sich bei der Untersuchung von DNA-Liganden-Komplexen auf die Berechnung von Normalmoden [75] oder harmonischen Moden [79]. MD-Simulationen dieser Komplexe wurden bisher nur mit sehr kurzen Zeitskalen durchgeführt [84]. Und da die Equilibrierung der Komplexe in MD-Simulationen derzeit nicht erreichbar sein dürfte [84], bietet sich auch für diese Systeme die Anwendung der in dieser Arbeit vorgestellten Simulationstechnik an.

### 2.5.2. DNA-Protein-Komplexe

Die Komplexbildung von Nukleinsäuren mit Proteinen besitzt eine noch weitaus größere biologische Bedeutung [6]. So steuern Mechanismen der molekularen Erkennung die Regulation der Transkription. Bei der molekularen Erkennung der Target-DNA spielen direkte und indirekte Erkennungsmechanismen eine Rolle, wobei direkte Erkennung über spezifische Wechselwirkungen wie Wasserstoffbrücken erfolgt und indirekte Erkennung auf Konformationsänderungen der Targetmoleküle beruht.

Erwähnt werden sollte, daß die Wechselwirkung der DNA mit Proteinen in Komplexen wie beispielsweise dem TBP/TATA-Box-Komplex [85] die Konformation der DNA dramatisch verändern kann [86, 87]. Die indirekte Erkennung der DNA durch Proteine erfordert Veränderungen der DNA-Konformation wie Krümmungen oder lokale Veränderungen struktureller Parameter, die spezifisch von der Basensequenz der Target-DNA abhängig sind. Beispiele dafür sind die Erkennung der Trp-Operator-Bindungsstelle der DNA durch das Repressor-Protein [88], der E2-Bindungsstelle der DNA durch das für die Entstehung von Krebs relevante menschliche Papillomavirus-E2-Protein [89] oder der DNA-Bindungsstelle durch den bei der Leukämie eine Rolle spielenden RUNX1-

Transkriptionsfaktor [90].

Theoretische Methoden können dabei der Untersuchung der Spezifität der Wechselwirkungen innerhalb eines bestimmten DNA-Protein-Komplexes dienen, die bei der direkten Erkennung durch spezifische Wasserstoffbrücken erklärt werden kann [91, 92]. Darüberhinaus ermöglichen theoretische Methoden auch eine energetische Analyse der Stabilisierung von Protein-DNA-Komplexen durch die Berechnung freier Energien der Komplexe [57, 58, 93].

Da die DNA-Protein-Komplexe eine komplizierte Architektur besitzen [6] und Proteine im Gegensatz zu kleineren Liganden durch eine deutlich höhere Zahl innerer Freiheitsgrade beschrieben werden müssen, sollen die theoretischen Simulationen zunächst an Komplexen von DNA-Molekülen mit kleineren Liganden Anwendung finden, auch wenn bereits an Simulationen von DNA-Protein-Komplexen gedacht ist.

### 2.6. Bedeutung und Anwendung der Erkenntnisse

Die theoretischen und experimentellen Erkenntnisse über Konformationen von Nukleinsäuren besitzen eine Bedeutung, die über die Strukturbiologie, deren Ziel das Verständnis der Struktur- und Funktionsbeziehungen von Makromolekülen ist, hinaus und weit in die Biologie, Biochemie und Medizin hineinreicht. So werden mit Hilfe der genannten Erkenntnisse Erfolge im molekularen Verständnis von Krankheiten wie der Entstehung von Krebs [89] und Leukämie [90] erzielt. Bei der Entwicklung neuer Pharmaka wie bestimmter Antibiotika und Antitumor-Wirkstoffe wird durch die Kenntnis der energetischen und strukturellen Grundlagen der Komplexbildung ein zielgerichtetes Design geeigneter Liganden erst möglich [94].

In der molekularen Gentherapie werden Erfolge erzielt, da durch sequenzspezifische Wechselwirkungen Manipulationen am genetischen Material erfolgen können. So können mit Hilfe eines Singulett-Sauerstoff generierenden Farbstoffs die DNA geschnitten, krankhafte Gene entfernt und durch Enzyme repariert werden [95]. Als Beispiel eines auf diese Weise in der photodynamischen Tumortherapie Anwendung findenden Farbstoffs wird in dieser Arbeit die Komplexbildung von Methylenblau mit der DNA untersucht [96, 97]. Methylenblau generiert Singulett-Sauerstoff nach Einwirkung von Laserlicht durch einen Triplett-Triplett-Energietransfer vom als Sensibilisator wirkenden Methylenblau zum molekularen Sauerstoff in einer Typ-II-Photoreaktion [98, 99]. Der Singulett-Sauerstoff kann zu photooxidativen Schädigungen der DNA führen, so daß die gezielte Bindung des Methylenblau an die DNA ein Schneiden der DNA-Stränge ermöglicht [96]. Die genaue Kenntnis der Methylenblau-DNA-Bindung ist Voraussetzung für das Erzielen einer effektiven Singulett-Sauerstoff-Quantenausbeute [40]. Die aus spektroskopischen Untersuchungen bekannte Sequenzspezifität dieser Bindung [100] kann durch die kovalente Kopplung des Farbstoffs an einen DNA-Einzelstrang und die Ausbildung einer Tripelhelix mit der Target-DNA soweit erhöht werden, daß die Idee einer alternativen, für den Patienten weniger risikobehafteten Gentherapie mit Laserlicht erfolversprechend erscheint. Nur sind dafür wie für die gezielte Wirkstoffentwicklung genaue Kenntnisse makromolekularer Strukturen einschließlich ihrer dynamischen Eigenschaften eine essentielle Voraussetzung.