Structure and dynamics of fluorophore-labelled DNA helices probed by NMR-spectroscopy

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

Dr. rer. nat.
im Fach Chemie
eingereicht an derMathematisch-Wissenschaftlichen Fakultät IHumboldt-Universität zu Berlin

von
Dipl.-Chem. André Dallmann
19.09.1980 in Berlin

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Dr. h.c. Christoph Markschies

Dekan der Mathematisch-Wissenschaftlichen Fakultät I:
Prof. Dr. Lutz-Helmut Schön

Gutachter:
1. Prof. Nikolaus P. Ernsting
2. Prof. Christian Griesinger
3. Prof. Clemens Mügge

eingereicht am: 28.10.2009
To my wife

Your love and support wrote this thesis
Abstract

Structural and dynamic perturbations in DNA upon incorporation of either fluorophore, 2-Aminopurine (2AP) or 2-Hydroxy-7-nitrofluorene (HNF), are characterized by NMR spectroscopy. For this purpose the NMR solution structures of the modified DNA duplexes with the sequence 5'-GCTGCAXACGTCG-3' are solved. For X=2AP (13mer2AP) the partner base in the complementary strand is T, while for X=HNF (13merHNF) an abasic site is introduced to avoid steric strain.

By comparing results on 13mer2AP with the corresponding unmodified DNA duplex (13merRef, X=A), any perturbation can be unambiguously assigned to 2AP incorporation. For the NMR solution structure of 13merRef and 13mer2AP small but significant changes in helical parameters are found throughout the helix. Imino proton exchange measurements reveal an extended, distributed effect of 2AP incorporation on the lifetimes of the central seven base pair. However, the reduced base pair lifetime of 2AP:T cannot fully account for the rapid water exchange observed with saturation transfer experiments in the absence of base catalyst. This indicates enhanced intrinsic catalysis. As a possible catalytic site the T O4 atom opposite 2AP is discussed, which is easily accessible through the major groove and lacks a hydrogen bonding partner within the base pair.

The overall NMR solution structure is found to be B-DNA. However the NOE cross-peaks involving the HNF residue can only be accounted for by two different orientations of the HNF inside the DNA helical stack. Their population ratio is estimated to be 1:1. Dynamical perturbation is indicated by the increased linewidth and strong upfield shift of the T residue to the 5'-side of the abasic site.
Zusammenfassung

Mittels NMR-Spektroskopie werden Störungen in Struktur und Dynamik von DNA untersucht, die durch den Einbau jeweils eines der beiden Fluorophore 2-Aminopurin (2AP) und 2-Hydroxy-7-nitrofluoren (HNF) hervorgerufen werden. Zu diesem Zweck werden die NMR-Strukturen der modifizierten Duplexe mit der Sequenz 5'-GCTGCAXACGTCG-3' berechnet. Im Fall X=2AP (13mer2AP) ist die Partnerbase im Komplementärstrang ein T, während gegenüber X=HNF (13mer-HNF) eine abasische Stelle eingeführt wird.


Die übergeordnete Struktur des 13merHNF ist eine B-Form DNA Helix. Die NOE Kreuzpeaks zu den Protonen im HNF können jedoch nur durch zwei verschiedene Orientierungen des HNFs in der helikalen Anordnung beschrieben werden. Das Verhältnis der beiden Orientierungen untereinander wird als 1:1 abgeschätzt. Störungen in der Basenpaardynamik werden durch die höhere Linienbreite und die starke Hochfeldverschiebung des T auf der 5'-Seite ausgehend von der abasischen Stelle angedeutet.
Inhaltsverzeichnis

1 Introduction 1

2 Theoretical background 7
   2.1 Structural aspects of DNA ................................. 7
   2.2 Base pair dynamics in DNA ................................. 11
      2.2.1 Imino proton exchange theory ......................... 11
      2.2.2 Inversion recovery experiments ....................... 16
   2.3 Solution structure determination of DNA ................ 21
      2.3.1 Nuclear Overhauser Effect spectroscopy .............. 21
      2.3.2 Structural information from Nuclear Overhauser Enhancement effects ................................. 25
      2.3.3 Residual Dipolar Couplings ............................ 29
      2.3.4 Simulated Annealing calculations ..................... 35

3 Experimental details 39
   3.1 2-Aminopurine ............................................ 39
      3.1.1 Sample preparation ................................. 39
      3.1.2 Measurements ....................................... 41
      3.1.3 Restraint generation ............................... 44
      3.1.4 Structure calculation ............................... 50
   3.2 2-Hydroxy-7-nitrofluorene ............................... 53
      3.2.1 Sample preparation ................................. 53
      3.2.2 Measurements ....................................... 54
      3.2.3 Restraint generation ............................... 55
      3.2.4 Structure calculation ............................... 56

4 Results and discussion 57
## Inhaltsverzeichnis

4.1 2-Aminopurine ................................................. 57  
   4.1.1 Introduction ............................................. 57  
   4.1.2 Results .................................................. 62  
      4.1.2.1 Chemical shift analysis ........................... 62  
      4.1.2.2 Structural comparison ............................. 65  
      4.1.2.3 Comparison of base pair dynamics ................. 71  
      4.1.2.4 Comparison of duplex melting ..................... 75  
   4.1.3 Discussion ................................................ 76  

4.2 2-Hydroxy-7-nitrofluorene .................................. 86  
   4.2.1 Introduction ............................................. 86  
   4.2.2 Results .................................................. 88  
      4.2.2.1 Chemical shift analysis ........................... 88  
      4.2.2.2 NMR solution structure ............................ 91  
      4.2.2.3 Duplex melting ................................... 94  
   4.2.3 Discussion ................................................ 95  

Summary ....................................................... 99  

Zusammenfassung .............................................. 101  

Script code .................................................... 147  
   .1 Force field parameter and topology files ................ 147  
   .2 Xplor-NIH calculation input files ........................ 196  
   .3 Xplor-NIH calculation restraints files and structures ... 212  
   .4 Lua scripts written for data export from CARA .......... 212  
   .5 Utility scripts ............................................ 253  

Chemical shifts .............................................. 261  
   .6 Chemical shifts of 13merHNF ............................... 261  
   .7 Chemical shifts of 13merRefGC ............................ 265  
   .8 Chemical shifts of 13merRef ................................ 267  
   .9 Chemical shifts of 13mer2AP ............................... 270  
  .10 Chemical shift differences for 13merRef and 13mer2AP ... 273  

Helical parameter ............................................ 275
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>.11</td>
<td>Helical parameters for 13merRef</td>
<td>275</td>
</tr>
<tr>
<td>.12</td>
<td>Helical parameters for 13mer2AP</td>
<td>278</td>
</tr>
<tr>
<td></td>
<td>List of abbreviations</td>
<td>281</td>
</tr>
<tr>
<td></td>
<td>Index of figures</td>
<td>283</td>
</tr>
<tr>
<td></td>
<td>Index of tables</td>
<td>285</td>
</tr>
</tbody>
</table>
Acknowledgement


Herrn Professor Doktor Christian Griesinger danke ich für seine großzügige Unterstützung und seine vielen hilfreichen Anregungen, die mir wichtige Impulse für meine Arbeit lieferten.

Herr Professor Doktor Clemens Mügge hat mir viel am und um das NMR-Spektrometer und mit vielen hilfreichen Hinweisen und Diskussionen geholfen. Ohne seinen Einsatz und seine Flexibilität bei der Meßzeiteinteilung, hätte ich meine Arbeit nicht in so kurzer Zeit durchführen können.

Mein Dank gilt ferner Frau Doktor Jennifer Tuma, die mir den Einstieg in die Thematik der NMR-Strukturanalytik sowie die Handhabung eines NMR-Gerätes sehr erleichterte.

Ich möchte mich auch ganz herzlich bei Herrn Diplom-Chemiker Lars Dehmel bedanken, der - trotz eigener Doppelbelastung durch Kind und Diplomarbeit - die Zeit fand, mit mir mehrere Tage hintereinander im Institut zu “campieren”, um die NMR-Titration zu optimieren. Desweiteren machte er sich um die Auswertung dieser Messungen durch das Schreiben mehrerer Skripte sehr verdient.

Herrn Doktor Horst Hennig danke ich für die große Hilfe in vielen formalen Angelegenheiten und darüber hinaus. Frau Iris Suter bzw. Frau Sabrina Penn und Frau Heiderose Steingräber möchte ich herzlich danken, da sie stets sehr hilfsbereit waren und ein offenes Ohr für die kleinen Probleme des (wissenschaftlichen) Alltags hatten.

Bei Herrn Diplom-Chemiker Patrick Wilke möchte ich mich für die Unterstützung im Rahmen seines Forschungspraktikums und darüber hinaus bedanken.


Amélie und Tristan - ihr seid meine größten Schätze. Ich liebe Euch.
1 Introduction

In 1953 Watson and Crick (WC) solved the puzzle of the deoxyribonucleic acid (DNA) structure [Watson and Crick, 1953]. Aided by the results of Chargaff [1950] and Franklin and Gosling [1953] they proposed the double helical structure. Since then the double helical arrangement of DNA has become common knowledge. But although the overall conformation had been deduced, questions remained. The idealized WC-model could not explain certain sequence-specific effects; an atomic picture was lacking.

In 1980 Dickerson and coworkers were the first to solve the crystal structure of a DNA duplex [Wing et al., 1980]. They refined the WC-model and showed that the ideal helical parameters (as theoretically predicted by Watson and Crick) are true on average but can deviate substantially dependent on sequence [Dickerson and Drew, 1981b,a, Drew et al., 1981]. Subsequently, more crystal structures of DNA sequences were solved, revealing large variations in helical parameters for B-DNA [Dickerson et al., 1982, Kopka et al., 1983, Heinemann and Alings, 1989] and giving structural insights into other conformations like A-DNA [Shakked et al., 1981, Conner et al., 1982, 1984] or Z-DNA [Drew et al., 1980, Wang et al., 1981, Drew and Dickerson, 1981, Rich et al., 1984]. A detailed atomic picture of DNA helical structure was now provided.

However it was found that crystallization can have profound effects on the conformation of DNA. In the crystalline state DNA strongly favours the A-form double helix, whereas DNA in solution occurs predominantly in the B-DNA form [Bloomfield et al., 2000]. Furthermore, the helical parameters strongly depend on the crystallization conditions [Jain and Sundaralingam, 1989, Shakked et al., 1989, Johansson et al., 2000]. Thus single crystal X-ray crystallography is of limited use when studying biological problems,
1 Introduction

particularly those involving DNA.

Parallel to the breakthrough of Dickerson and coworkers [Wing et al., 1980] the advent of 2-dimensional techniques [Jeener et al., 1979, Kumar et al., 1980a, Macura and Ernst, 1980] extended Nuclear Magnetic Resonance Spectroscopy (NMR) methods in a way that studying large biological molecules was now feasible. First efforts centered on protein structure determination [Wagner et al., 1981, Wuthrich et al., 1982, Zuiderweg et al., 1983], but eventually application to DNA structure determination in solution followed [Hare et al., 1983, Feigon et al., 1984, Clore and Gronenborn, 1984, 1985, Hosur et al., 1986]. These early studies focused on the sequential assignment of DNA or protein resonances, the structural content however was discussed only qualitatively. With the advent of powerful computers, techniques were developed that allowed for determination of 3-dimensional structures of biomolecules with NMR structural information as restraints [Williamson et al., 1985, Zuiderweg et al., 1985].

The development and subsequent refinement of the solid-phase phosphor-amidite approach for oligonucleotide synthesis [Sinha et al., 1984, Dahl et al., 1987, Schulhof et al., 1987, Caruthers et al., 1987] allowed for relatively cheap and easy access to large quantities of nucleic acids with a defined primary sequence. This marked a breakthrough for nucleic acids research since it was now possible to vary specific base positions in a predefined sequence. Thereby studying the effect of base mismatches on the helical arrangement of the duplex was facilitated [Kalnik et al., 1988, Roongta et al., 1990, Moe and Russu, 1992]. Furthermore, a means for introduction of arbitrary artificial nucleotides at any position in the duplex was provided by automated solid-phase synthesis.

Artificial DNA double strand structures have been investigated since the late 1980s [Li et al., 1987, Evans and Levine, 1988]. Several different types of modifications have to be distinguished. These involve: backbone modification [Pieters et al., 1989, Betts et al., 1995, Nielsen et al., 2009], fluorophores covalently linked to natural bases [Krugh et al., 1989, Schwartz et al., 1997, Subramaniam et al., 2001], fluorophores substituting a natural base [Nordlund et al., 1989, Guckian et al., 1998, Engman et al., 2004] or even a base pair [Matray and Kool, 1998, Guckian et al., 2000, Smirnov et al., 2002],
and intercalators, which bind to DNA through stacking and/or electrostatic interactions [Fede et al., 1993, Spielmann et al., 1995, Davies et al., 1997]. All of the latter studies focus on the introduction of chromophores, since their fluorescent properties can be exploited for studying DNA [Wojczewski et al., 1999].

The spectroscopic properties of covalently attached DNA modifications are utilized in numerous ways. Over the past decades different strategies have been developed for detecting single nucleotide polymorphisms (SNPs). These strategies employ fluorophore-quencher systems (molecular beacons) [Tan et al., 2004], DNA-mediated electron transfer (DETEQ) [Wagenknecht, 2008] or forced intercalation probes (FIT) [Koehler et al., 2005]. Furthermore, fluorescent molecules are introduced at different locations into DNA in order to get long-range structural information by exploiting fluorescence resonance energy transfer (FRET) [Lilley and Wilson, 2000]. In addition it has been demonstrated that transient absorption spectroscopy of fluorophore-modified DNA can be used to follow dynamics on the pico- to femtosecond timescale [Zewail, 2000].

Supramolecular vibrational modes of biological molecules are important for their function, and many have frequencies below 200 cm$^{-1}$ or 6 THz. Examples are the primary event of vision (60 cm$^{-1}$) [Wang et al., 1994], oxygen acceptance of hemoglobin (39 cm$^{-1}$) [Klug et al., 2002], chemical reactions in myoglobin (51 cm$^{-1}$) [Austin et al., 1989], and conformational change of bacteriorhodopsin (115 cm$^{-1}$) [Xie et al., 2001]. For DNA transcription the double helix must be opened to expose the coding bases to chemical reactions. Thermal melting of double-stranded oligonucleotides is similar because it starts with a “denaturation bubble” [Prohofsky et al., 1979]. The latter is reached through collective modes between 60 and 140 cm$^{-1}$ which compress and stretch the interbase H-bonds [Cocco and Monasson, 2000]. However, such resonances in the low-frequency region are difficult to detect due to mixing of the DNA modes with those of hydration water. Resolving such collective vibrational modes of a biological molecule by molecular THz spectroscopy is a new but potent application for chromophores in DNA.

Here the chromophore functions like a THz light source when its charge distribution is suddenly altered by femtosecond optical excitation $S_1 \leftarrow S_0$. The electric field around
the probe is changed instantly and acts on nearby groups with partial charges. Most of these change their nuclear position in an overdamped fashion but some may oscillate briefly. Altogether a reaction field \( R(t) \) is created which is reported by the polar probe molecule, through an emission frequency which depends on \( R(t) \). The probe molecule is therefore not only light source but also detector. A response function can be obtained which is related to the local THz absorption spectrum. In this way the low-frequency vibrational structure of biomolecules can be accessed.

Molecular THz spectroscopy should reduce inhomogeneous broadening because the perturbing electric field and the reaction field are local. Only those modes will interact which have oscillator strength in the region, at the right direction. The obvious disadvantage is the need to embed a probe molecule inside double-stranded DNA as an artificial nucleobase. The probe has to be free of internal modes which are active below \( \approx 300 \text{ cm}^{-1} \), since they would mix with the macromolecular dynamics to be reported. For this reason the best-studied polarity probes, coumarins [Horng et al., 1995, Zhao et al., 2005], are not eligible. Instead one must use chromophors which have been shown to report the far infrared spectrum of pure liquids such as acetonitrile [Ruthmann et al., 1998, Karunakaran et al., 2008] or water [Lustres et al., 2005]. Required is bio-organic development of suited chromophores guided by optical femtosecond spectroscopy in the condensed phase. However, excellent suitability from an optical point of view is to no avail when the helical structure is severely disrupted, since duplex features are to be probed.

Thus, for all aforementioned spectroscopic techniques which investigate DNA features, it is advantageous or even imperative to know the exact orientation of the fluorophore inside the DNA double helix. Moreover it is instructive to have information on the stacking interactions of the fluorophore with the adjacent base pairs. Finally, it can be decisive - especially for biologically motivated hybridization studies - to be able to characterize the structural and dynamic perturbation of the DNA helical structure in terms of helical parameters and base pair lifetimes.

In order to solve above question, this work utilizes NMR spectroscopy for the structure
determination of the DNA sequences given in Fig. 1.1 and 1.2 in order to solve above question. In both sequences the central base or base pair is modified in order to have the single modification site and its adjacent base pairs unperturbed by fraying effects at the helix termini [Nonin et al., 1995]. Two different modifications are examined, 2-Aminopurine (2AP) and 2-Hydroxy-7-nitrofluorene (HNF).

**Duplex DNA with 2-Hydroxy-7-nitrofluorene**

The probe HNF is incorporated opposite to an abasic site (1’-2’-didesoxyribose) to avoid steric strain which might otherwise disrupt the overall B-DNA conformation or force the fluorophore into an extrahelical position. However, deletion of the partner base introduces increased flexibility into the DNA duplex at the modification site [Lin and de los Santos, 2001, Smirnov et al., 2002]. In conjunction with the different electronic properties of the HNF residue as compared to a natural base pair, HNF is expected to introduce large local perturbations compared to a more native modification. A first purpose of this work is to find out whether the DNA duplex adopts an overall B-DNA conformation
1 Introduction

with the HNF stacked inside the double helix. This information is essential since HNF has been designed to report on macromolecular vibrational modes in DNA via transient absorption spectroscopy, which cannot be observed in case of an extrahelical orientation of HNF.

Duplex DNA with 2-Aminopurine

The 2AP-containing duplex represents a substantially different case. 2AP causes only a slight perturbation of the directly adjacent base pairs as suggested in earlier works by Lycksell et al. [1987] and Nordlund et al. [1989]. 2AP is commonly used to monitor base stacking-unstacking events in biologically relevant sequences [Allan and Reich, 1996, Reddy and Rao, 2000, Bernards et al., 2002, Daujotyte et al., 2004, Neely et al., 2005, Lenz et al., 2007]. Therefore it is crucial to characterize the perturbation induced upon 2AP incorporation structurally (in terms of helical parameters) and dynamically (in terms of reliable base pair lifetimes). This is the second aim of this work.

Since 2AP is structurally isomeric to adenine (A), it closely resembles the latter in size and shape. It has been found that it can also form stable base pairs with thymine (T) [Ronen, 1979, Sowers et al., 1986]. The two duplex structures of 13merRef with X=A and 13mer2AP with X=2AP differ only in the location of the amino group of the central residue (see Fig. 1.1). Thus any structural or dynamic differences that are observed between the two corresponding solution structures can be directly attributed to the incorporation of 2AP. This is the major difference compared to pertinent works; these investigate a ten base pair palindromic DNA duplex with two 2AP residues incorporated [Lycksell et al., 1987, Nordlund et al., 1989].
2 Theoretical background

2.1 Structural aspects of DNA

DNA is composed of four naturally occurring nucleobases, adenine (A), guanine (G), thymine (T) and cytosine (C). While A and G are derived from purine, T and C are pyrimidine derivatives. Base pairs are formed between A:T and G:C. The structure and nomenclature of these two base pair motifs is depicted in Fig. 2.1. The higher thermodynamic stability of G:C compared to A:T base pairs [Xia et al., 1998] originates from the fact that the latter base pair forms only two instead of three hydrogen bonds (Fig. 2.1).

By attaching A, G, C, or T to the C1’-atom of a 2’’-deoxy-β-D-ribose the nucleosides adenosine, guanosine, cytidine and thymidine are formed, respectively. The structure and nomenclature of 2’’-deoxy-β-D-ribose, which is called sugar in the following, is shown in Fig. 2.2. All atoms of the sugar are marked with a “ ’ ” to distinguish them from nucleobase atoms. The sugar conformation is defined by five dihedral angles \( \nu_0 - \nu_4 \) (Fig. 2.3). The latter are interdependent and thus can be described by only two parameters

\[
\nu_j = \Psi_m \cdot \cos(P + 144(j - 2)) \quad j = 0, 1, 2, 3, 4
\]

(2.1)

the pucker angle P and the pucker amplitude \( \Psi_m \). Note that this simple relation is valid only for cyclopentane, but deviations can be accounted for by introducing correction terms [Altona and Sundaralingam, 1972].

The glycosidic torsion angle \( \chi \) and the backbone dihedral angles \( \alpha - \zeta \) can be used to characterize the helical structure of DNA (Fig. 2.3). While \( \alpha - \zeta \) define the sugar-
phosphate backbone of DNA, $\chi$ determines the position of the nucleobase relative to the sugar. Two different orientations are sampled, the anti range which centers around $\chi = -135^\circ$ and the syn range which samples values around $\chi = +45^\circ$ [Bloomfield et al., 2000]. The latter is less stable as the nucleobase is located above the sugar, which leads to steric clashes. Thus the syn-orientation of nucleobases is only found for a special helical arrangement of DNA, the left-handed Z-DNA, which needs external stabilization, e.g. by high salt concentration [Rich et al., 1984].

Two main helical arrangements are found for DNA, the A-form and the B-form. While the former is more often found in crystal structures, B-DNA is the dominant conforma-
2.1 Structural aspects of DNA

Fig. 2.2: Nomenclature and structure of 2'-deoxy-β-D-ribose [Roberts, 1993].

Fig. 2.3: Dihedral angles in the sugar-phosphate backbone of DNA. The picture is taken from Roberts [1993].

<table>
<thead>
<tr>
<th></th>
<th>A-DNA</th>
<th>B-DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helix handedness</td>
<td>Right</td>
<td>Right</td>
</tr>
<tr>
<td>bpt/repeating unit</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>bpt/turn</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Helix twist, (°)</td>
<td>32.7</td>
<td>36.0</td>
</tr>
<tr>
<td>Rise/bp, (Å)</td>
<td>2.9</td>
<td>3.4</td>
</tr>
<tr>
<td>Helix pitch, (Å)</td>
<td>32</td>
<td>34</td>
</tr>
<tr>
<td>Base pair inclination, (°)</td>
<td>12</td>
<td>2.4</td>
</tr>
<tr>
<td>P distance from helix axis, (Å)</td>
<td>9.5</td>
<td>9.4</td>
</tr>
<tr>
<td>X displacement from bp to helix axis, Å</td>
<td>-4.1</td>
<td>0.8</td>
</tr>
<tr>
<td>Glycosidic bond orientation</td>
<td>anti</td>
<td>anti</td>
</tr>
<tr>
<td>Sugar conformation</td>
<td>C3'-endo</td>
<td>C2'-endo</td>
</tr>
<tr>
<td>Major groove depth</td>
<td>13.5</td>
<td>8.5</td>
</tr>
<tr>
<td>width, (Å)</td>
<td>2.7</td>
<td>11.7</td>
</tr>
<tr>
<td>Minor groove depth</td>
<td>2.8</td>
<td>7.5</td>
</tr>
<tr>
<td>width, (Å)</td>
<td>11.0</td>
<td>5.7</td>
</tr>
</tbody>
</table>

(a) Important parameters
(b) Helical backbone

Fig. 2.4: Comparison of A- and B-form DNA. The pictures are taken from Bloomfield et al. [2000].

tion in solution. The most important parameters for both helical arrangements are compiled in Fig. 2.4. While the overall arrangement is similar (right-handedness, glycosidic bond orientation), differences exist. One that is commonly used to distinguish between the two helical arrangements is the sugar conformation, with C3'-endo (P around 18°) dominant for A-DNA and C2'-endo (P around 162°) for B-DNA. The widths of the minor and major grooves, which are depicted on the molecular and atomic level in Fig. 2.1 and Fig. 2.4 respectively, also differ significantly. The larger helical rise and twist va-
Fig. 2.5: Helical parameters that describe the orientation of base pairs relative to the molecular frame (upper row), base pair partners (middle row) and consecutive base pairs (lower row) relative to each other. The picture is taken from Roberts [1993].

Values for B-DNA lead to an elongated shape of the latter, while A-DNA is much more compressed (Fig. 2.4).

The helical parameters, which are visualized in Fig. 2.5, describe the arrangement of oligonucleotides in a double helix. They can be divided into three subgroups. Parameters in the upper row yield information on the orientation of base pairs relative to the molecular frame. The middle row parameters report on the orientation of the two base pair partners relative to each other. The lower row is most often used to characterize the helical arrangement, since these parameters give information about the orientation of two consecutive base pairs relative to each other. Especially the rise and twist-values determine the overall shape of the double helix (Fig. 2.4).
2.2 Base pair dynamics in DNA

2.2.1 Imino proton exchange theory

The imino protons of G and T are located near the helical axis at the center of each base pair and thus are effectively shielded against solvent or catalyst attack. Consequently, the central assumption in imino proton exchange theory of DNA is that exchange with bulk water can only proceed via a transient opening of the base pair [Kochoyan et al., 1987, 1988, Leijon and Graslund, 1992]. For Watson-Crick duplexes, base pair lifetimes do not depend on the nature of the adjacent pairs. This suggests that opening involves single base pairs only; the possibility of collective opening motions is ruled out [Leroy et al., 1985, Gueron et al., 1987].

Fig. 2.6 depicts a kinetic scheme of the processes involved in imino proton exchange theory [Leijon and Graslund, 1992]. Exchange from the closed state is not possible [Nonin et al., 1995], thus the first step must be the opening of the base pair with rate constants $k_{op}$ and $k_{cl}$ for opening and closing, respectively. From the open state, where the imino proton is assumed to be fully accessible [Kochoyan et al., 1988], two different processes can occur: exchange via an external base catalyst (rate constant $k_{ext}^{cat}$) and via an intrinsic pathway ($k_{int}^{cat}$). As a possible intrinsic catalyst the endocyclic nitrogen of the complementary base has been proposed [Leroy et al., 1985, Kochoyan et al., 1988]. Frequently used external base catalysts are Trishydroxymethylaminomethane (TRIS) or ammonia [Kochoyan et al., 1988, Moe and Russu, 1990, Bhattacharya et al., 2002].

Under conditions of a stable structure ($k_{op} << k_{cl}$), which can be safely assumed for duplex lengths $> 10$ base pairs, the concentration of base pairs in the open state $[NH^*]$ is quasistationary, i.e. the kinetics are pseudo-first-order.

\[
\frac{d[NH^*]}{dt} = 0 = k_{op} [NH^* \cdots N] - (k_{cl} + k_{int}^{cat} + k_{ext}^{cat}) [NH^*] \quad (2.2a)
\]

\[
(k_{cl} + k_{int}^{cat} + k_{ext}^{cat}) [NH^*] = k_{op} [NH^* \cdots N] \quad (2.2b)
\]

\[
[NH^*] = \frac{k_{op}}{(k_{cl} + k_{int}^{cat} + k_{ext}^{cat})} [NH^* \cdots N] \quad (2.2c)
\]
The effective imino proton exchange rate $k_{ex}$ is defined by

$$k_{ex} [NH^* \cdots N] = (k_{int}^\text{cat} + k_{ext}^\text{cat}) [NH^*]$$

(2.3)

where $[NH^* \cdots N]$ is the concentration of the base pair in the closed state. Substituting eq. (2.2c) into eq. (2.3) and cancelling of $[NH^* \cdots N]$ gives

$$k_{ex} = \frac{k_{op} k_{int}^\text{cat}}{k_{cl} + k_{int}^\text{cat} + k_{ext}^\text{cat}}$$

(2.4a)

$$\frac{1}{k_{ex}} = \frac{1}{k_{op}} \left( 1 + \frac{k_{cl} k_{int}^\text{cat}}{k_{int}^\text{cat} + k_{ext}^\text{cat}} \right)$$

(2.4b)

which can be reduced to

$$\frac{1}{k_{ex}} = \frac{1}{k_{op}} + \frac{1}{K_{diss} (k_{int}^\text{cat} + k_{ext}^\text{cat})}$$

(2.5)

by substituting $K_{diss} = \frac{k_{op}}{k_{cl}}$. Introducing $\tau_{ex} = \frac{1}{k_{ex}}$, the imino proton exchange time, and $\tau_{op} = \frac{1}{k_{op}}$, the base pair lifetime, gives

$$\tau_{ex} = \tau_{op} + \frac{1}{k_{int}^\text{cat} K_{diss} + k_{ext}^\text{cat} K_{diss}}$$

(2.6)

The transfer rate of the imino proton to an external catalyst in the isolated mono-
2.2 Base pair dynamics in DNA

The base pair dynamics in DNA nucleoside \( k_i \) should be comparable to \( k_{\text{ext cat}} \). The former is given by

\[
k_i = \frac{k_{\text{coll}} [B]}{1 + 10^{pK_a^n - pK_a^c}} = k_{\text{iso}} [B]
\]

(2.7)

where \( k_{\text{coll}} \) is the collision rate constant and \( pK_a^n \) and \( pK_a^c \) are the \( pK_a \)-values for the nucleotide and the catalyst respectively [Eigen, 1964]. Since the latter are all constant, they can be substituted by introducing \( k_{\text{iso}} \), the constant of proportionality between \( k_i \) and the concentration of the external base catalyst [B]. Thus \( k_{\text{ext}} \) can be expressed as

\[
k_{\text{ext cat}} = \alpha k_{\text{iso}} [B]
\]

(2.8)

where restricted accessibility of the imino proton in the open state as compared to the isolated nucleoside is taken into account by the parameter \( \alpha \), which ranges from 0 (not accessible) to 1 (unrestricted accessibility). For natural base pairs and all the commonly used base catalysts \( \alpha \) was found to be approximately 1 [Kochoyan et al., 1988]. Consequently, differences in accessibility might become important for modified nucleotides, but can be safely neglected for natural ones. Substituting eq. (2.8) into eq. (2.6) gives

\[
\tau_{\text{ex}} = \tau_{\text{op}} + \frac{1}{k_{\text{cat}}^\text{int} K_{\text{diss}} + \alpha k_{\text{iso}} K_{\text{diss}} [B]}
\]

(2.9)

which is an accurate description when \( k_{\text{cat}}^\text{int} \approx k_{\text{ext}} \). At high external base catalyst concentrations, \( k_{\text{cat}}^\text{ext} \) dominates \( k_{\text{cat}}^\text{int} \) and eq. (2.9) simplifies to

\[
\tau_{\text{ex}} = \tau_{\text{op}} + \frac{1}{(\alpha k_{\text{iso}} K_{\text{diss}})[B]}
\]

(2.10)

and the imino proton exchange time becomes a linear function of the inverse of the external catalyst concentration. In the other extreme case, when no external catalyst is available, \( k_{\text{cat}}^\text{int} >> k_{\text{ext}} \) and consequently eq. (2.9) reduces to

\[
\tau_{\text{ex}} = \tau_{\text{op}} + \frac{1}{k_{\text{cat}}^\text{int} K_{\text{diss}}}
\]

(2.11)
which means that changes in the imino proton exchange can be either due to different base pair lifetimes (which can be determined via eq. (2.10)) or altered intrinsic exchange.

Early works on imino proton exchange considered the opening of the base pair to be rate-limiting, i.e. $k_{cl} << (k_{cal}^{int} + k_{cal}^{ext})$ [Teitelbaum and Englander, 1975b,a, Patel and Hilbers, 1975, Hilbers and Patel, 1975]. In that case, eq. (2.4a) simplifies to

$$k_{op} \frac{k_{cal}^{int} + k_{cal}^{ext}}{k_{cal}^{int} + k_{cal}^{ext}} = k_{op} = k_{ex}$$

which implies that imino proton exchange with water directly measures the lifetime of the closed base pair ($\tau_{op}$) since the latter is no longer a function of $[B]$. Exchange occurs every time the base pair opens, thus $\tau_{ex}$ and consequently $\tau_{op}$ are maximal. As a result, base pair lifetimes published before 1985 were overestimated by approximately one order of magnitude. In that year Gueron and coworkers showed that $\tau_{ex}$ depends on the external catalyst concentration [Leroy et al., 1985, Gueron et al., 1987], demonstrating that imino proton exchange in polynucleotides is not opening-limited. Instead, eq. (2.10) is validated by their results. By introducing the apparent dissociation constant $\alpha K_d$ (eq. (2.13a)) [Kochoyan et al., 1988] one obtains the commonly used expression for $\tau_{ex}$ (eq. (2.13b)):

$$\alpha K_d = \alpha k_{iso} K_{diss}$$  \hspace{1cm} (2.13a)

$$\tau_{ex} = \tau_{op} + \frac{1}{(\alpha K_d)} \frac{1}{[B]}$$  \hspace{1cm} (2.13b)

A plot of $[B]^{-1}$ vs $\tau_{ex}$ allows for determination of the apparent dissociation constant from the slope of the linear fit. However, interpretation of $\alpha K_d$ is not straightforward and hampered by the approximations detailed above. Much more informative of base pair dynamics is $\tau_{op}$, which can be determined from the intercept with the ordinate (in the limit of infinite $[B]$) of the plot $[B]^{-1}$ vs $\tau_{ex}$.

Based on the reaction $[BH^+] + [H_2O] \rightleftharpoons [B] + [H_3O^+]$ and the definition of the acidity constant ($K_s$)

$$K_s = K_{eq} [H_2O] = \frac{[B][H_3O^+]}{[BH^+]}$$ \hspace{1cm} (2.14)
2.2 Base pair dynamics in DNA

an expression is obtained, which relates the concentration of free external base catalyst $[B]$ and the total added base concentration $[B_0]$.

\[
\frac{1}{[B]} = (1 + 10^{(pK_a^{ca} - pH)}) \frac{1}{[B_0]}
\]  

(2.15)

It follows that $[B]$ is a function of $[B_0]$ and the pH-value.

The imino proton exchange time $\tau_{ex}$ can be determined by measuring the line broadening of the imino proton resonance due to addition of base catalyst [Lycksell et al., 1987].

\[
\frac{1}{\tau_{ex}} = \pi \Delta
\]  

(2.16)

Line broadening provides an easy and time-efficient way to experimentally determine $\tau_{ex}$. However, since line broadening can also have other sources apart from the exchange process (e.g. different shim settings for each titration point) results obtained with this method tend to be inaccurate.

An alternative way is to calculate $\tau_{ex}$ from the difference of the spin-lattice relaxation time with ($T_1^{ext}$) and without external base catalyst ($T_1^{int}$) [Bhattacharya et al., 2002].

\[
\frac{1}{\tau_{ex}} = \frac{1}{T_1^{ext}} - \frac{1}{T_1^{int}}
\]  

(2.17)

$T_1^{ext}$ and $T_1^{int}$ can be measured by NMR spectroscopy with high precision - errors below 2% can be achieved [Sass and Ziessow, 1977] - and thus allow for more precise values of $\tau_{ex}$ to be determined.
2 Theoretical background

2.2.2 Inversion recovery experiments

Base pair dynamics of DNA can be followed by NMR spectroscopy. As detailed above (section 2.2.1), the key observable is chemical exchange of imino protons with bulk water as a function of external catalyst concentration. According to eq. (2.17) \( \tau_{ex} \) can be measured by determining \( T_1^{ext} \) and \( T_1^{int} \). However, a difference in magnetization between the imino proton and water protons has to be created in order to make the exchange observable by NMR.


Although the majority of the former studies used the saturation recovery method, more recent works rely on inversion recovery experiments. Saturation recovery might be preferable for measurements of long \( T_1 \)-values (> 5 s) [Levy and Peat, 1975] or when the choice of optimal delay times is difficult due to a large range of \( T_1 \)-values of interest [Roscher et al., 1996]. But these conditions are clearly not fulfilled for \( T_1 \)-values of imino protons which are on a ms-timescale [Moe and Russu, 1990]. Becker et al. [1980] compared the efficiency of inversion and saturation recovery in determining \( T_1 \)-values with a certain precision. They report inversion recovery to be much more efficient, with saturation recovery requiring 8-times more scans to achieve comparable signal-to-noise ratio [Weiss et al., 1980, Becker et al., 1980]. Thus inversion recovery should generally be preferred over saturation recovery experiments.

As to the question whether to invert the water or the imino proton region, no com-
prehensive study has yet been made. Regarding water inversion, Mihailescu and Russu [2001] point out that in order to obtain values for $\tau_{ex}$, the magnetization of the imino protons as a function of the exchange delay $\tau$ has to be fitted to the equation

$$M(\tau) = M(0) - \frac{[M(0) - M^0]}{(T_1 + k_{ex})^{-1}} - \frac{k_{ex} M_0 T_{1\text{water}} \tau}{(q - 1)^{-1}}$$

(2.18)

where the values of $T_{1\text{water}}$ and $q$ (efficiency of inversion for the water signal) have to be determined in a separate experiment. Furthermore, the value of $(T_1 + k_{ex})$ is determined for each imino proton resonance of interest by selective saturation [Mihailescu and Russu, 2001]. These additional experiments are time-consuming and thus inversion of the water signal is not to be preferred. In consequence, the method of choice for measuring base pair lifetimes is selective inversion of the imino proton region.

In the standard inversion recovery experiment the $z$-magnetization of one or multiple spins is inverted by an initial $180^\circ$-pulse ($+M_z \rightarrow -M_z$). After a variable delay time ($\tau$) a $90^\circ$-pulse is used to detect the recovered magnetization (Fig. 2.7). The plot of signal intensity vs delay time is fitted exponentially with three instead of two parameters ($A$, $B$ and $T_1$) to account for imperfect inversion of the signal

$$I(\tau) = A + B \exp(-\tau/T_1)$$

(2.19)

[Sass and Ziessow, 1977, Kowalewski et al., 1977]. To create differences in magnetization between water and the imino protons, the initial $180^\circ$-pulse has to be applied selectively to the imino proton region.

Selective inversion can be achieved in numerous ways. In principle, the commonly used constant-amplitude (rectangular), unselective (hard) pulse can be applied selectively by varying the pulse length and adjusting the pulse power. Rectangular pulses have the advantage of relatively straightforward implementation and optimization due to their analytical solution to the Bloch equations [Hajduk et al., 1993]. However, the excitation profile, which can be calculated by the latter equations, is imperfect. While there is a central lobe of strong excitation, the sharp leading and trailing edges of the pulse give rise
to a set of sidelobes (cf. Fig. 2.8), whose amplitude decreases with offset from the centre frequency [Freeman, 1992]. The sidelobes lead to considerable off-resonance excitation.

Similar reasoning can be applied to inversion. In order to reduce off-resonance excitation, the transition region at the edges of the pulse must be smoothed.

A number of shaped pulses for selective bandwidth inversion were created when - with the advent of the computer - numerical calculation of the Bloch equations became possible. Among the first shapes to be proposed was the Gaussian pulse envelope, which is still popular [Bhattacharya et al., 2002]. The advantage of the latter is that in the linear response region (pulse length $\tau_p \ll T_1, T_2$) the excitation function is another Gaussian (see Fig. 2.8). Thus sidelobes are effectively avoided [Bauer et al., 1984]. However, a Gaussian excitation function is far from the ideal of a rectangle, consequently the completely inverted region is quite small [McDonald and Warren, 1991]. Other pulse shapes like Hermitian [Warren, 1984], Gaussian cascades [Emsley and Bodenhausen, 1990] and quaternion pulses [Emsley and Bodenhausen, 1992] try to extend the region of complete inversion with increasing success.

However, the advantages, which were outlined above are achieved at the cost of longer pulse durations. But when the the relaxation times $T_1$ and $T_2$ become comparable to (or even shorter than) the pulse duration, relaxation during the 180°-pulse is no longer negligible. This can have a profound effect on the excitation profile and hamper applicability of shaped pulses for selective excitation [Hajduk et al., 1993]. Furthermore, most shaped pulses were derived on the basis of the Bloch equations neglecting such effects as radiation damping, relaxation and coupling [McDonald and Warren, 1991]. While each of these effects can be accounted for separately [Warren et al., 1989, Hajduk et al., 1993, Ewing et al., 1990], to account for all of them at once is difficult.

The adiabatic sweep is a completely different approach to selective bandwidth inversion. The usually employed pulsed NMR experiment is operated at a static magnetic field $B_o$ and uses a pulsed magnetic field $B_1$ to simultaneously invert all frequencies of
interest with a constant, bandwidth-centered carrier frequency. In the adiabatic passage experiment, the carrier frequency is modulated over the whole bandwidth and thus the frequencies of interest are inverted successively [Tannús and Garwood, 1997]. An alternative, equivalent approach is to modulate the phase of the center frequency [Garwood and DelaBarre, 2001]. Adiabatic pulses have two main advantages over standard pulses. They are rather insensitive to $B_1$ field inhomogeneity [Bohlen and Bodenhausen, 1993] and require much less radiofrequency power for inverting nuclear spins over a wide range of chemical shifts [Kupce and Freeman, 1995]. Many different pulse shapes, which enhance one or the other advantage, were created: the hyperbolic secant pulse [Silver et al., 1984, 1985], the chirped pulse [Bohlen and Bodenhausen, 1993, Fu and Bodenhausen, 1995] or the WURST pulse [Kupce and Freeman, 1995, 1997]. Some are now utilized in broadband heteronuclear spin decoupling where radiofrequency power is limited [Kupce and Freeman, 2007] or in in vivo NMR where surface coils with a spatially inhomoge-
Theoretical background

Theoretical background

neous $B_1$ field are used [Tannús and Garwood, 1997]. Yet there are also disadvantages. One of them is a distinct phase distortion which is introduced for large bandwidths due to the different times at which the frequencies are inverted. This disadvantage can be circumvented, however, by introducing a second adiabatic pulse which cancels the phase distortion introduced by the first [Kupce and Freeman, 1997]. Further disadvantages are the requirement that adiabatic rotations must be accomplished rapidly relative to $T_1$ and $T_2$ of the nuclear spins of interest [Garwood and DelaBarre, 2001]. This limits application of adiabatic pulses to spins with quite long relaxation times and those which span a large number of chemical shifts, e.g. heteronuclei like $^{13}C$. 


2.3 Solution structure determination of DNA

2.3.1 Nuclear Overhauser Effect spectroscopy

NMR spectroscopy has become the method of choice for structure determination of biomolecules in solution. This success is primarily based on the possibility to directly extract distance information utilizing the Nuclear Overhauser Enhancement effect (NOE). Overhauser [1953] was the first to discover that the intensity of the signal of one resonance is changed upon perturbation of another due to cross-relaxation. Initially the NOE was observed by selective saturation of one line and subsequent recording of the 1D spectrum over the whole spectral region of interest. This approach proved to be very useful for the structure elucidation of small molecules [Colson et al., 1967, Woods et al., 1968, Schirmer et al., 1970, Schirmer and Noggle, 1972]. Its applicability to biomolecules however, was limited by the long accumulation time and severe spectral overlap which leads to poor selectivity of saturation [Kumar et al., 1981]. The development of the 2D-NMR spectroscopy finally allowed to acquire the complete set of NOE effects for a macromolecule with a single experiment [Macura and Ernst, 1980]. Furthermore, the 2D Nuclear Overhauer Enhancement Spectroscopy (NOESY)-experiment could be readily repeated in $H_2O$ [Kumar et al., 1980b].

The pulse sequence of a general 2D NOESY experiment is shown in Fig. 2.9. With an initial 90°-pulse transverse magnetization is created. The latter is allowed to precess freely during the evolution time $t_1$, thereby frequency-labelling the magnetization components. A second 90°-pulse rotates the magnetization onto the negative z-axis. During the subsequent mixing period with variable length $\tau_m$, z-magnetization components exchange through dipole-dipole cross-relaxation. A third 90°-pulse again creates transverse magnetization which is finally detected. All three pulses are applied non-selectively [Macura and Ernst, 1980].

In the absence of scalar spin-spin-interactions, cross-relaxation of the longitudinal
magnetization components $M_{ij}$ can be described with the equation:

$$\dot{m} = R \times m \quad (2.20)$$

$R$ is the relaxation matrix comprising the cross relaxation rates $R_{ij}$ and the external relaxation (leakage) rates $R_i$ (Fig. 2.10). The vector $m$ comprises the deviation of $M_{zi}$ from thermal equilibrium for $i$ spins

$$m_i = M_{zi} - \frac{n_i}{N} M_o \quad (2.21)$$

where $M_o$ is the total equilibrium magnetization of the $N$ nuclei. After the evolution period the initial $z$-magnetization components are encoded by the precession frequencies
of the components \( \omega_i \)

\[
m_i(0) = M_o \frac{n_i}{N} \left[ \cos(\omega_i t_1) \exp \left( -\frac{t_1}{T_{2i}} \right) - 1 \right]
\] (2.22)

such that the cross-relaxation pathway can be traced back to its origin. The recovery of the magnetization back to equilibrium during the mixing period can be described by the following solution to eq. (2.20) [Macura and Ernst, 1980]

\[
m(\tau_m) = \exp[-\mathbf{R} \tau_m] \mathbf{m}(0)
\] (2.23)

where \( \mathbf{m}(\tau_m) \) is the matrix comprising the magnetization components after the mixing period, \( \mathbf{m}(0) \) represents the intensities of the diagonal peaks at \( \tau_m = 0 \). The diagonal \( (R_{ii}) \) and off-diagonal \( (R_{ij}) \) relaxation matrix elements are given as [Roberts, 1993]:

\[
R_{ii} = q_{ij} \sum_{i,j} \left\{ J_{0,ij}(\omega_i - \omega_j) + 3 [J_{1,ij}(\omega_i) + J_{1,ij}(\omega_j)] + 6 J_{2,ij}(\omega_i + \omega_j) + R_{1i} \right\}
\]

\[
R_{ij} = q_{ij} \left[ 6 J_{2,ij}(\omega) - J_{0,ij}(\omega) \right]
\] (2.24a, 2.24b)

where \( R_{1i} \) is the leakage rate, which can usually be neglected in the absence of paramagnetic nuclei. The factor \( q_{ij} \) comprises all constant values

\[
q_{ij} = \frac{\gamma_i^2 \gamma_j^2 \hbar^2 \mu_0^2}{160 \pi^2}
\] (2.25)

where \( \gamma_i \) and \( \gamma_j \) are the gyromagnetic ratios for spins \( i \) and \( j \) respectively, \( \hbar \) is the Planck constant (divided by \( 2\pi \)) and \( \mu_0 \) is the magnetic constant or vacuum permeability. \( J_{n,ij}(\omega) \) represents the spectral densities for the zero, single and double quantum transitions (\( n=0,1,2 \) resp.)

\[
J_{n,ij}(\omega) = \frac{\tau_c^{ij}}{1 + (n \omega_0 \tau_c^{ij})^2} \frac{1}{\tau_{tj}^n}
\] (2.26)
where $\tau_{ij}^c$ is the rotational correlation time of the vector between spins i and j and $r_{ij}$ is the distance between the latter. Since the differences in resonance frequency for various spins are negligibly small compared to the value of the resonance frequency itself, $\omega_i$ and $\omega_j$ are approximated by the center frequency $\omega_0$. The factor n in the denominator is given by the number of quanta involved in the transition. With the help of some assumptions the intensities of NOE cross-peaks can be directly related to distances in the molecule under investigation, as is seen in the next section.
2.3.2 Structural information from Nuclear Overhauser Enhancement effects

The NOE effect is an invaluable tool for structure elucidation, since it is correlated with the distance between the interacting nuclei. There are different methods to extract structural information out of NOE cross-peak intensities. A very popular method is to classify peaks according to their peak intensities into strong, medium and weak and setting up loose distance restraints of 1.8-2.8 Å, 1.8-3.3 Å and 1.8-5.0 Å, respectively. This approach has the clear advantage of simplicity and abolishes the need to integrate NOE cross-peaks in the 2D-spectrum. However, intensities are only representative of peak volumes when the lineshape for all peaks is the same. This is true only when the digital resolution is larger than the peakwidth [Roberts, 1993], which is clearly not fulfilled for modern high-resolution spectra of macromolecules. A second disadvantage is that such loose distance restraints do not have much restraining power and consequently the structure is ill-defined. In particular for nucleic acids, the classification on the basis of peak intensities is unsuitable, since the proton density is much less than in proteins (0.35 protons/atom versus 0.52 protons/atom in proteins) and thus yields fewer distance restraints. Additionally, long-range peaks as detected in proteins cannot be observed in DNA due to its rod-like shape. The need for more accurate distance restraints in DNA structure determination by NMR is thus intensified [MacDonald and Lu, 2002].

Very accurate distance restraints can be extracted using the Full Matrix Relaxation Approach. Here distances are calculated directly from the off-diagonal elements of the relaxation matrix, which in turn is obtained via

$$R = \frac{\ln[m(0)] - \ln[m(\tau_m)]}{\tau_m}$$

[Roberts, 1993]. To obtain the complete intensity matrix $m$ is not possible in practice. Thus two different algorithms have been proposed to fill the “gaps” in $m$. Schematic representations for both algorithms are given in Fig. 2.11. Both algorithms rely on a reasonable starting structure from which an intensity matrix is calculated. The latter is then combined with the intensity matrix derived from the experimental NOEs and
Two different approaches to obtain distances from NOE experiments. While one approach is based on the self-consistency of the relaxation matrix (right part), the other relies on external structure calculation (left part). Popular implementations of these approaches are the programs MARDIGRAS [Borgias and James, 1990] and IRMA [Boelens et al., 1988]. This picture is taken from Roberts [1993].

The relaxation matrix is calculated using eq. (2.27). At this step the two algorithms differ. One approach checks the relaxation matrix for self-consistency and produces a reconciled relaxation matrix (MARDIGRAS [Borgias and James, 1990]) from which the intensity matrix is in turn calculated. The other approach (IRMA [Boelens et al., 1988]) produces a set of distance restraints from the relaxation matrix. The restraints are used to refine the starting structure. The refined structure is then used to calculate an intensity matrix and the process starts all over again. The main advantage of either approach
is that by measuring NOESY spectra at different mixing times, spin diffusion can be accounted for and very accurate distance restraints are produced. However, as pointed out by Lane [1996] and Tonelli and James [1998] error limits are often too small since conformational averaging leads to considerable errors in NOE intensity. Furthermore, acquiring and especially accurately integrating NOE data for several mixing times is exceedingly time-consuming.

A third possibility to derive distance restraints from NOE cross-peak intensities is the Isolated Spin Pair Approximation (ISPA). Several assumptions are made. First, a single correlation time ($\tau_c$) for the whole molecule is introduced, with which $\tau_{ij}^c$ in eq. (2.26) is replaced.

$$J_{n,ij}(\omega) = \frac{\tau_c}{1 + n^2 \omega_0^2 \tau_c^2} \frac{1}{r_{ij}^6}$$  \hspace{1cm} (2.28)

In some cases local mobility of residues must be taken into account and a modified spectral density function has to be used [Lipari and Szabo, 1982a,b]

$$J_{n,ij}(\omega) = \left( \frac{S^2 \tau_e}{1 + n \omega_0^2 \tau_e^2} + \frac{S^2 \tau_e}{1 + n \omega_0^2 \tau_e^2} \right) \frac{1}{2 r_{ij}^6}$$  \hspace{1cm} (2.29)

where $\tau_e$ is the effective correlation time of the local mobility site and $S^2$ is the generalized order parameter, which is a measure for the flexibility of the site with values ranging from 0 (unrestricted motion) to 1 (fully restricted motion). In proteins, order parameters range from 1 to as low as 0.6 for flexible side chains [Flynn et al., 2001], while $S^2$ in DNA is on the order of 0.8 for all proton pairs [Lane, 1993, 1996]. In the ISPA approach all distances are referenced to a fixed distance (both of which have $S^2 \approx 0.8$). Thus the contribution of local mobility is cancelled. The assumption of a single correlation time for the whole molecule is valid since correlation times for base and sugar protons are comparable [Reid et al., 1989] and oligonucleotides shorter 15 base pairs in length can be assumed isotropic rotors [Birchall and Lane, 1990].

In the ISPA approach the matrix exponential of eq. (2.23) is approximated with a Taylor expansion

$$exp[-R \tau_m] = 1 - R \tau_m + \frac{1}{2} R^2 \tau_m^2 - \ldots$$  \hspace{1cm} (2.30)
whereby the peak intensities in the NOESY spectrum are given as

\[ A_{ij} = \delta_{ij} - R_{ij} \tau_m + \frac{1}{2} \sum_k R_{ik} R_{jk} \tau_m^2 - \ldots \]  

(2.31)

The central assumption in ISPA is that for short mixing times, the Taylor expansion can be truncated after the linear term. Thus any effects of spin diffusion, which is magnetization transfer over third atoms (represented by the quadratic term), are neglected and the cross-peak intensity \((i \neq j)\) becomes a linear function of \(r_{ij}^{-6}\).

\[ A_{ij} = R_{ij} \tau_m = q_{ij} \tau_c \tau_m \left( \frac{5 - 4 \omega_0^2 \tau_c^2}{1 + 4 \omega_0^2 \tau_c^2} \right) \frac{1}{r_{ij}^6} \]  

(2.32)

Consequently, distances can be derived by referencing to a known, fixed distance whereby all constant terms are cancelled.

\[ \frac{A_{ref}}{A_{ij}} = \left( \frac{r_{ij}}{r_{ref}} \right)^6 \quad \text{or} \quad r_{ij} = r_{ref} \sqrt[6]{\frac{A_{ref}}{A_{ij}}} \]  

(2.33)

A commonly used reference distance is the C-H5-H6 distance which is fixed at appr. 2.5 Å [Reid et al., 1989]. While many cross-peaks can be referenced with this distance, it is necessary to introduce several others to account for fast rotation in methyl groups (e.g. C7-H7) or solvent exchange with amino and imino protons (e.g. H42-H5).
2.3 Solution structure determination of DNA

Fig. 2.12: Panel (a) shows the steric interaction of DNA with Pf1 which prevents the DNA from tumbling isotropically and thus induces residual order. The latter gives rise to Residual Dipolar Couplings (RDC) which can be determined by measuring the difference in dipolar coupling of i.e. an N-H bond vector with and without alignment (panel (b)). This picture is taken from MacDonald and Lu [2002].

2.3.3 Residual Dipolar Couplings

In combination with NOE data, Residual Dipolar Couplings (RDC) are now routinely employed for the NMR structure determination of macromolecules as they compensate for the drawbacks of NOE data. Due to the \( r_{ij}^{-6} \)-dependence of the NOE (cf. section 2.3.1) only information on closely spaced spins (\( r_{ij} \approx 5 \text{ Å} \)) is provided. In proteins, long-range peaks between residues far apart in the primary sequence can be observed owing to the tertiary fold. Due to the rod-like shape of short and medium-size DNA, only information on directly adjacent base pairs is available. Long-range effects like kinking in A-tract DNA could thus not be described prior to the development of RDC measurement [MacDonald et al., 2001, Stefl et al., 2004].

RDCs yield information on the orientation of bond vectors relative to the molecular frame. Thus also distant parts in a macromolecule can be characterized relative to each other, which significantly improves results from structure calculations, especially with regard to the global fold [Mauffret et al., 2002].

The first account of measuring anisotropic interactions was published more than 45
years ago by Saupe and Englert [1963]. Only after the advent of high-resolution NMR spectrometers and methods, this technique could be applied to biological macromolecules with success. Bax and Tjandra [1997] were able to demonstrate that dissolving the protein ubiquitin in a very dilute solution of bicelles induces residual order while retaining the high resolution of NMR spectra. This marked a breakthrough in the field of biomolecular NMR since measurement of residual order effects such as RDCs, were not limited to molecules with natively high magnetic susceptibility anisotropy anymore [Tolman et al., 1995, Tjandra et al., 1996]. Initially application of RDCs to NMR structure determination centered on proteins [Bax and Tjandra, 1997, Tjandra and Bax, 1997, Tjandra et al., 1997], but subsequently the importance of RDCs for the structure determination of DNA was demonstrated extensively in theory [Vermeulen et al., 2000, Mauffret et al., 2002] and in practice [Tjandra et al., 2000, Zidek et al., 2001, MacDonald and Lu, 2002, Wu et al., 2004, Stefl et al., 2004].

Inducing just enough alignment for reliable measurement of RDCs, while retaining the high resolution and unambiguousness of the spectra is central to the success of RDC measurements. Tjandra and Bax [1997] showed that the degree of alignment can easily be adjusted by varying the concentration of the bicelles. In the following, diverse alignment media were developed in order to extend the applicability of the approach: bicelle-based alignment [Tjandra and Bax, 1997, Tjandra et al., 1997, Ottiger and Bax, 1999a, Barrientos et al., 2000, Al-Hashimi et al., 2000, Ruckert and Otting, 2000], filamentous phage [Hansen et al., 1998], stretched gels [Chou et al., 2001, Ma et al., 2008] or paramagnetic tagging [Wohnert et al., 2003]. The bacteriophage Pf1 proved to be particularly suited for aligning oligonucleotides since it is stable over a wide range of temperatures and Pf1-DNA interaction is minimized due to electrostatic repulsion between the two negatively charged macromolecules [Hansen et al., 1998, Prestegard et al., 2000]. The steric interaction of Pf1 with DNA is illustrated in Fig. 2.12a.

RDCs ($D_{ij}$) are determined by measuring the difference of the dipolar coupling in the
2.3 Solution structure determination of DNA

![Diagram of magnetic field and internuclear vector angles](image)

**Fig. 2.13:** Orientation of the magnetic field, defined by angles $\xi_x, \xi_y, \xi_z$ and the internuclear vector, described by angles $\zeta_x, \zeta_y, \zeta_z$, in the macromolecular frame. This picture is taken from Blackledge [2005].

The presence ($^1J_{ij}^{ani}$) and absence of molecular alignment ($^1J_{ij}^{iso}$), as is shown in Fig. 2.12b.

\[
^1J_{ij}^{ani} = ^1J_{ij}^{iso} + D_{ij}
\]  

(2.34)

The RDC term can be described by the following equation

\[
D_{ij} = -\gamma_i \gamma_j \mu_0 \hbar \frac{3}{4 \pi^2} \left(\frac{3 \cos^2 \alpha_{ij}(t) - 1}{2 r_{ij}^2(t)}\right)
\]  

(2.35)

where $\alpha$ denotes the angle of the internuclear vector between atoms i and j with the applied magnetic field and $\langle \rangle$ designates the time and ensemble average. When no preferred alignment is induced, all values of $\alpha$ are sampled with equal probability over time and thus $D_{ij}$ is averaged to zero. If residual order is induced by aligning the molecule, information on the orientation of the internuclear vector can be extracted. However, in order to have a fixed reference system, it is desirable to relate the orientation of the internuclear vector to the macromolecular frame rather than the magnetic field. To achieve this, two conditions must be met: 1) the distance between the atoms which give rise to $D_{ij}$ does not change significantly over time (or else its distribution is known), such that $r_{ij}$ can be substituted in eq. (2.35) with an effective distance $r_{ij}^{eff}$ which is known and where averaging is already taken into account. 2) It is assumed that for macromolecules the time average of $\alpha$ can be expressed by two convoluted motions, the macromolecule tumbling with respect to the magnetic field vector ($\xi_x, \xi_y, \xi_z$) and the internuclear vector.
moving inside the macromolecular frame \((\zeta_x, \zeta_y, \zeta_z)\) as illustrated in Fig. 2.13.

\[
\cos \alpha_{ij} = \begin{pmatrix} \cos \xi_x \\ \cos \xi_y \\ \cos \xi_z \end{pmatrix} \begin{pmatrix} \cos \zeta_x \\ \cos \zeta_y \\ \cos \zeta_z \end{pmatrix} = \sum_{k=x,y,z} \cos \xi_k \cos \zeta_k \tag{2.36}
\]

To a first approximation, the internuclear vector is considered to be fixed within the macromolecular frame and thus time-averaging has no effect on \(\zeta\). Introducing the alignment tensor \(A\) with dimensionless units

\[
A_{kl} = \langle \cos \xi_k \cos \xi_l \rangle \tag{2.37}
\]

eq (2.35) can then be rewritten as

\[
D_{ij} = -\frac{3 \gamma_i \gamma_j \mu_0 \hbar S_{\text{flex}}}{8 \pi^2 (\tau_{ij}^\text{eff})^3} x_{y,z} \sum_{k,l=x,y,z} \left( A_{kl} \cos \xi_k \cos \xi_l - \frac{1}{9} \delta_{kl} \right) \tag{2.38}
\]

where \(S_{\text{flex}}\) represents an order parameter to account for local flexibility of the internuclear vector. Here the assumption is that the local motion does not influences the overall alignment of the macromolecule, which is reasonable in the absence of large amplitude fluctuations. In most cases, the local motion can be approximated with the diffusion-in-a-cone model, where the order parameter \(S_{\text{flex}}\) is related to the generalized order parameter \(S\), which scales down the measured RDCs linearly [de Alba and Tjandra, 2002].

In the present framework the alignment tensor \(A\) has all elements non-zero. It is however desirable to find a specific molecular frame in which all off-diagonal elements of \(A\) are zero. Such a frame, called the principal axis system (PAS) can be found by a 3D Euler rotation of the current molecular frame with parameters \(\alpha, \beta, \gamma\) [Blackledge, 2005]. Eq. (2.38) can then be rewritten in terms of the polar angles \(\theta, \phi\) (Fig. 2.14, left panel), which describe the orientation of the internuclear vector in the eigenframe of the
2.3 Solution structure determination of DNA

Fig. 2.14: Orientation of the internuclear vector with polar coordinates $\theta$ and $\phi$ in the eigenframe of the alignment tensor with eigenvalues $A_{xx}$, $A_{yy}$, and $A_{zz}$ (left panel). The right panel illustrates the orientational degeneracy of RDCs. As can be clearly seen a large number of different orientations can be sampled when the RDC takes intermediate values. Only extreme RDC value define unique orientations of the internuclear vector. This picture is taken from Blackledge [2005].

alignment tensor with eigenvalues $|A_{xx}| \leq |A_{yy}| \leq |A_{zz}|$ [Lipsitz and Tjandra, 2004]

$$D_{ij} = \frac{\gamma_i \gamma_j \mu_0 \hbar S_{\text{flex}}}{4 \pi^2 (r_{ij}^{eff})^3} \left[ A_a (3 \cos^2 \theta - 1) + A_r \sin^2 \theta \cos 2\phi \right] \quad (2.39)$$

where $A_a$ and $A_r$ are the axial and rhombic components of the alignment tensor $A$ [Prestegard et al., 2000], which are given by

$$A_a = \frac{1}{2} A_{zz} \quad \text{and} \quad A_r = \frac{1}{2} (A_{xx} - A_{yy}) \quad (2.40)$$

In consequence, the alignment tensor $A$ is determined by 5 independent parameters: the three Euler angles needed for rotation of the reference frame $(\alpha, \beta, \gamma)$ and the two components of the alignment tensor $A$ ($A_a, A_r$). The alignment tensor can then be unambiguously determined from a minimum set of 5 experimental RDCs by singular value decomposition [Losonczi et al., 1999].

With $A$ at hand, the orientation of any internuclear vector with respect to the macromolecular frame can be calculated. Unfortunately, there exists a multitude of orientations of an internuclear vector that is consistent with a given intermediate RDCs value, as is illustrated in Fig. 2.14, right panel. Only extreme values correspond to unambiguous
orientations. The orientational degeneracy for intermediate RDC values clearly limits their value for structure determination. Thus it is desirable to remove this degeneracy by either measuring more RDCs per residue, or RDCs for the same internuclear vectors but with a different alignment tensor [de Alba and Tjandra, 2002, Blackledge, 2005].
2.3 Solution structure determination of DNA

2.3.4 Simulated Annealing calculations

The main idea behind Simulated Annealing (SA) calculations is coupling simple energy minimization to Molecular Dynamics. Thus the problem that molecules converge to a local minimum instead of the global can be circumvented. The Newtonian equation of motion

\[ F_i(t) = -\frac{\partial V}{\partial r_i} = m_i a_i(t) \]  

relates the acceleration \( a_i(t) \) of each atom \( i \) at time \( t \) to the derivative of the potential energy \( V \) with respect to the atom position \( r_i \). \( F_i(t) \) represents the force which acts upon the atom \( i \) with mass \( m_i \) [Bloomfield et al., 2000]. The atoms are moved according to the force that is exerted upon them for a given time-step (typically between 1-5 fs). From the knowledge of the last and current atom positions and velocities, a new force is calculated which in turn acts on the atoms. This cycle is repeated until a convergence criterion (e.g. a minimum change in the gradient of the potential energy) is met. Initially, atom velocities are computed using a Gaussian or Maxwell distribution. The atom coordinates are derived from a starting structure. Since the initial coordinates and velocities determine all subsequent ones, it is important to start from a reasonable structure. This is usually achieved by starting from already known crystal or NMR structures.

Temperature-coupling of Molecular Dynamics is achieved by introducing an average kinetic energy which is computed via Boltzmann statistics

\[ \langle E_{\text{kin}} \rangle = \langle \frac{1}{2} m \nu^2 \rangle = \frac{3}{2} k_b T \]  

where \( k_b \) is the Boltzmann factor, \( m \) the atom mass and \( \nu \) the atom velocity. The higher the temperature (\( T \)) is chosen, the higher the kinetic energy of the system. Thus at high temperatures kinetic barriers can be overcome, and the global minimum should be accessible.

The number of atoms in macromolecules such as DNA or proteins is on the order of thousands with three times as much cartesian coordinates to be calculated at each step. In order to make computation of macromolecules feasible, Molecular Dynamics calculati-
ons rely on the predefinition of atom types. For these atom types many parameters such as bond lengths, bond angles, dihedral angles, partial charges etc. are assumed to be fixed and are comprised in the force field. In the present work the program XPLOR-NIH [Schwieters et al., 2003] was used, which employs the CHARMM force field [Weiner et al., 1984, MacKerell et al., 2000]. The total potential energy \( V_{\text{tot}} \) consists of two components [Brünger, 1996]

\[
V_{\text{tot}} = E_{\text{emp}} + E_{\text{eff}}
\]  

(2.43)

where the empirical \( E_{\text{emp}} \) and the effective energy term \( E_{\text{eff}} \) are given as [Brünger, 1996].

\[
E_{\text{emp}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihe}} + E_{\text{vdW}} + E_{\text{Coulomb}}
\]  

(2.44a)

\[
E_{\text{eff}} = E_{\text{noe}} + E_{\text{rdc}} + E_{\text{plan}} + E_{\text{cdih}}
\]  

(2.44b)

\( E_{\text{bond}}, E_{\text{angle}}, E_{\text{dihe}} \) and all energy terms of \( E_{\text{eff}} \) are calculated as the product of a scaling factor (the force constant) and the deviation of the observed value from the equilibrium one, e.g.

\[
E_{\text{bond}} = k_{\text{bond}} (r_{ij}^{\text{obs}} - r_{ij}^{\text{equ}})
\]  

(2.45)

The equilibrium values for \( E_{\text{noe}} \) and \( E_{\text{rdc}} \) are taken from experiment while these of \( E_{\text{cdih}} \) and \( E_{\text{plan}} \) are averages from the literature [Brünger, 1996]. The corresponding scaling factors are defined in the calculation input and thus can be used to increase the restraining power of selected energy terms. Equilibrium values and force constants of \( E_{\text{bond}}, E_{\text{angle}}, E_{\text{dihe}} \) constitute one part of the force field. \( E_{\text{vdW}} \) is given as

\[
E_{\text{vdW}} = \sum_{i,j} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right)
\]  

(2.46)

with \( A_{ij} = 2 \sqrt{\varepsilon_{ii} \varepsilon_{jj}} (\sigma_{ii} - \sigma_{jj}) \) and \( A_{ij} = 2 \sqrt{\varepsilon_{ii} \varepsilon_{jj}} (\sigma_{ii} - \sigma_{jj}) \). The atomic permitivities \( (\varepsilon_{ii/jj}) \) and van-der-Waals radii \( (r_{ii/jj}) \) set up another part of the force field. The partial
2.3 Solution structure determination of DNA

atomic charges \((q_i, q_j)\) in the \(E_{\text{Coulomb}}\)-term constitute the last part of the force field.

\[
E_{\text{Coulomb}} = \sum_{ij} \frac{q_i q_j}{\varepsilon_0 r_{ij}} \tag{2.47}
\]

When explicit treatment of water is not feasible due to restrictions on the calculation time, the solvent screening effect can be approximated by introducing a distance dependent permittivity of free space \(\varepsilon_0(r_{ij})\).

Ultimately, force field parameters are based on crystal structures and ab-initio calculations of small model molecules (bond lengths, bond angles and dihedral angles), infrared spectroscopy data (force constants) and empirical testing (where no experimental source is available) [Weiner et al., 1984]. When chemically modified nucleotides are to be incorporated into calculations, parameters for this nucleotide have to be added to the existing force field. In the absence of experimental data, these parameters are calculated using ab-initio methods.
3 Experimental details

3.1 2-Aminopurine

3.1.1 Sample preparation

NMR sample preparation DNA strands were obtained (from BIOTEZ, Berlin, and BIOSPRING, Frankfurt/M, Germany) already purified by reverse-phase high-pressure liquid chromatography (HPLC). After hybridization they were subjected to size exclusion chromatography (SEPHADEX G25) and ammonia treatment (lyophylization with ≈3% NH₃-solution) to remove residual, low molecular weight impurities (mainly Et₃N-buffer from HPLC). Equivalent amounts of complementary single strands were hybridized by rapid heating to 90°C and subsequent gradual cooling to room temperature at a rate of 1°C per minute. NMR samples were prepared at 3 mM duplex concentration in D₂O (D₂O 99.98%) and H₂O (H₂O:D₂O/90:10) at pH 6.5 in 10 mM Na₂HPO₄/NaH₂PO₄ and 150 mM NaCl solution.

RDC sample preparation Samples for measuring residual dipolar couplings (RDC) were prepared in D₂O as described above with the addition of 20 mg/ml Pf1 (obtained from ASLA BIOTECH, Riga) suspended in the same buffer. This necessitates rebuffering of Pf1 since it is obtained in a different and non-deuterated buffer. Rebuffering is achieved by ultracentrifuging 100 µl of Pf1 two times with 600 µl phosphate buffer (as described above) and subsequently two times with 600 µl deuterated phosphate buffer at 60000 rpm. Each ultracentrifuging step is performed at 4°C for 2 h. Next the sediment is suspended in the DNA sample. The high viscosity of Pf1 complicates sample handling and thus the suspension has to be stirred until a viscose, clear, gel-like sample is obtained. After
transferring it into the NMR tube, bubbles have to be removed by slow centrifugation of
the NMR tube (max. 500 rpm). The degree of orientation can be checked by measuring
the quadrupolar deuterium splitting of the HOD signal [Lipsitz and Tjandra, 2004].
In case of degradation or non-complete suspension of Pf1 in the sample, the expected
symmetric doublet with splitting of $10\pm5\,\text{Hz}$ is asymmetric, extremely broadened or
even non-observable.
3.1.2 Measurements

**Duplex melting** Melting of the 13mer2AP duplex was monitored by optical absorption at 260 nm and by the fluorescence quantum yield (due to the 2AP nucleobase) from 310 nm excitation [Evans et al., 1992]. For comparison, melting of the 13merRef duplex was measured by absorption only. The solutions had a total single-strand concentration of 23.7 mM and the optical path length was 1 cm. Temperature was varied between 25 and 85 °C; the standard error of transfer between absorption and emission temperatures was ± 0.011 °C. Following changes of typically 1 °C, equilibration was allowed for at least 15 minutes. The relative fluorescence yield of pure 2AP in buffered water was measured in the same manner for reference. Measurements were corrected for density changes.

**Imino proton exchange** Exchange rates of the imino protons were obtained from inversion recovery experiments at 298 K on a Bruker Avance 600 NMR spectrometer. For this purpose the standard 1D 1H-WATERGATE (water suppression by gradient tailored excitation) pulse program was modified to start with a selective inversion pulse. As discussed in section 2.2.2 the pulse shapes least sensitive to relaxation effects are rectangular and gaussian pulses. Dornberger et al. [1999] and Bhattacharya et al. [2002] use millisecond gaussian-shape pulses for inversion, but such long pulse durations again facilitate relaxation effects. Additionally, the imino protons are shifted upfield (≈13 ppm) of the bulk of the DNA (1-9 ppm) and especially the HOD (≈4.80 ppm) NMR signals and thus off-resonance excitation effects can be ruled out. In consequence, a selective rectangular inversion pulse of 327 µs duration centered on the imino proton region was employed in this work. After a variable delay with 21 settings ranging from τ = 20 µs up to 2 ms, a 1D spectrum was recorded. The settings for the variable delay have been optimized to give accurate results for slowly as well as fast exchanging imino protons.

Samples were prepared in H2O with the buffer described earlier. In order to cancel the effect of intrinsic catalysis (cf. section 2.2.1), the spin-lattice relaxation time $T_1^{int}$ of all imino protons was measured prior to titration with base catalyst. Subsequently, each duplex was titrated with 1 M TRIS buffer as base catalyst for proton exchange. Test runs
with ammonia as base catalyst gave inferior results, since the latter is highly volatile. This makes high-concentration salt solutions of ammonia, which are necessary in order to keep the pH stable upon addition of the catalyst, inaccessible. Another set of test runs, where pure ammonia and HCl (for pH compensation) were added separately to the NMR sample, resulted in temporary but extreme pH changes (pH values ranging from 1 to 11). These changes led to gradual degradation of the DNA sample. Furthermore, the error in concentration introduced by evaporation of ammonia is intolerable when high precision results are to be obtained. In order to keep the pH as stable as possible, TRIS-hydrochloride at pH 7.4 was used. At every titration point the pH was measured in order to determine the amount of free base catalyst precisely (cf. eq. (2.15 in section 2.2.1). All pH values centered around 7.1. To estimate the error in $T_1$-determination, three sets of measurements were carried out at each titration point, resulting in 693 1D spectra (21 delay settings, 11 titration points).

Each spectrum was fitted (with the MATHEMATICA program environment) to give the integrals of the imino proton peaks. Complete spectral fitting became necessary since standard region integration as implemented in TOPSPIN is sensitive to peak overlap. Due to continued broadening of the imino proton signals upon base catalyst addition, integrals obtained by the latter method were inconsistent and thus gave rise to large errors (up to 100%) in $T_1$-values. In contrast, complete spectral fitting in MATHEMATICA yielded $T_1$-values with errors of typically 0.5-5%.

**Structural NMR**  NMR measurements for structure calculation were carried out on a BRUKER AVANCE 600 NMR spectrometer. The optimum temperature of 298 K was determined by monitoring the imino proton signal intensity (cf. Fig. 3.1). For each duplex, NOESY-, DQF-COSY- (double quantum filtered correlated spectroscopy), TOCSY- (total correlation spectroscopy) and HMQC-spectra (heteronuclear multiple quantum coherence spectroscopy) in $D_2O$, a WATERGATE-NOESY-spectrum in $H_2O$ and an HMQC-spectrum in $D_2O/Pf1$ were recorded. The quadrupolar splitting of the $^2H$ NMR signal after addition of Pf1-phage was 15.72 Hz and 13.65 Hz for 13merRef and 13mer2AP, re-
3.1 2-Aminopurine

Fig. 3.1: Imino proton signal intensities for the temperature range 283-328K. in 13mer2AP.

pectively. Upon addition of Pf1, the sample became very viscose and standard shimming procedures failed to produce reasonable linewidths. Two workarounds were investigated. One was to optimize the shim for a BRUKER standard sample and use these shim settings on the DNA/Pf1 sample. The second workaround involved optimization of the lineshape of the HOD signal. This procedure demands that after every change of the shim settings, a 1D spectrum of the sample had to be acquired. Although the latter procedure is more time-consuming than the other one, it yielded supreme results and thus was employed.

For DQF-COSY- (TOCSY-) spectra, 16 (32) transients were acquired, with 2048×256 points in F2 and F1 dimensions. For NOESY-spectra in both solvents, 16 transients were acquired with 4096×2048 points at a mixing time of 150 ms. For the HMQC-spectra with and without Pf1, 192 transients were acquired with 8192×512 points. The high number of data points was necessary to obtain sufficient resolution in the proton dimension in order to determine the RDC-values with a precision below 1 Hz. The optimal d2-delay, where both the non-aromatic and aromatic region are equally enhanced, was determined to be 2.5 ms. All spectra were processed with the BRUKER TOPSPIN-software, and signals were assigned with the help of CARA [Keller, 2004].
3 Experimental details

3.1.3 Restraint generation

**Force field parametrization**  Density functional theory (DFT) calculations of the 2AP-nucleotide were performed with GAUSSIAN03 employing a triple zeta valence plus polarization (TZVP) basis set. Due to the structural proximity of 2AP and A, force field parameter and topology entries of the latter were adopted and changed where necessary. Special emphasis was placed on the charge distribution since this was shown to be the main difference between A and 2AP [Broo, 1998, Mishra et al., 2000, Jean and Krueger, 2006]. For this purpose several methods for obtaining point charges were compared. The fastest method is the Mulliken population analysis, which, however, gives unrealistic results [Breneman and Wiberg, 1990]. This is due to the equal distribution of overlap populations between the two involved atoms. Hence another approach has been developed, which tries to derive atomic charges via fitting of point charges to the Molecular Electrostatic Potential (MEP) [Connolly, 1983, Singh and Kollman, 1984]. Crucial to the success of this technique is the algorithm which is used for fitting. Three different algorithms were tested; the CHelp-[Chirlian and Francl, 1987], the CHelpG-[Breneman and Wiberg, 1990] and the MK-algorithm [Besler et al., 1990]. Out of these three the MK- and CHelpG-algorithms gave consistent and similar results. Partial charges derived with the CHelpG-algorithm were employed in the force field. Although the sugar moiety was simulated as well, only the partial charges for the 2AP residue were integrated into the force field. Due to the similarity of A and 2AP the influence of the latter on the partial charges of the sugar can be considered negligible (as the same values for partial charges of sugar atoms are also used for G,C and T residues). Thus the partial charge values for the sugar were taken as already defined for the native residues in the force field. To retain the neutrality of the 2AP residue the partial charge at the N7 atom of 2AP was increased from -0.61 to -0.45 (cf. section 3.1.3).

**Distance restraints**  The assigned NOE cross-peaks were converted to distance restraints by referencing their integrals to the integrals of known distances employing the Isolated Spin Pair Approximation. The NOE cross-peaks were integrated with the pro-
gram CARA [Keller, 2004] using the sum-over-rectangle method. As reference distances Me-H6 T (3.1 Å) for all NOE cross-peaks involving methyl protons, H42-H5 C (2.4 Å) for all NOE cross-peaks involving exchangeable protons and H5-H6 C (2.5 Å) for the remaining NOE cross-peaks were used (bond lengths adapted from the force field parameters). For the purpose of exporting the integral values obtained by CARA [Keller, 2004] to an XPLOR-NIH [Schwieters et al., 2003] restraints file, a LUA script was written (see section .4). This script classifies the integrated peaks according to the overlap with other peaks and scales their volume integrals accordingly. Additionally, uncertainties for the NOE restraints are automatically calculated from the standard deviation of the reference peaks’ volume integrals. The estimated uncertainty is then increased according to the classification of each peak. Furthermore, this classification is printed into a separate file, which can be used to assess whether or not peak overlap might prevent a reliable estimation of the peak volume.

Residual Dipolar Coupling restraints The experimentally determined RDCs (see tables 3.1 and 3.2) of C-H bond vectors can be input into the structure calculations with the measurement precision as error bounds. From the experimentally determined RDC-values the orientation of the corresponding internuclear vector is determined via eq. (2.39) in section 2.3.3. All constants in the latter equation are comprised within the factor $D_a$

$$D_a = -\frac{\gamma_i \gamma_j \mu_0 \hbar S_{\text{flex}}}{4\pi^2 (r_{ij})^3}$$

(3.1)

for which only one value is used throughout the calculation. This necessitates scaling of different sets of RDCs, for example when C-H as well as C-C RDCs have been measured. Scaling is achieved by introducing prefactors to the $D_a$-factor, which are defined as the ratio of the two $D_a$-factors involved. For example, when C-C RDCs are to be scaled to a set of C-H RDCs, the corresponding prefactor would be calculated as:

$$D_{a}^{\text{pre}}(CC) = \frac{D_a(CH)}{D_a(CC)} = \frac{\gamma_H}{\gamma_C} \left(\frac{r_{CC}}{r_{CH}}\right)^3 = \frac{42.576}{10.705} \left(\frac{1.496}{1.090}\right)^3 \approx 10.28 .$$

(3.2)
where $\frac{2\gamma}{\gamma_C}$ is the change in gyromagnetic ratio when working with C-H or C-C RDCs and $r_{CC}$ and $r_{CH}$ are the C-C and C-H bond lengths between the atoms defining the internuclear vector. Although in this work only C-H RDCs were measured, scaling issues become important for the implementation of the experimentally determined C-H RDCs of the T methyl groups.

Due to the fast rotation of the latter, only an averaged value for all three C-H bond vectors can be measured. This fast rotation scales the corresponding RDC values by

$$P_2(\cos\beta) = \frac{3}{2} \cos^2 \beta - \frac{1}{2}$$

(3.3)

where $\beta$ is the C5-C7-H7* angle. The latter is usually assumed to be the ideal tetrahedral angle of 109.5°, but was experimentally determined for methyl groups to be 110.9° by Ottiger and Bax [1999b]. Although small, the deviation from the ideal angle has a strong impact on methyl RDC scaling since $P_2(\cos\beta)$ is a steep function $\beta$. Thus with $P_2(\cos 109.5)=-0.3329$ and $P_2(\cos 110.9)=-0.3089$, conversion factors between C5-C7 and C7-H7* methyl RDCs can be calculated as -3.42 and -3.17, respectively, according to

$$\frac{D_{CH\_Me}}{D_{CC\_Me}} = P_2(\cos\beta) \ D_{a\CC}^{\text{pre}}(CC) .$$

(3.4)

The value of -3.17 was determined by measuring the correlation of experimentally determined C-H and C-C methyl RDCs [Ottiger and Bax, 1999b].

In order to implement the methyl C-H RDCs into the structure calculations, they have to be associated with a unique internuclear vector. Thus it is necessary to convert the C7-H7* RDCs into the corresponding C5-C7 RDCs. Since the latter must now be input as C-C RDCs, a separate restraints file has to be used. There are in principle two ways how this implementation can be achieved in XPLOR-NIH v2.20 [Schwieters et al., 2003]. One way is to convert all experimentally determined methyl C-H RDC values by hand to the corresponding C-C values with the factor $1/(-3.17)=-0.3155$. This is done prior to implementing them into the calculation. In that case, the prefactor $D_{a\CC}^{\text{pre}}(CC)$ can be used to scale the C-C RDC input file to the C-H RDC input. According to the manual,
3.1 2-Aminopurine

this scaling is handled automatically by the scale_toCH routine of the program XPLOR-NIH v2.20 [Schwieters et al., 2003]. Due to a bug however, the automatic scaling is only effective when the methyl C-C RDCs are input as type “CACO”. The other way to input C-H methyl RDCs into structure calculations will be implemented in the new version of XPLOR-NIH (v2.24) but was made available to the author in advance (personal communication of Dr. Charles Schwieters). In that new version, the methyl RDCs are input as a separate restraints file too, but their values do not have to be scaled by hand. They are automatically recognized by the program as methyl RDCs and scaled to the C-H RDC input by

$$D_{Me}^{pre} = -3.17 \cdot D_a^{pre}(CC) = 32.59 \text{ .}$$

(3.5)

However, scaling of the $D_a$-factor and scaling the RDC values is not equivalent, since the energy for the RDC potential term is given by [Brünger, 1996]

$$E_{RDC} = k_{RDC} (D_{calc} - D_{obs})^2$$

(3.6)

where $k_{RDC}$ is the scale factor for the RDC energy term. The latter must be modified for different sets of RDCs according to their respective errors with a weighting factor $\omega_{ij}$. When using the prefactor $D_{Me}^{pre}$ for implementation of methyl RDCs, the energy of the C-C methyl RDCs has to be scaled by

$$\omega_{CC} = \frac{1}{(-3.17)^2} \omega_{CH} = 0.0995 \omega_{CH}$$

(3.7)

due to dependence of the energy on the square of the difference between calculated ($D_{calc}$) and observed ($D_{obs}$) RDC.
### 3 Experimental details

**Tabelle 3.1:** Experimentally determined RDCs for 13merRef. The RDCs were measured with a precision of ±0.6 Hz.

<table>
<thead>
<tr>
<th>Res Vector</th>
<th>(J_{(CH)}) (Hz)</th>
<th>(J_{(CH)}) (aligned) (Hz)</th>
<th>RDC (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A6 C2-H2</td>
<td>201.6</td>
<td>225.6</td>
<td>24.0</td>
</tr>
<tr>
<td>A7 C2-H2</td>
<td>203.4</td>
<td>224.4</td>
<td>21.0</td>
</tr>
<tr>
<td>A8 C2-H2</td>
<td>202.2</td>
<td>223.8</td>
<td>21.6</td>
</tr>
<tr>
<td>A16 C2-H2</td>
<td>202.8</td>
<td>214.8</td>
<td>12.0</td>
</tr>
<tr>
<td>A24 C2-H2</td>
<td>201.6</td>
<td>224.4</td>
<td>22.8</td>
</tr>
<tr>
<td>C9 C5-H5</td>
<td>171.6</td>
<td>190.2</td>
<td>18.6</td>
</tr>
<tr>
<td>C17 C5-H5</td>
<td>170.4</td>
<td>183.6</td>
<td>13.2</td>
</tr>
<tr>
<td>T3 C7-H7</td>
<td>126.6</td>
<td>122.4</td>
<td>-4.2</td>
</tr>
<tr>
<td>T11 C7-H7</td>
<td>126.0</td>
<td>118.8</td>
<td>-7.2</td>
</tr>
<tr>
<td>T19 C7-H7</td>
<td>123.0</td>
<td>112.8</td>
<td>-10.2</td>
</tr>
<tr>
<td>T20 C7-H7</td>
<td>127.2</td>
<td>118.2</td>
<td>-9.0</td>
</tr>
<tr>
<td>T21 C7-H7</td>
<td>126.6</td>
<td>118.2</td>
<td>-8.4</td>
</tr>
<tr>
<td>T3 C1'-H1'</td>
<td>155.4</td>
<td>167.4</td>
<td>12.0</td>
</tr>
<tr>
<td>C9 C1'-H1'</td>
<td>162.0</td>
<td>169.2</td>
<td>7.2</td>
</tr>
<tr>
<td>C17 C1'-H1'</td>
<td>162.6</td>
<td>166.8</td>
<td>4.2</td>
</tr>
<tr>
<td>T20 C1'-H1'</td>
<td>163.8</td>
<td>178.2</td>
<td>14.4</td>
</tr>
<tr>
<td>T21 C1'-H1'</td>
<td>153.6</td>
<td>174.0</td>
<td>20.4</td>
</tr>
<tr>
<td>T3 C6-H6</td>
<td>177.0</td>
<td>195.0</td>
<td>18.0</td>
</tr>
<tr>
<td>C5 C6-H6</td>
<td>180.0</td>
<td>201.0</td>
<td>21.0</td>
</tr>
<tr>
<td>T20 C6-H6</td>
<td>177.0</td>
<td>192.6</td>
<td>15.6</td>
</tr>
<tr>
<td>T21 C6-H6</td>
<td>175.8</td>
<td>190.8</td>
<td>15.0</td>
</tr>
<tr>
<td>C23 C6-H6</td>
<td>175.8</td>
<td>201.0</td>
<td>25.2</td>
</tr>
<tr>
<td>A7 C8-H8</td>
<td>213.0</td>
<td>235.8</td>
<td>22.8</td>
</tr>
<tr>
<td>A16 C8-H8</td>
<td>213.0</td>
<td>230.4</td>
<td>17.4</td>
</tr>
</tbody>
</table>
### 3.1 2-Aminopurine

Tabelle 3.2: Experimentally determined RDCs for 13mer2AP. The RDCs were measured with a precision of ±0.6 Hz.

<table>
<thead>
<tr>
<th>Res</th>
<th>Vector</th>
<th>$J_{(CH)}$ (Hz)</th>
<th>$J_{(CH)}$ (aligned) (Hz)</th>
<th>RDC (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A6</td>
<td>C2-H2</td>
<td>202.2</td>
<td>222.0</td>
<td>19.8</td>
</tr>
<tr>
<td>2AP7</td>
<td>C6-H6</td>
<td>179.4</td>
<td>200.4</td>
<td>21.0</td>
</tr>
<tr>
<td>A8</td>
<td>C2-H2</td>
<td>202.8</td>
<td>218.4</td>
<td>15.6</td>
</tr>
<tr>
<td>A16</td>
<td>C2-H2</td>
<td>201.6</td>
<td>213.6</td>
<td>12.0</td>
</tr>
<tr>
<td>A24</td>
<td>C2-H2</td>
<td>201.6</td>
<td>219.0</td>
<td>17.4</td>
</tr>
<tr>
<td>C5</td>
<td>C5-H5</td>
<td>170.4</td>
<td>184.2</td>
<td>13.8</td>
</tr>
<tr>
<td>C9</td>
<td>C5-H5</td>
<td>171.0</td>
<td>189.6</td>
<td>18.6</td>
</tr>
<tr>
<td>C17</td>
<td>C5-H5</td>
<td>170.4</td>
<td>183.0</td>
<td>12.6</td>
</tr>
<tr>
<td>T3</td>
<td>C7-H7</td>
<td>127.2</td>
<td>121.2</td>
<td>-6.0</td>
</tr>
<tr>
<td>T11</td>
<td>C7-H7</td>
<td>127.2</td>
<td>121.2</td>
<td>-6.0</td>
</tr>
<tr>
<td>T19</td>
<td>C7-H7</td>
<td>126.6</td>
<td>118.2</td>
<td>-8.4</td>
</tr>
<tr>
<td>T20</td>
<td>C7-H7</td>
<td>127.2</td>
<td>120.6</td>
<td>-6.6</td>
</tr>
<tr>
<td>T21</td>
<td>C7-H7</td>
<td>126.6</td>
<td>119.4</td>
<td>-7.2</td>
</tr>
<tr>
<td>T3</td>
<td>C1'-H1'</td>
<td>160.8</td>
<td>172.2</td>
<td>11.4</td>
</tr>
<tr>
<td>2AP7</td>
<td>C1'-H1'</td>
<td>156.6</td>
<td>170.4</td>
<td>13.8</td>
</tr>
<tr>
<td>C9</td>
<td>C1'-H1'</td>
<td>160.8</td>
<td>168.6</td>
<td>7.8</td>
</tr>
<tr>
<td>C17</td>
<td>C1'-H1'</td>
<td>162.6</td>
<td>166.2</td>
<td>3.6</td>
</tr>
<tr>
<td>T21</td>
<td>C1'-H1'</td>
<td>161.4</td>
<td>182.4</td>
<td>21.0</td>
</tr>
<tr>
<td>T3</td>
<td>C6-H6</td>
<td>175.2</td>
<td>193.8</td>
<td>18.6</td>
</tr>
<tr>
<td>C5</td>
<td>C6-H6</td>
<td>175.2</td>
<td>195.6</td>
<td>20.4</td>
</tr>
<tr>
<td>T20</td>
<td>C6-H6</td>
<td>176.4</td>
<td>190.2</td>
<td>13.8</td>
</tr>
<tr>
<td>T21</td>
<td>C6-H6</td>
<td>176.4</td>
<td>192.0</td>
<td>15.6</td>
</tr>
<tr>
<td>C23</td>
<td>C6-H6</td>
<td>174.0</td>
<td>199.2</td>
<td>25.2</td>
</tr>
<tr>
<td>2AP7</td>
<td>C8-H8</td>
<td>214.2</td>
<td>234.6</td>
<td>20.4</td>
</tr>
<tr>
<td>A16</td>
<td>C8-H8</td>
<td>213.6</td>
<td>229.8</td>
<td>16.2</td>
</tr>
</tbody>
</table>
3 Experimental details

3.1.4 Structure calculation

Calculation input Structure calculations were performed with XPLOR-NIH v2.20 [Schwitters et al., 2003]. A total of 340 (333) NOE distance restraints and 24 (25) Residual Dipolar Couplings were used for 13merRef (13mer2AP). The experimental data were supplemented with 144 backbone dihedral restraints, 78 hydrogen bond distance restraints and 28 planarity restraints (see Table 3.3).

Initial MD-calculations were performed with dihedral restraints allowing both A-form and B-form conformations (with error bars of ±50°). B-form conformation was experimentally confirmed by $^3$J coupling constants for H1'-H2' derived by P.E.COSY (primitive exclusive correlated spectroscopy) and NOESY-cross-peak intensities characteristic of B-DNA. Consequently, regular dihedral values from the literature [Roberts, 1993] were included in the calculations.

<table>
<thead>
<tr>
<th>Tabelle 3.3: Overview of structural statistics for 13merRef and 13mer2AP calculations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NOE restraints</strong></td>
</tr>
<tr>
<td>- total</td>
</tr>
<tr>
<td>- interresidue</td>
</tr>
<tr>
<td>- intraresidue</td>
</tr>
<tr>
<td>- 2AP to DNA</td>
</tr>
<tr>
<td><strong>RDC restraints</strong></td>
</tr>
<tr>
<td><strong>Dihedral angle restraints</strong></td>
</tr>
<tr>
<td><strong>H-bond restraints</strong></td>
</tr>
<tr>
<td><strong>Base pair planarity restr.</strong></td>
</tr>
<tr>
<td><strong>NOE viol. (&gt; 0.5 Å)</strong></td>
</tr>
<tr>
<td><strong>RDC viol. (&gt; 0.4 Hz)</strong></td>
</tr>
<tr>
<td><strong>Dihe viol. (&gt; 5°)</strong></td>
</tr>
<tr>
<td><strong>RMSD to ave. struct. in Å</strong></td>
</tr>
</tbody>
</table>

The structures were calculated in two steps. First, a reasonable starting structure with well defined local conformation was computed. To ensure that no bias is introduced towards local energy minima, the calculation started from an elongated and equilibrated structure. The resulting structure, which is mainly defined by NOE restraint data, was used as input for Simulated Annealing calculations including RDC data. The need for locally well defined starting structures in order to calculate reasonable structures which
satisfy NOE as well as RDC data is documented in the literature [Vermeulen et al., 2000, Mauffret et al., 2002].

**Simulated Annealing protocol**  The two complete MD protocols are given in the Appendix, section .2. Only the protocols for the 13mer2AP calculation is given, since the other calculations were carried out with the same protocol (with the exception of the input file names). The input scripts are based on the example files of the Xplor-NIH package (refine_full.py and sa.inp) but were substantially modified in the course of this work.

The MD protocol with only NOE restraints as experimental input consisted of an initial minimization (50 steps) followed by 48 ps of high-temperature cartesian coordinate dynamics at 3000 K, subsequent gradual cooling to 25 K in 120 steps of 0.05 ps length and a final minimization (3000 steps).

The MD protocol including the experimental RDC restraints consisted of an initial cartesian coordinate minimization (1000 steps) followed by 50 ps of high-temperature cartesian coordinate torsion angle dynamics at 20000 K, subsequent gradual cooling to 25 K in 154 steps of 0.5 ps length (34 steps to cool down to 3000 K, followed by 120 steps to reach the end temperature) and a final minimization (3000 steps). The alignment tensor values were allowed to float during the calculations, as implemented in Xplor-NIH (v2.20) [Schwieters et al., 2003].

For each run an ensemble of 100 structures was computed. The 10 minimum energy structures without violation of restraints were chosen to compute an averaged structure which was energy-minimized to yield the final structure. The root-mean-squares-deviation (RMSD) of the 10 minimum energy structures to the average, minimized structure is a measure for the precision of the calculation.

**Structure validation**  To check the accuracy of the structures, NOESY-spectra were back-calculated from the average structures with the Full Matrix Relaxation Approach implemented in Xplor-NIH (v2.20) [Schwieters et al., 2003]. The back-calculated spectra were visualized with the program GIFA [Pons et al., 1996] and overlayed with the
3 Experimental details

experimental ones. Furthermore, RDCs were predicted from the average structure using the program PALES [Zweckstetter and Bax, 2000] and compared to the experimental ones. Finally, helical parameters were calculated with the help of the program 3DNA [Lu and Olson, 2003]. These were checked for consistency with values for regular B-DNA helices.

A number of convenience scripts were written to automate data conversion between PALES [Zweckstetter and Bax, 2000] and XPLOR-NIH (v2.20) [Schwieters et al., 2003], the latter program and GIFA [Pons et al., 1996] and for quick access to energies or number of restraints in structure calculations output files. These utility scripts are comprised in the Appendix in section .5.
3.2 2-Hydroxy-7-nitrofluorene

Since many experimental details for HNF are equivalent to the ones for 2AP, the description for the latter is referenced where appropriate in order to avoid redundancy.

3.2.1 Sample preparation

2-Hydroxy-7-nitro-fluorenewas synthesized (Matthias Pfaffe) in four steps. The 2'-deoxyriboside dRi-HNF was prepared by reaction with 1'α-Chloro-3',5'-di-O-toluoyl-2'-deoxy-D-ribose in the presence of activated molecular sieve. Quantification of the H1'-H2'/H2" and H3'-H2'/H2" NOE cross-peaks in the NOESY-spectra of the main product revealed that the α-glycoside was formed predominantly. After purification of the latter by column chromatography, the corresponding phosphor-amidite was reached by standard methods. Fixed-phase synthesis of the labelled DNA strand (BIOTEZ) required a fourfold increase over the normal reaction time for coupling dRi-HNF. Further sample preparation is equivalent to the one described in section 3.1.1.
3 Experimental details

3.2.2 Measurements

**Duplex melting**  Melting of the 13merHNF duplex was monitored by optical absorption at 260 nm and by the red-shift of the weak absorption band of HNF at 380 nm. The solution had a total single-strand concentration of 23.5 mM and the optical path length was 1 cm. The following procedures are equivalent to the ones described in section 3.1.2.

**Structural NMR**  NMR measurements for structure calculation were carried out on a Bruker AVANCE 600 NMR spectrometer at 298 K. For each duplex, NOESY-, DQF-COSY-, TOCSY- and HMQC-spectra in D$_2$O, a WATERGATE-NOESY-spectrum in H$_2$O and an HMQC-spectrum in D$_2$O/Pf1 were recorded. The quadrupolar splitting of the $^2$H NMR signal after addition of Pf1-phage was 10.55 Hz. The following procedures and settings are equivalent to the ones described in section 3.1.2.
3.2 2-Hydroxy-7-nitrofluorene

3.2.3 Restraint generation

Distance restraints, RDC restraints and force field parameters were generated as described in section 3.1.3. The experimental RDC values that were used for restraint generation are given in Table 3.4. In the case of HNF however, no topology information for a structurally similar compound was available. Thus the corresponding parameter and topology input files had to be generated from the structure calculated by DFT methods as described in section 3.1.3. The latter can be achieved utilizing the program Xplo2D [Kleywegt and Jones, 1998].

The influence of the HNF on the partial charges of the sugar residue was considered to be non-negligible for C1’ and the directly bonded atoms due to the strong structural and electronic differences between native nucleobases and HNF. Thus, in contrast to 2AP, for the atoms C1’, C4’, H4’, H1”, and O4’ of HNF the partial charges derived from the DFT-calculations were used. The charge at O4’ was increased by +0.13 in order to retain neutrality of the residue (cf. section 3.1.3).

Tabelle 3.4: Experimentally determined RDCs for 13merHNF. The RDCs were measured with a precision of ±0.6 Hz.

<table>
<thead>
<tr>
<th>Res</th>
<th>Vector</th>
<th>( J_{(CH)} ) (Hz)</th>
<th>( J_{(CH)} ) (aligned) (Hz)</th>
<th>RDC (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A6</td>
<td>C2-H2</td>
<td>202.2</td>
<td>228.0</td>
<td>25.8</td>
</tr>
<tr>
<td>A8</td>
<td>C2-H2</td>
<td>201.0</td>
<td>223.8</td>
<td>22.8</td>
</tr>
<tr>
<td>A16</td>
<td>C2-H2</td>
<td>201.6</td>
<td>220.8</td>
<td>19.2</td>
</tr>
<tr>
<td>A24</td>
<td>C2-H2</td>
<td>201.0</td>
<td>229.8</td>
<td>28.8</td>
</tr>
<tr>
<td>T3</td>
<td>C7-H7</td>
<td>124.8</td>
<td>120.0</td>
<td>-1.6</td>
</tr>
<tr>
<td>T11</td>
<td>C7-H7</td>
<td>125.4</td>
<td>117.0</td>
<td>-2.8</td>
</tr>
<tr>
<td>HNF7</td>
<td>C1’-H1”</td>
<td>169.2</td>
<td>182.4</td>
<td>13.2</td>
</tr>
<tr>
<td>C9</td>
<td>C1’-H1’</td>
<td>163.2</td>
<td>172.2</td>
<td>9.0</td>
</tr>
<tr>
<td>C17</td>
<td>C1’-H1’</td>
<td>161.4</td>
<td>166.2</td>
<td>4.8</td>
</tr>
<tr>
<td>T3</td>
<td>C6-H6</td>
<td>176.4</td>
<td>198.0</td>
<td>21.6</td>
</tr>
<tr>
<td>A6</td>
<td>C8-H8</td>
<td>213.6</td>
<td>244.2</td>
<td>30.6</td>
</tr>
<tr>
<td>A8</td>
<td>C8-H8</td>
<td>215.4</td>
<td>238.8</td>
<td>23.4</td>
</tr>
<tr>
<td>A16</td>
<td>C8-H8</td>
<td>213.6</td>
<td>234.0</td>
<td>20.4</td>
</tr>
<tr>
<td>A24</td>
<td>C8-H8</td>
<td>214.8</td>
<td>240.6</td>
<td>25.8</td>
</tr>
<tr>
<td>HNF7</td>
<td>C6-H6</td>
<td>157.8</td>
<td>189.0</td>
<td>31.2</td>
</tr>
<tr>
<td>HNF7</td>
<td>C8-H8</td>
<td>162.0</td>
<td>191.4</td>
<td>20.4</td>
</tr>
<tr>
<td>ABA20</td>
<td>C1’-H1’</td>
<td>147.0</td>
<td>147.0</td>
<td>0.0</td>
</tr>
<tr>
<td>ABA20</td>
<td>C1’-H1”</td>
<td>145.8</td>
<td>148.2</td>
<td>2.4</td>
</tr>
</tbody>
</table>
3 Experimental details

3.2.4 Structure calculation

The Simulated Annealing protocol and Structure validation were carried out as described in section 3.1.4.

Calculation input Structural calculations were performed with XPLOR-NIH v2.20 [Schwitters et al., 2003]. As the HNF was found to exist in two different orientation with in the duplex, two calculation runs were performed with the HNF methylene group pointing towards the major or minor groove for face-up and face-down orientations, respectively. A total of 401 (403) NOE distance restraints and 19 Residual Dipolar Couplings were used for the calculation of the face-up (face-down) orientation. The experimental data were supplemented with 124 backbone dihedral restraints, 72 hydrogen bond distance restraints and 24 planarity restraints (see Table 3.5).

Initial MD-calculations were performed with dihedral restraints, allowing both A-form and B-form conformations (with error bars of ±50°). B-form conformation was experimentally confirmed by NOESY-cross-peaks intensities characteristic of B-DNA. Consequently, regular dihedral values from the literature [Roberts, 1993] were included in the calculations for all but the HNF residue.

| Table 3.5: Overview of structural statistics for 13merHNF calculations |
|-----------------------------------------------|-----|-----|
| NOE restraints                              | face-up | face-down |
| - total                                     | 401  | 403  |
| - interresidue                              | 225  | 226  |
| - intraresidue                              | 176  | 177  |
| - 2AP to DNA                                | 33   | 31   |
| RDC restraints                              | 19   | 19   |
| Dihedral angle restraints                    | 124  | 124  |
| H-bond restraints                           | 72   | 72   |
| Base pair planarity restr.                  | 24   | 24   |
| NOE viol. (> 0.5 Å)                         | 0    | 0    |
| RDC viol. (> 0.4 Hz)                        | 0    | 0    |
| Dihe viol. (> 5°)                           | 0    | 0    |
| RMSD to ave. struct. in Å                    | 0.63 | 0.46 |
4 Results and discussion

4.1 2-Aminopurine

4.1.1 Introduction

2AP is known since the 1950s for its role in the mutagenic transition of A:T to G:C base pairs [Freese, 1959, Rogan and Bessman, 1970, Ronen, 1979]. As a structural isomer of A it can form stable WC-type base pairs with T (see Fig. 4.1) [Ronen, 1979, Sowers et al., 1986]. Base pairs with C, A or G are much weaker [Sowers et al., 1986, Fazakerley et al., 1987, Hochstrasser et al., 1994, Law et al., 1996]. Thus 2AP is incorporated instead of A by DNA polymerase opposite to T [Freese, 1959], though at a lower frequency than A [Rogan and Bessman, 1970, Watanabe and Goodman, 1982].

The mutagenic transition of A:T to G:C occurs during replication of 2AP:T base pairs [Bernstein et al., 1976, Goodman et al., 1977]. By adjusting the type of polymerase [Grossberger and Clough, 1981] or the concentration of nuclease [Clayton et al., 1979] the mutagenic rate can be tuned. The influence of nucleotide sequence on the mutagenic rate has no discernable logical pattern [Pless and Bessman, 1983]. Another way to stimulate the mutagenic transition is to increase the dCTP- or decrease the TTP-level
in the cell [Caras et al., 1982]. In 1981 Watanabe et al. [Watanabe and Goodman, 1981] demonstrated that 2AP:C mispairs form at 320-fold higher frequency than A-C mispairs. They concluded that 2AP:C mispairing is central to the mutagenicity of 2AP.

The stabilization of the 2AP:C base pair in solution therefore had to be investigated. Four models have been proposed for the geometry of this mispair. Ronen [1979] suggested a WC geometry with one hydrogen bond, while Goodman and Ratliff [1983] presented evidence for two hydrogen bonds involving one of the bases in the rare tautomeric imino form. Sowers et al. put forth the other two pairing schemes. Based on NMR imino proton data and pH-titration experiments, they proposed WC geometry through protonation of the mispair at neutral pH [Sowers et al., 1986]. A $^{15}$N-enriched NMR-study later suggested a wobble geometry at high pH [Sowers et al., 1989], corroborated by fluorescence anisotropy measurements which showed pH-dependent conformational changes in the geometry of the 2AP:C mispair [Guest et al., 1991]. $^{15}$N-enriched NMR-studies by Fagan et al. [1996] and Sowers et al. [2000] reproduced this pH-dependence and could finally demonstrate that the wobble geometry is predominant at neutral pH while protonation occurs sequence-dependently at lower pH. This conclusion is supported by a theoretical study [Sherer and Cramer, 2001].

The fluorescence properties of 2AP differ markedly from those of the natural nucleobases [Longworth et al., 1966, Callis, 1979, Serrano-Andres et al., 2006]. This was recognized originally by Ward et al. [1969] who observed a decrease in quantum yield of fluorescence by two orders of magnitude upon stacking of 2AP in a DNA helix. The effect was first exploited to determine thermodynamic parameters of stacking associations by Bierzynski et al. [Bierzynski et al., 1977b,a, Gajewska et al., 1982].

Fluorescence quenching of 2AP can also be used to study structural transitions in biological systems [Ward et al., 1969]. At the beginning such applications were thwarted by the lack of site-specific incorporation techniques. After chemical synthesis of 2AP-containing DNA-duplexes was realized [Eritja et al., 1986, McLaughlin et al., 1988], fluorescence quenching of 2AP was employed to study structural transitions in DNA [Lycksell et al., 1987, Patel et al., 1992]. In 1993 Bloom et al. [Bloom et al., 1993, 1994] were the
first to use 2AP fluorescence quenching to follow the kinetics of polymerase-catalyzed 2AP-insertion on a millisecond timescale. But on the whole, fluorescence quenching of a 2AP nucleobase surrogate is used to detect base stacking-unstacking transitions.

The local dynamics in the vicinity of 2AP, and its variation with sequence, can be studied by monitoring the fluorescence anisotropy decay [Nordlund et al., 1989]. This technique was used by Guest et al. [1991] to characterize the dynamic behaviour of mismatched base pairs. Hochstrasser et al. [1994] combined both techniques - fluorescence quenching and anisotropy decay - to study DNA double strand melting upon binding to the Klenow fragment. Xu et al. [1994] utilized 2AP fluorescence to show that DNA duplex melting proceeds via a highly flexible, yet B-DNA type transition state.


A new possibility to study structural transitions in biomolecules has recently been devised which makes use of unique properties of 2AP. Johnson et al. [2004] demonstrated
that a low-energy Circular Dichroism band is observed for 2AP dinucleotides incorporated into DNA or RNA double strands. They used this method to monitor structural changes in RNA hairpin loops [Johnson et al., 2005b] or breathing fluctuations at replication forks [Johnson et al., 2005a]. This new technique and the high number of 2AP-related publications show the significance of 2AP for the study of biological macromolecules.

Structure perturbation upon incorporation of 2AP into a DNA-helix was investigated in detail only once, by Nordlund et al. [1989]. A palindromic sequence containing two 2AP:T base pairs was investigated by NMR spectroscopy, fluorescence decay and Molecular Dynamics simulations. Base pairing is observed in the duplex, but the melting temperature of the duplex was considerably lower than in the reference with WC A:T base pairs. The missing of NOE-crosspeaks to the 5’-side of the 2AP-substitution was interpreted as a local disturbance of the overall B-type helix [Nordlund et al., 1989]. Studies on minor groove binding drugs support this conclusion. They indicate that 2AP incorporation into A-T tracts reduces the binding affinities of these drugs to the level of G-C tracts [Loontiens et al., 1991, Patel et al., 1992]. Minor groove distortions by 2AP have also been proposed by a theoretical study [Lankas et al., 2002].

Dynamical effects due to incorporation of 2AP were examined even more scarcely. Nordlund et al. [1989] reported efficient recognition and cleavage of 2AP-containing duplexes by EcoRI, and concluded that dynamics are not perturbed. On the contrary, Lycksell et al. [1987] observed a 1 ms lifetime for the 2AP:T base pair, lower than for the corresponding A:T base pair, but state that the accuracy is correct probably only within a factor of 2. A destabilizing effect on the neighbouring pairs was not observed.

When structural transitions in biological systems are studied with a molecular probe, the assumption is that the modified system behaves like the natural one. Consequently the introduction of the probe must leave the structure unchanged. Disturbance of the original helix when 2AP occupies an A site would limit or even prevent the use of 2AP-fluorescence in some studies of biological systems. Therefore the above-mentioned hypothesis of a local disturbance has to be tested.

For this purpose a detailed analysis of the solution structure of a DNA-duplex in which
an A is substituted by 2AP is presented. The only change introduced into the helix is the difference in the position of the amino group of A and 2AP (cf. Fig. 4.1). In contrast to previous studies [Lycksell et al., 1987, Nordlund et al., 1989] the nonpalindromic nature of the sequence allows to have a single perturbation site. Thus any change in structure or base pair dynamics can be directly attributed to the 2AP incorporation. The central base pair was modified in order to detect possible long-range effects (up to three base pairs) of 2AP incorporation, which otherwise might be rendered ambiguous due to base pair fraying effects [Nonin et al., 1995]. A symmetric 13 base pair sequence was chosen to dismiss the possibility of mispairing, loop formation and fraying effects.
4 Results and discussion

4.1 Results

4.1.2 Chemical shift analysis

The NOESY-spectra were assigned with the sequential approach [Roberts, 1993, Bloomfield et al., 2000]. The sequential assignments of the sugar H1’ and base H6/H8 protons for 13merRef and 13mer2AP are shown in Fig. 4.2a and 4.2b, respectively. Cross peaks expected for a regular B-DNA helix are present in the NOESY-spectra of both samples. Based on the assignment of the H1’ and H6/H8 protons, the remaining base and sugar protons could be assigned by combining the information from the COSY-, TOCSY-, HMQC-, and NOESY-spectra. Assignment of the exchangeable protons was done in the WATERGATE-NOESY-spectrum. The imino proton of T20 was easily identified, because of the symmetry of the sequence, but the remaining imino, amino, and H2 protons had to be referenced to the already assigned non-exchangeable C H5 protons via the H42/H41-H5 NOE cross-peaks.

$^1$H Chemical Shift Differences (CSDs) between 13merRef and 13mer2AP are negligible (<0.10 ppm) for all but the H1’, imino, and H2 protons of the central three base pairs (see Table 4.1 and Table .13 in the Appendix, section .10). There is no significant trend that either 3’- or 5’-neighbours exhibit larger CSDs. No CSDs for the H6 and H8 protons
**Tabelle 4.1:** Selected $^1H$ Chemical shift differences (CSD) (reference HOD at 4.80 ppm, calculated as $X_{13merRef} - X_{13mer2AP}$) for 13merRef and 13mer2AP. Only the inner 7 base pairs are shown. In the case of X7 H2 the difference between the chemical shifts of H2 (A) and H6 (2AP) is taken.

<table>
<thead>
<tr>
<th>Residue</th>
<th>H1'</th>
<th>H1/H3/H41/H42</th>
<th>H2/H5/H7</th>
<th>H6/H8</th>
</tr>
</thead>
<tbody>
<tr>
<td>G4</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-</td>
<td>-0.03</td>
</tr>
<tr>
<td>C5</td>
<td>-0.08</td>
<td>0.03/-0.01</td>
<td>-0.02</td>
<td>-0.04</td>
</tr>
<tr>
<td>A6</td>
<td>-0.17</td>
<td>-</td>
<td>-0.16</td>
<td>-0.01</td>
</tr>
<tr>
<td>X7</td>
<td>0.29</td>
<td>-</td>
<td>-0.67</td>
<td>0.16</td>
</tr>
<tr>
<td>A8</td>
<td>-0.12</td>
<td>-</td>
<td>-0.13</td>
<td>-0.04</td>
</tr>
<tr>
<td>C9</td>
<td>0.00</td>
<td>-0.03/0.02</td>
<td>-0.06</td>
<td>-0.04</td>
</tr>
<tr>
<td>G10</td>
<td>-0.01</td>
<td>-0.05</td>
<td>-</td>
<td>-0.02</td>
</tr>
<tr>
<td>C17</td>
<td>-0.04</td>
<td>-0.01/-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td>G18</td>
<td>0.00</td>
<td>0.03</td>
<td>-</td>
<td>-0.03</td>
</tr>
<tr>
<td>T19</td>
<td>-0.03</td>
<td>0.21</td>
<td>-0.04</td>
<td>-0.04</td>
</tr>
<tr>
<td>T20</td>
<td>0.09</td>
<td>0.53</td>
<td>-0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>T21</td>
<td>0.11</td>
<td>0.13</td>
<td>0.00</td>
<td>-0.02</td>
</tr>
<tr>
<td>G22</td>
<td>-0.04</td>
<td>-0.05</td>
<td>-</td>
<td>0.00</td>
</tr>
<tr>
<td>C23</td>
<td>0.01</td>
<td>-0.02/-0.02</td>
<td>-0.02</td>
<td>0.01</td>
</tr>
</tbody>
</table>

for purine and pyrimidine bases with the exception of residues A7/2AP7 are observed.

In Fig. 4.3 the sum of the absolute values of the CSDs of all protons belonging to one residue is given for 13merRef to 13mer2AP (blue columns) and 13merRef to 13merRefGC (red columns). In 13merRefGC the central base pair of 13merRef is substituted by G:C. With the exception of the modification site, absolute per-residue CSDs from 13merRef to 13mer2AP are smaller than to 13merRefGC. Per-residue CSDs are significant two bases in each direction from the modification site in the 2AP-containing strand. In the unmodified counterstrand only the central three T bases exhibit significant CSDs. The chemical shifts of all assigned protons and carbons for 13merRef and 13mer2AP are given in the Appendix, sections .8 and .9, respectively.
Fig. 4.3: The sum of the absolute values of the CSDs of all protons belonging to one residue is given for: 13mer2AP to 13merRef (blue columns) and 13merRef to 13merRefGC (red columns) for comparison.
4.1.2.2 Structural comparison

From a family of 100 calculated structures the 10 minimum-energy, violation-free structures are shown in Fig. 4.4 for 13merRef (left) and 13mer2AP (right). The latter were used to calculate an average structure each. The accuracy of the calculations was checked by back-calculating the NOESY-spectrum from the average structure, followed by comparison with the experimental spectrum (Fig. 4.5a and 4.5b). Additionally, the RDCs were back-calculated from the average structure with the program PALES [Zweckstetter, 2008]. The correlation plots of experimentally determined vs predicted RDCs are shown in Fig. 4.6a and 4.6b and yielded correlation factors \( R \) of 1.000 and q-factors of 0.002 and 0.003 for 13merRef and 13mer2AP, respectively. The precision of the calculations can be assessed by the RMSD of the 10 best structures from their average, 0.30 Å for 13merRef and 0.33 Å for 13mer2AP. An overlay between the averaged structures is shown in Fig. 4.4 (middle part); their RMSD is 0.46 Å.

Helical parameters were calculated for the 10 minimum-energy, violation-free structu-
4 Results and discussion

Fig. 4.5: Overlay of the experimental NOESY-spectrum at 150 ms mixing time (green) and the NOESY-spectrum back-calculated from the average structure (red). Arrows point to NOE cross-peaks involving the modification site.

Correlation of experimental and back-calculated Residual Dipolar Couplings

Fig. 4.6: Plot of the experimental vs predicted RDCs for 13merRef and 13mer2AP.
4.1 2-Aminopurine

Fig. 4.7: 13merRef and 13mer2AP: Differences $(X(13merRef) - X(13mer2AP))$ in translational helical parameters between base pair partners [Å]. Blue columns represent “Shear”-values, green ones “Stretch” and red ones “Stagger”. Error bars indicate the estimated uncertainty.

Fig. 4.8: 13merRef and 13mer2AP: Differences $(X(13merRef) - X(13mer2AP))$ in rotational helical parameter between base pair partners [°]. Blue columns represent “Buckle”-values, green ones “Propeller Twist” and red ones “Opening”. Error bars indicate the estimated uncertainty.
4 Results and discussion

Fig. 4.9: 13merRef and 13mer2AP: Differences \((X(13merRef) - X(13mer2AP))\) in translational helical parameter between base pairs [Å]. Blue columns represent “Shift”-values, green ones “Slide” and red ones “Rise”. Error bars indicate the estimated uncertainty.

Fig. 4.10: 13merRef and 13mer2AP: Differences in rotational helical parameter between base pairs [°]. Blue columns represent “Tilt”-values, green ones “Roll” and red ones “Twist”. Error bars indicate the estimated uncertainty.
4.1 2-Aminopurine

res of 13merRef and 13mer2AP with the program 3DNA [Lu and Olson, 2003]. Fig. 4.7 and 4.8 depict differences between 13merRef and 13mer2AP in translational and rotational helical parameters between base pair partners, while Fig. 4.9 and 4.10 display differences in translational and rotational helical parameters between base pairs. The value \(X\) represented by a column in Fig. 4.7-4.10 is calculated as:

\[
X = \frac{1}{10} \left( \sum_{n=1}^{10} X_n(13\text{merRef}) - \sum_{n=1}^{10} X_n(13\text{mer2AP}) \right)
\] (4.1)

where \(X_n(13\text{merRef})\) and \(X_n(13\text{mer2AP})\) represent the helical parameter values calculated for the 10 minimum-energy, violation-free structures of 13merRef and 13mer2AP, respectively. The RMSD of these values to their average is shown as an error bar. Thus a measure for the precision of a given helical parameter is obtained. In general, for translational parameters only values which deviate by at least 0.1 Å (with the corresponding uncertainties taken into account) and for rotational parameters > 5° are interpreted. Deviations below these thresholds are too weak to be reliably described by the force field (deviations of 0.1 Å and 5° from the equilibrium value are allowed for bond lengths and angles respectively).

In the following, the observed perturbations in helical parameters are presented. For definition of the helical parameters please see section 2.1, Fig. 2.5.

- The “Stagger”-values (red columns in Fig. 4.7) differ significantly for base pairs G4:C23 (-0.23 ± 0.09 Å), A8:T19 (-0.18 ± 0.07 Å) and G10:C17 (0.24 ± 0.06 Å), while values for “Shear” and “Stretch” (blue and green columns in Fig. 4.7) do not show significant deviations.

- Of the rotational helical parameters within base pairs only the “Propeller Twist”-values (green columns in Fig. 4.8) of base pairs X7:T20 (-7.0 ± 1.6°) and A8:T19 (-8.1 ± 2.1°) exhibit significant deviations, while “Buckle”- and “Opening”-values are equal within the estimated error (blue and red columns in Fig. 4.8 respectively).

- The rotational helical parameters between base pairs “Tilt”, “Roll” and “Twist” (blue, green and red columns in Fig. 4.10 respectively) do not show significant variations within
the estimated error.

- The “Shift”-values (blue columns in Fig. 4.9) differ significantly for base pair steps C2:T3 (-0.26 ± 0.10 Å), A6:X7 (0.20 ± 0.03 Å) and A8:C9 (0.26 ± 0.03 Å). The “Slide”-values (green columns in Fig. 4.9) differ significantly for base pair steps T3:G4 (0.28 ± 0.15 Å), A6:X7 (0.22 ± 0.11 Å) and A8:C9 (-0.26 ± 0.12 Å). The “Rise”-values (red columns in Fig. 4.9) do not differ within the estimated confidence interval.

In summary, significant deviation in helical parameters are observed throughout the duplex, with the central three base pairs exhibiting the largest deviations. The average helical parameters calculated from the values for the 10 minimum-energy, violation-free structures of 13merRef and 13mer2AP are listed in the Appendix in sections .11 and .12, respectively.
4.1.2.3 Comparison of base pair dynamics

The T20 imino proton resonance behaves substantially different depending on the counter base 2AP or A. In 13mer2AP this resonance can be observed in the 1D-spectrum, though it is broadened substantially (cf. Fig. 4.12). The corresponding diagonal signal in the NOESY-spectrum however vanishes with increasing mixing time beyond 100 ms. Fig. 4.11 illustrates this unusual decay behaviour by comparing signal intensities for the diagonal and representative cross-peak signals of T20 (●) and T3 (▲) in 13mer2AP. Diagonal signals T20 H3 and T3 H3 (solid lines) are normalized at 50 ms to show their different decay behaviour. Without this normalization, the diagonal peak of T20 H3 reaches only 20% of the T3 H3 signal intensity at 50 ms. Cross-peaks T20H3–2APH6 and T3H3–A24H2 (dashed lines) are given relative to the corresponding diagonal signal at 50 ms. The diagonal T20 H3 peak almost vanishes when going from 50 to 100 ms mixing time. The intensity of the NOE cross-peak to 2AP H6 decays slowly, in contrast to the T3H3–A24H2 cross-peak where the intensity increases for the first 200 ms.

To test whether H2O exchange is responsible for the unusual decay behaviour of T20 H3 we performed water saturation transfer experiments. In the latter experiments 1D-spectra in H2O are acquired with two different methods of attenuating the water signal. With the presaturation method the water signal is saturated before the actual pulse sequence (Fig. 4.12 (B,D): irradiation time 3 s, strength 55 dB). Consequently, exchange of now saturated water protons with unsaturated imino protons leads to attenuation of the latter signal. This cosaturation can be circumvented by using the WATERGATE
4 Results and discussion

Fig. 4.12: Saturation transfer experiments in H\textsubscript{2}O. Resonances from exchanging protons are reduced in the 1D presaturation spectra (lower black lines, B,D), reflecting the relative rates by which imino protons exchange with the solvent. Water saturation is avoided in 1D WATERGATE spectra which are shown for comparison (upper gray lines, A,C).

pulse sequence (Fig. 4.12 (A,C)), which does not excite the water protons and thus prevents their detection. Here water exchange with imino protons leaves their signal unperturbed. Fig. 4.12 depicts the imino proton region with both attenuation techniques, presaturation (B,D) and WATERGATE (A,C) for 13merRef, 13mer2AP respectively. In 13mer2AP the cosaturation of the T20 imino proton is so strong that the signal is not observable in presaturation experiments. The signals of the directly adjacent bases T19 and T21 are reduced in intensity when compared to the corresponding signals in 13merRef.

Base pair dynamics can be studied by measuring the base pair lifetime $\tau_{op}$. The latter is obtained by extrapolating a plot of imino proton exchange $\tau_{ex}$ vs inverse base concentration $1/[B]$ to infinite catalyst concentration (for explanation see section 2.2.1). Fig. 4.13 shows the corresponding data and linear fits for the central 7 base pairs and the extrapolated lifetimes for 13merRef and 13mer2AP. Lifetimes in green refer to overlap-
Fig. 4.13: Basepair lifetime determination. The inversion recovery of imino proton signals is affected by TRIS base which catalyzes exchange with water. The exchange time $\tau_{ex}$ depends on the inverse base concentration $1/[B]$, with different ranges for G:C and A:T pairs. Extrapolation to infinite concentration gives the lifetimes which are collected in the right panel. Lifetimes in green could not be determined separately due to spectral overlap.

Ped resonances in the imino proton region. Their overlap is complete and signal recovery identical; therefore only averaged lifetimes can be given. The $R^2$-values for all linear fits are above 0.996. Confidence intervals for each lifetime are also determined from the fits. The lifetimes of the terminal base pairs could not be measured due to “base pair fraying” which is commonly observed at the helical termini. Here base pair lifetimes are considerably shortened, broadening the signal from terminal imino protons to the point of vanishing and weakening the signal from neighbouring ones [Leijon and Graslund, 1992, Nonin et al., 1995]. Lifetimes for the semiterminal G:C (0.5 ms) and A:T base pairs (1.1 ms) are found to be identical for 13merRef and 13mer2AP within the estimated error (the corresponding linear plots are shown separately in Fig. 4.15 for reasons of clarity). Contrary to the outer base pairs, lifetimes for the inner seven pairs differ between 13merRef and 13mer2AP. The substitution A→2AP reduces $\tau_{op}$ severely for the central A:T pairs (by factors 0.55, 0.20 and 0.70 for T21, T20, T19) and less so for the more remote G:C pairs (0.8-0.7).
4 Results and discussion

Fig. 4.14: Melting curves via 13mer2AP absorbance (black) and 2AP fluorescence yield (blue points - after heating, red points - after cooling).

Fig. 4.15: Linear fits for the imino proton exchange times of T3, T11, G15 and G25 vs the inverse base catalyst concentration.
4.1.2.4 Comparison of duplex melting

Melting curves of 13mer2AP are shown in Fig. 4.14. Black points depict the melting point as determined by following the temperature-dependent absorption at 260 nm. Blue and red points represent duplex melting as determined by monitoring temperature-dependent 2AP fluorescence yield for heating and cooling, respectively. With UV absorption the melting behavior of the entire duplex is monitored whereas the 2AP fluorescence yield is sensitive to the local environment only. Melting points of 59.7 and 59.2 °C for the duplex as a whole and the centre are measured respectively. Thus $\Delta T_1 = 0.5$ K can be determined for duplex vs centre melting. The melting point of 13merRef was measured as 63.2 °C, with $\Delta T_2 = 3.5$ K for 13merRef vs 13mer2AP melting.
4 Results and discussion

4.1.3 Discussion

Structure perturbation

Nordlund et al. [1989] studied 2AP in a palindromical decamer sequence. They reported that NOE cross-peaks expected for a regular B-DNA helix, were missing to the 5′-side of the 2AP incorporation site. From that they conclude that the structure is perturbed locally with the perturbation extending only to the adjacent base pairs [Nordlund et al., 1989]. In contrast, all NOE cross-peaks expected for a regular B-DNA helix have been identified for 13mer2AP. This indicates that 2AP incorporation induces no pronounced perturbation in 13mer2AP. A possible explanation might be the higher detection limit of the 600 MHz NMR spectrometer (instead of 300 MHz) and the 5-fold higher duplex concentration employed in this work. Another reason might be the low stability of the decamer duplex, whose melting point decreases by 8.6 K (to 32.8 °C) upon introduction of the two 2AP residues [Nordlund et al., 1989].

The per-atom CSD data analysis for 13merRef and 13mer2AP (cf. section 4.1.2.1) supports the above conclusion that no pronounced perturbation is induced upon 2AP incorporation. Significant per-atom CSDs are observed exclusively between protons spatially adjacent to 2AP in 13mer2AP or A in 13merRef. This suggests that the observed CSDs are due to the different ring current effects induced by 2AP and A [Wijmenga et al., 1997]. Thus per-atom CSD data does not hint at structural differences between the two samples, but instead points to electronic differences between 2AP and A, which are evidenced by quantum chemical calculations [Broo, 1998, Mishra et al., 2000, Jean and Krueger, 2006].

On the contrary, absolute per-residue CSDs indicates that also structural differences exist between 13merRef and 13mer2AP which propagate at least two base pairs in each direction from the modification site. Comparison with the total CSDs between 13merRef and 13merRefGC suggests that the substitution of an A:T by a G:C base pair has a stronger impact on the helical structure than substitution of A by 2AP since total CSDs for the latter are equal or less for the base pairs adjacent to the modification site (cf. Fig. 4.3). This indicates that the helical structures of 13merRef and 13mer2AP both
adopt a regular B-DNA conformation. Yet very subtle structural changes between these two structures exist, which effect the significant CSDs observed for the bases C5 and C9.

A comparison to CSD data on other single mismatch sites is instructive (Note that \(^1\)H-CSD data for A→2AP substitution, other than reported here, is not available.). Bhattacharya et al. [2002] examined the effect of single mismatches A:A, G:G, C:C on the imino proton chemical shift. CSDs for the adjacent base pairs were significantly more perturbed, than in our case, reflecting the disturbed helical structure resulting from the non-WC-geometries of the mismatch sites. Klewer et al. [2000] studied incorporation of 3-nitropyrrrole into DNA. They report CSDs to the unperturbed structure that are comparable to these of 13merRef/13mer2AP. This is surprising given the fact that 3-nitropyrrrole is in a syn conformation. Despite this perturbation, the overall helical arrangement is found to be the B-form. This is in line with the results for 13mer2AP, where a B-type helix with small but significant deviations to the structure of 13merRef is suggested by the CSD data.

The overall B-DNA helical structure of 13merRef and 13mer2AP is supported by the calculation results. Slightly smaller RMSDs among the 10 best-energy structures (0.30 and 0.33 Å for 13merRef and 13mer2AP) as compared to the RMSD between the corresponding average structures (0.46 Å) are observed. This indicates that while the overall conformation is identical, minor but significant differences exist. These differences can be visualized by analysing and comparing the helical parameters for 13merRef and 13mer2AP. All of the helical parameters are within the range typically observed for B-DNA (cf. section 2.1). However, significant deviations between 13merRef and 13mer2AP structures exist throughout the duplex (see section 4.1.2.2) for the parameters “Stagger” and “Propeller Twist” between base pair partners, and “Slide” and “Shift” between two base pairs. The largest deviations are observed for the central three base pairs. This is shown in Fig. 4.16 which displays the latter pairs enlarged and from a tilted angle in order to improve visualization of the helical parameters of interest. In conclusion, the structural differences between 13merRef and 13mer2AP are small but detectable.

The structural perturbation induced upon 2AP incorporation is very weak compared
Fig. 4.16: Comparison of the three central base pairs (average structures, center of Fig. 4.4 enlarged) for 13mer2AP (yellow) and 13merRef (red). The backbone was omitted for clarity. In the right panel the view was tilted to better visualize the differences in “Shift”- and “Slide”-values.
to other base analogues. Engman et al. investigated the structure of tC containing DNA. They find that the overall B-DNA conformation is preserved. However, the DNA is slightly bent and several NOE cross-peaks involving tC are not consistent with a B-DNA conformation. Furthermore, the second ring of tC, which is not involved in base pairing, extends into the major groove, thus preventing protein docking [Engman et al., 2004]. An even larger fluorophore, pyrene, was used by Smirnov et al. [2002] as a base pair analogue opposite to an abasic site. The overall B-DNA conformation is again preserved, but local mobility or even an alternative orientation of the 5'-adjacent base pair is introduced. This is indicated by the presence of two imino proton signals and very weak or lacking NOE cross-peaks for this base pair [Smirnov et al., 2002]. 2,4-Difluorotoluene - a steric mimic of T devoid of hydrogen bonding sites - was introduced by Guckian et al. [1998]. Although a strong thermodynamic destabilization is found ($\Delta T=11$ K), no significant deviation of B-DNA conformation is observed as indicated by uninterrupted sequential connectivities. However, a detailed analysis of helical parameters and base pair dynamics is lacking. A whole base pair devoid of hydrogen bonding was used in a related study [Guckian et al., 2000]. It was demonstrated that overall B-DNA conformation is retained. Local mobility however is increased as sensed by broadening of the imino proton resonance at position 5. None of these works use a joint (structure, thermodynamics and base pair kinetics) approach to study perturbation upon incorporation of fluorophores. Moreover, a detailed analysis of helical parameters and a comparison with a control with WC base pairs only is lacking. Thus the present work constitutes a novel approach to the topic of DNA perturbation studies.

**Basepair lifetimes**

Although the surrounding structure in 13merRef and 13mer2AP is almost identical, large dynamic differences are indicated by the unusually fast decay of the T20 imino proton diagonal signal in the NOESY-spectrum of 13mer2AP (cf. Fig. 4.11). Inspection of the diagonal signal in NOESY spectra was used by Engman et al. [2004] to qualitatively check for increased water exchange in a modified DNA duplex. The lack of any uncommon phenomena led to the conclusion that base pair dynamics are not perturbed by
incorporation of the fluorophore tC. Following that reasoning, the fast decay observed with the T20 imino proton may indicate increased chemical exchange with water. This interpretation is corroborated by the results of the water saturation transfer experiments. Here the signal of the T20 imino proton is completely cosaturated in 13mer2AP, while the signals from T3 in the same duplex and T20 in the reference duplex are only slightly reduced (Fig. 4.12). Interestingly though, intensities of the adjacent imino protons (T19,T21) are more reduced compared to 13merRef, which indicates faster water exchange for these protons, too. This in turn suggests that 2AP incorporation also effects the dynamics of the adjacent base pairs.

The latter hypothesis has been validated by the results from base pair lifetime measurements. Upon 2AP incorporation the lifetimes of the central seven base pairs seem to be reduced. But the lifetimes of the G4, G10 and G18 imino protons for 13mer2AP and T19 and T20 for 13merRef (shown in green in Fig. 4.13) could not be determined separately due to complete spectral overlap. However their lifetimes should be similar since their recovery behaviour is identical (as assessed by inspection of the 21 1D-spectra resulting from the 21 delay settings after the inversion pulse). Thus the averaged lifetimes can be interpreted for the overlapped protons. Interpretation of differences in base pair lifetimes for the G18, T19 and T21 imino protons of 13merRef and 13mer2AP is complicated as the estimated confidence intervals overlap slightly. However, differences between these lifetimes can be assumed due to the observed reduction of imino proton signal intensity in the water saturation transfer experiment. Additionally, significantly reduced lifetimes for the G4 and G10 imino protons indicate an extended effect of 2AP incorporation on the base pair dynamics of the three adjacent base pairs in each direction. An alternative explanation for the reduced base pair lifetimes of G4 and G10 might be different sample conditions for 13merRef and 13mer2AP. This can be a significant source of errors, when comparing two different samples. However, lifetimes of the base pairs C2:G25, T3:A24,T11:A16, and C12:G15 are equal within the estimated error although these base pairs are influenced by base pair fraying effects. Since the latter effect is sensitive to solution properties (pH, buffer, temperature), different sample conditions
can be ruled out as an explanation. Additionally, the excellent agreement between these lifetimes shows that the lifetime measurements are reliable even below 1 ms. In summary, one can conclude that 2AP incorporation has an extensive, distributed effect on the base pair dynamics of the three adjacent pairs in each direction.

Base pair lifetimes in 2AP-containing DNA have been examined by Lycksell et al. [1987]. Their result for the modification site, of $1 \pm 2$ ms, agrees qualitatively with the one reported here but is less precise. In contrast, these authors do not observe an extensive, distributed effect on base pair dynamics due to the incorporation of 2AP into the DNA helix. In their palindromic decamer sequence, the melting point is lowered upon incorporation of 2AP from $41.4^\circ C$ to $32.8^\circ C$ [McLaughlin et al., 1988]. This suggests that at $25^\circ C$, the temperature at which base pair lifetimes were measured, the latter are already affected by duplex melting. This is corroborated by the fact that the opening motion of some inner base pairs is too rapid to be observed. Further support stems from the fact that the central two A:T pairs in their reference sequence constitute an A-tract motif, which would suggest a base pair lifetime for the core A:T pair of over 10 ms [Leroy et al., 1988, Moe and Russu, 1990] rather than 6 ms as observed by Lycksell et al. [1987].

Its length of 13 base pairs makes the non-palindromic duplex studied in the present work thermodynamically more stable. Thus it is less susceptible to destabilization due to base pair fraying. Influence of non-native modifications on the base pair dynamics of adjacent base pairs has been reported in the literature. Moe and Russu [1992] investigated a palindromic dodecamer sequence, with a G:T mismatch introduced at position 4. An increased base pair opening rate is demonstrated for position 5. A similar effect at position 3 could not be resolved due to the high error for the obtained base pair lifetime (60%) in the reference and signal overlap in the mismatch sequence [Moe and Russu, 1992]. These observations support the finding that 2AP incorporation effects the base pair lifetimes of the three adjacent pairs in each direction.

The extensive, distributed effect of 2AP incorporation on base pair opening dynamics can be explained with transition state theory. For base pair opening to occur, the stabilizing enthalpy contribution from the hydrogen bonds between the partners and the
4 Results and discussion

stacking interaction with the adjacent base pairs have to be overcome when the transition state is formed. Thus a reduction in strength of either interaction would favor base pair opening and thereby lead to a reduced base pair lifetime. Similar reasoning is used to explain the severely reduced base pair lifetimes of the helical termini [Nonin et al., 1995].

Reduced stacking interactions and hydrogen bonding energies throughout the duplex are indicated by the 3.5 K reduction in overall duplex melting temperature upon 2AP incorporation. This is supported by the results of the structural analysis. The higher melting point suggests that stacking interactions and hydrogen bonding in 13merRef are stronger than for 13mer2AP. Thus any change of the position of base pairs (stacking) or base pairing partners (hydrogen bonding) relative to each other must lead to the reduction of either stabilizing energy term. Consequently, the observed deviations in “Stagger” and “Propeller Twist” (defined between base pair partners) on the one hand and “Shift” and “Slide” (defined between base pairs) on the other hand indicate reduced hydrogen bonding and stacking interactions, respectively. Locally, the strong impact of 2AP incorporation on the central three base pairs is indicated by the 0.5 K reduction in duplex melting temperature when monitored via 2AP fluorescence. Since the latter is sensitive only to the directly adjacent base pairs, the observed reduction implies considerable premelting and thus destabilization of the duplex structure around the 2AP modification site as a whole. Premelting around the incorporation site has also been observed by Law et al. [1996] in a 2AP modified undecamer, where the destabilization was found to be even stronger (1.6 K). As pointed out before, deviations in helical parameters are strongest for the central three base pairs, which corroborates the results of the melting study. The latter results are further supported by the fact, that base pair lifetimes are considerably shortened and water saturation transfer is increased for the central three pairs. Thus one can conclude that 2AP incorporation destabilizes the duplex structure, by reducing the stacking and hydrogen bond interactions. This in turn leads to a lower activation enthalpy for the transition state of the base pair opening reaction. Thereby the extensive distributed effect of 2AP incorporation on the base pair lifetimes can be
4.1 2-Aminopurine

explained.

The effect of 2AP incorporation on the base pair dynamics of remote pairs can be used to explain results from DNA-protein interaction studies. The latter indicate that although protein activity is reduced, binding affinities are enhanced upon incorporation of 2AP into DNA [Brennan et al., 1986, Petrauskene et al., 1995, Malygin et al., 1999, Reddy and Rao, 2000]. Methyltransferases bind to DNA by flipping the target base out of the helix [Allan and Reich, 1996, Allan et al., 1998, Malygin et al., 1999, Reddy and Rao, 2000]. Thus higher base pair opening rates would enhance this binding process. The latter assumption is supported by the observation of binding enhancement upon introduction of mismatches [Moe and Russu, 1992]. Reddy and Rao [2000] investigated EcoP15I DNA Methyltransferase binding to DNA. They found that binding was enhanced for 2AP-containing DNA as compared to native DNA, regardless, whether 2AP replaces A in the recognition sequence or outside. Similar results were reported by Malygin et al. [1999] for the T4 DAM Methyltransferase. Enhancement of the binding process when 2AP is substituted for an A just outside the recognition center can only be explained by the influence of 2AP substitution on the base pair opening dynamics of several pairs in each direction as was observed for 13mer2AP. Similar observations were made with 2AP-containing DNA binding to restriction endonucleases [Petrauskene et al., 1995, Brennan et al., 1986], which can be explained analogously.

However, the reduced base pair lifetime of 2AP:T cannot account for the increased water exchange in the absence of added catalyst. Despite the fact that in 13mer2AP the base pair lifetimes of A16:T11 (1.1 ms) and A24:T3 (1.0 ms) are significantly smaller than for 2AP:T (1.6 ms), the corresponding imino resonances are still detectable in the presaturation experiment, while that of 2AP:T is missing (see Fig. 4.12). This observation suggests that the short base pair lifetime of 2AP:T accounts only in part for the increased water exchange. Since \( \tau_{op} \) is measured in the limit of high catalyst concentration, but the presaturation experiments were carried out in the absence of catalyst, intrinsic catalysis seems to be involved. One possible catalytic site could be the N1 atom in 2AP, which may have higher basicity compared to A. Electronic differences are indeed
indicated by the CSD data, but it is unlikely that the pK is raised sufficiently high to account for the results. An alternative mechanism involves the T20 O4 atom, which lacks a hydrogen bond to its partner base and can be easily accessed by solvent or base catalyst via the major groove. Thus, by forming a strong hydrogen bond with a potential, effective catalyst, it brings the latter within reach of the imino proton. In this way the catalytic ensemble is preorganised, increasing the probability of the catalytic exchange. Additionally, lack of the sterically demanding amino group in the major groove suggests that for the imino proton to be accessed by the catalyst, full opening of the 2AP:T base pair is not required. Thus exchange catalysis is more effective for 2AP:T as compared to A:T base pairs. Support for this explanation comes from a comparison of T1-values in the absence of added catalyst. The latter is shorter by roughly a factor three or more for T20 in 13mer2AP (29.5 ms) when compared to the other T imino protons in the duplex (T3: 199.7 ms, T11: 83.6 ms, T19: 98.8 ms, T21: 141.0 ms).

Another possible explanation for our observations would be exchange from the closed state. Arguments against this possibility are given by Nonin et al. [1995] who find that exchange from the closed state is negligible even for terminal base pairs, since the imino protons are not accessible sideways. The increased exchange rates of terminal imino protons are explained by reduced stacking interactions [Nonin et al., 1995]. Leroy et al. [1993] find that in C:C+ mispairs the amino protons of C have different exchange times, depending on whether or not they are hydrogen bonded. This supports the conclusion that exchange from the closed state is strongly inhibited when the imino protons are not accessible sideways. The latter is true however for G:T mismatches, where possible exchange from the closed state was discussed by Moe and Russu [1992]. Since sideways accessibility is not the case for 2AP:T base pairs, exchange from the closed state can be ruled out.

The conclusion that the hydrogen bond of the T O4 atom with bulk water or base catalyst increases the efficiency of imino proton exchange is supported by results of Strekowski et al. [1987]. They studied the interaction of DNA with different intercalators and minor groove binders. Regardless of G:C or A:T binding specificity, all molecules contai-
ning an amino side chain substituent selectively catalyzed A:T imino proton exchange with bulk water. Analogously to the 2AP:T case, this A:T specificity of can be explained by the formation of a hydrogen bond between the amino side chain of the drug and the T non-hydrogen-bonded O2 atom located in the minor groove. The possibility for this interaction is lacking in G:C base pairs since the formation of three hydrogen bonds leaves no hydrogen-bonding partner available. Thus the formation of this hydrogen bond preorganizes the catalytic ensemble, analogously to what was described for 2AP:T.
4 Results and discussion

4.2 2-Hydroxy-7-nitrofluorene

4.2.1 Introduction

HNF is a fluorene derivative which has been synthesized and incorporated into a DNA double strand only recently [Dallmann et al., 2009]. Consequently, no reports on HNF or HNF-DNA constructs can be found in the literature. For 2AP subtle changes in structure and dynamics had to be analyzed in order to answer the question of applicability to biological systems. In contrast, HNF has been designed for the proof-of-principle that macromolecular vibrational modes can be measured via transient absorption spectroscopy with polarity probes [Dallmann et al., 2009]. Thus the focus of this part of the work is on the position of the fluorophore inside the DNA helix. A comparison of helical parameters with a native structure (as in case of 2AP) is not instructive since HNF is not as close to a natural base pair as 2AP:T.

Compared to 2AP, HNF is much larger in size and thus functions as a base pair surrogate. Due to synthetic reasons the HNF could only be obtained in sufficient purity in the α-glycosidic form. Preliminary molecular modelling of the HNF-containing duplex with the program HyperChem v7.5 indicated that the hydroxyl linkage introduces orientational flexibility of the HNF moiety. Thus with both, α- and β-glycosidic form of the HNF nucleotide, stacking of the HNF with its adjacent base pairs can be realized. The same test calculations suggested that the introduction of an abasic site imposes less steric strain on HNF and prevents the latter (or its partner base) from being flipped out of the helical stack. Thus the same sequence as for 13mer2AP was studied, with the central base pair substituted by X=HNF in the left.
hand strand and Y=ABA (abasic site) in the other one (cf. Fig. 4.17). A symmetric, nonpalindromic 13 base pair sequence was chosen to dismiss the possibility of mispairing, loop formation and fraying effects. As in 13mer2AP, the central base pair was chosen for modification, in order to avoid fraying effects. The structures and nomenclature of HNF and the abasic site are also depicted in Fig. 4.17. For clarity only hydrogens which are important for the definition of the structure are shown. The nomenclature of the sugar protons follows the one described in section 2.1, with the hydrogen in the β-position of C1’ symbolized by H1”. Grey numbers indicate the numbering scheme of the carbon atoms and the corresponding hydrogens of the fluorene body (Fig. 4.17).
4 Results and discussion

Fig. 4.18: Sequential assignment in the sugar $H1'$ base $H6/H8$ region of the NOESY-spectrum in $D_2O$ for 13merHNF. Only the intraresidual cross-peaks $H1'$-$H6/H8$ are marked for clarity. Starting points are marked with blue, end points with red circles. Termination points of the sequential assignment within one strand due to the modification site are marked with purple circles.

4.2.2 Results

4.2.2.1 Chemical shift analysis

The assignment of the DNA signals was achieved following standard procedures as described in Roberts [1993] and Bloomfield et al. [2000]. In contrast to 13mer2AP however, the sequential assignment was interrupted at the modification site due to the lack of base protons at the abasic site (cf. Fig. 4.18, indicated by purple circles). Assignment at the modification site was however possible through several inter- and intrastrand NOE cross-peaks from the HNF $H5$, $H6$, $H8$ and $H1$, $H3$, $H4$ protons to residues T19, ABA20, T21 and A6, A8, respectively. As an example, the NOE cross-peaks A8-$H8$:HNF7-$H1$ and A8-$H8$:HNF7-$H3$ are marked in Fig. 4.18.

Based on the assignment of the $H1'$ and $H6/H8$ protons, the remaining base and sugar protons (including HNF and ABA) could be assigned by combining the information from
4.2 2-Hydroxy-7nitrofluorene

**Fig. 4.19:** Comparison of the imino proton region of 13merRef, 13merRefGC and 13merHNF. Red solid lines follow the chemical shift assignments of a given imino proton resonance starting from 13merHNF (a) through 13merRef (b) to 13merRefGC (c). The solid green line indicates the chemical shift of the T20 imino proton which does not exist in the 13merHNF duplex.

the COSY-, TOCSY-, HMQC-, and NOESY-spectra. Assignment of the exchangeable protons was done in the WATERGATE-NOESY-spectrum. Because of the symmetry of the sequence the imino, amino, and H2 protons had to be referenced to the already assigned non-exchangeable H5 protons via the H42/H41-H5 NOE cross-peaks. Assignment of the imino protons of T19 and T21 was complicated by the strong upfield shift of both protons when compared to either corresponding protons in 13merRef and 13merRefGC (see Fig. 4.19).

CSD analysis of 13merHNF and the corresponding native structures 13merRef (with a central A:T base pair) and 13merRefGC (with a central G:C base pair) is shown in Fig. 4.20, red and blue columns respectively. For comparison the CSDs between the to native structures 13merRef and 13merRefGC are shown (green columns). The sum of the absolute values of the CSDs for all protons belonging to one residue is given in order to avoid cancelling of negative and positive shifts. This results in a loss of information regarding the direction of the chemical shift differences, but allows for comparison between the different residues. With the exception of the three central base pairs, absolute per-residue CSDs upon introduction of HNF are comparable or only slightly larger than for the exchange of an A:T vs a G:C base pair (Fig. 4.20). Residues that are more
4 Results and discussion

![Bar chart showing total chemical shift differences of all protons belonging to one residue for 13merRef to 13merRefGC (green columns), 13merHNF to 13merRef (red columns) and 13merHNF to 13merRefGC (blue columns). 13merRefGC has the same sequence as 13merRef but with a central G:C base pair.]

than two base pairs removed from the modification site do not exhibit significant chemical shift perturbations. The chemical shifts of all assigned proton and carbon atoms for 13merHNF, 13merRefGC and 13merRef are given in the Appendix, sections .6, .7 and .8.
4.2 2-Hydroxy-7-nitrofluorene

4.2.2 NMR solution structure

The NMR solution structure is determined from experimental NOE and RDC data as described in section 3.1.4. While we see only one NOE data set for the duplex as a whole, the subset relating to the HNF chromophore cannot be described by a single orientation. Instead two structures of the same duplex are needed with different orientations of the chromophore: one where the fluorene methylene group points towards the major groove (face-up) and one where it points to the minor groove (face-down). Interestingly, the RDC restraints allow both orientations equally well. Simulated Annealing calculations for the two orientations produced two families of structures. The best (minimum-energy, violation-free) 10 of each, which are shown in Fig. 4.21, are used for generating the average structures. The latter are depicted in Fig. 4.22. The HNF probe fits into the helical fold (Fig. 4.21 and 4.22) and stacks with residues 6, 8, 21 and partly 19. This is illustrated in Fig. 4.23, where the central three base pairs of each orientation are zoomed and rotated to visualize stacking interactions.

The calculation results were validated by back-calculation of the NOESY-spectra (red and blue for face-up and face-down orientations respectively) and comparison with the experimental one (green). Fig. 4.24a and 4.24b differ only in the displayed region. While
4 Results and discussion

**Fig. 4.23:** The central three base pairs of the average structures of 13merHNF in face-up (left panel) and face-down (right panel) orientation are depicted.

**Fig. 4.24:** Overlay of the experimental NOESY-spectrum at 150 ms mixing time (green) and the NOESY-spectra back-calculated from the averaged, minimized structures of the HNF in the face-up (blue) and face-down (red) orientation. Arrows point to NOE cross-peaks involving the modification site. Panel (a) depicts the H1’-H1/H3 NOE cross-peaks, which can be only accounted for with the two different orientations of HNF within the double helix. Panel (b) shows the H1’-H6/H8 region, where predicted spectra for both orientations fit the experimental one equally well.
the latter shows the H1’-H6/H8 proton region (the same as in Fig. 4.18), Fig. 4.24a centers on two characteristic intraregional NOE cross-peaks, namely H1/H3-H1” of the HNF residue. The latter peaks can only be accounted for by the two different orientations of the HNF inside the double helix, as is demonstrated by the back-calculated spectra. As either peak would be non-observable with only one orientation sampled, the population ratio of the face-up and face-down orientation can be estimated from the integral ratio of these two peaks to be 1:1. The accuracy of the calculations has been further validated by back-calculation of the RDCs from the average structure. The correlation plots of experimentally determined vs predicted RDCs are shown in Fig. 4.25a (4.25b) and yielded correlation factors \( R \) of 1.000 and \( q \)-factors of 0.004 (0.005) for the face-up (face-down) orientations. The precision of the calculations is assessed by the RMSDs of 10 minimum-energy, violation-free structures to their average structure, which are 0.66 Å and 0.47 Å for the face-up and face-down structures, respectively. An analysis of the helical parameters of the duplex structure was not possible since the modification of the central base pair made the definition of the helical axis impossible for the commonly used programs 3DNA [Lu and Olson, 2003] and CURVES+ [Lavery et al., 2009].
4 Results and discussion

**Fig. 4.26**: Absorption changes upon hybridisation, when lowering temperature from 85 to 25°C. The well-known hyperchromism of the UV absorbance (blue arrow) is accompanied by a 1190 cm\(^{-1}\) red-shift of the HNF absorption band.

**Fig. 4.27**: Temperature-dependence of the spectra in Fig. 4.26. The amplitude of the absorption peak around 260 nm (blue points, right scale) yields a melting point of 64°C. The peak position of the HNF absorption band (red points, left scale) shows a melting point which is lower by 3°C.

### 4.2.2.3 Duplex melting

UV/vis absorption spectra of the DNA-HNF duplex are shown in Fig. 4.26 as a function of temperature. The nucleobases absorb intensely at 260 nm while the HNF chromophor is seen by a weak band around 380 nm, corresponding to the S1→S0 transition. At 85°C only single strands are present. As the temperature is lowered to 25°C, the HNF absorption band experiences a red-shift while the nucleobase absorption decreases.

Spectral change with temperature is quantified by "melting curves" shown in Fig. 4.27. The relative UV absorption amplitude decreases from 1 in the single strands to 0.74 in the duplex (right scale, blue points). A melting point \(T_m\) of 64°C is found for the total concentration \(c_T=23.5 \text{ mM}\) of the single strands. The HNF peak position is shown as red points (left scale) in Fig. 4.27. With this measure the melting point is located 3 K lower.
4.2.3 Discussion

Orientational flexibility of the HNF residue is indicated by the existence of two equally strong NOE cross-peaks between the H1" of the sugar and the H1 and H3 protons, respectively (cf. Fig. 4.24a). From their integrals a 1:1 population ratio could be estimated. Another possible explanation for the observation of a 1:1 integral ratio for these peaks might be spin diffusion effects. This possibility can be ruled out, as the integral ratio exhibited no dependence on the NOESY mixing time parameter. Calculations on both orientations and subsequent spectrum prediction support this conclusion. Interchange between the two orientations of HNF involves a 180° flip around the C1'-O2-bond which links the HNF to the sugar (cf. Fig. 4.17 and 4.22). This can take place only during transient opening of the formal HNF-abasic site base pair. For natural base pairs such opening motions are observed on a millisecond timescale. Another argument for orientational flip on that timescale is that NOE cross-peak signals would appear as an average of both orientations. This might explain why only one NOE data set is observed with some NOE cross-peaks involving the HNF chromophore being exclusively consistent with either orientation and others which can be used to describe both orientations but only with increased error bounds. The consistency of the RDC data with either orientation can be explained by the orientational degeneracy of the small experimental RDC set measured in a single orientational medium (cf. section 2.3.3). Thus one can conclude that the HNF chromophore undergoes orientational flip on a ms timescale.

Possible causes for the orientational degeneracy might be the introduction of the abasic site. The latter has already been reported in the literature to severely increase the flexibility of the helical structure in the vicinity [Coppel et al., 1997, Lin et al., 1998, Hoehn et al., 2001, Lin and de los Santos, 2001, Smirnov et al., 2002, Chen et al., 2007, 2008]. Another reason might be that the HNF chromophore is a conjugated π-system which is largely devoid of functional groups (with the exception of the hydroxyl and the nitro groups at position 2 and 7, respectively). The charge distribution obtained by the DFT calculations described in section 3.1.3 is roughly symmetrical to a 180°-flip around the long HNF axis. Thus no preference in terms of electronic interaction with
4 Results and discussion

the adjacent base pairs can occur.

In both orientations the HNF chromophore intercalates into the DNA double helix despite its α-glycosidic attachment to the sugar (instead of β for natural nucleotides). This is indicated by the existence of interstrand NOE cross-peaks from the T19, ABA20 and T21 residues to the H5, H6, and H8 protons of HNF as well as the intrastrand ones from the A6 and A8 residues to the H1, H3 and H4 protons. The comparison of the CSD data for 13merHNF, 13merRef and 13merRefGC shows that significant CSDs occur exclusively for the central five base pairs. This suggests that by HNF incorporation the helical structure of the DNA is mainly perturbed two base pairs in each direction while the overall B-DNA conformation remains intact.

Stacking interactions with residues 6, 8, 21 and partly 19 are illustrated in Fig. 4.23. While T21 is fully stacked with the HNF moiety, the T19 residue is turned outward, away from HNF. Thus stacking interactions do not stabilize the T19:A8 base pair as well as the other base pairs. As a consequence the T19 imino proton is shifted up-field and broadened, similar to what is observed with semi-terminal imino protons [Nonin et al., 1995]. This is illustrated by the NMR spectra in Fig. 4.19 where the imino protons of the HNF-substituted duplex (a) are compared to those of the same duplex containing a central A:T (b) or G:C (c) base pair. The T21 imino proton on the other hand, though also shifted upwards significantly, exhibits a very sharp resonance due to the ring current induced by HNF and the stabilizing effect of strong stacking interactions with the latter.

The thermodynamics of duplex formation for 13merHNF has been examined, to probe whether HNF intercalates and how stable the local and global helical arrangement is. When monitored via the temperature-dependent 260 nm absorbance of the nucleobases, a melting point of 64.0 °C is found. A comparison with the melting point of 13merRefGC is instructive. Under the same conditions, it is determined to be T<sub>m</sub>=69.35 °C. From a comparison of standard hybridization enthalpies and with the assumption of a common hybridization entropy of -1.2 kJ/mol [Xia et al., 1998], 13merHNF is found to be less stable by 6.9 kJ/mol compared to 13merRefGC. This is equivalent to the lack of enthalpy from hydrogen bonding. When following the red-shift of the absorption band of HNF
at 380 nm, the melting point is estimated at 61.0 °C. The difference in melting points of 3.5 K can be compared to the 0.5 K which are found for the 13mer2AP duplex (see section 4.1.2.4). The large difference found for 13merHNF indicates considerable premelting around the HNF chromophore. This is in line with the results from the structure calculations and chemical shift analysis, which show orientational exchange and less favorable stacking interactions of HNF with T19.

In the past, base pair mimics devoid of hydrogen bonding have been demonstrated to be incorporated by DNA polymerases with comparable efficiency and even higher selectivity than natural bases due to steric complementarity [Morales and Kool, 1998, Guckian et al., 1998, Matray and Kool, 1999, Guckian et al., 2000]. The pyrene nucleotide, for example, can sterically mimic a WC base pair and is incorporated into DNA duplexes opposite to an abasic site without disruption of structure or decrease in duplex stability [Matray and Kool, 1998, Singh et al., 2002, Smirnov et al., 2002]. But pyrene has a pronounced effect on the local dynamics of adjacent base pairs, indicated by the presence of two interconverting resonances for the thymine imino proton to the 5’-side and broadening of the imino proton of the adjacent G:C base pair [Smirnov et al., 2002]. The fact that for the HNF-containing DNA duplex we do not observe interconverting signals, indicates that the perturbation of local dynamics is weaker than for a pyrene residue. However, substantial broadening of the T19 imino proton and orientational flip of the HNF indicate that local flexibility is also induced.
Summary

Structural and dynamic perturbations in DNA upon incorporation of either fluorophore, 2-Aminopurine (2AP) or 2-Hydroxy-7-nitrofluorene (HNF), are characterized by NMR spectroscopy. For this purpose the NMR solution structures of the modified DNA duplexes with the sequence 5’-GCTGCAXACGTCG-3’ are solved. For X=2AP (13mer2AP) the partner base in the complementary strand is T, while for X=HNF (13merHNF) an abasic site is introduced to avoid steric strain.

As a structural isomer of A, the fluorescence properties of 2AP are commonly utilized to monitor stacking-unstacking transitions in molecular biology. By comparing results on 13mer2AP with the corresponding unmodified DNA duplex (13merRef, X=A), any perturbation can be unambiguously assigned to 2AP incorporation. For the NMR solution structure of 13merRef and 13mer2AP small but significant changes in helical parameters are found throughout the helix. Imino proton exchange measurements reveal an extended, distributed effect of 2AP incorporation on the lifetimes of the central seven base pair. This effect is explained by decreased activation enthalpy for base pair opening due to weakened stacking interactions and hydrogen bonding. The latter are indicated by the reduced melting point of 13mer2AP compared to 13merRef. Local melting around the modification site, as sensed by 2AP fluorescence, further supports the results from base pair dynamics. However, the reduced base pair lifetime of 2AP:T cannot fully account for the rapid water exchange observed with saturation transfer experiments in the absence of base catalyst. This indicates enhanced intrinsic catalysis. As a possible catalytic site the T O4 atom opposite 2AP is discussed, which is easily accessible through the major groove and lacks a hydrogen bonding partner within the base pair.
HNF is a fluorene derivative which was designed as a molecular probe for THz vibrational activity in biomolecules. Due to its elongated shape it is introduced opposite to an abasic site, which is known to induce orientational flexibility into DNA helices. The overall NMR solution structure is found to be B-DNA. However the NOE cross-peaks involving the HNF residue can only be accounted for by two different orientations of the HNF inside the DNA helical stack. Their population ratio is estimated to be 1:1. Dynamical perturbation is indicated by the increased linewidth and strong upfield shift of the T residue to the 5'-side of the abasic site. This can be explained with the help of the solution structure, which shows that this T residue cannot stack efficiently with the HNF. Lack of one stacking partner leads to destabilization of the base pair, as is observed for the helical termini.

The dynamic as well as the structural perturbation due to HNF incorporation is large compared to the perturbations induced upon 2AP incorporation. This is not surprising considering the rather artificial shape of HNF as compared to 2AP. Furthermore, 2AP incorporation does not require introduction of an abasic site, which causes orientational flexibility within the double helix. In conclusion, HNF is ill-suited for application to biological problems. Changes observed with 2AP are much smaller. Structural differences are weak but the extended, distributed effect of 2AP incorporation on base pair opening rates limits its value for biologically significant applications such as monitoring stacking-unstacking transitions.

Future development should concentrate on the design of a fluorophore with the demonstrated spectral properties of HNF [Dallmann et al., 2009] but the more native shape of 2AP. For example, the introduction of an electron withdrawing group at the 6-position of 2AP would be an interesting future goal. While a relevant structure (6-Cyano-2-Aminopurine) has already been designed [Hocek and Holý, 1995], attachment to the sugar and incorporation into a DNA sequence have not been demonstrated yet. The introduction of another functional group to 2AP at the 6-position would circumvent the problem of solvent attack and thus reduce the dynamic as well as thermodynamic destabilization.
Zusammenfassung

Mittels NMR-Spektroskopie werden Störungen in Struktur und Dynamik von DNA untersucht, die durch den Einbau jeweils eines der beiden Fluorophore 2-Aminopurin (2AP) und 2-Hydroxy-7-nitrofluoren (HNF) hervorgerufen werden. Zu diesem Zweck werden die NMR-Strukturen der modifizierten Duplexe mit der Sequenz 5’-GCTGCAXACGTCG-3’ berechnet. Im Fall X=2AP (13mer2AP) ist die Partnerbase im Komplementärstrang ein T, während gegenüber X=HNF (13merHNF) eine abasische Stelle eingeführt wird.

4 Results and discussion

im Sättigungstransfer-Experiment ohne Zugabe von Basenkatalysator erklären. Als Erklärung für diese Diskrepanz wird eine effizientere intrinsische Katalyse vermutet. Als mögliche, katalytisch aktive Stelle wird das T O4 Atom diskutiert, welches über die große Furche leicht zugänglich ist und das keine Wasserstoffbrückenbindung innerhalb des Basenpaares ausbilden kann.


4.2 2-Hydroxy-7nitrofluorene

Die zukünftige Entwicklung neuartiger Fluorophore sollte sich an den bereits demons-
strierten spektralen Eigenschaften des HNF [Dallmann et al., 2009] und der deutlich
natürlicheren Struktur des 2AP orientieren. Die Einführung einer elektronenziehenden
Gruppe in der 6-Position des 2AP stellt ein attraktives Syntheseziel dar. Während eine
relevante Struktur (6-Cyano-2-Aminopurine) eines entsprechenden Fluorophores bereits
beschrieben wurde [Hocek and Holý, 1995], konnte die glycosidische Verknüpfung und
der Einbau in DNA noch nicht realisiert werden. Durch das Einführen einer weiteren
funktionellen Gruppe in der 6-Position des 2AP könnte sowohl die dynamische als auch
die thermodynamische Destabilisierung verhindert werden.
Literaturverzeichnis


R.P. Bandwar and S.S. Patel. Peculiar 2-aminopurine fluorescence monitors the dynamics


Martin Blackledge. Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. *Progress in Nucle-


Jingyang Chen, Francois-Yves Dupradeau, David A. Case, Christopher J. Turner, and JoAnne Stubbe. Nuclear magnetic resonance structural studies and molecular mode-


Literaturverzeichnis


Agostino Fede, Martin Billeter, Werner Leupin, and Kurt Wuethrich. Determination of the NMR solution structure of the Hoechst 33258-d(gtgaattccac)2 complex and


Literaturverzeichnis


Literaturverzeichnis


Mikael Leijon. Proton exchange rates measured by saturation transfer using delayed randomization of the solvent magnetization. *Journal of Magnetic


Literaturverzeichnis


Junhe Ma, Gregory I. Goldberg, and Nico Tjandra. Weak alignment of biomacromolecules in collagen gels: An alternative way to yield residual dipolar couplings for


Literaturverzeichnis


Jakob T. Nielsen, Khalil Arar, and Michael Petersen. Solution structure of a locked nucleic acid modified quadruplex: Introducing the v4 folding topology13. *Ange-


D. J. Patel and C. W. Hilbers. Proton nuclear magnetic resonance investigations of


Markus Ruckert and Gottfried Otting. Alignment of biological macromolecules in novel nonionic liquid crystalline media for nmr experiments. *Journal of the Ame-
J. Ruthmann, S. A. Kovalenko, N. P. Ernsting, and D. Ouw. Femtosecond relaxation of
2-amino-7-nitrofluorene in acetonitrile: Observation of the oscillatory contribution to

M. Sass and D. Ziessow. Error analysis for optimized inversion recovery spin-lattice rela-
B7GXD-4CRG70F-GK/2/793c66bf441c7e295cb2135d932d5f7e.

A. Saupe and G. Englert. High-resolution nuclear magnetic resonance spectra of orien-
link.aps.org/abstract/PRL/v11/p462.

Roger E. Schirmer and Joseph H. Noggle. Quantitative application of the nuclear
overhauser effect to the determination of molecular structure. Journal of the
American Chemical Society, 94(9):2947–2952, May 1972. ISSN 0002-7863. URL
http://dx.doi.org/10.1021/ja00764a009.

Roger E. Schirmer, Joseph H. Noggle, Jeffrey Paul. Davis, and Phillip A. Hart. Determina-
tion of molecular geometry by quantitative application of the nuclear overhauser
0002-7863. URL http://dx.doi.org/10.1021/ja00714a005.

J. C. Schulhof, D. Molko, and R. Teoule. The final deprotection step in oligonucleotide
synthesis is reduced to a mild and rapid ammonia treatment by using labile base-

Jerina, and D. G. Gorenstein. Solution structure of the minor conformer of a dna


Literaturverzeichnis


Literaturverzeichnis


Nico Tjandra, Shin-ichi Tate, Akira Ono, Masatsune Kainosho, and Ad Bax. The nmr structure of a dna dodecamer in an aqueous dilute liquid crystalline phase. *Journal of


Gerhard Wagner, Anil Kumar, and Kurt Wuethrich. Systematic application of two-dimensional <sup>1</sup>h nuclear-magnetic-resonance techniques for studies of


Richard Wing, Horace Drew, Tsunehiro Takano, Chris Broka, Shoji Tanaka, Keiichi Itakura, and Richard E. Dickerson. Crystal structure analysis of a complete turn of b-
Literaturverzeichnis


Jens Wohnert, Katherine J. Franz, Mark Nitz, Barbara Imperiali, and Harald Schwalbe. Protein alignment by a coexpressed lanthanide-binding tag for the measurement of residual dipolar couplings. Journal of the American Chemical Society, 125(44):13338–13339, November 2003. ISSN 0002-7863. URL http://dx.doi.org/10.1021/ja036022d.


Tianbing Xia, Jr. John SantaLucia, Mark E. Burkard, Ryszard Kierzek, Susan J. Schroeder, Xiaoqi Jiao, Christopher Cox, and Douglas H. Turner. Thermodynamic para-


.1 Force field parameter and topology files

In this section the force field parameter and topology files, respectively, that were used throughout the calculations are presented. Changes or additions that were introduced by the author are marked as “mod by anda”.

```
! checkversion 1.0

evaluate ($kchbond = 2000)
evaluate ($kchangle = 1000)
evaluate ($kchimpr = 1000)
```

```
! **************************** mod by anda -- HNF***************************
BOND PX1 OX2 1489.209 1.485 ! Nobs = 1
BOND PX1 OX3 1489.209 1.485 ! Nobs = 1
BOND PX1 OX4 3350.720 1.593 ! Nobs = 1
BOND OX4 CX5 1709.551 1.427 ! Nobs = 1
BOND OX4 CX5 1709.551 1.427 ! Nobs = 1
BOND OX4 CX5 1709.551 1.427 ! Nobs = 1
BOND CX5 HX6 $kchbond 1.090 ! Nobs = 1
BOND CX5 HX7 $kchbond 1.090 ! Nobs = 1
BOND CX5 CX8 5235.500 1.511 ! Nobs = 1
BOND CX8 HX9 $kchbond 1.090 ! Nobs = 1
BOND CX8 OX10 2769.190 1.446 ! Nobs = 1
BOND CX8 CX12 3350.720 1.528 ! Nobs = 1
BOND OX10 CX11 1982.674 1.420 ! Nobs = 1
```
Script code

BOND CX11 CX14 1709.551 1.521 ! Nobs = 1
BOND CX11 OX18 1982.674 1.420 ! mod by anda, taken from C1′−O4′ distance
BOND CX11 HX43 $kchbond 1.090 ! Nobs = 1
BOND CX12 CX14 3350.720 1.518 ! Nobs = 1
BOND CX12 OX17 1982.674 1.431 ! Nobs = 1
BOND CX14 HX15 $kchbond 1.090 ! Nobs = 1
BOND CX14 HX16 $kchbond 1.090 ! Nobs = 1
BOND OX18 CX19 2769.190 1.446 ! mod by anda, taken from C4′−O4′ distance
BOND CX19 CX24 1500.000 1.391 ! mod by anda, extrapolated from AGCT C5−C6 values, mean
distance from calc
BOND CX21 CX24 1500.000 1.391 ! mod by anda, extrapolated from AGCT C5−C6 values, mean
distance from calc
BOND CX21 HX35 $kchbond 1.090 ! Nobs = 1
BOND CX22 CX23 1500.000 1.391 ! mod by anda, extrapolated from AGCT C5−C6 values, mean
distance from calc
BOND CX22 CX25 1500.000 1.451 ! mod by anda, extrapolated from AGCT C5−C6 values
BOND CX23 CX24 1500.000 1.391 ! mod by anda, extrapolated from AGCT C5−C6 values, mean
distance from calc
BOND CX25 CX27 1000.000 1.509 ! mod by anda, extrapolated from AGCT C5−C6 values
BOND CX25 HX40 $kchbond 1.090 ! Nobs = 1
BOND CX25 CX26 1500.000 1.391 ! mod by anda, extrapolated from AGCT C5−C6 values, mean
distance from calc
BOND CX25 CX28 1500.000 1.391 ! mod by anda, extrapolated from AGCT C5−C6 values, mean
distance from calc
BOND CX26 CX27 1500.000 1.509 ! mod by anda, extrapolated from AGCT C5−C6 values
BOND CX26 CX31 1500.000 1.391 ! mod by anda, extrapolated from AGCT C5−C6 values, mean
distance from calc
BOND CX26 HX41 $kchbond 1.090 ! Nobs = 1
BOND CX27 HX42 $kchbond 1.090 ! Nobs = 1
BOND CX28 CX29 1500.000 1.391 ! mod by anda, extrapolated from AGCT C5−C6 values, mean
distance from calc
BOND CX28 HX37 $kchbond 1.090 ! Nobs = 1
BOND CX29 CX30 1500.000 1.391 ! mod by anda, extrapolated from AGCT C5−C6 values, mean
distance from calc
BOND CX29 HX38 $kchbond 1.090 ! Nobs = 1
BOND CX30 CX31 1500.000 1.391 ! mod by anda, extrapolated from AGCT C5−C6 values, mean
distance from calc
BOND CX30 NX32 1370.370 1.471 ! mod by anda, k taken from C C4−N4, length from calc
BOND CX31 HX39 $kchbond 1.090 ! Nobs = 1
BOND NX32 OX33 1734.375 1.227 ! mod by anda, taken from T C2 ON
BOND NX32 OX34 1734.375 1.227 ! mod by anda, taken from T C2 ON

!!!!!!!!!!!!!!!!!!!! end mod by anda − HNF **************

! the generic bonds were taken from param11.dna with 3×kq
BOND C5R OH 876.000 1.4300 ! 5′ end
BOND C5D OH 876.000 1.4300 ! 5′ end
BOND C3R OH 876.000 1.4300 ! 3′ end
BOND C3D OH 876.000 1.4300 ! 3′ end
BOND O2R HO 1350.000 0.9572
.1 Force field parameter and topology files

! Phos. - combined RNA/DNA statistics used

\begin{tabular}{|c|c|c|c|}
\hline
BOND & P & O1P & 1489.209 1.485 0.015 Phos \\
\hline
BOND & P & O2P & 1489.209 1.485 0.015 P \\
\hline
BOND & P & O5R & 3350.720 1.593 0.010 P \\
\hline
BOND & P & O8R & 3350.720 1.593 0.010 P ! For 5pho patch \\
\hline
BOND & P & O3R & 2326.889 1.607 0.012 P ! mod by anda \\
\hline
BOND & P & OX17 & 2326.889 1.607 0.012 P ! mod by anda \\
\hline
\end{tabular}

! Sugars

 RNA statistics

\begin{tabular}{|c|c|c|c|}
\hline
BOND & O5R & C5R & 1709.551 1.425 0.014 Sugar \\
\hline
BOND & C5R & C4R & 1982.674 1.510 0.013 S \\
\hline
BOND & C4R & C3R & 2769.190 1.524 0.011 S \\
\hline
BOND & C3R & C2R & 2769.190 1.525 0.011 S \\
\hline
BOND & C2R & C1R & 3350.720 1.528 0.010 S \\
\hline
BOND & O4R & C1R & 2326.888 1.414 0.012 S \\
\hline
BOND & O4R & C4R & 2326.888 1.453 0.012 S \\
\hline
BOND & O3R & C3R & 1982.674 1.423 0.013 S \\
\hline
BOND & C2R & O2R & 1982.674 1.413 0.013 S \\
\hline
\end{tabular}

 DNA statistics

\begin{tabular}{|c|c|c|c|}
\hline
BOND & O5R & C5D & 1709.551 1.427 0.014 Sugar \\
\hline
BOND & C5D & C4D & 5235.500 1.511 0.008 S \\
\hline
BOND & C4D & C3D & 3350.720 1.528 0.010 S \\
\hline
BOND & C3D & C2D & 3350.720 1.518 0.010 S \\
\hline
BOND & C2D & C1D & 1709.551 1.521 0.014 S \\
\hline
BOND & O4D & C1D & 1982.674 1.420 0.013 S \\
\hline
BOND & O4D & C4D & 2769.190 1.446 0.011 S \\
\hline
BOND & O3R & C3D & 1982.674 1.431 0.013 S \\
\hline
\end{tabular}

 hydrogen/carbon

\begin{tabular}{|c|c|}
\hline
BOND & C4R & H \( kchbond \) 1.09 \\
\hline
BOND & C3R & H \( kchbond \) 1.09 \\
\hline
BOND & C2R & H \( kchbond \) 1.09 \\
\hline
BOND & C1R & H \( kchbond \) 1.09 \\
\hline
BOND & C5R & H \( kchbond \) 1.09 \\
\hline
BOND & C4D & H \( kchbond \) 1.09 \\
\hline
BOND & C3D & H \( kchbond \) 1.09 \\
\hline
BOND & C2D & H \( kchbond \) 1.09 \\
\hline
BOND & C1D & H \( kchbond \) 1.09 \\
\hline
BOND & C5D & H \( kchbond \) 1.09 \\
\hline
\end{tabular}

! Bases

 base specific bonds taken from param11 . dna , 3+kq

\begin{tabular}{|c|c|c|c|}
\hline
BOND & O2U & HO & 1350.000 0.957 UR \\
\hline
BOND & HN & NNA & 1416.000 1.010 URA \\
\hline
BOND & HN & N1T & 1416.000 1.010 Infer. \\
\hline
BOND & HN & N1C & 1416.000 1.010 \\
\hline
BOND & HN & N9G & 1416.000 1.010 \\
\hline
BOND & HN & N9A & 1416.000 1.010 \\
\hline
BOND & HN & N9P & 1416.000 1.010 \\
\hline
\end{tabular}

149
### Script code

<table>
<thead>
<tr>
<th>Bond</th>
<th>Atom1</th>
<th>Atom2</th>
<th>kq</th>
<th>x_eq</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>HN</td>
<td>N3U</td>
<td>1416.000</td>
<td>1.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HN</td>
<td>N3T</td>
<td>1416.000</td>
<td>1.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2</td>
<td>N2</td>
<td>1416.000</td>
<td>1.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2</td>
<td>N4C</td>
<td>1416.000</td>
<td>1.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2</td>
<td>N2G</td>
<td>1416.000</td>
<td>1.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2</td>
<td>N6A</td>
<td>1416.000</td>
<td>1.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HO</td>
<td>OH</td>
<td>1350.000</td>
<td>0.960</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

! Base sugar joint bonds (scale from sugar)

<table>
<thead>
<tr>
<th>Bond</th>
<th>Atom1</th>
<th>Atom2</th>
<th>kq</th>
<th>x_eq</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1R</td>
<td>N1T</td>
<td>1709.551</td>
<td>1.473</td>
<td>! 0.014 Base</td>
<td></td>
</tr>
<tr>
<td>C1R</td>
<td>N1U</td>
<td>4136.691</td>
<td>1.469</td>
<td>! 0.009 B</td>
<td></td>
</tr>
<tr>
<td>C1R</td>
<td>N1C</td>
<td>2326.889</td>
<td>1.470</td>
<td>! 0.012 B</td>
<td></td>
</tr>
<tr>
<td>C1R</td>
<td>N9G</td>
<td>4136.691</td>
<td>1.459</td>
<td>! 0.009 B</td>
<td></td>
</tr>
<tr>
<td>C1R</td>
<td>N9A</td>
<td>3350.720</td>
<td>1.462</td>
<td>! 0.010 B</td>
<td></td>
</tr>
<tr>
<td>C1R</td>
<td>N9P</td>
<td>3350.720</td>
<td>1.462</td>
<td>! 0.010 B</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bond</th>
<th>Atom1</th>
<th>Atom2</th>
<th>kq</th>
<th>x_eq</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1D</td>
<td>N1T</td>
<td>1709.551</td>
<td>1.473</td>
<td>! 0.014 B</td>
<td></td>
</tr>
<tr>
<td>C1D</td>
<td>N1U</td>
<td>4136.691</td>
<td>1.469</td>
<td>! 0.009 B</td>
<td></td>
</tr>
<tr>
<td>C1D</td>
<td>N1C</td>
<td>2326.889</td>
<td>1.470</td>
<td>! 0.012 B</td>
<td></td>
</tr>
<tr>
<td>C1D</td>
<td>N9G</td>
<td>4136.691</td>
<td>1.459</td>
<td>! 0.009 B</td>
<td></td>
</tr>
<tr>
<td>C1D</td>
<td>N9A</td>
<td>3350.720</td>
<td>1.462</td>
<td>! 0.010 B</td>
<td></td>
</tr>
<tr>
<td>C1D</td>
<td>N9P</td>
<td>3350.720</td>
<td>1.462</td>
<td>! 0.010 B</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bond</th>
<th>Atom1</th>
<th>Atom2</th>
<th>kq</th>
<th>x_eq</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2C</td>
<td>ON</td>
<td>1370.370</td>
<td>1.240</td>
<td>! 0.009 B</td>
<td></td>
</tr>
<tr>
<td>C4C</td>
<td>N4C</td>
<td>1370.370</td>
<td>1.335</td>
<td>! 0.009 B</td>
<td></td>
</tr>
<tr>
<td>N1C</td>
<td>C2C</td>
<td>1110.000</td>
<td>1.397</td>
<td>! 0.010 B</td>
<td></td>
</tr>
<tr>
<td>N1C</td>
<td>C6C</td>
<td>3083.333</td>
<td>1.367</td>
<td>! 0.006 B</td>
<td></td>
</tr>
<tr>
<td>C2C</td>
<td>NC</td>
<td>1734.375</td>
<td>1.353</td>
<td>! 0.008 B</td>
<td></td>
</tr>
<tr>
<td>NC</td>
<td>C4C</td>
<td>2265.306</td>
<td>1.335</td>
<td>! 0.007 B</td>
<td></td>
</tr>
<tr>
<td>C4C</td>
<td>C5C</td>
<td>1734.375</td>
<td>1.425</td>
<td>! 0.008 B</td>
<td></td>
</tr>
<tr>
<td>C5C</td>
<td>C6C</td>
<td>1734.375</td>
<td>1.339</td>
<td>! 0.008 B</td>
<td></td>
</tr>
<tr>
<td>C5C</td>
<td>H</td>
<td>$kchbond</td>
<td>1.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6C</td>
<td>H</td>
<td>$kchbond</td>
<td>1.09</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bond</th>
<th>Atom1</th>
<th>Atom2</th>
<th>kq</th>
<th>x_eq</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1T</td>
<td>C2T</td>
<td>1734.375</td>
<td>1.376</td>
<td>! 0.008 B</td>
<td></td>
</tr>
<tr>
<td>C2T</td>
<td>N1T</td>
<td>1734.375</td>
<td>1.373</td>
<td>! 0.008 B</td>
<td></td>
</tr>
<tr>
<td>N3T</td>
<td>C4T</td>
<td>1734.375</td>
<td>1.382</td>
<td>! 0.008 B</td>
<td></td>
</tr>
<tr>
<td>C4T</td>
<td>C5T</td>
<td>1730.370</td>
<td>1.445</td>
<td>! 0.009 B</td>
<td></td>
</tr>
<tr>
<td>C5T</td>
<td>C6T</td>
<td>2265.306</td>
<td>1.339</td>
<td>! 0.007 B</td>
<td></td>
</tr>
<tr>
<td>C6T</td>
<td>N1T</td>
<td>2265.306</td>
<td>1.378</td>
<td>! 0.007 B</td>
<td></td>
</tr>
<tr>
<td>C2T</td>
<td>ON</td>
<td>1734.375</td>
<td>1.220</td>
<td>! 0.008 B</td>
<td></td>
</tr>
<tr>
<td>C4T</td>
<td>ON</td>
<td>1370.370</td>
<td>1.228</td>
<td>! 0.009 B</td>
<td></td>
</tr>
<tr>
<td>C5T</td>
<td>C3E</td>
<td>3083.333</td>
<td>1.496</td>
<td>! 0.006 B</td>
<td></td>
</tr>
<tr>
<td>C6T</td>
<td>H</td>
<td>$kchbond</td>
<td>1.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3E</td>
<td>H</td>
<td>$kchbond</td>
<td>1.09</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bond</th>
<th>Atom1</th>
<th>Atom2</th>
<th>kq</th>
<th>x_eq</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC</td>
<td>C2A</td>
<td>1370.370</td>
<td>1.339</td>
<td>! 0.009 B</td>
<td></td>
</tr>
<tr>
<td>C2A</td>
<td>N3A</td>
<td>1370.370</td>
<td>1.331</td>
<td>! 0.009 B</td>
<td></td>
</tr>
<tr>
<td>N3A</td>
<td>C4A</td>
<td>3083.333</td>
<td>1.344</td>
<td>! 0.006 B</td>
<td></td>
</tr>
<tr>
<td>C4A</td>
<td>C5A</td>
<td>2265.306</td>
<td>1.383</td>
<td>! 0.007 B</td>
<td></td>
</tr>
<tr>
<td>C5A</td>
<td>C6A</td>
<td>1370.370</td>
<td>1.406</td>
<td>! 0.009 B</td>
<td></td>
</tr>
</tbody>
</table>
.1 Force field parameter and topology files

<table>
<thead>
<tr>
<th>Bond</th>
<th>Atom 1</th>
<th>Atom 2</th>
<th>BondLength</th>
<th>BondAngle</th>
</tr>
</thead>
<tbody>
<tr>
<td>C6A</td>
<td>N6A</td>
<td>2265.306</td>
<td>1.351</td>
<td>16.007</td>
</tr>
<tr>
<td>C5A</td>
<td>N7A</td>
<td>3083.333</td>
<td>1.388</td>
<td>16.006</td>
</tr>
<tr>
<td>N7A</td>
<td>C8A</td>
<td>2265.306</td>
<td>1.311</td>
<td>16.007</td>
</tr>
<tr>
<td>C8A</td>
<td>N9A</td>
<td>1734.375</td>
<td>1.373</td>
<td>16.008</td>
</tr>
<tr>
<td>N9A</td>
<td>C4A</td>
<td>3083.333</td>
<td>1.374</td>
<td>16.006</td>
</tr>
<tr>
<td>C6A</td>
<td>N6A</td>
<td>1734.375</td>
<td>1.351</td>
<td>16.007</td>
</tr>
<tr>
<td>C8A</td>
<td>H</td>
<td>$kchbond$</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td>C2A</td>
<td>H</td>
<td>$kchbond$</td>
<td>1.09</td>
<td></td>
</tr>
</tbody>
</table>

*purine*

<table>
<thead>
<tr>
<th>Bond</th>
<th>Atom 1</th>
<th>Atom 2</th>
<th>BondLength</th>
<th>BondAngle</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC</td>
<td>C2P</td>
<td>1370.370</td>
<td>1.333</td>
<td>16.009</td>
</tr>
<tr>
<td>C2P</td>
<td>N3P</td>
<td>1370.370</td>
<td>1.333</td>
<td>16.009</td>
</tr>
<tr>
<td>N3P</td>
<td>C4P</td>
<td>3083.333</td>
<td>1.344</td>
<td>16.006</td>
</tr>
<tr>
<td>C4P</td>
<td>C5P</td>
<td>2265.306</td>
<td>1.383</td>
<td>16.007</td>
</tr>
<tr>
<td>C5P</td>
<td>C6P</td>
<td>1370.370</td>
<td>1.406</td>
<td>16.009</td>
</tr>
<tr>
<td>C6P</td>
<td>NC</td>
<td>2265.306</td>
<td>1.351</td>
<td>16.007</td>
</tr>
<tr>
<td>C5P</td>
<td>N7P</td>
<td>3083.333</td>
<td>1.388</td>
<td>16.006</td>
</tr>
<tr>
<td>N7P</td>
<td>C8P</td>
<td>2265.306</td>
<td>1.311</td>
<td>16.007</td>
</tr>
<tr>
<td>C6P</td>
<td>N9P</td>
<td>1734.375</td>
<td>1.373</td>
<td>16.008</td>
</tr>
<tr>
<td>N9P</td>
<td>C4P</td>
<td>3083.333</td>
<td>1.374</td>
<td>16.006</td>
</tr>
<tr>
<td>C6P</td>
<td>H</td>
<td>$kchbond$</td>
<td>1.09</td>
<td>16.008</td>
</tr>
<tr>
<td>C8P</td>
<td>H</td>
<td>$kchbond$</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td>C2P</td>
<td>H</td>
<td>$kchbond$</td>
<td>1.09</td>
<td></td>
</tr>
</tbody>
</table>

*guanine*

<table>
<thead>
<tr>
<th>Bond</th>
<th>Atom 1</th>
<th>Atom 2</th>
<th>BondLength</th>
<th>BondAngle</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNA</td>
<td>C2G</td>
<td>1734.375</td>
<td>1.373</td>
<td>16.008</td>
</tr>
<tr>
<td>C2G</td>
<td>N3G</td>
<td>1734.375</td>
<td>1.323</td>
<td>16.008</td>
</tr>
<tr>
<td>N3G</td>
<td>C4G</td>
<td>2265.306</td>
<td>1.350</td>
<td>16.007</td>
</tr>
<tr>
<td>C4G</td>
<td>C5G</td>
<td>2265.306</td>
<td>1.379</td>
<td>16.007</td>
</tr>
<tr>
<td>C5G</td>
<td>C6G</td>
<td>1110.000</td>
<td>1.419</td>
<td>16.010</td>
</tr>
<tr>
<td>C6G</td>
<td>NNA</td>
<td>2265.306</td>
<td>1.391</td>
<td>16.007</td>
</tr>
<tr>
<td>C5G</td>
<td>N7G</td>
<td>3083.333</td>
<td>1.388</td>
<td>16.006</td>
</tr>
<tr>
<td>N7G</td>
<td>C8G</td>
<td>3083.333</td>
<td>1.305</td>
<td>16.006</td>
</tr>
<tr>
<td>C8G</td>
<td>N9G</td>
<td>2265.306</td>
<td>1.374</td>
<td>16.007</td>
</tr>
<tr>
<td>N9G</td>
<td>C4G</td>
<td>1734.375</td>
<td>1.375</td>
<td>16.008</td>
</tr>
<tr>
<td>C2G</td>
<td>N2G</td>
<td>1110.000</td>
<td>1.341</td>
<td>16.010</td>
</tr>
<tr>
<td>C8G</td>
<td>H</td>
<td>$kchbond$</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td>C2G</td>
<td>H</td>
<td>$kchbond$</td>
<td>1.09</td>
<td></td>
</tr>
</tbody>
</table>

*uracil*

<table>
<thead>
<tr>
<th>Bond</th>
<th>Atom 1</th>
<th>Atom 2</th>
<th>BondLength</th>
<th>BondAngle</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2U</td>
<td>ON</td>
<td>1370.370</td>
<td>1.219</td>
<td>16.009</td>
</tr>
<tr>
<td>C4U</td>
<td>ON</td>
<td>1734.375</td>
<td>1.232</td>
<td>16.008</td>
</tr>
<tr>
<td>N1U</td>
<td>C2U</td>
<td>1370.370</td>
<td>1.381</td>
<td>16.009</td>
</tr>
<tr>
<td>N1U</td>
<td>C6U</td>
<td>1370.370</td>
<td>1.375</td>
<td>16.009</td>
</tr>
<tr>
<td>C2U</td>
<td>N3U</td>
<td>2265.306</td>
<td>1.373</td>
<td>16.007</td>
</tr>
<tr>
<td>N3U</td>
<td>C4U</td>
<td>1370.370</td>
<td>1.380</td>
<td>16.009</td>
</tr>
<tr>
<td>C4U</td>
<td>C5U</td>
<td>1370.370</td>
<td>1.431</td>
<td>16.009</td>
</tr>
<tr>
<td>C5U</td>
<td>C6U</td>
<td>1370.370</td>
<td>1.337</td>
<td>16.009</td>
</tr>
<tr>
<td>C5U</td>
<td>H</td>
<td>$kchbond$</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td>C6U</td>
<td>H</td>
<td>$kchbond$</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td>C2D</td>
<td>NX29</td>
<td>2265.306</td>
<td>1.479</td>
<td>check param, added for pyr</td>
</tr>
</tbody>
</table>

ANGLES

<table>
<thead>
<tr>
<th>Angle</th>
<th>Atom 1</th>
<th>Atom 2</th>
<th>Atom 3</th>
<th>AngleLength</th>
<th>AngleAngle</th>
</tr>
</thead>
<tbody>
<tr>
<td>OX2</td>
<td>PX1</td>
<td>OX3</td>
<td>1337.074</td>
<td>119.600</td>
<td>Nobs = 1</td>
</tr>
<tr>
<td>OX2</td>
<td>PX1</td>
<td>OX4</td>
<td>357.719</td>
<td>108.100</td>
<td>Nobs = 1</td>
</tr>
</tbody>
</table>

************** mod by anda -- HNF ***************

151
Script code

\[
\text{ANGLe OX3 PX1 OX4} & \quad 412.677 \quad 108.300 \quad \text{Nobs} = 1 \\
\text{ANGLe PX1 OX4 CX3} & \quad 1175.163 \quad 120.900 \quad \text{Nobs} = 1 \\
\text{ANGLe OX4 CX5 HX6} & \quad \text{#changle} \quad 109.70 \quad \text{Nobs} = 1 \\
\text{ANGLe OX4 CX5 HX7} & \quad \text{#changle} \quad 109.70 \quad \text{Nobs} = 1 \\
\text{ANGLe OX4 CX5 CX8} & \quad 1534.906 \quad 110.200 \quad \text{Nobs} = 1 \\
\text{ANGLe HX6 CX5 HX7} & \quad \text{#changle} \quad 109.17 \quad \text{Nobs} = 1 \\
\text{ANGLe HX6 CX5 CX8} & \quad \text{#changle} \quad 109.17 \quad \text{Nobs} = 1 \\
\text{ANGLe HX7 CX5 CX8} & \quad \text{#changle} \quad 109.17 \quad \text{Nobs} = 1 \\
\text{ANGLe CX5 CX8 HX9} & \quad \text{#changle} \quad 107.78 \quad \text{Nobs} = 1 \\
\text{ANGLe CX5 CX8 OX10} & \quad 429.678 \quad 109.400 \quad \text{Nobs} = 1 \\
\text{ANGLe CX5 CX8 CX12} & \quad 488.878 \quad 114.700 \quad \text{Nobs} = 1 \\
\text{ANGLe HX9 CX5 CX8} & \quad \text{#changle} \quad 112.98 \quad \text{Nobs} = 1 \\
\text{ANGLe HX9 CX5 CX8} & \quad \text{#changle} \quad 112.98 \quad \text{Nobs} = 1 \\
\text{ANGLe CX10 CX8 CX12} & \quad 1099.976 \quad 105.600 \quad \text{Nobs} = 1 \\
\text{ANGLe CX8 CX10 CX11} & \quad 650.874 \quad 109.700 \quad \text{Nobs} = 1 \\
\text{ANGLe OX10 CX11 CX14} & \quad 909.671 \quad 106.100 \quad \text{Nobs} = 1 \\
\text{ANGLe OX10 CX11 OX18} & \quad 357.719 \quad 107.016 \quad \text{Nobs} = 1 \\
\text{ANGLe OX10 CX11 HX43} & \quad \text{#changle} \quad 112.29 \quad \text{Nobs} = 1 \\
\text{ANGLe OX18 CX11 HX43} & \quad \text{#changle} \quad 112.29 \quad \text{Nobs} = 1 \\
\text{ANGLe OX18 CX19 CX20} & \quad 621.574 \quad 110.300 \quad \text{Nobs} = 1 \\
\text{ANGLe CX19 CX20 CX24} & \quad 621.574 \quad 110.300 \quad \text{Nobs} = 1 \\
\text{ANGLe OX18 CX19 CX24} & \quad 1457.566 \quad 120.259 \quad \text{Nobs} = 1 \\
\text{ANGLe CX20 CX19 CX24} & \quad 1457.566 \quad 120.259 \quad \text{Nobs} = 1 \\
\text{ANGLe CX11 CX14 CX12} & \quad 1457.566 \quad 120.259 \quad \text{Nobs} = 1 \\
\text{ANGLe CX11 CX14 CX12} & \quad 1457.566 \quad 120.259 \quad \text{Nobs} = 1 \\
\text{ANGLe CX12 CX14 HX15} & \quad \text{#changle} \quad 112.29 \quad \text{Nobs} = 1 \\
\text{ANGLe CX11 CX14 HX15} & \quad \text{#changle} \quad 112.29 \quad \text{Nobs} = 1 \\
\text{ANGLe CX11 CX14 HX15} & \quad \text{#changle} \quad 112.29 \quad \text{Nobs} = 1 \\
\text{ANGLe CX12 CX14 HX15} & \quad \text{#changle} \quad 111.36 \quad \text{Nobs} = 1 \\
\text{ANGLe CX12 CX14 HX15} & \quad \text{#changle} \quad 111.36 \quad \text{Nobs} = 1 \\
\text{ANGLe CX12 CX14 HX15} & \quad \text{#changle} \quad 111.36 \quad \text{Nobs} = 1 \\
\text{ANGLe HX15 CX14 HX16} & \quad \text{#changle} \quad 107.52 \quad \text{Nobs} = 1 \\
\text{ANGLe OX11 OX18 CX19} & \quad 650.874 \quad 120.224 \quad \text{mod by anda, gaussian value, kforce c1-O4-C4} \\
\text{ANGLe OX18 CX19 CX20} & \quad 621.574 \quad 124.456 \quad \text{mod by anda, gaussian value, kforce C4-C5-C6} \\
\text{ANGLe OX18 CX19 CX24} & \quad 621.574 \quad 115.284 \quad \text{mod by anda, gaussian value, kforce C4-C5-C6} \\
\text{ANGLe CX20 CX19 CX24} & \quad 1457.566 \quad 120.259 \quad \text{mod by anda, mean gaussian value, kforce C,A,G C4-C5-C6} \\
\text{ANGLe CX11 CX14 CX20 CX21} & \quad 1457.566 \quad 120.215 \quad \text{mod by anda, mean gaussian value, kforce C,A,G C4-C5-C6} \\
\text{ANGLe CX19 CX20 CX21} & \quad 1457.566 \quad 120.215 \quad \text{mod by anda, mean gaussian value, kforce C,A,G C4-C5-C6} \\
\text{ANGLe CX19 CX20 HX36} & \quad \text{#changle} \quad 120.775 \quad \text{mod by anda, gaussian value} \\
\text{ANGLe CX21 CX20 HX36} & \quad \text{#changle} \quad 119.007 \quad \text{mod by anda, gaussian value} \\
\text{ANGLe CX20 CX21 HX36} & \quad \text{#changle} \quad 119.007 \quad \text{mod by anda, gaussian value} \\
\text{ANGLe CX20 CX21 HX36} & \quad \text{#changle} \quad 119.007 \quad \text{mod by anda, gaussian value} \\
\text{ANGLe CX21 CX22 CX23} & \quad 1457.566 \quad 119.780 \quad \text{mod by anda, mean gaussian value, kforce C,A,G C4-C5-C6} \\
\text{ANGLe CX21 CX22 CX23} & \quad 1457.566 \quad 119.780 \quad \text{mod by anda, mean gaussian value, kforce C,A,G C4-C5-C6} \\
\text{ANGLe CX21 CX22 CX25} & \quad 1457.566 \quad 131.505 \quad \text{mod by anda, mean gaussian value, kforce C,A,G C4-C5-C6} \\
\text{ANGLe CX22 CX23 CX25} & \quad 1457.566 \quad 108.714 \quad \text{mod by anda, mean gaussian value, kforce C,A,G C4-C5-C6} \\
\text{ANGLe CX22 CX23 CX24} & \quad 1457.566 \quad 120.959 \quad \text{mod by anda, mean gaussian value, kforce C,A,G C4-C5-C6} \\
\text{ANGLe CX22 CX23 CX27} & \quad 1457.566 \quad 109.983 \quad \text{mod by anda, mean gaussian value, kforce C,A,G C4-C5-C6}
.1 Force field parameter and topology files

\[ \text{C4–C5–C6} \]
\[ \text{ANGLE CX24 CX23 CX27} \quad 1457.566 \quad 129.058 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{C4–C5–C6} \]
\[ \text{ANGLE CX19 CX24 CX23} \quad 1457.566 \quad 119.094 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX23 CX24 CX40} \quad \text{kchange} \quad 118.576 \quad \text{mod by anda, gaussian value} \]
\[ \text{ANGLE CX22 CX25 CX26} \quad 1457.566 \quad 108.593 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX22 CX25 CX28} \quad 1457.566 \quad 131.102 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX26 CX25 CX27} \quad 1457.566 \quad 110.022 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX25 CX26 CX28} \quad 1457.566 \quad 120.807 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX25 CX26 CX31} \quad 1457.566 \quad 129.171 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX23 CX27 CX26} \quad 1457.566 \quad 108.593 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX25 CX26 CX31} \quad 1457.566 \quad 120.807 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX27 CX26 CX28} \quad 1457.566 \quad 110.022 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX22 CX25 CX28} \quad 1457.566 \quad 120.305 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX23 CX27 CX26} \quad 1457.566 \quad 129.171 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX25 CX26 CX28} \quad 1457.566 \quad 119.272 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX29 CX28 CX30} \quad 1457.566 \quad 119.411 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX28 CX29 CX30} \quad 1457.566 \quad 122.394 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{ANGLE CX29 CX30 NX32} \quad 569.362 \quad 118.868 \quad \text{mod by anda, mean gaussian value, kforce A} \]
\[ \text{C5–C6–N6} \]
\[ \text{ANGLE CX29 CX30 NX32} \quad 569.362 \quad 118.738 \quad \text{mod by anda, mean gaussian value, kforce A} \]
\[ \text{C5–C6–N6} \]
\[ \text{ANGLE CX26 CX31 CX30} \quad 1457.566 \quad 117.812 \quad \text{mod by anda, mean gaussian value, kforce C,A,G} \]
\[ \text{C4–C5–C6} \]
\[ \text{ANGLE CX26 CX31 CX30} \quad \text{kchange} \quad 122.587 \quad \text{mod by anda, gaussian value} \]
\[ \text{ANGLE CX26 CX31 CX30} \quad \text{kchange} \quad 119.602 \quad \text{mod by anda, gaussian value} \]
\[ \text{ANGLE CX30 NX32 OX33} \quad 210.000 \quad 117.818 \quad \text{mod by anda, mean gaussian value, kforce A C–N–H double} \]
\[ \text{ANGLE CX30 NX32 OX34} \quad 210.000 \quad 117.978 \quad \text{mod by anda, mean gaussian value, kforce A C–N–H double} \]
\[ \text{ANGLE OX33 NX32 OX34} \quad 1337.074 \quad 124.204 \quad \text{mod by anda, mean gaussian value, kforce O–P–O} \]

!************************************************ end of mod by anda – HNF **************

! Phos.

! the ANGLEs were taken from param11.dna with 3*kq

ANGLE HO O8R C5R 139.500 107.300
ANGLE HO O8R C5R 139.500 107.300
Script code

ANGLE HO OH C5D 139.500 107.300
ANGLE HO O5R C5D 139.500 107.300
ANGLE HO O3R P 139.500 107.300
ANGLE HO OH P 139.500 107.300 ! For 5pho patch
ANGLE HO O2R C2R 139.500 107.300
ANGLE OH P O3R 144.300 102.600 !
ANGLE OH P O5R 144.300 102.600 !
ANGLE OH P O1P 296.700 108.230 !
ANGLE OH P O2P 296.700 108.230 !
ANGLE OH C5R C4R 210.000 112.000 !
ANGLE OH C5D C4D 210.000 112.000 !
ANGLE C4D C3D OH 139.500 111.000 !
ANGLE C4R C3R OH 139.500 111.000 !
ANGLE C2D C3D OH 139.500 111.000 !
ANGLE C2R C3R OH 139.500 111.000 !
ANGLE C3R OH HO 139.500 107.300 !
ANGLE C3D OH HO 139.500 107.300 !

! Phos. - combined RNA/DNA statistics used

<table>
<thead>
<tr>
<th>kq</th>
<th>x_eq</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1P</td>
<td>P</td>
<td>O2P</td>
</tr>
<tr>
<td>O5R</td>
<td>P</td>
<td>O1P</td>
</tr>
<tr>
<td>O5R</td>
<td>P</td>
<td>O2P</td>
</tr>
<tr>
<td>O3R</td>
<td>P</td>
<td>O5R</td>
</tr>
<tr>
<td>O17</td>
<td>P</td>
<td>O5R</td>
</tr>
<tr>
<td>O3R</td>
<td>PX1</td>
<td>OX4</td>
</tr>
<tr>
<td>O2P</td>
<td>P</td>
<td>O3R</td>
</tr>
<tr>
<td>O1P</td>
<td>P</td>
<td>OIR</td>
</tr>
<tr>
<td>OX1</td>
<td>PX1</td>
<td>OIR</td>
</tr>
<tr>
<td>OX2</td>
<td>PX1</td>
<td>OIR</td>
</tr>
<tr>
<td>O2P</td>
<td>P</td>
<td>OX17</td>
</tr>
<tr>
<td>O1P</td>
<td>P</td>
<td>OX17</td>
</tr>
<tr>
<td>O5R</td>
<td>C5R</td>
<td>C4R</td>
</tr>
<tr>
<td>P</td>
<td>O5R</td>
<td>C5R</td>
</tr>
<tr>
<td>P</td>
<td>O3R</td>
<td>C3R</td>
</tr>
<tr>
<td>O5R</td>
<td>C5D</td>
<td>C4D</td>
</tr>
<tr>
<td>P</td>
<td>OIR</td>
<td>C5D</td>
</tr>
<tr>
<td>P</td>
<td>OIR</td>
<td>C3D</td>
</tr>
<tr>
<td>PX1</td>
<td>OIR</td>
<td>C3D</td>
</tr>
<tr>
<td>P</td>
<td>OX17</td>
<td>CX12</td>
</tr>
</tbody>
</table>

! Sugars

! RNA statistics

<table>
<thead>
<tr>
<th>kq</th>
<th>x_eq</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>O4R</td>
<td>C4R</td>
<td>C3R</td>
</tr>
<tr>
<td>C5R</td>
<td>C4R</td>
<td>C3R</td>
</tr>
<tr>
<td>C5R</td>
<td>C4R</td>
<td>C4R</td>
</tr>
<tr>
<td>C1R</td>
<td>O4R</td>
<td>C4R</td>
</tr>
<tr>
<td>C4R</td>
<td>C3R</td>
<td>C2R</td>
</tr>
<tr>
<td>C3R</td>
<td>C2R</td>
<td>C1R</td>
</tr>
<tr>
<td>O4R</td>
<td>C1R</td>
<td>C2R</td>
</tr>
<tr>
<td>N1T</td>
<td>C1R</td>
<td>C2R</td>
</tr>
</tbody>
</table>

154
1 Force field parameter and topology files

\[\text{ANGLE N1C C1R C2R} \quad 429.678 \quad 113.400 \quad 11.6 \; \text{S}\]
\[\text{ANGLE N1U C1R C2R} \quad 429.678 \quad 113.400 \quad 11.6 \; \text{S}\]
\[\text{ANGLE N9G C1R C2R} \quad 429.678 \quad 113.400 \quad 11.6 \; \text{S}\]
\[\text{ANGLE N9A C1R C2R} \quad 429.678 \quad 113.400 \quad 11.6 \; \text{S}\]
\[\text{ANGLE N9P C1R C2R} \quad 429.678 \quad 113.400 \quad 11.6 \; \text{S}\]
\[\text{ANGLE O4R C1R N1T} \quad 1099.976 \quad 108.200 \quad 11.0 \; \text{S}\]
\[\text{ANGLE O4R C1R N1C} \quad 1099.976 \quad 108.200 \quad 11.0 \; \text{S}\]
\[\text{ANGLE O4R C1R N1U} \quad 1099.976 \quad 108.200 \quad 11.0 \; \text{S}\]
\[\text{ANGLE O4R C1R N9A} \quad 1099.976 \quad 108.200 \quad 11.0 \; \text{S}\]
\[\text{ANGLE O4R C1R N9P} \quad 1099.976 \quad 108.200 \quad 11.0 \; \text{S}\]
\[\text{ANGLE O4R C1R N9G} \quad 1099.976 \quad 108.200 \quad 11.0 \; \text{S}\]
\[\text{ANGLE C1R C2R O2R} \quad 357.719 \quad 110.600 \quad 12.9 \; \text{S} \quad \text{scale from phos.}\]
\[\text{ANGLE C3R C2R O2R} \quad 357.719 \quad 113.300 \quad 12.9 \; \text{S} \quad \text{scale from phos.}\]
\[\text{ANGLE C4R C3R O3R} \quad 445.032 \quad 110.500 \quad 12.6 \; \text{S} \quad \text{scale from phos.}\]
\[\text{ANGLE C2R C3R O3R} \quad 383.726 \quad 111.000 \quad 12.8 \; \text{S} \quad \text{scale from phos.}\]

!DNA statistics

\[\text{ANGLE O4D C4D C3D} \quad 1099.976 \quad 105.600 \quad 11.0 \; \text{S}\]
\[\text{ANGLE C5D C4D C3D} \quad 488.878 \quad 114.700 \quad 11.5 \; \text{S}\]
\[\text{ANGLE C5D C4D O4D} \quad 429.678 \quad 109.400 \quad 11.6 \; \text{S}\]
\[\text{ANGLE C1D O4D C4D} \quad 650.874 \quad 109.700 \quad 11.3 \; \text{S}\]
\[\text{ANGLE C4D C3D C2D} \quad 1099.976 \quad 103.200 \quad 11.0 \; \text{S}\]
\[\text{ANGLE C3D C2D C1D} \quad 650.874 \quad 107.200 \quad 11.3 \; \text{S}\]
\[\text{ANGLE O4D C1D C2D} \quad 909.071 \quad 106.100 \quad 11.1 \; \text{S}\]
\[\text{ANGLE N1T C1D C2D} \quad 488.878 \quad 114.200 \quad 11.5 \; \text{S}\]
\[\text{ANGLE N1C C1D C2D} \quad 488.878 \quad 114.200 \quad 11.5 \; \text{S}\]
\[\text{ANGLE N1U C1D C2D} \quad 488.878 \quad 114.200 \quad 11.5 \; \text{S}\]
\[\text{ANGLE N9G C1D C2D} \quad 488.878 \quad 114.200 \quad 11.5 \; \text{S}\]
\[\text{ANGLE N9A C1D C2D} \quad 488.878 \quad 114.200 \quad 11.5 \; \text{S}\]
\[\text{ANGLE N9P C1D C2D} \quad 488.878 \quad 114.200 \quad 11.5 \; \text{S}\]
\[\text{ANGLE O4D C1D N1T} \quad 1357.996 \quad 107.800 \quad 0.9 \; \text{S}\]
\[\text{ANGLE O4D C1D N1C} \quad 1357.996 \quad 107.800 \quad 0.9 \; \text{S}\]
\[\text{ANGLE O4D C1D N1U} \quad 1357.996 \quad 107.800 \quad 0.9 \; \text{S}\]
\[\text{ANGLE O4D C1D N9A} \quad 1357.996 \quad 107.800 \quad 0.9 \; \text{S}\]
\[\text{ANGLE O4D C1D N9G} \quad 1357.996 \quad 107.800 \quad 0.9 \; \text{S}\]
\[\text{ANGLE O4D C1D N9P} \quad 1357.996 \quad 107.800 \quad 0.9 \; \text{S}\]
\[\text{ANGLE C4D C3D O3R} \quad 621.574 \quad 110.300 \quad 12.2 \; \text{S} \quad \text{scale from phos.}\]
\[\text{ANGLE C2D C3D O3R} \quad 412.677 \quad 110.600 \quad 12.7 \; \text{S} \quad \text{scale from phos.}\]

!Ribose terms involving non-exchageables

\[\text{ANGLE OH C5R H} \quad \$kchangle \quad 109.83\]
\[\text{ANGLE O5R C5R H} \quad \$kchangle \quad 109.83\]
\[\text{ANGLE H C5R H} \quad \$kchangle \quad 109.11\]
\[\text{ANGLE C4R C5R H} \quad \$kchangle \quad 109.11\]
\[\text{ANGLE C5R C4R H} \quad \$kchangle \quad 107.93\]
\[\text{ANGLE H C4R C3R} \quad \$kchangle \quad 107.13\]
\[\text{ANGLE H C4R O4R} \quad \$kchangle \quad 113.74\]
\[\text{ANGLE H C3R C4R} \quad \$kchangle \quad 111.35\]
\[\text{ANGLE H C3R O3R} \quad \$kchangle \quad 105.87\]
\[\text{ANGLE H C3R OH} \quad \$kchangle \quad 105.87\]
\[\text{ANGLE H C3R C2R} \quad \$kchangle \quad 112.27\]
\[\text{ANGLE H C2R C3R} \quad \$kchangle \quad 111.41\]
\[\text{ANGLE H C2R O2R} \quad \$kchangle \quad 113.07\]
\[\text{ANGLE H C2R C1R} \quad \$kchangle \quad 112.38\]
\[\text{ANGLE H C1R C2R} \quad \$kchangle \quad 111.95\]
\[\text{ANGLE H C1R N1C} \quad \$kchangle \quad 107.70\]
ANGlE  H  CIR  NIU  $kchangle  107.70
ANGlE  H  CIR  NIT  $kchangle  107.70
ANGlE  H  CIR  N9A  $kchangle  107.70
ANGlE  H  CIR  N9P  $kchangle  107.70
ANGlE  H  CIR  N9G  $kchangle  107.70
ANGlE  H  CIR  O4R  $kchangle  106.86

! Deoxyribose terms involving non-exchangeables!

ANGlE  O5H  C5D  H  $kchangle  109.70
ANGlE  O5R  C5D  H  $kchangle  109.70
ANGlE  H  C5D  H  $kchangle  109.17
ANGlE  C4D  C5D  H  $kchangle  109.17
ANGlE  C5D  C4D  H  $kchangle  107.78
ANGlE  H  C4D  C5D  $kchangle  106.91
ANGlE  H  C4D  O4D  $kchangle  112.98
ANGlE  H  C2D  C4D  $kchangle  111.16
ANGlE  H  C2D  OIR  $kchangle  109.94
ANGlE  H  C2D  O8H  $kchangle  109.94
ANGlE  H  C2D  C2D  $kchangle  111.98
ANGlE  H  C2D  C3D  $kchangle  111.36
ANGlE  H  C2D  H  $kchangle  107.52
ANGlE  H  C2D  C1D  $kchangle  107.52 !mod by anda
ANGlE  H  C2D  C1D  $kchangle  112.29
ANGlE  H  C1D  C2D  $kchangle  110.94
ANGlE  H  C1D  N1C  $kchangle  108.25
ANGlE  H  C1D  N1U  $kchangle  108.25
ANGlE  H  C1D  N1T  $kchangle  108.25
ANGlE  H  C1D  N9A  $kchangle  108.25
ANGlE  H  C1D  N9P  $kchangle  108.25
ANGlE  H  C1D  N9G  $kchangle  108.25
ANGlE  H  C1D  O4D  $kchangle  107.95

! Bases!

Cytosine

<table>
<thead>
<tr>
<th>ANGlE</th>
<th>C6C</th>
<th>N1C</th>
<th>C2C</th>
<th>keq</th>
<th>x_eq</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANGlE</td>
<td>C6C</td>
<td>N1C</td>
<td>C2C</td>
<td>2277.447</td>
<td>120.300</td>
<td>0.40 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>C5C</td>
<td>N1C</td>
<td>C2C</td>
<td>743.656</td>
<td>119.200</td>
<td>0.70 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>C2C</td>
<td>NC</td>
<td>C4C</td>
<td>1457.566</td>
<td>119.900</td>
<td>0.50 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>NC</td>
<td>C4C</td>
<td>C5C</td>
<td>2277.447</td>
<td>121.900</td>
<td>0.40 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>C4C</td>
<td>C5C</td>
<td>NC</td>
<td>1457.566</td>
<td>117.400</td>
<td>0.50 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>C5C</td>
<td>C6C</td>
<td>NC</td>
<td>1457.566</td>
<td>121.000</td>
<td>0.50 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>C6C</td>
<td>NC</td>
<td>C2C</td>
<td>1012.199</td>
<td>118.900</td>
<td>0.60 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>NC</td>
<td>C2C</td>
<td>ON</td>
<td>743.656</td>
<td>121.900</td>
<td>0.70 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>NC</td>
<td>C4C</td>
<td>N4C</td>
<td>743.656</td>
<td>118.000</td>
<td>0.70 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>C5C</td>
<td>C4C</td>
<td>N4C</td>
<td>743.656</td>
<td>120.200</td>
<td>0.70 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>C6C</td>
<td>N1C</td>
<td>CIR</td>
<td>763.873</td>
<td>120.800</td>
<td>0.120 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>C2C</td>
<td>CIR</td>
<td>C4C</td>
<td>909.071</td>
<td>118.800</td>
<td>0.10 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>C6C</td>
<td>N1C</td>
<td>CID</td>
<td>763.873</td>
<td>120.800</td>
<td>0.120 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>C2C</td>
<td>N1C</td>
<td>CID</td>
<td>909.071</td>
<td>118.800</td>
<td>0.10 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>C6C</td>
<td>N1C</td>
<td>H2</td>
<td>105.000</td>
<td>120.000</td>
<td>0.120 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>H2</td>
<td>N4C</td>
<td>H2</td>
<td>105.000</td>
<td>120.000</td>
<td>0.120 B</td>
</tr>
<tr>
<td>ANGlE</td>
<td>N1C</td>
<td>C6C</td>
<td>H</td>
<td>$kchangle  119.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANGlE</td>
<td>C5C</td>
<td>C6C</td>
<td>H</td>
<td>$kchangle  119.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANGlE</td>
<td>C6C</td>
<td>C5C</td>
<td>H</td>
<td>$kchangle  119.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANGlE</td>
<td>C6C</td>
<td>C5C</td>
<td>H</td>
<td>$kchangle  121.54</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

156
.1 Force field parameter and topology files

<table>
<thead>
<tr>
<th>Thymine</th>
<th>kq</th>
<th>x_eq</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANGLE</td>
<td>C6T N1T C2T C6T N1T C2T C4T</td>
<td>1457.566</td>
<td>121.300</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C4T C5T C6T C4T C5T C6T N1T</td>
<td>1457.566</td>
<td>121.300</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N1T C2T ON N1T C2T ON N3T C3A</td>
<td>569.362</td>
<td>123.100</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C6T N1T C1R C6T N1T C1D C2A</td>
<td>488.878</td>
<td>120.400</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C2T C6A NC C2A</td>
<td>1012.199</td>
<td>118.600</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C2A N3A C4A</td>
<td>1457.566</td>
<td>110.600</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N3A C4A C5A</td>
<td>743.656</td>
<td>126.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C4A C5A C6A</td>
<td>1457.566</td>
<td>117.000</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C5A C6A NC</td>
<td>1457.566</td>
<td>117.700</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C5A C6A N6A</td>
<td>569.362</td>
<td>123.700</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C4A C5A N7A</td>
<td>1457.566</td>
<td>110.700</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N7A C8A N9A</td>
<td>1457.566</td>
<td>113.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C8A N9A C4A</td>
<td>2277.447</td>
<td>105.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C9A C10A C5A</td>
<td>2277.447</td>
<td>105.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N3A C4A N9A</td>
<td>569.362</td>
<td>127.400</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C6A C5A N7A</td>
<td>743.656</td>
<td>132.300</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N3A C4A N9A</td>
<td>569.362</td>
<td>127.300</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C4A C5A N7A</td>
<td>339.499</td>
<td>127.700</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C6A N6A H2</td>
<td>105.000</td>
<td>120.000</td>
</tr>
<tr>
<td>ANGLE</td>
<td>H2 N6A H2</td>
<td>105.000</td>
<td>120.000</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N7A C8A N9A</td>
<td>1457.566</td>
<td>113.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C8A N9A C4A</td>
<td>2277.447</td>
<td>105.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C9A C10A C5A</td>
<td>2277.447</td>
<td>105.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N3A C4A N9A</td>
<td>569.362</td>
<td>127.400</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C6A C5A N7A</td>
<td>743.656</td>
<td>132.300</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N3A C4A N9A</td>
<td>569.362</td>
<td>127.300</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C4A C5A N7A</td>
<td>339.499</td>
<td>127.700</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C6A N6A H2</td>
<td>105.000</td>
<td>120.000</td>
</tr>
<tr>
<td>ANGLE</td>
<td>H2 N6A H2</td>
<td>105.000</td>
<td>120.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Purine</th>
<th>kq</th>
<th>x_eq</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANGLE</td>
<td>C6P NC C2P</td>
<td>1012.199</td>
<td>118.600</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N7A C8A N9A</td>
<td>1457.566</td>
<td>113.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C8A N9A C4A</td>
<td>2277.447</td>
<td>105.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C9A C10A C5A</td>
<td>2277.447</td>
<td>105.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N3A C4A N9A</td>
<td>569.362</td>
<td>127.400</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C6A C5A N7A</td>
<td>743.656</td>
<td>132.300</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N3A C4A N9A</td>
<td>569.362</td>
<td>127.300</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C4A C5A N7A</td>
<td>339.499</td>
<td>127.700</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C6A N6A H2</td>
<td>105.000</td>
<td>120.000</td>
</tr>
<tr>
<td>ANGLE</td>
<td>H2 N6A H2</td>
<td>105.000</td>
<td>120.000</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N7A C8A N9A</td>
<td>1457.566</td>
<td>113.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C8A N9A C4A</td>
<td>2277.447</td>
<td>105.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C9A C10A C5A</td>
<td>2277.447</td>
<td>105.800</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N3A C4A N9A</td>
<td>569.362</td>
<td>127.400</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C6A C5A N7A</td>
<td>743.656</td>
<td>132.300</td>
</tr>
<tr>
<td>ANGLE</td>
<td>N3A C4A N9A</td>
<td>569.362</td>
<td>127.300</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C4A C5A N7A</td>
<td>339.499</td>
<td>127.700</td>
</tr>
<tr>
<td>ANGLE</td>
<td>C6A N6A H2</td>
<td>105.000</td>
<td>120.000</td>
</tr>
<tr>
<td>ANGLE</td>
<td>H2 N6A H2</td>
<td>105.000</td>
<td>120.000</td>
</tr>
</tbody>
</table>
.1 Force field parameter and topology files

\begin{verbatim}
# Force field parameter and topology files

ANGLe  N1U  C2U  ON  743.656  122.800  10.70  B
ANGLe  N3U  C2U  ON  743.656  119.400  10.70  B
ANGLe  C5U  C4U  ON  1012.199  125.900  10.60  B
ANGLe  C6U  N1U  CIR  561.212  121.200  11.40  B
ANGLe  C6U  N1U  CID  561.212  121.200  11.40  B !DNA
ANGLe  C4U  C2U  ON  763.872  117.700  11.20  B

{ !

\end{verbatim}
DIHEdral CX20 CX19 CX24 HX40 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.73
DIHEdral CX19 CX20 CX24 HX40 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.26
DIHEdral CX19 CX20 CX21 HX35 2400.0 0 0.00 ! Nobs = 1 . . . Value = −179.84
DIHEdral HX36 CX20 CX21 HX35 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.67
DIHEdral CX20 CX21 CX22 CX23 2400.0 0 0.00 ! Nobs = 1 . . . Value = −0.02
DIHEdral CX20 CX21 CX22 CX25 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.24
DIHEdral HX35 CX21 CX22 CX23 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.92
DIHEdral HX35 CX21 CX22 CX27 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.03
DIHEdral CX21 CX22 CX23 CX24 2400.0 0 0.00 ! Nobs = 1 . . . Value = −0.02
DIHEdral CX21 CX22 CX23 CX28 2400.0 0 180.00 ! Nobs = 1 . . . Value = 179.89
DIHEdral CX25 CX22 CX23 CX24 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.98
DIHEdral CX25 CX22 CX23 CX27 2400.0 0 0.00 ! Nobs = 1 . . . Value = −0.07
DIHEdral CX25 CX22 CX25 CX26 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.90
DIHEdral CX25 CX22 CX25 CX28 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.02
DIHEdral CX23 CX22 CX25 CX26 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.06
DIHEdral CX23 CX22 CX25 CX28 2400.0 0 180.00 ! Nobs = 1 . . . Value = 179.98
DIHEdral CX22 CX23 CX24 CX19 2400.0 0 0.00 ! Nobs = 1 . . . Value = −0.18
DIHEdral CX22 CX23 CX24 HX40 2400.0 0 180.00 ! Nobs = 1 . . . Value = 179.97
DIHEdral CX27 CX23 CX24 CX19 2400.0 0 180.00 ! Nobs = 1 . . . Value = 179.93
DIHEdral CX27 CX23 CX24 HX40 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.08
DIHEdral CX22 CX23 CX27 CX26 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.05
DIHEdral CX22 CX23 CX27 HX41 2400.0 0 120.00 ! Nobs = 1 . . . Value = 117.87
DIHEdral CX22 CX23 CX27 HX42 2400.0 0 −120.00 ! Nobs = 1 . . . Value = −117.81
DIHEdral CX24 CX23 CX27 CX26 2400.0 0 180.00 ! Nobs = 1 . . . Value = 179.95
DIHEdral CX24 CX23 CX27 HX41 2400.0 0 −60.00 ! Nobs = 1 . . . Value = −62.23
DIHEdral CX24 CX23 CX27 HX42 2400.0 0 60.00 ! Nobs = 1 . . . Value = 62.09
DIHEdral CX22 CX25 CX26 CX27 2400.0 0 0.00 ! Nobs = 1 . . . Value = −0.03
DIHEdral CX22 CX25 CX26 CX31 2400.0 0 180.00 ! Nobs = 1 . . . Value = 179.97
DIHEdral CX28 CX25 CX26 CX27 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.96
DIHEdral CX28 CX25 CX26 CX31 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.04
DIHEdral CX22 CX25 CX28 CX29 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.92
DIHEdral CX22 CX25 CX28 HX37 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.06
DIHEdral CX26 CX25 CX28 CX29 2400.0 0 0.00 ! Nobs = 1 . . . Value = −0.01
DIHEdral CX26 CX25 CX28 HX37 2400.0 0 180.00 ! Nobs = 1 . . . Value = 179.97
DIHEdral CX25 CX26 CX27 CX23 2400.0 0 0.00 ! Nobs = 1 . . . Value = −0.01
DIHEdral CX25 CX26 CX27 HX41 2400.0 0 −120.00 ! Nobs = 1 . . . Value = −117.79
DIHEdral CX25 CX26 CX27 HX42 2400.0 0 120.00 ! Nobs = 1 . . . Value = 117.79
DIHEdral CX31 CX26 CX27 CX23 2400.0 0 180.00 ! Nobs = 1 . . . Value = 179.99
DIHEdral CX31 CX26 CX27 HX41 2400.0 0 60.00 ! Nobs = 1 . . . Value = 62.22
DIHEdral CX31 CX26 CX27 HX42 2400.0 0 −60.00 ! Nobs = 1 . . . Value = −62.20
DIHEdral CX25 CX26 CX31 CX30 2400.0 0 0.00 ! Nobs = 1 . . . Value = −0.06
DIHEdral CX25 CX26 CX31 HX39 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.96
DIHEdral CX27 CX26 CX31 CX30 2400.0 0 180.00 ! Nobs = 1 . . . Value = 179.94
DIHEdral CX27 CX26 CX31 HX39 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.04
DIHEdral CX25 CX28 CX29 CX30 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.01
DIHEdral CX25 CX28 CX29 HX38 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.95
DIHEdral CX37 CX28 CX29 CX30 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.97
DIHEdral CX37 CX28 CX29 HX38 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.07
DIHEdral CX28 CX29 CX30 CX31 2400.0 0 0.00 ! Nobs = 1 . . . Value = −0.03
DIHEdral CX28 CX29 CX30 HX39 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.98
DIHEdral CX38 CX29 CX30 CX31 2400.0 0 180.00 ! Nobs = 1 . . . Value = 179.93
DIHEdral CX38 CX29 CX30 HX32 2400.0 0 0.00 ! Nobs = 1 . . . Value = −0.02
DIHEdral CX29 CX30 CX31 CX26 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.06
DIHEdral CX29 CX30 CX31 HX39 2400.0 0 180.00 ! Nobs = 1 . . . Value = 179.95
DIHEdral CX32 CX30 CX31 CX26 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.99
.1 Force field parameter and topology files

DHEdral NX32 CX30 CX31 HX39 2400.0 0 0.00 ! Nobs = 1 . . . Value = −0.09
DHEdral CX29 CX30 NX32 OX33 2400.0 0 0.00 ! Nobs = 1 . . . Value = 0.03
DHEdral CX29 CX30 NX32 OX34 2400.0 0 180.00 ! Nobs = 1 . . . Value = 179.92
DHEdral CX31 CX30 NX32 OX33 2400.0 0 180.00 ! Nobs = 1 . . . Value = −179.93
DHEdral CX31 CX30 NX32 OX34 2400.0 0 0.00 ! Nobs = 1 . . . Value = −0.04

! ****************** end of mod by anda — HNF ***************

{
  ! Dihedrals from param11.dna (included for terminal residues)
  !
  $DHEdral$ X C2R C3R X 4.50 3 0.000
  $DHEdral$ X C4R C3R X 4.50 3 0.000
  $DHEdral$ X C2R C1R X 4.50 3 0.000
  $DHEdral$ X C3R O5R X 1.50 3 0.000
  $DHEdral$ X C3R O3R X 1.50 3 0.000
  $DHEdral$ X C3R O3R X 1.50 3 0.000
  $DHEdral$ X C5R OH X 1.50 3 0.000
  $DHEdral$ X C5R OH X 1.50 3 0.000
  $DHEdral$ X C2R O2R X 1.50 3 0.000
  $DHEdral$ X O5R P X 2.25 3 0.000
  $DHEdral$ X OH P X 2.25 3 0.000
  $DHEdral$ OH C5R C4R O4R 4.50 3 0.000
  $DHEdral$ OH C5R C4R C3R 4.50 3 0.000 ! gamma
  $DHEdral$ C3R O3R P OH 2.25 3 0.000 ! added by infer
  $DHEdral$ C3R O3R P OH 2.25 2 0.000 ! ATB, 7-SEP-84
  $DHEdral$ C5R O5R P OH 2.25 3 0.000 ! added by infer
  $DHEdral$ C5R O5R P OH 2.25 2 0.000 ! ATB, 7-SEP-84
  !
  $DHEdral$ X C2D C3D X 4.50 3 0.000
  $DHEdral$ X C4D C3D X 4.50 3 0.000 ! DNA
  $DHEdral$ X C2D C1D X 4.50 3 0.000
  $DHEdral$ X C3D O5R X 1.50 3 0.000
  $DHEdral$ X C3D O3R X 1.50 3 0.000
  $DHEdral$ X C3D O3R X 1.50 3 0.000
  $DHEdral$ X C3D O3R X 1.50 3 0.000
  $DHEdral$ X C5D OH X 1.50 3 0.000
  $DHEdral$ X C5D OH X 1.50 3 0.000
  $DHEdral$ X C5D O4D X 4.50 3 0.000
  $DHEdral$ X C5D O4D X 4.50 3 0.000
  $DHEdral$ C3D O3R P OH 2.25 3 0.000
  $DHEdral$ C3D O3R P OH 2.25 2 0.000
  $DHEdral$ C5D O5R P OH 2.25 3 0.000
  $DHEdral$ C5D O5R P OH 2.25 2 0.000
}

! Base hydrogen DHEdralis taken from param11.dna
DHEdral X C2G N2G X 18.0 2 180.000
DHEdral X C6A N6A X 18.0 2 180.000
DHEdral X C6A N4C X 18.00 2 180.000
DHEdral X C4C N4C X 18.00 2 180.000

!*********************** mod by anda — HNF **********************

!!! IMPRoper PX1 OX2 OX3 OX4 750.0 0 −35.000 ! Nobs = 1 . . . Value = −36.223
IMPRoper CX5 OX4 HX6 HX7 $kchimpr$ 0 32.868 ! mod by anda, gaussian value
IMPRoper CX8 CX5 HX9 OX10 94.5 0 −36.228 ! mod by anda, gaussian value, k C5D X X
Script code

C2D

IMPRoper CX11 OX10 CX14 OX18 94.5 0 -35.563 ! mod by anda, gaussian value
IMPRoper CX12 CX8 HX13 CX14 94.5 0 41.651 ! mod by anda, gaussian value
IMPRoper CX14 CX11 CX12 HX15 94.5 0 28.144 ! mod by anda, gaussian value
IMPRoper CX19 OX18 CX20 CX24 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX20 CX19 CX21 HX36 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX21 CX20 CX22 HX35 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX22 CX21 CX23 CX25 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX23 CX22 CX24 CX27 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX24 CX19 CX23 HX40 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX25 CX22 CX26 CX28 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX26 CX25 CX27 CX31 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX27 CX23 CX26 HX41 94.5 0 28.808 ! mod by anda, gaussian value
IMPRoper CX28 CX25 CX29 HX37 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX29 CX28 CX30 HX38 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX30 CX29 CX31 NX32 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX31 CX26 CX30 HX39 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper NX32 CX30 OX33 OX34 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX29 CX31 CX24 CX20 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX28 CX26 CX23 CX21 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper OX33 NX32 CX30 CX29 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper OX34 NX32 CX30 CX31 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper HX38 HX39 HX40 HX36 2400.0 0 0.000 ! mod by anda, gaussian value
IMPRoper CX31 CX25 CX23 CX20 2400.0 0 180.000 ! mod by anda, planarity
IMPRoper CX29 CX26 CX22 CX19 2400.0 0 180.000 ! mod by anda, planarity
IMPRoper CX26 CX25 CX22 CX23 2400.0 0 0.000 ! mod by anda, planarity
IMPRoper CX28 CX27 CX22 CX19 2400.0 0 180.000 ! mod by anda, planarity
IMPRoper CX31 CX25 CX27 CX21 2400.0 0 180.000 ! mod by anda, planarity
IMPRoper NX32 CX25 CX27 CX21 2400.0 0 180.000 ! mod by anda, planarity
IMPRoper CX30 CX25 CX27 CX21 2400.0 0 180.000 ! mod by anda, planarity
IMPRoper CX31 CX25 CX24 CX20 2400.0 0 180.000 ! mod by anda, planarity
IMPRoper HX38 CX26 CX21 HX40 2400.0 0 180.000 ! mod by anda, planarity
IMPRoper HX39 CX28 CX23 HX36 2400.0 0 180.000 ! mod by anda, planarity
IMPRoper CX31 CX28 CX23 CX20 2400.0 0 180.000 ! mod by anda, planarity

!**************************** end of mod by anda – HNF ***************

!*IMPRopers to keep the two purine rings parallel:
!*guanine
!IMPRoper C8G C4G C5G NNA 250.0 2 180.000
IMPRoper C8G C5G C4G C2G 250.0 2 180.000
IMPRoper N3G C4G C5G N7G 250.0 2 180.000
IMPRoper C6G C5G C4G N9G 250.0 2 180.000
!* adenine
IMPRoper C8A C4A C5A N9A 250.0 2 180.000 ! WYE AND PATCHED RESIDUES
IMPRoper C8A C5A C4A C2A 250.0 2 180.000
IMPRoper C8A C4A C5A NC 250.0 2 180.000
IMPRoper N3A C4A C5A N7A 250.0 2 180.000
IMPRoper C6A C5A C4A N9A 250.0 2 180.000
!*purine
IMPRoper C8P C4P C5P N9P 250.0 2 180.000 ! WYE AND PATCHED RESIDUES
IMPRoper C8P C5P C4P C2P 250.0 2 180.000
IMPRoper C8P C4P C5P NC 250.0 2 180.000
IMPRoper N3P C4P C5P N7P 250.0 2 180.000
IMPRoper C6P C5P C4P N9P 250.0 2 180.000

162
.1 Force field parameter and topology files

! other base specific non-exch hydrogen IMPRopers
IMPRoper H C4C C6C C5C $kchimpr 0 0.000
IMPRoper H N1C C5C C6C $kchimpr 0 0.000
IMPRoper H C4U C6U C5U $kchimpr 0 0.000
IMPRoper H N1U C5U C6U $kchimpr 0 0.000
IMPRoper H N1T C5T C6T $kchimpr 0 0.000
IMPRoper H N7A N9A C8A $kchimpr 0 0.000
IMPRoper H NC N3A C2A $kchimpr 0 0.000
IMPRoper H N7P N9P C8P $kchimpr 0 0.000
IMPRoper H NC N3P C2P $kchimpr 0 0.000
IMPRoper H N7G N9G C8G $kchimpr 0 0.000

! Impropers for ribose chirality
IMPRoper H C2R O4R N9A $kchimpr 0 −65.000°CIR
IMPRoper H C2R O4R N9P $kchimpr 0 −65.000°CIR
IMPRoper H C2R O4R N9G $kchimpr 0 −65.000°CIR
IMPRoper H C2R O4R N1C $kchimpr 0 −65.000°CIR
IMPRoper H C2R O4R N1U $kchimpr 0 −65.000°CIR
IMPRoper H C2R O4R N1T $kchimpr 0 −65.000°CIR
IMPRoper H C3R C1R O2R $kchimpr 0 65.000°C2R
IMPRoper H C4R C2R O3R $kchimpr 0 60.300°C3R
IMPRoper H C4R O4R C3R $kchimpr 0 70.300°C4R
IMPRoper H O5R H C4R $kchimpr 0 72.000°C5R

IMPRoper H C2D O4D N9A $kchimpr 0 −65.280°CID
IMPRoper H C2D O4D N9P $kchimpr 0 −65.280°CID
IMPRoper H C2D O4D N9G $kchimpr 0 −65.280°CID
IMPRoper H C2D O4D N1C $kchimpr 0 −65.280°CID
IMPRoper H C2D O4D N1T $kchimpr 0 −65.280°CID
IMPRoper H C3D C1D H $kchimpr 0 72.000°C3D

! Prop. — periodical potentials from combined RNA/DNA statistics

DIHedral O3R P O5R C5R 1.41 3 24 ! alpha !P (20.3)
DIHedral P O5R C5R C4R 3.45 0 178 ! beta !P (13.0)
DIHedral O5R C5R C4R C3R 12.24 3 18 ! gamma !S (6.9)
DIHedral O5R C5R C4R O4R 24.28 3 14 1 ! !S (4.9)
DIHedral C4R C3R O3R P 7.88 0 −153 ! eps !P (8.6)
DIHedral C3R O3R P O5R 1.75 3 33 ! zeta !P (18.3)

DIHedral O3R P O5R C5D 1.41 3 6.0 !DNA
DIHedral P O5R C5D C4D 3.45 0 183.5

163
**Script code**

DIHedra!l O5R C5D C4D C3D 12.42 3 18.3
DIHedra!l O5R C5D C4D O4D 24.28 3 14.1
DIHedra!l C4D C3D O3R P 7.88 0 214.0
DIHedra!l C3D O3R P O5R 1.75 3 0.3

Phos. – discrete values from combined RNA/DNA statistics

<table>
<thead>
<tr>
<th>keq</th>
<th>x_eq</th>
<th>(sigma in parenthesis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIHedra!l O3R P O5R C5R 6.07 0 285.3</td>
<td>(9.8) alphal !P</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l O3R P O5R C5R 3.98 0 81.1</td>
<td>(12.1) alpha2; alphal=180.</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l P O5R C5R C4R 3.44 0 183.5</td>
<td>(13.0) betal !P</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l O5R C5R C4R C3R 14.23 0 179.4</td>
<td>(6.4) gammal !S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C5R C4R C3R O3R 3.85 0 292.9</td>
<td>(12.3) gammal !S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R C3R O3R P 7.88 0 214.0</td>
<td>(8.6) eps !P</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C3R O3R P O5R 25.30 0 289.2</td>
<td>(4.8) zetal !P</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l O4R C4R C3R O3R 31.53 0 265.8</td>
<td>(4.3) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C1R C2R C3R C4R 44.99 0 326.9</td>
<td>(3.6) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C2R C3R C4R O4R 28.79 0 22.6</td>
<td>(4.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C3R C4R O4R C1R 15.67 0 357.7</td>
<td>(6.1) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N1T 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N1C 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N1U 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N9A 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N9G 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N9U 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N9P 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
</tbody>
</table>

Sugars

! c3'−endo conformation as the default for for RNA, c2'−endo for DNA,

RNA statistics, C3'−endo

<table>
<thead>
<tr>
<th>keq</th>
<th>x_eq</th>
<th>(sigma in parenthesis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIHedra!l O3R C5R C4R C3R O3R 30.12 0 81.1</td>
<td>(4.4) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l O4R C4R C3R C2R O3R 33.10 0 201.8</td>
<td>(4.2) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C1R C2R C3R C4R 74.36 0 35.9</td>
<td>(2.8) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C2R C3R C4R O4R 60.67 0 324.7</td>
<td>(3.1) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C3R O4R C1R N1T 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N1C 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N1U 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N9A 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N9G 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N9U 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4R O4R C1R N9P 13.80 0 241.4</td>
<td>(6.5) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l O4R C1R N1T C2T 13.38 0 195.7</td>
<td>(6.6) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l O4R C1R N1C C2C 13.38 0 195.7</td>
<td>(6.6) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l O4R C1R N1U C2U 13.38 0 195.7</td>
<td>(6.6) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l O4R C1R N9A C4A 2.97 0 193.3</td>
<td>(14.0) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l O4R C1R N9P C4P 2.97 0 193.3</td>
<td>(14.0) c3'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l O4R C1R N9G C4G 2.97 0 193.3</td>
<td>(14.0) c3'−endo S</td>
<td></td>
</tr>
</tbody>
</table>

DNA statistics (c2'−endo)

<table>
<thead>
<tr>
<th>keq</th>
<th>x_eq</th>
<th>(sigma in parenthesis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIHedra!l C5D C4D C3D O3R 36.44 0 145.2</td>
<td>(4.0) c2'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l O4D C4D C3D O3R 31.53 0 265.8</td>
<td>(4.3) c2'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C3D C2D C1D C4D 44.99 0 326.9</td>
<td>(3.6) c2'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C2D C3D C4D C1D 28.79 0 22.6</td>
<td>(4.5) c2'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C3D C4D C1D C2D 15.67 0 357.7</td>
<td>(6.1) c2'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4D O4D C1D C2D 14.69 0 340.7</td>
<td>(6.3) c2'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C5D C4D C3D C2D 34.68 0 262.0</td>
<td>(4.1) c2'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4D O4D C1D C2D 12.99 0 217.7</td>
<td>(6.7) c2'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4D O4D C1D N1C 12.99 0 217.7</td>
<td>(6.7) c2'−endo S</td>
<td></td>
</tr>
<tr>
<td>DIHedra!l C4D O4D C1D N1U 12.99 0 217.7</td>
<td>(6.7) c2'−endo S</td>
<td></td>
</tr>
</tbody>
</table>

164
.1 Force field parameter and topology files

**DIHedral**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIHedral C4D O4D CID N9G</td>
<td>12.99</td>
<td>217.7</td>
</tr>
<tr>
<td>DIHedral C4D O4D CID N9A</td>
<td>12.99</td>
<td>217.7</td>
</tr>
<tr>
<td>DIHedral C4D O4D CID N9P</td>
<td>12.99</td>
<td>217.7</td>
</tr>
</tbody>
</table>

**DNA c2'−endo sugar base joint torsions** (combined RNA/DNA statistics used)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIHedral O4D CID N1T C2T</td>
<td>1.72</td>
<td>229.8</td>
</tr>
<tr>
<td>DIHedral O4D CID N1C C2C</td>
<td>1.72</td>
<td>229.8</td>
</tr>
<tr>
<td>DIHedral O4D CID N1U C2U</td>
<td>1.72</td>
<td>229.8</td>
</tr>
<tr>
<td>DIHedral O4D CID N9A C4A</td>
<td>1.00</td>
<td>237.0</td>
</tr>
<tr>
<td>DIHedral O4D CID N9P C4P</td>
<td>1.00</td>
<td>237.0</td>
</tr>
</tbody>
</table>

**!RNA statistics (c2'−endo)**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIHedral C5R C4R C3R O3R</td>
<td>24.28</td>
<td>147.3</td>
</tr>
<tr>
<td>DIHedral O4R C1R C2R C3R</td>
<td>50.43</td>
<td>35.2</td>
</tr>
<tr>
<td>DIHedral O4R C1R C3R O3R</td>
<td>20.75</td>
<td>268.1</td>
</tr>
<tr>
<td>DIHedral C1R C2R C3R C4R</td>
<td>74.36</td>
<td>324.6</td>
</tr>
<tr>
<td>DIHedral C2R C3R C4R O4R</td>
<td>31.53</td>
<td>24.2</td>
</tr>
<tr>
<td>DIHedral C3R C4R O4R C1R</td>
<td>17.94</td>
<td>357.7</td>
</tr>
<tr>
<td>DIHedral C4R O4R C1R C2R</td>
<td>21.56</td>
<td>339.2</td>
</tr>
<tr>
<td>DIHedral C5R C4R C3R C2R</td>
<td>34.68</td>
<td>263.4</td>
</tr>
<tr>
<td>DIHedral O3R C3R C2R O2R</td>
<td>33.05</td>
<td>319.7</td>
</tr>
<tr>
<td>DIHedral O4R C4R C1R N1T</td>
<td>19.27</td>
<td>216.6</td>
</tr>
<tr>
<td>DIHedral O4R C4R C1R N1C</td>
<td>19.27</td>
<td>216.6</td>
</tr>
<tr>
<td>DIHedral O4R C4R C1R N1U</td>
<td>19.27</td>
<td>216.6</td>
</tr>
<tr>
<td>DIHedral O4R C4R C1R N9G</td>
<td>19.27</td>
<td>216.6</td>
</tr>
<tr>
<td>DIHedral O4R C4R C1R N9A</td>
<td>19.27</td>
<td>216.6</td>
</tr>
</tbody>
</table>

**DNA c3'−endo sugar base joint torsions** (combined RNA/DNA statistics used)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIHedral C5D C4D CID O3R</td>
<td>30.12</td>
<td>81.1</td>
</tr>
<tr>
<td>DIHedral O4D C4D CID O3R</td>
<td>33.10</td>
<td>201.8</td>
</tr>
<tr>
<td>DIHedral O4D CID C2D CID</td>
<td>24.28</td>
<td>335.4</td>
</tr>
<tr>
<td>DIHedral C1D C2D CID CID</td>
<td>74.36</td>
<td>35.9</td>
</tr>
<tr>
<td>DIHedral C2D CID C4D CID</td>
<td>60.67</td>
<td>324.7</td>
</tr>
<tr>
<td>DIHedral C3D C4D O4D CID</td>
<td>22.42</td>
<td>20.5</td>
</tr>
<tr>
<td>DIHedral C4D O4D CID CID</td>
<td>15.67</td>
<td>2.8</td>
</tr>
<tr>
<td>DIHedral C5D C4D CID CID</td>
<td>60.67</td>
<td>204.0</td>
</tr>
<tr>
<td>DIHedral C4D O4D CID N1T</td>
<td>13.80</td>
<td>241.4</td>
</tr>
<tr>
<td>DIHedral C4D O4D CID N1C</td>
<td>13.80</td>
<td>241.4</td>
</tr>
<tr>
<td>DIHedral C4D O4D CID N1U</td>
<td>13.80</td>
<td>241.4</td>
</tr>
<tr>
<td>DIHedral C4D O4D CID N9G</td>
<td>13.80</td>
<td>241.4</td>
</tr>
<tr>
<td>DIHedral C4D O4D CID N9A</td>
<td>13.80</td>
<td>241.4</td>
</tr>
</tbody>
</table>

**DNA c3'−endo sugar base joint torsions** (combined RNA/DNA statistics used)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIHedral O4D CID N1T C2T</td>
<td>13.38</td>
<td>195.7</td>
</tr>
</tbody>
</table>
Script code

! DIHEDRAL O4D C1D N1C C2C 13.38 0 195.7 ! 6.6 ! c3'-endo S
! DIHEDRAL O4D C1D N1U C2U 13.38 0 195.7 ! 6.6 ! c3'-endo S
! DIHEDRAL O4D C1D N9A C4A 2.97 0 193.3 ! 14.0 ! c3'-endo S
! DIHEDRAL O4D C1D N9G C4G 2.97 0 193.3 ! 14.0 ! c3'-endo S

! Impropers taken from param11 dna, 3*kg
IMPRoper C5R X X C2R 94.5 0 35.260
IMPRoper C5R X X C1R 94.5 0 35.260
IMPRoper OH X X C3R 94.5 0 35.260
IMPRoper OH X X C4R 94.5 0 35.260
IMPRoper OH X X C1R 94.5 0 35.260
IMPRoper O3R X X C3R 94.5 0 35.260
IMPRoper O3R X X C2R 94.5 0 35.260
IMPRoper O2R X X C2R 94.5 0 35.260
IMPRoper C4R O5R C1R N1T 94.5 0 35.260
IMPRoper C4R O5R C1R N1C 94.5 0 35.260
IMPRoper C4R O5R C1R N9G 94.5 0 35.260
IMPRoper C4R O5R C1R N9A 94.5 0 35.260
IMPRoper C4R O5R C1R N9P 94.5 0 35.260
IMPRoper C4R O5R C1R N1U 94.5 0 35.260
IMPRoper N1T C2R O4R C1R 94.5 0 35.260
IMPRoper N1C C2R O4R C1R 94.5 0 35.260
IMPRoper N9A C2R O4R C1R 94.5 0 35.260
IMPRoper N9P C2R O4R C1R 94.5 0 35.260
IMPRoper C4R O5R C1R N1U 94.5 0 35.260
IMPRoper N1U C2R O4R C1R 94.5 0 35.260
IMPRoper C5D X X C2D 94.5 0 35.260
IMPRoper C5D X X C1D 94.5 0 35.260
IMPRoper OH X X C3D 94.5 0 35.260
IMPRoper OH X X C4D 94.5 0 35.260
IMPRoper OH X X C1D 94.5 0 35.260
IMPRoper O3R X X C3D 94.5 0 35.260
IMPRoper O3R X X C2D 94.5 0 35.260
IMPRoper O5R X X C1D 94.5 0 35.260
IMPRoper C4D O5R C1D N1T 94.5 0 35.260
IMPRoper C4D O5R C1D N1C 94.5 0 35.260
IMPRoper C4D O5R C1D N9G 94.5 0 35.260
IMPRoper C4D O5R C1D N9A 94.5 0 35.260
IMPRoper C4D O5R C1D N9P 94.5 0 35.260
IMPRoper C4D O5R C1D N9A 94.5 0 35.260
IMPRoper C4D O5R C1D N9G 94.5 0 35.260
IMPRoper C4D O5R C1D N9P 94.5 0 35.260
IMPRoper C5D O4D C3D C4D 94.5 0 35.260
IMPRoper N1T C2D O4D C1D 94.5 0 35.260
IMPRoper N1C C2D O4D C1D 94.5 0 35.260
IMPRoper N9A C2D O4D C1D 94.5 0 35.260
IMPRoper N9P C2D O4D C1D 94.5 0 35.260

166
.1 Force field parameter and topology files

! the following impropers were taken from param11x.dna
! the higher kq was used to enforce the ring planarity

! cytosine
IMProper C4C X X ON 2400.0 0 0.000
IMProper C4C X X N1C 250.0 0 0.000
IMProper C6C X X NC 250.0 0 0.000
IMProper C4C X X N2 2400.0 0 0.000
IMProper C2C X X ON 2400.0 0 0.000

! infer
IMProper C1R C2C C6C N1C 2400.0 0 0.000
IMProper C1D C2C C6C N1C 2400.0 0 0.000
IMProper N4C NC C5C C4C 2400.0 0 0.000
IMProper C2C NC C4C C5C 250.0 0 0.000
IMProper C5C C6C N1C C2C 250.0 0 0.000
IMProper H2 C4C H2 N4C 250.0 0 0.000
IMProper C5C C4C N4C H2 2000.0 0 0.000

! uracil
IMProper C4U X X ON 2400.0 0 0.000
IMProper C4U X X N1U 250.0 0 0.000
IMProper C6U X X N3U 250.0 0 0.000
IMProper C4U X X N2 2400.0 0 0.000
IMProper C2U X X ON 2400.0 0 0.000
IMProper ON N3U C5U C4U 250.0 0 0.000
IMProper C2U N3U C4U C5U 250.0 0 0.000
IMProper C5U C6U N1U C2U 250.0 0 0.000
IMProper H2 C4U H2 ON 250.0 0 0.000
IMProper HN C2U C4U N3U 250.0 0 0.000
IMProper H C3D NX29 C1D 500.0 0 0.000

! thymidine
IMProper C4T X X ON 2400.0 0 0.000
IMProper C4T X X N1T 250.0 0 0.000
IMProper C6T X X N3T 250.0 0 0.000
IMProper C4T X X N2 2400.0 0 0.000
IMProper C2T X X ON 2400.0 0 0.000
IMProper C1R C2T C6T N1T 2400.0 0 0.000
IMProper C1D C2T C6T N1T 2400.0 0 0.000
IMProper ON N3T C5T C4T 2400.0 0 0.000
IMProper C2T N3T C4T C5T 250.0 0 0.000
IMProper C5T C6T N1T C2T 250.0 0 0.000
IMProper H2 C4T H2 ON 250.0 0 0.000
IMProper CC3E C4T C6T C5T 2400.0 0 0.000

! infer
IMProper HN C2T C4T N3T 250.0 0 0.000

! The ring-spanning impropers have been left out.

! adenine
IMProper N2A N3A NC C2A 250.0 0 0.000
IMPRoper H2 C2A H2 N2A 250.0 0 0.000
IMPRoper C4A C5A N7A C8A 250.0 0 0.000
IMPRoper C5A C4A N9A C8A 250.0 0 0.000
IMPRoper C4A X X NC 250.0 0 0.000
IMPRoper C2A X X N9A 250.0 0 0.000
IMPRoper C2A X X C5A 250.0 0 0.000
IMPRoper C6A C5A C4A N3A 250.0 0 0.000
IMPRoper C5A X X N9A 250.0 0 0.000
IMPRoper C6A X X N6A 2400.0 0 0.000
IMPRoper H2 X X N6A 250.0 0 0.000

IMPRoper C1R C4A C8A N9A 2400.0 0 0.000
IMPRoper C1D C4A C8A N9A 2400.0 0 0.000
IMPRoper N9A C4A C5A N7A 250.0 0 0.000
IMPRoper N7A C8A N9A C4A 250.0 0 0.000
IMPRoper N3A C2A NC C6A 250.0 0 0.000
IMPRoper C5A C6A N6A H2 2000.0 0 0.000

IMPRoper C1R C4P C8P N9P 2400.0 0 0.000
IMPRoper C1D C4P C8P N9P 2400.0 0 0.000
IMPRoper N9P C4P C5P N7P 250.0 0 0.000
IMPRoper N7P C8P N9P C4P 250.0 0 0.000
IMPRoper N3P C2P NC C6P 250.0 0 0.000

IMPRoper C1R C4G C8G N9G 2400.0 0 0.000
IMPRoper C1D C4G C8G N9G 2400.0 0 0.000
IMPRoper N9G C4G C5G N7G 250.0 0 0.000
IMPRoper N7G C8G N9G C4G 250.0 0 0.000
IMPRoper N3G C2G NNA C6G 250.0 0 0.000
IMPRoper H2 H2 C2G N2G 250.0 0 0.000

! The ring-spanning impropers have been left out.

purine

IMPRoper C6P C5P C4P N3P 250.0 0 0.000
IMPRoper C5P X X N9P 250.0 0 0.000
IMPRoper C2P X X N9P 250.0 0 0.000
IMPRoper C6G X X N9G 250.0 0 0.000
IMPRoper C6G X X O6G 2400.0 0 0.000
IMPRoper C5G X X N9G 250.0 0 0.000
IMPRoper C1R C4P C8P N9P 2400.0 0 0.000
IMPRoper C1D C4P C8P N9P 2400.0 0 0.000

IMPRoper C6P C5P C4P N3P 250.0 0 0.000
IMPRoper C5P X X N9P 250.0 0 0.000
IMPRoper C2P X X N9P 250.0 0 0.000
IMPRoper C6G X X N9G 250.0 0 0.000
IMPRoper C6G X X O6G 2400.0 0 0.000
IMPRoper C5G X X N9G 250.0 0 0.000
IMPRoper C1R C4G C8G N9G 2400.0 0 0.000
IMPRoper C1D C4G C8G N9G 2400.0 0 0.000

IMPRoper C6G X X N9G 250.0 0 0.000
IMPRoper C6G X X O6G 2400.0 0 0.000
IMPRoper C2G X X N2G 2400.0 0 0.000

! The ring-spanning impropers have been left out.

guanine

IMPRoper C4G C5G N7G C8G 250.0 0 0.000
IMPRoper C5G C4G N9G C8G 250.0 0 0.000
IMPRoper C4G X X NNA 250.0 0 0.000
IMPRoper C2G X X N9G 250.0 0 0.000
IMPRoper C2G X X C5G 250.0 0 0.000
IMPRoper C6G C5G C4G N3G 250.0 0 0.000
IMPRoper C5G X X N9G 250.0 0 0.000
IMPRoper C6G X X O6G 2400.0 0 0.000
IMPRoper C2G X X N2G 2400.0 0 0.000

IMPRoper C1R C4G C8G N9G 2400.0 0 0.000
IMPRoper C1D C4G C8G N9G 2400.0 0 0.000
IMPRoper N9G C4G C5G N7G 250.0 0 0.000
IMPRoper N7G C8G N9G C4G 250.0 0 0.000
IMPRoper N3G C2G NNA C6G 250.0 0 0.000
IMPRoper H2 H2 C2G N2G 250.0 0 0.000
### 1 Force field parameter and topology files

<table>
<thead>
<tr>
<th>IMPRoper</th>
<th>Parameter</th>
<th>Topology</th>
<th>Value 1</th>
<th>Value 2</th>
<th>Value 3</th>
<th>Value 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMPRoper</td>
<td>HN</td>
<td>C2G</td>
<td>C6G</td>
<td>NNA</td>
<td>2000.0</td>
<td>0</td>
</tr>
</tbody>
</table>

Lennard-Jones parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value 1</th>
<th>Value 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varepsilon )</td>
<td>( \sigma )</td>
<td>( \varepsilon )</td>
</tr>
<tr>
<td>(Kcal/mol)</td>
<td>(Å)</td>
<td>(Kcal/mol)</td>
</tr>
</tbody>
</table>

Taken from Rossky Karplus and Rahman BIOPOLY (1979)

0.05 ADDED TO RADII TO IMPRoperOVE ON NUCL.ACID STACKING/LN

NONBonded

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value 1</th>
<th>Value 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>FX1</td>
<td>0.5849</td>
<td>3.3854</td>
</tr>
<tr>
<td>FX2</td>
<td>0.2304</td>
<td>2.7290</td>
</tr>
<tr>
<td>FX3</td>
<td>0.2304</td>
<td>2.7290</td>
</tr>
<tr>
<td>FX4</td>
<td>0.2304</td>
<td>2.7290</td>
</tr>
<tr>
<td>FX5</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX6</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX7</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX8</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX9</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX10</td>
<td>0.2304</td>
<td>2.7290</td>
</tr>
<tr>
<td>FX11</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX12</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX13</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX14</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX15</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX16</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX17</td>
<td>0.2304</td>
<td>2.7290</td>
</tr>
<tr>
<td>FX18</td>
<td>0.2304</td>
<td>2.7290</td>
</tr>
<tr>
<td>FX19</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX20</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX21</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX22</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX23</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX24</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX25</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX26</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX27</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX28</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX29</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX30</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX31</td>
<td>0.0900</td>
<td>3.2970</td>
</tr>
<tr>
<td>FX32</td>
<td>0.1600</td>
<td>2.8591</td>
</tr>
<tr>
<td>FX33</td>
<td>0.2304</td>
<td>2.7290</td>
</tr>
<tr>
<td>FX34</td>
<td>0.2304</td>
<td>2.7290</td>
</tr>
<tr>
<td>FX35</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX36</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX37</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX38</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX39</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX40</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX41</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
<tr>
<td>FX42</td>
<td>0.0045</td>
<td>2.6160</td>
</tr>
</tbody>
</table>
### Script code

```plaintext
NONBonded HX43 0.0045 2.6160 0.0045 2.6160 ! mod by anda, std value

END mod by anda

NONBonded C5R 0.0900 3.2970 0.0900 3.2970
NONBonded C1R 0.0900 3.2970 0.0900 3.2970
NONBonded C2R 0.0900 3.2970 0.0900 3.2970
NONBonded C3R 0.0900 3.2970 0.0900 3.2970
NONBonded C4R 0.0900 3.2970 0.0900 3.2970

NONBonded C5D 0.0900 3.2970 0.0900 3.2970 !DNA
NONBonded C1D 0.0900 3.2970 0.0900 3.2970
NONBonded C2D 0.0900 3.2970 0.0900 3.2970
NONBonded C3D 0.0900 3.2970 0.0900 3.2970
NONBonded C4D 0.0900 3.2970 0.0900 3.2970

NONBonded HN 0.0045 2.6160 0.0045 2.6160
NONBonded H2 0.0045 1.6040 0.0045 1.6040
NONBonded H 0.0045 2.6160 0.0045 2.6160

! give it the same as th Hn from RKR
NONBonded HO 0.0045 1.6040 0.0045 1.6040

! THIS STILL IS AN EXTENDED ATOM
NONBonded O3R 0.2304 2.7290 0.2304 2.7290
NONBonded O4R 0.2304 2.7290 0.2304 2.7290
NONBonded O4D 0.2304 2.7290 0.2304 2.7290
NONBonded O5R 0.2304 2.7290 0.2304 2.7290
NONBonded O1P 0.2304 2.7290 0.2304 2.7290
NONBonded O2P 0.2304 2.7290 0.2304 2.7290
NONBonded P 0.5849 3.3854 0.5849 3.3854

! bases
NONBonded C2 0.0900 3.2970 0.0900 3.2970
NONBonded C3 0.0900 3.2970 0.0900 3.2970
NONBonded CB 0.0900 3.2970 0.0900 3.2970
NONBonded CE 0.0900 3.2970 0.0900 3.2970
NONBonded CH 0.0900 3.2970 0.0900 3.2970

NONBonded N2 0.1600 2.8591 0.1600 2.8591
NONBonded N3U 0.1600 2.8591 0.1600 2.8591
NONBonded N3T 0.1600 2.8591 0.1600 2.8591
NONBonded NNA 0.1600 2.8591 0.1600 2.8591
NONBonded NB 0.1600 2.8591 0.1600 2.8591
NONBonded NC 0.1600 2.8591 0.1600 2.8591

NONBonded NH2E 0.1600 3.0291 0.1600 3.0291
NONBonded NS 0.1600 2.8591 0.1600 2.8591
NONBonded NIT 0.1600 2.8591 0.1600 2.8591
NONBonded NIC 0.1600 2.8591 0.1600 2.8591
NONBonded N9A 0.1600 2.8591 0.1600 2.8591
NONBonded N9P 0.1600 2.8591 0.1600 2.8591
NONBonded NSG 0.1600 2.8591 0.1600 2.8591
NONBonded N1U 0.1600 2.8591 0.1600 2.8591
NONBonded ON 0.2304 2.7290 0.2304 2.7290
```

170
| NONBonded  | O2R | 0.2304 | 2.7290 | 0.2304 | 2.7290 |
| NONBonded  | OB  | 0.2304 | 2.5508 | 0.2304 | 2.5508 |
| NONBonded  | SD  | 0.3515 | 2.6727 | 0.3515 | 2.6727 |
| NONBonded  | G2  | 0.2304 | 2.7290 | 0.2304 | 2.7290 |

| NEW |
| NONBonded  | C6C | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C5C | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C4C | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C2C | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C6U | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C5U | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C4U | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C2U | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C6A | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C5A | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C4A | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C2A | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C6P | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C5P | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C4P | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C2P | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C8G | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C6G | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C5G | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C4G | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C2G | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C6T | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C5T | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C4T | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | C2T | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | N4C | 0.1600 | 2.8591 | 0.1600 | 2.8591 |
| NONBonded  | O4U | 0.2304 | 2.7290 | 0.2304 | 2.7290 |
| NONBonded  | N7G | 0.1600 | 2.8591 | 0.1600 | 2.8591 |
| NONBonded  | N3G | 0.1600 | 2.8591 | 0.1600 | 2.8591 |
| NONBonded  | N2G | 0.1600 | 2.8591 | 0.1600 | 2.8591 |
| NONBonded  | N3A | 0.1600 | 2.8591 | 0.1600 | 2.8591 |
| NONBonded  | N7A | 0.1600 | 2.8591 | 0.1600 | 2.8591 |
| NONBonded  | N6A | 0.1600 | 2.8591 | 0.1600 | 2.8591 |
| NONBonded  | O6G | 0.2304 | 2.7290 | 0.2304 | 2.7290 |
| NONBonded  | CC3E | 0.0900 | 3.2970 | 0.0900 | 3.2970 |
| NONBonded  | N2A | 0.1600 | 2.8591 | 0.1600 | 2.8591 |
| NONBonded  | N2P | 0.1600 | 2.8591 | 0.1600 | 2.8591 |
| NONBonded  | N3P | 0.1600 | 2.8591 | 0.1600 | 2.8591 |
| NONBonded  | NTP | 0.1600 | 2.8591 | 0.1600 | 2.8591 |

| special solute-solute hydrogen bonding potential parameters |
| AEXP 4 |
| REXP 6 |
| HAE 4 |
| AAEX 2 |

| all possible combinations of HB-pairs in nucleic acids:
WELL DEPTHS DEEPENED BY 0.5 KCAL TO IMPROVE BASEPAIR ENERGIES /LN
 AND DISTANCES INCREASED BY 0.05

\[ E_{\text{min}} \quad R_{\text{min}} \]

\[ \text{(Kcal/mol)} \quad \text{(A)} \]

\[ h_{\text{bond}} \quad N^* \quad O^* \quad -14.0 \quad 2.95 \]

\[ h_{\text{bond}} \quad N^* \quad N^* \quad -14.5 \quad 3.05 \]

\[ h_{\text{bond}} \quad O^* \quad O^* \quad -15.75 \quad 2.80 \]

\[ h_{\text{bond}} \quad O^* \quad N^* \quad -15.50 \quad 2.90 \]

THE FOLLOWING NBFIXES ARE FOR DNA-DNA HYDROGEN BONDING

\[ \begin{array}{cccc}
\text{nbfix} & \text{HO} & \text{ON} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{HO} & \text{O3R} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{HO} & \text{O5R} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{HO} & \text{OH} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{HO} & \text{O2R} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{HO} & \text{NC} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{ON} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{O2} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{O5R} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{O4R} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{O4D} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{O3R} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{O2R} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{OH} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{N7A} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{N7P} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{N7G} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{N3A} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{N3P} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H} & \text{N3G} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{HN} & \text{ON} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{HN} & \text{O2R} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{HN} & \text{OH} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{HN} & \text{NC} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H2} & \text{ON} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H2} & \text{O2R} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H2} & \text{OH} & 0.05 & 0.1 & 0.05 & 0.1 \\
\text{nbfix} & \text{H2} & \text{NC} & 0.05 & 0.1 & 0.05 & 0.1 \\
\end{array} \]

\[ \text{!mod by anda} \]

\[ \text{\texttt{NB\textdagger} AP} \]

\[ \text{BOND} \quad \text{C2P} \quad \text{N2G} \quad 1110.000 \quad 1.341 \quad 0.010 \quad \text{B} \]

\[ \text{ANGLE} \quad \text{NC} \quad \text{C2P} \quad \text{N2G} \quad 449.866 \quad 116.20 \quad 0.90 \quad \text{B} \]

172
.1 Force field parameter and topology files

```
ANGLe N3P C2P N2G 743.656 119.900 ! 0 . 7 0 B
ANGLe C2P N2G H2 105.000 120.000 ! from param11 . dna , 3
DIHedral X C2P N2G X 18.0 2 180.000
IMProper N3P C2P N2G H2 2000.0 0 0.000
IMProper C2P X X N2G 2400.0 0 0.000
! end of mod by

set echo=on message=on end

!RNA TOPOLOGY FILE 'FRAMEWORK' FROM TOPALLHDG . DNA AND ATOM NAMES
! INCLUDES ALL NONEXCHANGEABLE HYDROGENS AND TERMS FOR BOND, ANGLE, AND
! IMPROPER. NONEXCHANGEABLE HYDROGEN CHARGES WERE ASSIGNED 0 . 0 3 5 .
! CARBON CHARGES WERE REDUCED 0 . 035 FOR EACH ATTACHED HYDROGEN.
! CREATED 2/24/96-- JASON P. RIFE AND PETER B. MOORE
! DNA-RNA-ALLATOM . TOP

set echo=false end

! checkversion 1.0

AUTOGENERATE ANGLES=TRUE END
```

```
MASS PX1 30 . 97400 ! assuming P -> 30 . 97400 + 1 . 008 * 0 (Hs)
MASS OX2 15 . 99990 ! assuming O -> 15 . 99990 + 1 . 008 * 0 (Hs)
MASS OX3 15 . 99990 ! assuming O -> 15 . 99990 + 1 . 008 * 0 (Hs)
MASS OX4 15 . 99990 ! assuming O -> 15 . 99990 + 1 . 008 * 0 (Hs)
MASS CX5 12 . 01100 ! assuming C -> 12 . 01100 + 1 . 008 * 0 (Hs)
MASS HX6 1 . 00800 ! assuming H -> 1 . 00800 + 1 . 008 * 0 (Hs)
MASS HX7 1 . 00800 ! assuming H -> 1 . 00800 + 1 . 008 * 0 (Hs)
MASS CX8 12 . 01100 ! assuming C -> 12 . 01100 + 1 . 008 * 0 (Hs)
MASS HX9 1 . 00800 ! assuming H -> 1 . 00800 + 1 . 008 * 0 (Hs)
MASS OX10 15 . 99990 ! assuming O -> 15 . 99990 + 1 . 008 * 0 (Hs)
MASS CX11 12 . 01100 ! assuming C -> 12 . 01100 + 1 . 008 * 0 (Hs)
MASS CX12 12 . 01100 ! assuming C -> 12 . 01100 + 1 . 008 * 0 (Hs)
MASS HX13 1 . 00800 ! assuming H -> 1 . 00800 + 1 . 008 * 0 (Hs)
MASS CX14 12 . 01100 ! assuming C -> 12 . 01100 + 1 . 008 * 0 (Hs)
MASS HX15 1 . 00800 ! assuming H -> 1 . 00800 + 1 . 008 * 0 (Hs)
MASS HX16 1 . 00800 ! assuming H -> 1 . 00800 + 1 . 008 * 0 (Hs)
MASS OX17 15 . 99990 ! assuming O -> 15 . 99990 + 1 . 008 * 0 (Hs)
MASS OX18 15 . 99990 ! assuming O -> 15 . 99990 + 1 . 008 * 0 (Hs)
MASS CX19 12 . 01100 ! assuming C -> 12 . 01100 + 1 . 008 * 0 (Hs)
MASS CX20 12 . 01100 ! assuming C -> 12 . 01100 + 1 . 008 * 0 (Hs)
MASS CX21 12 . 01100 ! assuming C -> 12 . 01100 + 1 . 008 * 0 (Hs)
MASS CX22 12 . 01100 ! assuming C -> 12 . 01100 + 1 . 008 * 0 (Hs)
MASS CX23 12 . 01100 ! assuming C -> 12 . 01100 + 1 . 008 * 0 (Hs)

173
<table>
<thead>
<tr>
<th>Mass Symbol</th>
<th>Mass Value</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>MASS CX24</code></td>
<td>12.01100</td>
<td>assuming C → 12.01100 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS CX25</code></td>
<td>12.01100</td>
<td>assuming C → 12.01100 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS CX26</code></td>
<td>12.01100</td>
<td>assuming C → 12.01100 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS CX27</code></td>
<td>12.01100</td>
<td>assuming C → 12.01100 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS CX28</code></td>
<td>12.01100</td>
<td>assuming C → 12.01100 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS CX29</code></td>
<td>12.01100</td>
<td>assuming C → 12.01100 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS CX30</code></td>
<td>12.01100</td>
<td>assuming C → 12.01100 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS CX31</code></td>
<td>12.01100</td>
<td>assuming C → 12.01100 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS NX32</code></td>
<td>14.00700</td>
<td>assuming N → 14.00700 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS OX33</code></td>
<td>15.99900</td>
<td>assuming O → 15.99900 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS OX34</code></td>
<td>15.99900</td>
<td>assuming O → 15.99900 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS HX35</code></td>
<td>1.00800</td>
<td>assuming H → 1.00800 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS HX36</code></td>
<td>1.00800</td>
<td>assuming H → 1.00800 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS HX37</code></td>
<td>1.00800</td>
<td>assuming H → 1.00800 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS HX38</code></td>
<td>1.00800</td>
<td>assuming H → 1.00800 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS HX39</code></td>
<td>1.00800</td>
<td>assuming H → 1.00800 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS HX40</code></td>
<td>1.00800</td>
<td>assuming H → 1.00800 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS HX41</code></td>
<td>1.00800</td>
<td>assuming H → 1.00800 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS HX42</code></td>
<td>1.00800</td>
<td>assuming H → 1.00800 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS HX43</code></td>
<td>1.00800</td>
<td>assuming H → 1.00800 + 0.008 * 0 (Hs)</td>
</tr>
<tr>
<td><code>MASS P</code></td>
<td>30.97400</td>
<td>phosphorus</td>
</tr>
<tr>
<td><code>MASS O1P</code></td>
<td>15.99940</td>
<td>O in phosphate</td>
</tr>
<tr>
<td><code>MASS O2P</code></td>
<td>15.99940</td>
<td>O in phosphate</td>
</tr>
<tr>
<td><code>MASS O3R</code></td>
<td>15.99940</td>
<td>ester → P-O-C</td>
</tr>
<tr>
<td><code>MASS O4R</code></td>
<td>15.99940</td>
<td>ester → P-O-C</td>
</tr>
<tr>
<td><code>MASS O5R</code></td>
<td>15.99940</td>
<td>ester → P-O-C</td>
</tr>
<tr>
<td><code>MASS OH</code></td>
<td>15.99940</td>
<td>correspond to OH</td>
</tr>
<tr>
<td><code>MASS C5D</code></td>
<td>14.02700</td>
<td>correspond to CH2E</td>
</tr>
<tr>
<td><code>MASS C4D</code></td>
<td>13.01900</td>
<td>correspond to CH1E</td>
</tr>
<tr>
<td><code>MASS C3D</code></td>
<td>13.01900</td>
<td>correspond to CH1E</td>
</tr>
<tr>
<td><code>MASS C2D</code></td>
<td>13.01900</td>
<td>correspond to CH1E</td>
</tr>
<tr>
<td><code>MASS C1D</code></td>
<td>13.01900</td>
<td>correspond to CH1E</td>
</tr>
<tr>
<td><code>MASS O4D</code></td>
<td>15.99940</td>
<td>ester → P-O-C</td>
</tr>
<tr>
<td><code>MASS O3D</code></td>
<td>15.99940</td>
<td>ester → P-O-C</td>
</tr>
<tr>
<td><code>MASS OH</code></td>
<td>15.99940</td>
<td>correspond to OH</td>
</tr>
<tr>
<td><code>MASS N2</code></td>
<td>14.00670</td>
<td>nitrogen in →NH2</td>
</tr>
<tr>
<td><code>MASS NNA</code></td>
<td>14.00670</td>
<td>correspond to NH</td>
</tr>
<tr>
<td><code>MASS ON</code></td>
<td>15.99940</td>
<td>correspond to O</td>
</tr>
<tr>
<td><code>MASS NC</code></td>
<td>14.00670</td>
<td>correspond to NR</td>
</tr>
</tbody>
</table>
. Force field parameter and topology files

<table>
<thead>
<tr>
<th>MASS</th>
<th>NS</th>
<th>14.00670</th>
<th>nitrogen in ring &gt;N-</th>
</tr>
</thead>
</table>

! Insert 4 Bases

! GUA
<table>
<thead>
<tr>
<th>MASS</th>
<th>N9G</th>
<th>14.00670</th>
<th>nitrogen in ring &gt;N-</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASS</td>
<td>C2G</td>
<td>12.0111</td>
<td>(prev CE)</td>
</tr>
<tr>
<td>MASS</td>
<td>N3G</td>
<td>14.00670</td>
<td>(prev NC)</td>
</tr>
<tr>
<td>MASS</td>
<td>C4G</td>
<td>12.01100</td>
<td>(prev CB)</td>
</tr>
<tr>
<td>MASS</td>
<td>C5G</td>
<td>12.01100</td>
<td>(prev CB)</td>
</tr>
<tr>
<td>MASS</td>
<td>C6G</td>
<td>12.01100</td>
<td>(prev CN)</td>
</tr>
<tr>
<td>MASS</td>
<td>N7G</td>
<td>14.00670</td>
<td>(prev NB)</td>
</tr>
<tr>
<td>MASS</td>
<td>C8G</td>
<td>12.0111</td>
<td>(prev CE)</td>
</tr>
<tr>
<td>MASS</td>
<td>O6G</td>
<td>15.99940</td>
<td>(prev CE)</td>
</tr>
<tr>
<td>MASS</td>
<td>N2G</td>
<td>14.00670</td>
<td>nitrogen in -NH2</td>
</tr>
</tbody>
</table>

! ADE
<table>
<thead>
<tr>
<th>MASS</th>
<th>N9A</th>
<th>14.00670</th>
<th>nitrogen in ring &gt;N-</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASS</td>
<td>C2A</td>
<td>12.0111</td>
<td>(prev CE)</td>
</tr>
<tr>
<td>MASS</td>
<td>N3A</td>
<td>14.00670</td>
<td>(prev NC)</td>
</tr>
<tr>
<td>MASS</td>
<td>C4A</td>
<td>12.01100</td>
<td>(prev CB)</td>
</tr>
<tr>
<td>MASS</td>
<td>C5A</td>
<td>12.01100</td>
<td>(prev CB)</td>
</tr>
<tr>
<td>MASS</td>
<td>C6A</td>
<td>12.01100</td>
<td>(prev CA)</td>
</tr>
<tr>
<td>MASS</td>
<td>N7A</td>
<td>14.00670</td>
<td>(prev NB)</td>
</tr>
<tr>
<td>MASS</td>
<td>C8A</td>
<td>12.0111</td>
<td>(prev CE)</td>
</tr>
<tr>
<td>MASS</td>
<td>N6A</td>
<td>14.00670</td>
<td>nitrogen in -NH2</td>
</tr>
</tbody>
</table>

! PUR
<table>
<thead>
<tr>
<th>MASS</th>
<th>N9P</th>
<th>14.00670</th>
<th>nitrogen in ring &gt;N-</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASS</td>
<td>C2P</td>
<td>12.0111</td>
<td>(prev CE)</td>
</tr>
<tr>
<td>MASS</td>
<td>N3P</td>
<td>14.00670</td>
<td>(prev NC)</td>
</tr>
<tr>
<td>MASS</td>
<td>C4P</td>
<td>12.01100</td>
<td>(prev CB)</td>
</tr>
<tr>
<td>MASS</td>
<td>C5P</td>
<td>12.01100</td>
<td>(prev CB)</td>
</tr>
<tr>
<td>MASS</td>
<td>C6P</td>
<td>12.01100</td>
<td>(prev CA)</td>
</tr>
<tr>
<td>MASS</td>
<td>N7P</td>
<td>14.00670</td>
<td>(prev NB)</td>
</tr>
<tr>
<td>MASS</td>
<td>C8P</td>
<td>12.0111</td>
<td>(prev CE)</td>
</tr>
</tbody>
</table>

! CYT
<table>
<thead>
<tr>
<th>MASS</th>
<th>N1C</th>
<th>14.00670</th>
<th>nitrogen in ring &gt;N-</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASS</td>
<td>C2C</td>
<td>12.01100</td>
<td>(prev CN)</td>
</tr>
<tr>
<td>MASS</td>
<td>C4C</td>
<td>12.01100</td>
<td>(prev CA)</td>
</tr>
<tr>
<td>MASS</td>
<td>C5C</td>
<td>12.0111</td>
<td>(prev CF)</td>
</tr>
<tr>
<td>MASS</td>
<td>C6C</td>
<td>12.0111</td>
<td>(prev CF)</td>
</tr>
<tr>
<td>MASS</td>
<td>CC3E</td>
<td>12.01100</td>
<td>(prev CF)</td>
</tr>
</tbody>
</table>

! THY
<table>
<thead>
<tr>
<th>MASS</th>
<th>N1T</th>
<th>14.00670</th>
<th>nitrogen in ring &gt;N-</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASS</td>
<td>N3T</td>
<td>14.00670</td>
<td>nitrogen in ring &gt;N-</td>
</tr>
<tr>
<td>MASS</td>
<td>C2T</td>
<td>12.01100</td>
<td>(prev CN)</td>
</tr>
<tr>
<td>MASS</td>
<td>C4T</td>
<td>12.01100</td>
<td>(prev CN)</td>
</tr>
<tr>
<td>MASS</td>
<td>C5T</td>
<td>12.0111</td>
<td>(prev CS)</td>
</tr>
<tr>
<td>MASS</td>
<td>C6T</td>
<td>12.0111</td>
<td>(prev CF)</td>
</tr>
<tr>
<td>MASS</td>
<td>CC3E</td>
<td>12.01100</td>
<td>(prev CF)</td>
</tr>
</tbody>
</table>

! END

<table>
<thead>
<tr>
<th>MASS</th>
<th>H</th>
<th>1.00800</th>
<th>non-exchangeable Hydrogens</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASS</td>
<td>HN</td>
<td>1.00800</td>
<td>correspond to H</td>
</tr>
</tbody>
</table>

175
Script code

MASS  H2  1.00800! hydrogen in −NH2
MASS  HO  1.00800! hydroxy hydrogen

! URI
MASS  NIU  14.00670! nitrogen in ring >N−
MASS  C2U  12.01100! (prev CN)
MASS  C4U  12.01100! (prev CA)
MASS  C5U  12.011! (prev CF)
MASS  C6U  12.011! (prev CF)
MASS  N3U  14.00670!

!*************************************************************
!*************************************************************

!********************* mod by anda − HNF *********************

RESIdue HNF

{ Note: electrostatics should normally not be used in }
{ crystallographic refinement since it can produce }
{ artefacts. For this reason, all charges are set to }
{ zero by default. Edit them if necessary }

GROUP
ATOM  P   TYPE PX1  CHARGE=1.20 END
ATOM  O1P TYPE OX2  CHARGE=−0.40 END
ATOM  O2P TYPE OX3  CHARGE=−0.40 END
ATOM  O5' TYPE OX4  CHARGE=−0.36 END
GROUP
ATOM  C5' TYPE CX5  CHARGE=−0.070 END
ATOM  H5' TYPE HX6  CHARGE=0.035 END
ATOM  H5' ' TYPE HX7  CHARGE=0.035 END
GROUP
ATOM  C4' TYPE CX8  CHARGE=0.100 END !mod. to match DFT calc
ATOM  H4' TYPE HX9  CHARGE=0.105 END !mod. to match DFT calc
ATOM  O4' TYPE OX10 CHARGE=−0.263 END !mod. to match DFT calc, inc. by +0.13 for neutrality
ATOM  C1' TYPE CX11 CHARGE=0.302 END !mod. to match DFT calc
ATOM  H1' ' TYPE HX13 CHARGE=0.076 END !mod. to match DFT calc
ATOM  O2' TYPE OX18 CHARGE=−0.244 END !reduced charge by 0.076 (residual charge of HNF)
GROUP
ATOM  C2' TYPE CX14 CHARGE=−0.070 END
ATOM  H2' TYPE HX15 CHARGE=0.035 END
ATOM  H2' ' TYPE HX16 CHARGE=0.035 END
GROUP
ATOM  C3' TYPE CX12 CHARGE=−0.035 END
ATOM  H3' TYPE HX13 CHARGE=0.035 END
GROUP
ATOM  O3' TYPE OX17 CHARGE=−0.36 END

!HNF−base
GROUP
ATOM  C2  TYPE CX19 CHARGE 0.449 END
ATOM  C3  TYPE CX20 CHARGE −0.372 END
ATOM  C4  TYPE CX21 CHARGE −0.089 END
ATOM  C10 TYPE CX22 CHARGE −0.119 END
ATOM  C11 TYPE CX23 CHARGE 0.225 END

176
1 Force field parameter and topology files

ATOM  C1  TYPE  CX24  CHARge  -0.447  END
ATOM  C13  TYPE  CX25  CHARge  0.078  END
ATOM  C12  TYPE  CX26  CHARge  0.134  END
ATOM  C9  TYPE  CX27  CHARge  -0.191  END
ATOM  C5  TYPE  CX28  CHARge  -0.154  END
ATOM  C6  TYPE  CX29  CHARge  -0.230  END
ATOM  C7  TYPE  CX30  CHARge  0.078  END
ATOM  C8  TYPE  CX31  CHARge  -0.317  END
ATOM  N7  TYPE  NX32  CHARge  0.751  END
ATOM  O71  TYPE  OX33  CHARge  -0.453  END
ATOM  O72  TYPE  OX34  CHARge  -0.456  END
ATOM  H4  TYPE  HX35  CHARge  0.147  END
ATOM  H3  TYPE  HX36  CHARge  0.156  END
ATOM  H5  TYPE  HX37  CHARge  0.141  END
ATOM  H6  TYPE  HX38  CHARge  0.181  END
ATOM  H8  TYPE  HX39  CHARge  0.193  END
ATOM  H1  TYPE  HX40  CHARge  0.102  END
ATOM  H91  TYPE  HX41  CHARge  0.105  END
ATOM  H92  TYPE  HX42  CHARge  0.164  END

BOND  P  O1P  BOND  P  O2P  BOND  P  O5'  BOND  O5'  C5'
BOND  C5'  H5'  BOND  C5'  C4'  BOND  C4'  H4'
BOND  C4'  O4'  BOND  C4'  C3'  BOND  C3'  H3'
BOND  C3'  O3'  BOND  C3'  C2'  BOND  C2'  H2'
BOND  C2'  C2  BOND  C2  C1  BOND  C1  H1
BOND  C1  C11  BOND  C11  C9  BOND  C9  H91
BOND  C9  H92  BOND  C5  C6  BOND  C6  H6
BOND  C6  H6  BOND  C7  C8  BOND  C8  H8
BOND  N7  O71

(dihedrals taken out in accord with other bases)

( Note: edit these DHEdrals if necessary )

DHEdral  O2P  P  O5'  C5'  ! flat ? (180 degrees = trans)  172.75
! DHEdral  P  O5'  C5'  H5'  ! flexible dihedral ???  -63.48
! DHEdral  P  O5'  C5'  H5''  ! flexible dihedral ???  55.20
! DHEdral  P  O5'  C5'  C4'  ! flat ? (180 degrees = trans)  176.08
! DHEdral  O5'  C5'  C4'  H4'  ! flat ? (180 degrees = trans)  172.35
! DHEdral  O5'  C5'  C4'  O4'  ! flexible dihedral ???  -69.24
! DHEdral  O5'  C5'  C4'  C3'  ! flexible dihedral ???  52.07
! DHEdral  H5'  C5'  C4'  H4'  ! flexible dihedral ???  51.91
! DHEdral  H5'  C5'  C4'  O4'  ! flat ? (180 degrees = trans)  170.31
! DHEdral  H5'  C5'  C4'  C3'  ! flexible dihedral ???  -68.38
! DHEdral  H5''  C5'  C4'  H4'  ! flexible dihedral ???  -66.18
! DHEdral  H5''  C5'  C4'  O4'  ! flexible dihedral ???  52.23
! DHEdral  H5''  C5'  C4'  C3'  ! flat ? (180 degrees = trans)  173.54
! DHEdral  C5'  C4'  C3'  C2'  ! flexible dihedral ???  -129.37
! DHEdral  C5'  C4'  C3'  O3'  ! flexible dihedral ???  112.32
! DHEdral  H4'  O4'  C3'  C2'  ! flexible dihedral ???  110.44
! DHEdral  H4'  C4'  C3'  O3'  ! flat ? (0 degrees = cis)  -7.87
! DHEdral  O4'  C4'  C3'  H3'  ! flexible dihedral ???  111.59
DHEdral  O4'  C4'  C3'  C2'  ! flat ? (0 degrees = cis)  -5.92

177
! DIHEDRAL O4' C4' C3' O3' ! flexible dihedral ??? −124.23
! DIHEDRAL O2 C1' C2' H2' ! flexible dihedral ??? 80.69
! DIHEDRAL O4' C1' O2 C2' ! flat ? (180 degrees = trans) 177.32
! DIHEDRAL C2' C1' O2 C2' ! flexible dihedral ??? −67.60
! DIHEDRAL H1' ' C1' O2 C2' ! flexible dihedral ??? 56.28
! DIHEDRAL O2 C2' C3' C4' ! flexible dihedral ??? −88.75
! DIHEDRAL H3' C3' C2' C1' ! flexible dihedral ??? −90.38
! DIHEDRAL O3' C3' C2' H2' ! flexible dihedral ??? −93.06

DIHEDRAL O2 C2 C3 C4 ! flat ? (180 degrees = trans) 180.39
DIHEDRAL O2 C2 C3 H3 ! flat ? (0 degrees = cis) −0.12
DIHEDRAL C1 C2 C3 C4 ! flat ? (0 degrees = cis) −0.46
DIHEDRAL O2 C2 C3 H3 ! flat ? (180 degrees = trans) 179.03
DIHEDRAL O2 C2 C1 C11 ! flat ? (180 degrees = trans) 179.62
DIHEDRAL O2 C2 C1 H1 ! flat ? (0 degrees = cis) −0.53
DIHEDRAL C3 C2 C1 C11 ! flat ? (0 degrees = cis) 0.42
DIHEDRAL C3 C2 C1 H1 ! flat ? (180 degrees = trans) 180.27
DIHEDRAL C2 C3 C4 C10 ! flat ? (0 degrees = cis) 0.26
DIHEDRAL C2 C3 C4 H4 ! flat ? (180 degrees = trans) 180.16
DIHEDRAL H3 C3 C4 C10 ! flat ? (180 degrees = trans) 180.76
DIHEDRAL C3 C4 C10 C13 ! flat ? (180 degrees = trans) 179.94
DIHEDRAL H4 C4 C10 C11 ! flat ? (180 degrees = trans) 180.08
DIHEDRAL H4 C4 C10 C13 ! flat ? (0 degrees = cis) 0.03
DIHEDRAL C4 C10 C11 C1 ! flat ? (0 degrees = cis) −0.02
DIHEDRAL C4 C10 C11 C9 ! flat ? (180 degrees = trans) 179.89
DIHEDRAL C13 C10 C11 C1 ! flat ? (180 degrees = trans) 180.02
DIHEDRAL C13 C10 C11 C9 ! flat ? (0 degrees = cis) −0.07
DIHEDRAL C4 C10 C13 C12 ! flat ? (180 degrees = trans) 180.10
DIHEDRAL C4 C10 C13 C5 ! flat ? (0 degrees = cis) 0.02
DIHEDRAL C11 C10 C13 C12 ! flat ? (0 degrees = cis) 0.06
DIHEDRAL C11 C10 C13 C5 ! flat ? (180 degrees = trans) 179.98
DIHEDRAL C10 C11 C1 C2 ! flat ? (0 degrees = cis) −0.18
DIHEDRAL C10 C11 C1 H1 ! flat ? (180 degrees = trans) 179.97
DIHEDRAL C9 C11 C1 C2 ! flat ? (180 degrees = trans) 179.93
DIHEDRAL C9 C11 C1 H1 ! flat ? (0 degrees = cis) 0.08
DIHEDRAL C10 C11 C9 C12 ! flat ? (0 degrees = cis) 0.05
! DIHEDRAL C10 C11 C9 H91 ! flexible dihedral ??? 117.87
! DIHEDRAL C10 C11 C9 H92 ! flexible dihedral ??? −117.81
! DIHEDRAL C1 C11 C9 C12 ! flat ? (180 degrees = trans) 179.95
! DIHEDRAL C1 C11 C9 H91 ! flexible dihedral ??? −62.23
! DIHEDRAL C1 C11 C9 H92 ! flexible dihedral ??? 62.09
DIHEDRAL C10 C13 C12 C9 ! flat ? (0 degrees = cis) −0.03
DIHEDRAL C10 C13 C12 C8 ! flat ? (180 degrees = trans) 179.97
DIHEDRAL C5 C13 C12 C9 ! flat ? (180 degrees = trans) 180.04
DIHEDRAL C5 C13 C12 C8 ! flat ? (0 degrees = cis) 0.04
DIHEDRAL C10 C13 C5 C6 ! flat ? (180 degrees = trans) 180.08
DIHEDRAL C10 C13 C5 H5 ! flat ? (0 degrees = cis) 0.06
DIHEDRAL C12 C13 C5 C6 ! flat ? (0 degrees = cis) −0.01
DIHEDRAL C12 C13 C5 H5 ! flat ? (180 degrees = trans) 179.97
DIHEDRAL C13 C12 C9 C11 ! flat ? (0 degrees = cis) −0.01
! DIHEDRAL C13 C12 C9 H91 ! flexible dihedral ??? −117.79
! DIHEDRAL C8 C12 C9 C11 ! flat ? (180 degrees = trans) 179.99
! DIHEDRAL C8 C12 C9 H91 ! flexible dihedral ??? 62.22
! DIHEDRAL C8 C12 C9 H92 ! flexible dihedral ??? −62.20
.1 Force field parameter and topology files

DHDebral C13 C12 C8 C7 ! flat ? (0 degrees = cis) −0.06
DHDebral C13 C12 C8 H8 ! flat ? (180 degrees = trans) 180.04
DHDebral C9 C12 C8 C7 ! flat ? (180 degrees = trans) 179.94
DHDebral C9 C12 C8 H8 ! flat ? (0 degrees = cis) 0.04
DHDebral C13 C5 C6 C7 ! flat ? (0 degrees = cis) 0.01
DHDebral C13 C5 C6 C8 H8 ! flat ? (180 degrees = trans) 180.05
DHDebral H5 C5 C6 C7 ! flat ? (180 degrees = trans) 180.03
DHDebral H5 C5 C6 H6 ! flat ? (0 degrees = cis) 0.07
DHDebral C5 C6 C7 C8 ! flat ? (0 degrees = cis) −0.03
DHDebral C5 C6 C7 N7 ! flat ? (180 degrees = trans) 180.02
DHDebral H6 C6 C7 C8 ! flat ? (180 degrees = trans) 179.93
DHDebral H6 C6 C7 N7 ! flat ? (0 degrees = cis) −0.02
DHDebral C6 C7 C8 C12 ! flat ? (0 degrees = cis) 0.06
DHDebral C6 C7 C8 H8 ! flat ? (180 degrees = trans) 179.95
DHDebral N7 C7 C8 C12 ! flat ? (180 degrees = trans) 180.01
DHDebral N7 C7 C8 H8 ! flat ? (0 degrees = cis) −0.09
DHDebral C6 C7 N7 OT1 ! flat ? (0 degrees = cis) 0.03
DHDebral C6 C7 N7 OT2 ! flat ? (180 degrees = trans) 179.92
DHDebral C8 C7 N7 OT1 ! flat ? (180 degrees = trans) 180.07
DHDebral C8 C7 N7 OT2 ! flat ? (0 degrees = cis) −0.04

{ Note: edit these IMPRopers if necessary }

!!! IMPRoper P OIP O2P O5' ! chirality or flatness improper −36.22 ! taken out in
accordance with other bases

IMPRoper C5' O5' H5' H5'' ! chirality or flatness improper −34.21
IMPRoper C4' C5' H4' O4' ! chirality or flatness improper −37.05
IMPRoper C1' O4' C2' O2' ! chirality or flatness improper −33.04
IMPRoper C3' C4' H3' C2'' ! chirality or flatness improper 41.55
IMPRoper C2' C1' C3' H2' ! chirality or flatness improper 28.65
IMPRoper C2 O2 C3 C1 ! chirality or flatness improper 0.51
IMPRoper C3 C2 C4 H3 ! chirality or flatness improper −0.27
IMPRoper C4 C3 C10 H4 ! chirality or flatness improper −0.05
IMPRoper C10 C4 C11 C13 ! chirality or flatness improper −0.03
IMPRoper C11 C10 C1 C9 ! chirality or flatness improper −0.06
IMPRoper C1 C2 C11 H1 ! chirality or flatness improper −0.08
IMPRoper C13 C10 C12 C5 ! chirality or flatness improper −0.04
IMPRoper C12 C13 C9 C8 ! chirality or flatness improper 0.00
IMPRoper C9 C11 C12 H9! ! chirality or flatness improper −29.65
IMPRoper C5 C13 C8' H5' ! chirality or flatness improper −0.01
IMPRoper C6 C5 C7 H6' ! chirality or flatness improper 0.02
IMPRoper C7 C6 C8 N7 ! chirality or flatness improper 0.03
IMPRoper C8 C12 C7 H8 ! chirality or flatness improper 0.05
IMPRoper N7 C7 OT1 OT2 ! chirality or flatness improper −0.06
IMPRoper C6 C8 C1 C3 ! chirality or flatness improper −0.01
IMPRoper C5 C12 C11 C4 ! chirality or flatness improper 0.02
IMPRoper O71 N7 C7 C6 ! chirality or flatness improper 0.03
IMPRoper O72 N7 C8 ! chirality or flatness improper 0.05
IMPRoper H6 H8 H1 H3 ! chirality or flatness improper −0.06
IMPRoper C8 C13 C11 C3 ! mod by anda, planarit
IMPRoper C6 C12 C10 C2 ! mod by anda, planarit
IMPRoper C12 C13 C10 C11 ! mod by anda, planarit
IMPRoper C5 C9 C10 C2 ! mod by anda, planarit
IMPRoper C8 C13 C9 C4 ! mod by anda, planarit
IMPRoper N7 C13 C9 C4 ! mod by anda, planarit
IMPRoper C7 C13 C9 C4 ! mod by anda, planarit

179
Script code

IMPRoper C8  C13  C1  C3  ! mod by anda, planarity
IMPRoper H6  C12  C4  H1  ! mod by anda, planarity
IMPRoper H8  C5  C11  H3  ! mod by anda, planarity
IMPRoper C8  C5  C11  C3  ! mod by anda, planarity

! { Note: edit any DONOrs and ACCEptors if necessary }

! ! DONOr H? O1P ! only true if −OHx (x>0)
! ACCEptor O1P P
! ! DONOr H? O2P ! only true if −OHx (x>0)
! ACCEptor O2P P
! ACCEptor O5' P
! ACCEptor O4' C4'
! ! DONOr H? O3' ! only true if −OHx (x>0)
! ACCEptor O3' C3'
! ACCEptor O2 C1'
! ! DONOr H? O71 ! only true if −OHx (x>0)
! ACCEptor O71 N7
! ! DONOr H? O72 ! only true if −OHx (x>0)
! ACCEptor O72 N7

END { RESIdue HNF }

!************************************************ end of mod by anda − HNF ************************************************

RESIdue GUA
GROUp
ATOM P  TYPE=P  CHARGE=1.20  END
ATOM O1P TYPE=O1P CHARGE=−0.40 END
ATOM O2P TYPE=O2P CHARGE=−0.40 END
ATOM O5' TYPE=O5' CHARGE=−0.36 END
GROUp
ATOM C5' TYPE=C5' CHARGE=−0.070 END
ATOM H5' TYPE=H CHARGE=0.035 END !JPR
ATOM H5' ' TYPE=H CHARGE=0.035 END !JPR
GROUp
ATOM C4' TYPE=C4' CHARGE=0.065 END
ATOM H4' TYPE=H CHARGE=0.035 END !JPR
ATOM O4' TYPE=O4' CHARGE=−0.30 END
ATOM C1' TYPE=C1' CHARGE=0.165 END !JPR
ATOM H1' TYPE=H CHARGE=0.035 END !JPR

! Insert Base
GROUp
ATOM N9 TYPE=N9G CHARGE=−0.19 END
ATOM C4 TYPE=C4G CHARGE=0.19 EXCLUSION=( N1 ) END
GROUp
ATOM N3 TYPE=N3G CHARGE=−0.35 EXCLUSION=( C6 ) END
ATOM C2 TYPE=C2G CHARGE=0.35 EXCLUSION=( C5 ) END
GROUp
ATOM N2 TYPE=N2G CHARGE=−0.42 END
ATOM H21 TYPE=H2 CHARGE=0.21 END
ATOM H22 TYPE=H2 CHARGE=0.21 END
GROUp
ATOM N1 TYPE=NNA CHARGE=−0.26 END
ATOM H1 TYPE=HN CHARGE=0.26 END

180
GROUP

ATOM C6  TYPE=C6G  CHARGE=0.30  END
ATOM O6  TYPE=O6G  CHARGE=-0.30  END
GROUP

ATOM C5  TYPE=C5G  CHARGE=0.02  END
ATOM N7  TYPE=N7G  CHARGE=-0.25  END
ATOM C8  TYPE=C8G  CHARGE=-0.145  END
ATOM H8  TYPE=H  CHARGE=0.035  END

!

GROUP

ATOM C2’  TYPE=C2R  CHARGE=0.115  END
ATOM H2’  TYPE=H  CHARGE=0.035  END
ATOM O2’  TYPE=O2R  CHARGE=-0.40  END
ATOM HO2’  TYPE=HO  CHARGE=0.25  END

GROUP

ATOM C3’  TYPE=C3R  CHARGE=-0.035  END
ATOM H3’  TYPE=H  CHARGE=0.035  END
GROUP

ATOM O3’  TYPE=O3R  CHARGE=-0.36  END

BOND P  O1P  BOND P  O2P  BOND P  O5’
BOND O5’  C5’  BOND C5’  C4’  BOND C4’  O4’
BOND C4’  C3’  BOND O4’  C1’  BOND C1’  N9
BOND C1’  C2’  BOND N9  C4  BOND N9  C8
BOND C4  N3  BOND C4  C5  BOND N3  C2
BOND C2  N2  BOND C2  N1  BOND N2  H21
BOND N2  H22  BOND N1  H1  BOND N1  C6
BOND C6  O6  BOND C6  C5  BOND C5  N7
BOND N7  C8  BOND C2’  C3’  BOND C3’  O3’
BOND C2’  O2’  BOND C8  H8
BOND O2’  HO2’
BOND C5’  H5’  BOND C5’  H5  BOND C4’  H4’
BOND C3’  H3’  BOND C2’  H2’  BOND C1’  H1’

{  
DHEdral P  O5’  C5’  C4’  DHEdral O5’  C5’  C4’  O4’
DHEdral O5’  C5’  C4’  C3’
}

{  
DHEdral C3’  C4’  O4’  C1’
DHEdral C4’  O4’  C1’  C2’  DHEdral O4’  C1’  C2’  C3’
DHEdral C1’  C2’  C3’  C4’
DHEdral C5’  C4’  C3’  O2’
DHEdral C4’  O4’  C1’  N9  C4
DHEdral C3’  C2’  O2’  H2’
}

!

IMPRoper N3  C2  N2  H21  IMPRoper C1’  C4  C8  N9
IMPRoper N9  C4  C5  N7  IMPRoper C4  C5  N7  C8
IMPRoper C5  N7  C8  N9  IMPRoper N7  C8  N9  C4
IMPRoper C8  N9  C4  C5  IMPRoper N2  N3  N1  C2

181
Script code

IMPRoper H1  C2  C6  N1  IMPRoper O6  N1  C5  C6
IMPRoper C4  N3  C2  N1  IMPRoper N3  C2  N1  C6
IMPRoper C2  N1  C6  C5  IMPRoper N1  C6  C5  C4
IMPRoper C6  C5  C4  N3  IMPRoper C5  C4  N3  C2
IMPRoper H22 H21 C2  N2
IMPRoper H8  N7  N9  C8

IMPRoper to keep the two purine rings parallel:
IMPRoper C8  C4  C5  N1  IMPRoper C8  C5  C4  C2
IMPRoper N3  C4  C5  N7  IMPRoper C6  C5  C4  N9

RIBOSE IMPROPER
IMPRoper H1’ C2’ O4’ N9  ’C1’
IMPRoper H2’ C3’ C1’ O2’ ’C2’
IMPRoper H3’ C4’ C2’ O3’ ’C3’
IMPRoper H4’ C5’ C3’ O4’ ’C4’
IMPRoper H5’ O5’ H5’ ’C4’ ’C5’

END {GUA}

! Insert Base

RESIdue ADE

GROUP
ATOM P TYPE=P CHARGE=1.20 END
ATOM O1P TYPE=O1P CHARGE=−0.40 END
ATOM O2P TYPE=O2P CHARGE=−0.40 END
ATOM O5’ TYPE=OSR CHARGE=−0.36 END
GROUP
ATOM C5’ TYPE=C5R CHARGE=−0.070 END
ATOM H5’ TYPE=H CHARGE=0.035 END
ATOM H5’ ’ TYPE=H CHARGE=0.035 END
GROUP
ATOM C4’ TYPE=C4R CHARGE=0.065 END
ATOM H4’ TYPE=H CHARGE=0.035 END
ATOM O4’ TYPE=O4R CHARGE=−0.30 END
ATOM C1’ TYPE=C1R CHARGE=0.165 END
ATOM H1’ TYPE=H CHARGE=0.035 END

1 Insert Base

GROUP
ATOM N9 TYPE=N9A CHARGE=−0.19 END
ATOM C4 TYPE=C4A CHARGE=0.19 EXCLUSION=( C6 ) END
GROUP
ATOM N3 TYPE=N3A CHARGE=−0.26 EXCLUSION=( C6 ) END
ATOM C2 TYPE=C2A CHARGE=0.225 EXCLUSION=( C5 ) END
ATOM H2 TYPE=H CHARGE=0.035 END
GROUP
ATOM N1 TYPE=NC CHARGE=−0.28 END
ATOM C6 TYPE=C6A CHARGE=0.28 END
GROUP
ATOM N6 TYPE=N6A CHARGE=−0.42 END
.1 Force field parameter and topology files

! Force field parameter and topology files

ATOM H61 TYPE=H2 CHARGE=0.21 END
ATOM H62 TYPE=H2 CHARGE=0.21 END

GROUP
ATOM C5 TYPE=C5A CHARGE=0.02 END
ATOM N7 TYPE=N7A CHARGE=-0.25 END
ATOM C8 TYPE=C8A CHARGE=0.195 END
ATOM H8 TYPE=H CHARGE=0.035 END

GROUP
ATOM C2 TYPE=C2R CHARGE=0.115 END
ATOM H2 TYPE=H CHARGE=0.035 END
ATOM O2 TYPE=O2R CHARGE=-0.40 END
ATOM HO2 TYPE=HO CHARGE=0.25 END

GROUP
ATOM C3 TYPE=C3R CHARGE=-0.035 END
ATOM H3 TYPE=H CHARGE=0.035 END

GROUP
ATOM O3 TYPE=O3R CHARGE=-0.36 END

BOND P O1P BOND P O2P BOND P O5'
BOND O5' C5' BOND C5' C4' BOND C4' O4'
BOND C4' C3' BOND C3' C1' BOND C1' N9
BOND C1' C2' BOND N9 C4 BOND N9 C8
BOND C4 N3 BOND C4 C5 BOND N3 C2
BOND C2 N1 BOND N1 C6 BOND C6 N6

BOND N6 H61 BOND N6 H62 BOND C6 C5

BOND C5 N7 BOND N7 C8 BOND C2' C3'
BOND C2' O2' BOND C3' O3'
BOND C8 H8 BOND C2 H2
BOND O2' HO2'
BOND C5' H5' BOND C5' H5' BOND C4' H4'
BOND C3' H3' BOND C2' H2' BOND C1' H1'

{ DHHedral P O5' C5' C4' DHHedral O5' C5' C4' O4'
DHHedral C3' C4' O4' C1'
DHHedral C4' O4' C1' C2'
DHHedral O4' C1' C2' C3'
}

{ DHHedral C3' C4' O4' C1'
DHHedral C4' C3' C2'
DHHedral C3' C2' C3' C4'
DHHedral C4' C3' C2' C3'
DHHedral C3' C2' O2' H2'
}

! IMPRoper C5 C6 N6 H61 IMPRoper C1' C4 C8 N9
IMPRoper N9 C4 C5 N7 IMPRoper C4 C5 N7 C8
IMPRoper C5 N7 C8 N9 IMPRoper N7 C8 N9 C4
IMPRoper C8 N9 C4 C5 IMPRoper N6 N1 C5 C6
IMPRoper H62 C6 H61 N6 IMPRoper C4 N3 C2 N1
IMPRoper N3 C2 H1 H2 IMPRoper C2 N1 C6 C5
IMPRoper N1 C6 C5 C4 IMPRoper C6 C5 C4 N3
Script code

IMPRoper C5  C4  N3  C2
IMPRoper H2  N1  N3  C2  IMPRoper H8  N7  N9  C8
! IMPRoper to keep the two purine rings parallel:
IMPRoper C8  C4  C5  N1  IMPRoper C8  C5  C4  C2
IMPRoper N3  C4  C5  N7  IMPRoper C6  C5  C4  N9

! RIBOSE IMPROPERS
IMPRoper  H1’  C2’  O4’  N9  ’!C1’
IMPRoper  H2’  C3’  C1’  O2’  ’!C2’
IMPRoper  H3’  C4’  C2’  O3’  ’!C3’
IMPRoper  H4’  C5’  C3’  O4’  ’!C4’
IMPRoper  H5’  O5’  H5’  C4’  ’!C5’

END {ADE}

-------------
RESIdue PUR
GROUP
ATOM  P  TYPE=P  CHARGE=1.20  END
ATOM  O1P  TYPE=O1P  CHARGE=−0.40  END
ATOM  O2P  TYPE=O2P  CHARGE=−0.40  END
ATOM  O5’  TYPE=O5R  CHARGE=−0.36  END
GROUP
ATOM  C5’  TYPE=C5R  CHARGE=−0.070  END
ATOM  H5’  TYPE=H  CHARGE=0.035  END
ATOM  H5’’  TYPE=H  CHARGE=0.035  END
GROUP
ATOM  C4’  TYPE=C4R  CHARGE=0.065  END
ATOM  H4’  TYPE=H  CHARGE=0.035  END
ATOM  O4’  TYPE=O4R  CHARGE=−0.30  END
ATOM  C1’  TYPE=C1R  CHARGE=0.165  END
ATOM  H1’  TYPE=H  CHARGE=0.035  END

! Insert Base
GROUP
ATOM  N9  TYPE=N9P  CHARGE=−0.19  END
ATOM  C4  TYPE=C4P  CHARGE=0.19  EXCLUSION=( N1 )  END
GROUP
ATOM  N3  TYPE=N3P  CHARGE=−0.26  EXCLUSION=( C6 )  END
ATOM  C2  TYPE=C2P  CHARGE=0.225  EXCLUSION=( C5 )  END
ATOM  H2  TYPE=H  CHARGE=0.035  END
GROUP
ATOM  N1  TYPE=NC  CHARGE=−0.28  END
ATOM  C6  TYPE=C6P  CHARGE=0.28  END
ATOM  H6  TYPE=H  CHARGE=0.035  END
GROUP
ATOM  C5  TYPE=C5P  CHARGE=0.02  END
ATOM  N7  TYPE=N7P  CHARGE=−0.25  END
ATOM  C8  TYPE=C8P  CHARGE=0.195  END
ATOM  H8  TYPE=H  CHARGE=0.035  END
! END
.1 Force field parameter and topology files

GROUP
ATOM C2' TYPE=C2R  CHARGE=0.115  END
ATOM H2' TYPE=H    CHARGE=0.035  END
ATOM O2' TYPE=O2R  CHARGE=−0.40  END
ATOM HO2' TYPE=HO  CHARGE=−0.25  END
GROUP
ATOM C3' TYPE=C3R  CHARGE=−0.035 END
ATOM H3' TYPE=H    CHARGE=0.035  END
GROUP
ATOM O3' TYPE=O3R  CHARGE=−0.36  END

BOND P  O1P  BOND P  O2P  BOND P  O5'
BOND O5' C5'  BOND C5' C4'  BOND C4' O4'
BOND C4' C3'  BOND O4' C1'  BOND C1' N9
BOND C1' C2'  BOND N9 C4  BOND N9 C8
BOND C4 N3  BOND C4 C5  BOND N3 C2
BOND C2 N1  BOND N1 C6  BOND C6 H6

BOND C6 C5
BOND C5 N7  BOND N7 C8  BOND C2' C3'
BOND C2' O2'  BOND C3' O3'
BOND C8 H8  BOND C2 H2
BOND O2' HO2'
BOND C5' H5'  BOND C5' H5'  BOND C4' H4'
BOND C3' H3'  BOND C2' H2'  BOND C1' H1'

{  
DIHEDral P  O5' C5' C4'  
DIHEDral O5' C5' C4' C3'  
}

{  
DIHEDral C3' C4' O4' C1'  
DIHEDral C4' O4' C1' C2'  
DIHEDral C1' C2' C3' C4'  
DIHEDral O4' C1' N9 C4  
DIHEDral O4' C1' C2' O2' H2'  
}

!
IMPRoper H6 N1 C5 C6  IMPRoper C1' C4 C8 N9
IMPRoper N9 C4 C5 N7  IMPRoper C4 C5 N7 C8
IMPRoper C5 N7 C8 N9  IMPRoper N7 C8 N9 C4
IMPRoper C8 N9 C4 C5  IMPRoper N6 N1 C5 C6
IMPRoper C4 N3 C2 N1
IMPRoper N3 C2 N1 C6  IMPRoper C2 N1 C6 C5
IMPRoper N1 C6 C5 C4  IMPRoper C6 C5 C4 N3
IMPRoper C5 C4 N3 C2
IMPRoper H2 N1 N3 C2  IMPRoper H8 N7 N9 C8
!
IMPRoper to keep the two purine rings parallel:
IMPRoper C8 C4 C5 N1  IMPRoper C8 C5 C4 C2
IMPRoper N3 C4 C5 N7  IMPRoper C6 C5 C4 N9

RIBOSE IMPROPERs

IMPRoper C2' C3' C1' O2'
IMPProper  H1'  C2'  O4'  N9  'C1'  
IMPProper  H2'  C3'  C1'  O2'  'C2'  
IMPProper  H3'  C4'  C2'  O3'  'C3'  
IMPProper  H4'  C5'  C3'  O4'  'C4'  
IMPProper  H5'  O5'  H5'  'C4'  'C5'  

END {PUR}

! __________________________________________________________
RESIdue  ABA
GROUP
ATOM P  TYPE=P  CHARGE=1.20  END
ATOM O1P  TYPE=O1P  CHARGE=−0.40  END
ATOM O2P  TYPE=O2P  CHARGE=−0.40  END
ATOM O5'  TYPE=O5R  CHARGE=−0.36  END
GROUP
ATOM C5'  TYPE=C5R  CHARGE=−0.070  END
ATOM H5'  TYPE=H  CHARGE=0.035  END
ATOM H5'  '  TYPE=H  CHARGE=0.035  END
GROUP
ATOM C4'  TYPE=C4R  CHARGE=0.065  END
ATOM H4'  TYPE=H  CHARGE=0.035  END
ATOM O4'  TYPE=O4R  CHARGE=−0.30  END
ATOM C1'  TYPE=C1R  CHARGE=0.165  END
ATOM H1'  TYPE=H  CHARGE=0.018  END
ATOM H1'  '  TYPE=H  CHARGE=0.017  END
GROUP
ATOM C2'  TYPE=C2R  CHARGE=0.115  END
ATOM H2'  TYPE=H  CHARGE=0.035  END
ATOM O2'  TYPE=O2R  CHARGE=−0.40  END
ATOM HO2'  TYPE=HO  CHARGE=0.25  END
GROUP
ATOM C3'  TYPE=C3R  CHARGE=−0.035  END
ATOM H3'  TYPE=H  CHARGE=0.035  END
GROUP
ATOM O3'  TYPE=O3R  CHARGE=−0.36  END
BOND P  O1P  BOND P  O2P  BOND P  O5'
BOND O5'  C5'  BOND C5'  C4'  BOND C4'  O4'
BOND C4'  C3'  BOND O4'  C1'
BOND C1'  C2'  BOND C2'  C3'
BOND C3'  O3'  BOND C2'  O2'
BOND O2'  HO2'  BOND C1'  H1'  
BOND C5'  H5'  BOND C5'  H5'  '  BOND C4'  H4'
BOND C3'  H3'  BOND C2'  H2'  BOND C1'  H1'  

{  
DHEedral P  O5'  C5'  C4'  DHEedral O5'  C5'  C4'  O4'
DHEedral O5'  C5'  C4'  C3'  
}  

{  
DHEedral C3'  C4'  O4'  C1'  
DHEedral C4'  O4'  C1'  C2'  DHEedral O4'  C1'  C2'  C3'  
}  

186
.1 Force field parameter and topology files

DIHeDrAL C1' C2' C3' C4' DIHeDrAL O4' C4' C3' O3'
DIHeDrAL C5' C4' C3' C2' DIHeDrAL O2' C2' C3' O3'
DIHeDrAL O4' C1' H1' ' C2
DIHeDrAL C3' C2' O2' H2'

! New dihedrals
DIHeDrAL C5' C4' C3' O3' DIHeDrAL C4' O4' C1' H1'

! RIBOSE IMPROPERS
! IMPRoper H1' C2' O4' H1' ' C1' ! mod by anda
IMPRoper H2' C3' C1' O2' ! C2'
IMPRoper H3' C4' C2' O1' ! C5'
IMPRoper H4' C5' C3' O4' ! C4'
IMPRoper H5' O5' H5' ' C4' ! C5'

END (ABA)

! Residue CYT
GROUP
ATOM P  TYPE=P  CHARGE=1.20  END
ATOM O1P  TYPE=O1P  CHARGE=−0.40  END
ATOM O2P  TYPE=O2P  CHARGE=−0.40  END
ATOM O5'  TYPE=O5R  CHARGE=−0.36  END
GROUP
ATOM C5'  TYPE=C5R  CHARGE=−0.070  END
ATOM H5'  TYPE=H  CHARGE=0.035  END
ATOM H5' '  TYPE=H  CHARGE=0.035  END
GROUP
ATOM C4'  TYPE=C4R  CHARGE=0.065  END
ATOM H4'  TYPE=H  CHARGE=0.035  END
ATOM O4'  TYPE=O4R  CHARGE=−0.30  END
ATOM C1'  TYPE=C1R  CHARGE=0.165  END
ATOM H1'  TYPE=H  CHARGE=0.035  END

! Insert Base
GROUP
ATOM N1  TYPE=N1C  CHARGE=−0.19  EXCLUSION=( C4 )  END
ATOM C6  TYPE=C6C  CHARGE=0.155  EXCLUSION=( N3 )  END
ATOM H6  TYPE=H  CHARGE=0.035  END
GROUP
ATOM C2  TYPE=C2C  CHARGE=0.30  EXCLUSION=( C5 )  END
ATOM O2  TYPE=ON  CHARGE=−0.30  END
GROUP
ATOM N3  TYPE=NC  CHARGE=−0.28  END
ATOM C4  TYPE=C4C  CHARGE=0.28  END
GROUP
ATOM N4  TYPE=N4C  CHARGE=−0.42  END
ATOM H41  TYPE=H2  CHARGE=0.21  END
Script code

ATOM H42  TYPE=H2  CHARGE=0.21  END
GROUP
ATOM C5  TYPE=C5C  CHARGE=-0.035  END  %CHRG
ATOM H5  TYPE=H  CHARGE=0.035  END
GROUP

! END

GROUP
ATOM C2'  TYPE=C2R  CHARGE=0.115  END
ATOM H2'  TYPE=H  CHARGE=0.035  END
ATOM O2'  TYPE=O2R  CHARGE=-0.40  END
ATOM HO2' TYPE=HO  CHARGE=0.25  END
GROUP
ATOM C3'  TYPE=C3R  CHARGE=-0.035  END
ATOM H3'  TYPE=H  CHARGE=0.035  END
GROUP
ATOM O3'  TYPE=O3R  CHARGE=-0.36  END

BOND P  O1P  BOND P  O2P  BOND P  O5'
BOND O5' C5'  BOND C5' C4'  BOND C4' O4'
BOND C4' C3'  BOND O4' C1'  BOND C1' N1
BOND C1' C2'  BOND N1 C2  BOND N1 C6
BOND C2' N3  BOND N3 C4
BOND C4' N4  BOND N4 H41  BOND N4 H42
BOND C2' O2
BOND C4' C5  BOND C5 C6  BOND C2' C3'
BOND C3' O3'  BOND C2' O2'
BOND C6' H6  BOND C5 H5
BOND O2' HO2'
BOND C5' H5'  BOND C5' H5'  BOND C4' H4'
BOND C3' H3'  BOND C2' H2'  BOND C1' H1'

{ } DHEDRAL P  O5' C5' C4' DHEDRAL O5' C5' C4' O4'
DHEDRAL O5' C5' C4' C3'

} { DHEDRAL C3' C4' O4' C1' DHEDRAL C4' O4' C1' C2' C3'
DHEDRAL C1' C2' C3' C4' DHEDRAL C4' C1' C2' C3'
DHEDRAL C5' C4' C3' C2' DHEDRAL C4' C1' N1 C2
DHEDRAL C3' C2' O2' H2'

! New dihedrals
DHEDRAL C5' C4' C3' O3' DHEDRAL C4' O4' C1' N1

}

IMPRoper C5 C4 N4 H41  IMPRoper C1' C2 C6 N1
IMPRoper O2 N1 N3 C2  IMPRoper N4 N3 C5 C4
IMPRoper N1 C2 N3 C4  IMPRoper C2 N3 C4 C5
IMPRoper N3 C4 C5 C6  IMPRoper C4 C5 C6 N1
IMPRoper C5 O6 N1 C2  IMPRoper C6 N1 C2 N3
IMPRoper H42 C4 H41 N4
### Force field parameter and topology files

**!RIBOSE IMPROPERs**

- IMPRoper H1' C2' O4' N1 ' C1'
- IMPRoper H2' C3' C1' O2' ' C2'
- IMPRoper H3' C4' C2' O3' ' C3'
- IMPRoper H4' C5' C3' O4' ' C4'
- IMPRoper H5' O5' H5' ' C4' ' C5'

END (CYT)

---

**RESidue THY**

GROUP
- ATOM P TYPE=P CHARGE=1.20 END
- ATOM O1P TYPE=O1P CHARGE=−0.40 END
- ATOM O2P TYPE=O2P CHARGE=−0.40 END
- ATOM O5' TYPE=O5R CHARGE=−0.36 END

GROUP
- ATOM C5' TYPE=C5R CHARGE=−0.070 END
- ATOM H5' TYPE=H CHARGE=0.035 END
- ATOM H5 '' TYPE=H CHARGE=0.035 END

GROUP
- ATOM C4' TYPE=C4R CHARGE=0.065 END
- ATOM C2 TYPE=C2T CHARGE=0.35 EXCLUSION=( C5 ) END
- ATOM C4 TYPE=C4T CHARGE=−0.30 END
- ATOM C5 TYPE=C5T CHARGE=−0.035 END
- ATOM C7 TYPE=C7E CHARGE=−0.070 END ! name per IUPAC-IUB recomm.

GROUP
- ATOM H71 TYPE=H CHARGE=0.035 END ! name per IUPAC-IUB recomm.
- ATOM H72 TYPE=H CHARGE=0.035 END ! name per IUPAC-IUB recomm.
- ATOM H73 TYPE=H CHARGE=0.035 END ! name per IUPAC-IUB recomm.

GROUP
- ATOM N1 TYPE=N1T CHARGE=−0.19 EXCLUSION=( C4 ) END
- ATOM C6 TYPE=C6T CHARGE=0.155 EXCLUSION=( N3 ) END
- ATOM H6 TYPE=H CHARGE=0.035 END

GROUP
- ATOM C2 TYPE=C2T CHARGE=−0.35 EXCLUSION=( C5 ) END
- ATOM O2 TYPE=ON CHARGE=−0.35 END

GROUP
- ATOM C3 TYPE=C3T CHARGE=−0.26 END
- ATOM H3 TYPE=HN CHARGE=0.26 END

GROUP
- ATOM C4 TYPE=C4T CHARGE=0.30 END
- ATOM O4 TYPE=ON CHARGE=−0.30 END

GROUP
- ATOM C5 TYPE=C5T CHARGE=−0.035 END
- ATOM C7 TYPE=C7E CHARGE=−0.070 END ! name per IUPAC-IUB recomm.
- ATOM H71 TYPE=H CHARGE=0.035 END ! name per IUPAC-IUB recomm.
- ATOM H72 TYPE=H CHARGE=0.035 END ! name per IUPAC-IUB recomm.
- ATOM H73 TYPE=H CHARGE=0.035 END ! name per IUPAC-IUB recomm.

GROUP
- ATOM N1 TYPE=N1T CHARGE=−0.19 EXCLUSION=( C4 ) END
- ATOM C6 TYPE=C6T CHARGE=0.155 EXCLUSION=( N3 ) END
- ATOM H6 TYPE=H CHARGE=0.035 END

GROUP
- ATOM C2 TYPE=C2T CHARGE=−0.35 EXCLUSION=( C5 ) END
- ATOM O2 TYPE=ON CHARGE=−0.35 END

GROUP
- ATOM C3 TYPE=C3T CHARGE=−0.26 END
- ATOM H3 TYPE=HN CHARGE=0.26 END

GROUP
- ATOM C4 TYPE=C4T CHARGE=0.30 END
- ATOM O4 TYPE=ON CHARGE=−0.30 END

GROUP
- ATOM C5 TYPE=C5T CHARGE=−0.035 END
- ATOM C7 TYPE=C7E CHARGE=−0.070 END ! name per IUPAC-IUB recomm.
- ATOM H71 TYPE=H CHARGE=0.035 END ! name per IUPAC-IUB recomm.
- ATOM H72 TYPE=H CHARGE=0.035 END ! name per IUPAC-IUB recomm.
- ATOM H73 TYPE=H CHARGE=0.035 END ! name per IUPAC-IUB recomm.

GROUP
GROUP
ATOM C2' TYPE=C2R CHARGE=0.115 END
ATOM H2' TYPE=H CHARGE=0.035 END
ATOM O2' TYPE=O2R CHARGE=-0.40 END
ATOM HO2' TYPE=HO CHARGE=0.25 END
GROUP
ATOM C3' TYPE=C3R CHARGE=-0.035 END
ATOM H3' TYPE=H CHARGE=0.035 END
GROUP
ATOM O3' TYPE=O3R CHARGE=-0.36 END

BOND P O1P BOND P O2P BOND P O5'
BOND O5' C5' BOND O5' C4' BOND C5' O4'
BOND C4' C3' BOND C4' C1' BOND C1' N1
BOND C1' C2' BOND C1' C2 BOND N1 C6
BOND C2 C2 O2 BOND C2 N3 BOND N3 H3
BOND N3 C4 BOND C4 O4 BOND C4 C5
BOND C5 C7 BOND C5 C6 BOND C2' C3'
BOND C3' O3' BOND C2' O2'
BOND O2' HO2'
BOND C5' H5' BOND O5' C5' H5'
BOND C3' H3' BOND C2' H2' BOND C1' H1'
BOND C4' H4' BOND C7 H71 BOND C7 H72
BOND C7 H73 BOND C6 H6

{
  DIHedral P O5' C5' C4' DIHedral O5' C5' C4' O4'
  DIHedral O5' C5' C4' C3'
}

{ DIHedral C3' C4' O4' C1'
  DIHedral C4' O4' C1' C2'
  DIHedral C1' C2' C3' C4'
  DIHedral C5' C4' C3' C2'
  DIHedral O4' C1' N1 C2
  DIHedral C3' C2' O2' H2'

! New dihedrals
  DIHedral C5' C4' C3' O3'
  DIHedral C4' O4' C1' N1
}

IMPRoper O4 N3 C5 C4 IMPRoper C1' C2 C6 N1
IMPRoper O2 N1 N3 C2 IMPRoper C4 C5 C6 N1
IMPRoper N1 C2 N3 C4 IMPRoper C2 N3 C4 C5
IMPRoper N3 C4 C5 C6
IMPRoper C5 O4 N1 C2 IMPRoper C6 N1 C2 N3
IMPRoper H3 C2 C4 N3
IMPRoper C7 C4 C6 C5 IMPRoper H6 N1 C5 C6
.1 Force field parameter and topology files

IMProper H3' C4' C2' O3' ! C3'
IMProper H4' C5' C3' O4' ! C4'
IMProper H5' O5' H5' ' C4' ' C5'

END (THY)

RESIdue URI
GROUP
ATOM P TYPE =P CHARGE= 1.20 END
ATOM O1P TYPE =O1P CHARGE= -0.40 END
ATOM O2P TYPE =O2P CHARGE= -0.40 END
ATOM O5' TYPE =O5R CHARGE= -0.36 END
GROUP
ATOM C5' TYPE =C5R CHARGE= -0.070 END
ATOM H5' TYPE =H CHARGE= 0.035 END
ATOM H5' ' TYPE =H CHARGE= 0.035 END
GROUP
ATOM C4' TYPE =C4R CHARGE= 0.065 END
ATOM H4' TYPE =H CHARGE= 0.035 END
ATOM O4' TYPE =O4R CHARGE= -0.30 END
ATOM C1' TYPE =C1R CHARGE= 0.165 END
ATOM H1' TYPE =H CHARGE= 0.035 END

GROUP
ATOM N1 TYPE =N1U CHARGE= -0.19 EXCLUSION=( C4 ) END
ATOM C6 TYPE =C6U CHARGE= 0.155 EXCLUSION=( N3 ) END
ATOM H6 TYPE =H CHARGE= 0.035 END
GROUP
ATOM C2' TYPE =C2U CHARGE= 0.30 EXCLUSION=( C5 ) END
ATOM O2 TYPE =ON CHARGE= -0.30 END
GROUP
ATOM N3 TYPE =N3U CHARGE= -0.28 END
ATOM H3 TYPE =HN CHARGE= 0.26 END
GROUP
ATOM C4 TYPE =C4U CHARGE= 0.24 END
ATOM O4 TYPE =ON CHARGE= -0.30 END
GROUP
ATOM C5 TYPE =C5U CHARGE= -0.035 END !JPR
ATOM H5 TYPE =H CHARGE= 0.035 END !JPR

GROUP
ATOM C2' TYPE =C2R CHARGE= 0.115 END
ATOM H2' TYPE =H CHARGE= 0.035 END !
ATOM O2' TYPE =O2R CHARGE= -0.40 END
ATOM HO2 TYPE =HO CHARGE= 0.25 END
GROUP
ATOM C3 TYPE =C3R CHARGE= -0.035 END
ATOM H3' TYPE =H CHARGE= 0.035 END
GROUP
ATOM O3' TYPE =O3R CHARGE= -0.36 END

191
Script code

BOND P O1P  BOND P O2P  BOND P O5'
BOND O5' C5'  BOND C5' C4'  BOND C4' O4'
BOND C4' C3'  BOND O4' C1'  BOND C1' N1
BOND C1' C2'  BOND N1 C2  BOND N1 C6
BOND C2 O2  BOND C2 N3  BOND N3 H3
BOND N3 C4  BOND C4 O4  BOND C4 C5
BOND C5 C6  BOND C2' C3'  BOND C3' O3'
BOND C2' O2'
BOND C5 H5  BOND C6 H6
BOND O2' HO2'
BOND C5' H5'  BOND C5' H5'  BOND C4' H4'
BOND C3' H3'  BOND C2' H2'  BOND C1' H1'
{  DIHEDRAL P O5' C5' C4'  DIHEDRAL O5' C5' C4' O4'
  DIHEDRAL C3' C4' C1'  }  DIHEDRAL C3' C4' C3'
  DIHEDRAL C4' O4' C1' C2'  DIHEDRAL O4' C1' C2' C3'
  DIHEDRAL C1' C2' C3' C4'  DIHEDRAL O4' C4' C3' O3'
  DIHEDRAL C5' C4' C3' C2'  DIHEDRAL O2' C2' C3' O3'
  DIHEDRAL O4' C1' N1 C2
  DIHEDRAL C3' C2' O2' H2'
  DIHEDRAL P O3' C3' C2'  DIHEDRAL P O3' C3' C4'
  \ New dihedrals
  DIHEDRAL C5' C4' C3' O3'  DIHEDRAL C4' O4' C1' N1
}  IMPROPER C1' C2 C6 N1
IMPROPER O2 N1 N3 C2
IMPROPER O4 N3 C5 C4
IMPROPER C2 N3 C4 C5
IMPROPER C4 C5 C6 N1
IMPROPER C6 N1 C2 N3
IMPROPER H5 C4 C6 C5
{GENERAL RIBOSE IMPROVERS
IMPROPER H1' C2' O4' N1 'C1'
IMPROPER H2' C3' C1' O2' 'C2'
IMPROPER H3' C4' C2' O3' 'C3'
IMPROPER H4' C5' C3' O4' 'C4'
IMPROPER H5' O5' H5' 'C4' 'C5'

END {URI}
.1 Force field parameter and topology files

DELETE ATOM HO2' END

GROUP
MODIFY ATOM C2' TYPE=C2D CHARGE=-0.07 END
MODIFY ATOM C5' TYPE=C5D CHARGE=-0.07 END
MODIFY ATOM C4' TYPE=C4D CHARGE=0.065 END
MODIFY ATOM O4' TYPE=O4D CHARGE=-0.30 END
MODIFY ATOM C1' TYPE=C1D CHARGE=0.165 END
MODIFY ATOM C3' TYPE=C3D CHARGE=-0.035 END
ADD ATOM H2' TYPE=H CHARGE=0.035 END
ADD BOND C2' H2' END
ADD ANGLE C1' C2' H2' END
ADD ANGLE C3' C2' H2'' END
ADD IMPProper H2' C3' H2'' C1' C2' chirality term
END [DEOX]

PRESidue 3TER ! 3-terminus (without phosphate)
! should be used as "LAST 3TER HEAD - + END"
GROUP ! i.e. to be patched to the last RNA residue
MODIFY ATOM -C3' TYPE=C3R CHARGE=0.15 END
MODIFY ATOM -O3' TYPE=OH CHARGE=-0.40 END
ADD ATOM -H3T TYPE=HO CHARGE=0.25 END
ADD BOND -O3' -H3T
ADD ANGLE -C3' -O3' -H3T
ADD DHEDral -C4' -C3' -O3' -H3T
END [3TER]

PRESidue 5TER ! 5-terminus (without phosphate)
! should be used as "FIRST 5TER TAIL - + END"
GROUP ! i.e. to be patched to the first RNA residue
ADD ATOM +H5T TYPE=HO CHARGE=0.25 END
ADD ATOM +O5' TYPE=OH CHARGE=-0.40 END
ADD ATOM +C5' TYPE=C5R CHARGE=0.15 END
DELETE ATOM +P END
DELETE ATOM +O1P END
DELETE ATOM +O2P END
ADD DHEDral +H5T +O5' +C5' +C4'
END [5TER]

PRESidue NUC ! patch for nucleic acid backbone
! should be used as "LINK NUC HEAD - + TAIL + + END"
! i.e. it links the previous RNA residue (-) with
! the current one (+)
Script code

GROUP

MODIFY ATOM −O3’ END !
MODIFY ATOM +P END !
MODIFY ATOM +O1P END ! this should correctly define the electrostatic
group boundary

MODIFY ATOM +O2P END !
ADD BOND −O3’ +P
ADD ANGLE −C3’ −O3’ +P
ADD ANGLE −O3’ +P +O1P
ADD ANGLE −O3’ +P +O2P
ADD ANGLE −O3’ +P +O5’
ADD D1HEDRAL −O3’ +P +O5’ +C5’

! ADD D1HEDRAL −C4’ −C3’ −O3’ +P
! ADD D1HEDRAL −C3’ −O3’ +P +O5’

END {NUC}

! mod by anda-----------------------------------------------
PRESidue 2AP

GROUP

DELETE ATOM H2 END
ADD ATOM N2 TYPE=N2G CHARGE=−0.91 END
ADD ATOM H21 TYPE=H2 CHARGE=0.39 END
ADD ATOM H22 TYPE=H2 CHARGE=0.38 END

MODIFY ATOM N1 CHARGE=−0.72 END
MODIFY ATOM C2 CHARGE= 0.98 END
MODIFY ATOM N3 CHARGE=−0.80 END
MODIFY ATOM C4 CHARGE= 0.60 END
MODIFY ATOM C5 CHARGE= 0.08 END
MODIFY ATOM C6 CHARGE= 0.20 END
MODIFY ATOM N7 CHARGE=−0.45 END ! changed from −0.61 to −0.45 for neutrality
MODIFY ATOM C8 CHARGE= 0.27 END
MODIFY ATOM N9 CHARGE=−0.25 END

! MODIFY ATOM N2 CHARGE=−0.91 END
! MODIFY ATOM H21 CHARGE= 0.39 END
! MODIFY ATOM H22 CHARGE= 0.38 END
MODIFY ATOM H6 CHARGE= 0.12 END
MODIFY ATOM H8 CHARGE= 0.11 END

ADD BOND C2 N2
ADD BOND N2 H21
ADD BOND N2 H22

ADD ANGLE N1 C2 N2
ADD ANGLE C2 N2 H21
ADD ANGLE C2 N2 H22
ADD ANGLE H21 N2 H22
ADD ANGLE N3 C2 N2

194
1 Force field parameter and topology files

ADD IMPRoper N3 C2 N2 H21
ADD IMPRoper N2 N3 N1 C2
ADD IMPRoper H22 H21 C2 N2
END {2AP}

! mod by anda----------------------------------

set echo=true end
Script code

2 Xplor-NIH calculation input files

Since all calculations were carried out under the same conditions, only the input files for the 13mer2AP calculation is shown. The input scripts are based on the example files of the Xplor-NIH package (refine_full.py and sa.inp) but were substantially modified in the course of this work. The first script was used to calculate the starting structure for the second script.

```plaintext
language

remarks file nmr/sa.inp
remarks Simulated annealing protocol for NMR structure determination.
remarks The starting structure for this protocol can be any structure with
remarks a reasonable geometry, such as randomly assigned torsion angles or
remarks extended strands.
remarks Author: Michael Nilges

{==== >}
evaluate ($init_t = 3000) { /* Initial simulated annealing temperature */}
{==== >}
evaluate ($high_steps= 48000) { /* Total number of steps at high temp. */}
{==== >}
evaluate ($cool_steps = 6000) { /* Total number of steps during cooling */}

parameter { /* Read the parameter file */}

{==== >}
@13mer_2AP_c_dna.param
end

{==== >}
structure @13mer_2AP_c_dna.psf end { /* Read the structure file */}
{==== >}
coordinates @13mer_2AP_c_dna.pdb { /* Read the coordinates */}

noe

{==== >}
nres=3000 { /* Estimate greater than the actual number of NOEs */}
class all
{==== >}
@13mer2AP_xplor_all_3.tbl { /* Read NOE distance ranges */}
@hbond_13mer_2AP.tbl
end

{==== >}
restraints dihedral
nass = 1000
@dihedral_13mer_2AP_BDNA.tbl { /* Read dihedral angle restraints */}
end
@plane_13mer_2AP.inp

{ /* Reduce the scaling factor on the force applied to disulfide */}
```
.2 Xplor-NIH calculation input files

(* bonds and angles from 1000.0 to 100.0 in order to reduce computation instability. *)

parameter
  bonds [ name SG ] ( name SG ) 100. TOKEN
  angle [ name CB ] ( name SG ) ( name SG ) 50. TOKEN
end

flags exclude * include bonds angle impr vdw elec noe cdih plan end

(* Friction coefficient for MD heatbath, in 1/ps. *)
vector do (fbeta=10) (all)
  (*Uniform heavy masses to speed molecular dynamics.*)
vector do (mass=100) (all)

noe (*Parameters for NOE effective energy term.*)
  ceiling=1000
  averaging * cent
  potential * soft
  scale * 50.
  sqoffset * 0.0
  sqconstant * 1.0
  sqexponent * 2
  soexponent * 1
  asymptote * 0.1
  rsqswitch * 0.5
end

parameter (*Parameters for the repulsive energy term.*)
  nbonds
    repel=1. (*Initial value for repel—modified later.*)
    rexp=2 irexp=2 rcon=1.
    nbxmod=3
    wmin=0.01
    cutnb=4.5 ctonnb=2.99 ctofnb=3.
    tolerance=0.5
end

restraints dihedral
  scale=5.
end

(* Loop through a family of 100 structures.*)
evaluate ($end_count=100)

coor copy end
evaluate ($count = 0)
evaluate ($count2 = 0)
while ($count < $end_count ) loop main
  evaluate ($count=$count+1)
  evaluate ($count2=$count2+1)
Script code

```plaintext
coor swap end
coor copy end

{ Initial minimization. }

restrictions dihedral scale=5. end
noe asymptote = 0.1 end
parameter nbonds repel=1. end end
constraints interaction
    (all) (all) weights * 1 vdw 0.002 end end
minimize powell nstep=50 drop=10. nprint=25 end

{ High-temperature dynamics. }

constraints interaction (all) (all)
    weights * 1 angl 0.4 impr 0.1 vdw 0.002 end end

evaluate ($nstep1=int($high_steps * 2. / 3. ))
evaluate ($nstep2=int($high_steps * 1. / 3. ))

dynamics verlet
    nstep=$nstep1 timestep=0.003 iavel=maxwell firstt=$init_t
tcoupling=true tbath=$init_t nprint=50 ipfrq=0
end

{Tilt the asymptote and increase weights on geometry.}
noe asymptote = 1.0 end

constraints interaction
    (all) (all) weights * 1 vdw 0.002 end end

{ Bring scaling factor for S-S bonds back }

parameter
    bonds ( name SG ) ( name SG ) 1000. TOKEN
    angle ( name CB ) ( name SG ) ( name SG ) 500. TOKEN
end

dynamics verlet
    nstep=$nstep2 timestep=0.001 iavel=current tcoupling=true
    tbath=$init_t nprint=50 ipfrq=0
end

{ Cool the system. }

restrictions dihedral scale=200. end

evaluate ($final_t = 25) { K }
evaluate ($tempstep = 25) { K }
evaluate ($ncycle = ($init_t-$final_t)/$tempstep)
evaluate ($nstep = int($cool_steps/$ncycle))
evaluate ($ini_rad = 0.9) evaluate ($fin_rad = 0.75)
evaluate ($ini_con= 0.003) evaluate ($fin_con= 4.0)
```

198
\[2\] Xplor-NIH calculation input files

evaluate \( (b_{\text{bath}} = b_{\text{init}\_1}) \)
evaluate \( (k_{\text{vdw}} = k_{\text{ini}\_con}) \)
evaluate \( (k_{\text{vdwfact}} = (f_{\text{fin}\_con}/f_{\text{ini}\_con})^{(1/n_{\text{cycle}})}) \)
evaluate \( (r_{\text{radius}} = r_{\text{ini}\_rad}) \)
evaluate \( (r_{\text{radfact}} = (f_{\text{fin}\_rad}/f_{\text{ini}\_rad})^{(1/n_{\text{cycle}})}) \)
evaluate \( (i_{\text{cool}} = 0) \)
while \((i_{\text{cool}} < n_{\text{cycle}})\) loop cool
  evaluate \((i_{\text{cool}} = i_{\text{cool}} + 1)\)
evaluate \((b_{\text{bath}} = b_{\text{bath}} - t_{\text{tempstep}})\)
evaluate \((k_{\text{vdw}} = \min(f_{\text{fin}\_con}, k_{\text{vdw}} * k_{\text{vdwfact}}))\)
evaluate \((r_{\text{radius}} = \max(f_{\text{fin}\_rad}, r_{\text{radius}} * r_{\text{radfact}}))\)
  parameter nbonds repel=r_{\text{radius}} end end
  constraints interaction (all) (all)
    weights * 1. vdw end end
  dynamics verlet
    nstep=\$nstep time=0.001 iasvel=current firstt=$bath
tcoup=true tbath=$bath nprint=\$nstep iprfrq=0
end

{====>} \{*Abort condition.*\}
evaluate \((c_{\text{critical}} = t_{\text{temp}}/b_{\text{bath}})\)
if \((c_{\text{critical}} > 10.)\) then
display \&\&\&\&\&\&\&\& rerun job with smaller timestep (i.e., 0.003)
stop
end if
end loop cool

{Final minimization.*}

constraints interaction (all) (all) weights * 1. vdw 1. end end

parameter \{Parameters for the repulsive energy term.*\}
nbonds repel=0. \{Initial value for repel—modified later.*\}
SWITCH
VSWITCH
RDIE
cutnb=11.5
nbxmod=5
wmin=0.01
cutfnb=10.5
ctonnb=9.5
tolerance=0.5
end
flags exclude * include bonds angle impr vdw elec noe cdih plan end

minimize powell nstep=3000 drop=10.0 nprint=25 end

{Write out the final structure(s).*}
print threshold=0.5 noe

199
evaluate ($rms_noe=$result)
evaluate ($violations_noe=$violations)
print threshold=0.0 noe
evaluate ($rms_noe22=$result)
print threshold=5. cdih
evaluate ($rms_cdih=$result)
evaluate ($violations_cdih=$violations)
print threshold=0.05 bonds
evaluate ($rms_bonds=$result)
print threshold=5. angles
evaluate ($rms_angles=$result)
evaluate ($rms_impropers=$result)

remarks
remarks energies: $ener, $bond, $angl, $impr, $vdw, $noe, $cdih, $elec
remarks

evaluate ($filename ="z13mer_2AP_c_dna_test"+encode ($count) + ".pdb")
write coordinates output =$filename end
evaluate ($filename2 ="z13mer_2AP_c_dna_test"+encode ($count2) + ".noe")
set display=$filename2 end
@@@@picktbl_13mer2AP_all
close $filename2 end
set display=OUTPUT end

end loop main

stop

seed = 10
numberOfStructures = 100
startStructure = 1

# User-specific which has to be adjusted for each new sample
andaSampleName = "13mer_2AP" # specify sample name
andaNumResAll = 26 # Total number of Residues
andaDihedral = "B" # specify which dihedral constraint set should be used
.2 Xplor-NIH calculation input files

andaDiheBound = 30.0  # specify upper and
lower bound for dihedral constraints
andaDiheBound2 = 30.0  # specify loose upper
and lower bound for dihedral constraints
andaSeqStrand1 = "GUA CYT THY GUA CYT ADE ADE ADE CYT GUA THY CYT GUA"  # Sequence of Strand I
andaSeqStrand2 = "CYT GUA ADE CYT GUA THY THY THY GUA CYT ADE GUA CYT"  # Sequence of Strand II
outFilename = andaSampleName+"_STRUCTURE.pdb"  # pdb output filename
andaNOExpr = "NOE_13mer2AP_xplor_ohneH5_clean tbl"  # file for reading in
experimental NOE constraints
andaDOpoln = "RDC_13mer2AP_safe.xplor"  # file for reading in
experimental RDC constraints
andaDOpolnMe = "RDC_13mer2AP_safe_Me.xplor"  # file for reading in
experimental Me−RDC constraints

# User−specific data, which can be adjusted for external script,
# but which can also be generated from the information given above
andaNumRes = andaNumResAll/2
andaInitCoord = "start_"+andaSampleName+".pdb"  # file created with initial
extended structure coordinates
andaInitPSF = "start_"+andaSampleName+".psf"  # file created with initial
extended structure
andaPlan = "plane_13mer_2AP.inp"  # file for reading in planar
constraints
andaOrie = "dna_positional_anda.setup"  # file for reading in ORIE
constraints
andaHbonds = "hbond_13mer_2AP.tbl"  # file for reading in Hbond
constraints
andaDihe = None  # file for reading in ideal
dihedral constraints
andaRAMA = "nucleic"  # file for reading in ideal
RAMA constraints

# Store each residue in list andaRes!
# (Index+1) corresponds to residue number
andaRes = range(andaNumResAll)
andai = 0
andaii = 3
for andabla in range(andaNumRes):
    andaRes[andaRes[andabla]]=andaSeqStrand2[andai : andaii]
    andai = andai+4
    andaii = andaii+4

if andaDihe==None:
    # Generate dihedral input table with ideal values for A− or B-DNA
    # input values for dihedrals
    alphaA, alphaB = −50.0, −46.0
    betaA, betaB = 172.0, −147.0
    gammaA, gammaB = 41.0, 36.0
    deltaA, deltaB = 79.0, 157.0
    epsA, epsB = −146.0, 155.0
    zetaA, zetaB = −78.0, −96.0
fileHandle = open ( andaSampleName+'_dihe.tbl', 'w' )
#
# B-DNA conformation
#
# if andaDihedral=="B":
# Alpha angle

fileHandle.write ( "! Ideal dihedral values for B-DNA

 Values taken from Roberts: NMR of Macromolecules

Alpha dihedral

"

for andai in range(1, andaNumRes+1):
    fileHandle.write ( "assign ( resid %s and name O3' )
        ( resid %s and name P )
        ( resid %s and name O5' )
        ( resid %s and name C5' ) 1.0 %s %s 2

" % ( andai, andai +1, andai +1, andai +1, alphaB , andaDiheBound ) )

for andai in range(14, andaNumRes+13):
    fileHandle.write ( "assign ( resid %s and name O3' )
        ( resid %s and name P )
        ( resid %s and name O5' )
        ( resid %s and name C5' ) 1.0 %s %s 2

" % ( andai, andai +1, andai +1, andai +1, alphaB , andaDiheBound ) )

# Beta angle

fileHandle.write ( "Beta dihedral

"

for andai in range(2, andaNumRes+1):
    fileHandle.write ( "assign ( resid %s and name P )
        ( resid %s and name O5' )
        ( resid %s and name C5' )
        ( resid %s and name C4' ) 1.0 %s %s 2

" % ( andai, andai , andai , andai , betaB , andaDiheBound ) )

for andai in range(15, andaNumRes+14):
    fileHandle.write ( "assign ( resid %s and name P )
        ( resid %s and name O5' )
        ( resid %s and name C5' )
        ( resid %s and name C4' ) 1.0 %s %s 2

" % ( andai, andai , andai , andai , betaB , andaDiheBound ) )

# Gamma angle

fileHandle.write ( "Gamma dihedral

"

for andai in range(1, andaNumRes+1):
    fileHandle.write ( "assign ( resid %s and name O5' )
        ( resid %s and name C5' )
        ( resid %s and name C4' )
        ( resid %s and name C3' ) 1.0 %s %s 2

" % ( andai, andai , andai , andai , gammaB, andaDiheBound ) )

for andai in range(14, andaNumRes+14):
    fileHandle.write ( "assign ( resid %s and name O5' )
        ( resid %s and name C5' )
        ( resid %s and name C4' )
        ( resid %s and name C3' ) 1.0 %s %s 2

" % ( andai, andai , andai , andai , gammaB, andaDiheBound ) )

# Delta angle

fileHandle.write ( "Delta dihedral

"

for andai in range(1, andaNumRes+1):
    fileHandle.write ( "assign ( resid %s and name C5' )
        ( resid %s and name C4' )
        ( resid %s and name C3' )
        ( resid %s and name O3' ) 1.0 %s %s 2

" % ( andai, andai, andai, andai, deltaB , andaDiheBound ) )

for andai in range(14, andaNumRes+14):
    fileHandle.write ( "assign ( resid %s and name C5' )
        ( resid %s and name C4' )
        ( resid %s and name C3' )
        ( resid %s and name O3' ) 1.0 %s %s 2

" % ( andai, andai, andai, andai, deltaB , andaDiheBound ) )

# Epsilon angle, loose bound

fileHandle.write ( "Epsilon dihedral

"

202
for andai in range(2, andaNumRes):
    fileHandle.write("assign ( resid %s and name C4' )\n    ( resid %s and name C3' )\n    ( resid %s and name O3' )\n    ( resid %s and name P ) 1.0 %s %s 2\n    ( andai, andai, andai, andai+1, epsB, andaDiheBound2 ))
for andai in range(15, andaNumRes+13):
    fileHandle.write("assign ( resid %s and name C4' )\n    ( resid %s and name C3' )\n    ( resid %s and name O3' )\n    ( resid %s and name P  ) 1.0 %s %s 2\n    ( andai, andai, andai, andai+1, epsB, andaDiheBound2 ))

# Zeta angle, loose bound
fileHandle.write("\n! Zeta dihedral\n")
for andai in range(2, andaNumRes):
    fileHandle.write("assign ( resid %s and name C3' )\n    ( resid %s and name O3' )\n    ( resid %s and name P )\n    ( resid %s and name O5' ) 1.0 %s %s 2\n    ( andai, andai, andai, andai+1, zetaB, andaDiheBound2 ))
for andai in range(15, andaNumRes+13):
    fileHandle.write("assign ( resid %s and name C3' )\n    ( resid %s and name O3' )\n    ( resid %s and name P )\n    ( resid %s and name O5' ) 1.0 %s %s 2\n    ( andai, andai, andai, andai+1, zetaB, andaDiheBound2 ))

# A-BDNA conformation
#
eelif andaDiheIdeal=='A':
# Alpha angle
    fileHandle.write("! Ideal dihedral values for A-DNA\nValues taken from Roberts : NMR of Macromolecules\nAlpha dihedral\n")
for andai in range(1, andaNumRes):
    fileHandle.write("assign ( resid %s and name O3' )\n    ( resid %s and name P )\n    ( resid %s and name O5' )\n    ( resid %s and name C5' ) 1.0 %s %s 2\n    ( andai, andai+1, andai+1, andai+1, alphaA, andaDiheBound )")
    for andai in range(14, andaNumRes+13):
        fileHandle.write("assign ( resid %s and name O3' )\n    ( resid %s and name P )\n    ( resid %s and name O5' )\n    ( resid %s and name C5' ) 1.0 %s %s 2\n    ( andai, andai+1, andai+1, andai+1, alphaA, andaDiheBound )")
# Beta angle
    fileHandle.write("\n! Beta dihedral\n")
    for andai in range(2, andaNumRes+1):
        fileHandle.write("assign ( resid %s and name P  )\n    ( resid %s and name O5' )\n    ( resid %s and name C5' )\n    ( resid %s and name C4' ) 1.0 %s %s 2\n    ( andai, andai, andai, andai, betaA, andaDiheBound )")
    for andai in range(15, andaNumRes+14):
        fileHandle.write("assign ( resid %s and name P  )\n    ( resid %s and name O5' )\n    ( resid %s and name C5' )\n    ( resid %s and name C4' ) 1.0 %s %s 2\n    ( andai, andai, andai, andai, betaA, andaDiheBound )")
# Gamma angle
    fileHandle.write("\n! Gamma dihedral\n")
    for andai in range(1, andaNumRes+1):
        fileHandle.write("assign ( resid %s and name O5' )\n    ( resid %s and name C5' )\n    ( resid %s and name C4' )\n    ( resid %s and name O5' ) 1.0 %s %s 2\n    ( andai, andai, andai, andai, gammaA, andaDiheBound )")
    for andai in range(15, andaNumRes+14):
        fileHandle.write("assign ( resid %s and name O5' )\n    ( resid %s and name C5' )\n    ( resid %s and name C4' )\n    ( resid %s and name C5' ) 1.0 %s %s 2\n    ( andai, andai, andai, andai, gammaA, andaDiheBound )")
name C3
( andai, andai, andai, andai, gammaA, andaDiheBound)
for andai in range(14, andaNumRes+14):
  fileHandle.write("assign ( resid %s and name O5' )\n  ( resid %s and name C5' )\n  ( resid %s and name C4' )\n  ( resid %s and name C3' ) 1.0 %s %s 2
( andai, andai, andai, andai, gammaA, andaDiheBound))
# Delta angle
fileHandle.write("\n! Delta dihedral\n
")
for andai in range(1, andaNumRes+1):
  fileHandle.write("assign ( resid %s and name C5' )\n  ( resid %s and name C4' )\n  ( resid %s and name C3' )\n  ( resid %s and name O3' ) 1.0 %s %s 2
( andai, andai, andai, andai, deltaA, andaDiheBound)
for andai in range(14, andaNumRes+14):
  fileHandle.write("assign ( resid %s and name C5' )\n  ( resid %s and name C4' )\n  ( resid %s and name C3' )\n  ( resid %s and name O3' ) 1.0 %s %s 2
( andai, andai, andai, andai, deltaA, andaDiheBound)
# Epsilon angle, loose bound
fileHandle.write("\n! Epsilon dihedral\n
")
for andai in range(2, andaNumRes):
  fileHandle.write("assign ( resid %s and name C4' )\n  ( resid %s and name C3' )\n  ( resid %s and name O3' )\n  ( resid %s and name P ) 1.0 %s %s 2
( andai, andai, andai, andai, epsA, andaDiheBound2)
for andai in range(15, andaNumRes+13):
  fileHandle.write("assign ( resid %s and name C4' )\n  ( resid %s and name C3' )\n  ( resid %s and name O3' )\n  ( resid %s and name P ) 1.0 %s %s 2
( andai, andai, andai, andai, epsA, andaDiheBound2)
# Zeta angle, loose bound
fileHandle.write("\n! Zeta dihedral\n
")
for andai in range(2, andaNumRes):
  fileHandle.write("assign ( resid %s and name C3' )\n  ( resid %s and name O3' )\n  ( resid %s and name P ) 1.0 %s %s 2
( andai, andai, andai, andai, zetaA, andaDiheBound2)
for andai in range(15, andaNumRes+13):
  fileHandle.write("assign ( resid %s and name C3' )\n  ( resid %s and name O3' )\n  ( resid %s and name P ) 1.0 %s %s 2
( andai, andai, andai, andai, zetaA, andaDiheBound2)
fileHandle.close()
andaDihe=andaSampleName+'_dihe.tbl'

==================================================================== end of anda mod
====================================================================
xplor.parseArguments()  # check for typos on the command-line
simWorld.setRandomSeed(seed)
#
# Create the PSF and initial PDB files as an extended structure
# mod by anda
#
import protocol
protocol.initParams('anda_old_nucleic')
protocol.initTopology('anda_old_nucleic')
# import psfGen
# from psfGen import pdbToPSF
# pdbToPSF(andaInitCoord, psfFilename=andaInitPSF, andaPar='anda_old_nucleic', customRename=0)
# from psfGen import seqToPSF
## generates PSF information from Sequence
# seqToPSF(andaSeqStrand1, seqType='dna', customRename=0)
# pass
# seqToPSF(andaSeqStrand2, startResid=andaNumRes+1, seqType='dna', customRename=0)
# pass
## generates an extended structure
# protocol.genExtendedStructure(andaInitCoord, numerator=10, verbose=1, maxFixupIters=2000)
protocol.initStruct(andaInitPSF)

#
# starting coords
#
protocol.initCoords(andaInitCoord)
#
protocol.fixupCovalentGeom(verbos=1)
#
# list of potential terms used in refinement
from potList import PotList
potList = PotList()
crossTerms=PotList('cross terms')  # can add some pot terms which are not refined against — but included in analysis

# parameters to ramp up during the simulated annealing protocol
#
from simulationTools import MultRamp, StaticRamp, InitialParams
rampedParams=[]
highTempParams=[]

from varTensorTools import create_VarTensor, calcTensor
media={}
for medium in ['pf1']:
    media[medium] = create_VarTensor(medium)
    pass

from xplorPot import XplorPot

# planarity restraints
xplor.command('@%s' % andaPlan)
potList.append(XplorPot('plan', xplor.simulation))

# initialize the aa-aa positional database
```python
# Script code

# xplor command('@s % andArie')
# potList.append(XplorPot('orie'))

# NOE potentials
from noePotTools import create_NOEPot
noePots = PotList('noe')
noe = create_NOEPot('noeAll', andAExp)
noe.setPotType('hard')
noePots.append(noe)

# need to be satisfied by all structures
noeHB = create_NOEPot('noeNH', andHBond)
noeHB.setPotType('hard')
noeHB.setScale(1000)
noePots.append(noeHB)
potList.append(noePots)
rampedParams.append(StaticRamp('noePots.setScale(50)'))

protocol.initDihedrals(andADihe)
potList.append(XplorPot('CDIH'))
highTempParams.append(StaticRamp('potList ['CDIH'].setScale(200)'))
rampedParams.append(StaticRamp('potList ['CDIH'].setScale(200)'))
rampedParams.append(MultRamp(10.200, 'potList ['CDIH'].setScale(VALUE)'))

# protocol.initRamaDatabase(andARAMA)
potList.append(XplorPot('rama'))
highTempParams.append(StaticRamp('potList ['RAMA'].setScale(0)'))
rampedParams.append(MultRamp(1.0, 'potList ['RAMA'].setScale(VALUE)'))

from rdcPotTools import Da_prefactor, create_RDCPot, scale_toCH

# from csaPotTools import create_CSAPot
# csaPots = PotList('csa')
# for (name, medium, force, tbl) in [('POP', 'phg3', 1, 'justin_csa.tbl'),
#                                  ('POP2', 'bic2', 1, 'justin_csa_bcl.tbl')]:
#     term = create_CSAPot(name, oTensor=medium[medium], file=tbl)
#     term.setDaScale(-term.DaScale()) # switch sign
#     term.setScale(force)
#     csaPots.append(term)
# pass
#
# potList.append(csaPots)
rampedParams.append(MultRamp(0.01, 0.2, 'csaPots.setScale(VALUE)'))

# Rh same for jch2 and phos/pho2, but da is different
## let's add this later with a cosRatio2 potential term
# rdcPots = PotList('rdcs')
# weight is the relative weighting of expts, as determined by expt.error
for (name, medium, weight, files) in [('JCH', 'pf1', 5, andDipoInp), ('methyl', 'pf1', 0.5, andDipoInpMe)]:
    term = create_RDCPot(name, oTensor=medium[medium], defThreshold=1.9)
    if type(files) == type('string'):
        files = [files]
    term.setScale(weight, files)
rampedParams.append(MultRamp(0.01, 0.2, 'rdcPots.setScale(VALUE)'))
```

206
pass
for file in files:
    term.addRestraints(open(file).read())
pass
term.setShowAllRestraints(1)
term.setScale(weight)
#term.setAveType('average')
term.setAveType('sum')
print name
scale_toCH(term) # also sets useDistance
print term.info()
print term.gyroA()
rdcPots.append(term)
pass

# # proton setup
# for key in ('HABS', 'HAB2'): rdcPots[key].setUseSign(0)
#
# for key in ('HABS', 'HSIG', 'HAB2', 'HSI2'):
#    rdcPots[key].setPotType('square');
#    rdcPots[key].setAveType('sum')
#    pass
potList.append(rdcPots)
rampedParams.append(MultRamp(0.01,1,'rdcPots.setScale(VALUE)'))

from rdcPotTools import Da_prefactor
print 'factor:', Da_prefactor['CH'] / Da_prefactor['NH']
for medium in media.values():
    calcTensor(medium)
    print 'medium: ', medium.instanceName(), \
        'Da: ', medium.Da(), 'Rh: ', medium.Rh()
    pass

# let's try fixing Da, Rh:
print medium
for (medium.Da,Rh) in (('pf1',-41.02,0.321),):
    medium = media[medium]
    medium.setDa(Da)
    medium.setRh(Rh)
    pass

### set up J coupling
# from jCoupPotTools import create_JCoupPot
# jCoup = create_JCoupPot('jcoup','couplings.tbl',
#     A=15.3,B=−6.1,C=1.6,phase=0 )
# jCoup.setThreshold(0)
# jCoup.setScale(10)
# potList.append(jCoup))

# protocol.initNBond(chunk=4.5)
potList.append( XplorPot('VDW') )
Script code

```python
potList.append(XplorPot("elec"))
#rampedParams.append(MultRamp(0.99, 0.78, #xplor.command("param nbonds repel VALUE end end"))
#rampedParams.append(MultRamp(0.001, 4, #xplor.command("param nbonds rcon VALUE end end"))
rampedParams.append(StaticRamp("xplor.command("param nbonds atom repel=0 wmin=0.01 nbxmod=5 cutnb=58.5 ctonnb=56.5 ctofnb=57.5 tolerance=0.5 rdie vswitch switch end end")
#highTempParams.append(StaticRamp("protocol.initNBond(cutnb=100, #tolerance=45, #repel=1.2, #selStr='name P' #)"))
for name in ("bond", "angl", "impr"):
potList.append(XplorPot(name))
pass
rampedParams.append(MultRamp(0.4, 1.0, "potList["ANGL"].setScale(VALUE)")
rampedParams.append(MultRamp(0.1, 1.0, "potList["IMPR"].setScale(VALUE)")

from ivm import IVM
import varTensorTools
mini = IVM() #initial alignment of orientation tensor axes
for medium in ("pf1"):
    media[medium].setFreedom("fixDa, fixRh")
#for medium in ("bic2"):
#    media[medium].setFreedom("fixDa, fixRhou, fixAxisTo bic1")
#for medium in ("phg2", "phg3"):
#    media[medium].setFreedom("fixDa, fixRhou, fixAxisTo phg1")
varTensorTools.topologySetup(mini, media.values())
protocol.initMinimize(mini, numSteps=20)
mini.fix("not resname ANI")
mini.run() #this initial minimization is not strictly necessary

#uncomment to allow Da, Rh to vary
for medium in ("pf1"):
    media[medium].setFreedom("varyDa, varyRhou")
#for medium in ("bic2"):
#    media[medium].setFreedom("varyDa, varyRhou, fixAxisTo bic1")
#for medium in ("phg2", "phg3"):
#    media[medium].setFreedom("varyDa, fixAxisTo phg1, fixRhouTo phg1")
dyn = IVM()
protocol.initDynamics(dyn, potList=potList)
```

208
varTensorTools.topologySetup(dyn, media.values())
protocol.torsionTopology(dyn)

# Give atoms uniform weights, except for the anisotropy axis
from atomAction import SetProperty
AtomSel("not resname ANI").apply( SetProperty(‘mass’,100 ) )
varTensorTools.massSetup(media.values(),300)
AtomSel("all ").apply( SetProperty(‘fric ’,10. ) )

### minc used for final cartesian minimization
### from selectTools import IVM_groupRigidSidechain
minc = IVM()
protocol.initMinimize(minc, potList=potList)
IVM_groupRigidSidechain(minc)
protocol.cartesianTopology(minc,’not resname ANI’)
varTensorTools.topologySetup(minc, media.values())

init_t1 = 200000
init_t2 = 20000
init_t3 = 3000

from simulationTools import AnnealIVM
anneal1= AnnealIVM(initTemp =init_t1 ,
finalTemp=init_t2 ,
tempStep =5000 ,
ivm=dyn ,
rampedParams = rampedParams )
anneal2= AnnealIVM(initTemp =init_t2 ,
finalTemp=init_t3 ,
tempStep =500 ,
ivm=dyn ,
rampedParams = rampedParams )
anneal3= AnnealIVM(initTemp =init_t3 ,
finalTemp=25 ,
tempStep =25 ,
ivm=dyn ,
rampedParams = rampedParams )

# initialize parameters for initial minimization.
InitialParams( rampedParams )
# high-temp dynamics setup – only need to specify parameters which
# differ from initial values in rampedParams
InitialParams( highTempParams )

# initial minimization
protocol.initMinimize(dyn ,
                      potList=[‘CDIH ’, ‘IMPR’],
                      numSteps=50)
dyn.run()
Script code

```python
# initial minimization
protocol.initMinimize(dyn,  
    potList=potList,  
    numSteps=1000)
minc.run()

# from simulationTools import testGradient  
# testGradient(potList, eachTerm=1)

def calcOneStructure(loopInfo):
    ## mod by anda: first annealing loop, to overcome high energy barriers
    # # initialize parameters for high temp dynamics.
    # InitialParams(rampedParams )
    # # high-temp dynamics setup - only need to specify parameters which
    # # differ from initial values in rampedParams
    # InitialParams(highTempParams )
    #
    # protocol.initDynamics(dyn,  
    #    initVelocities=1,  
    #    bathTemp=init_t1,  
    #    potList=potList,  
    #    finalTime=10)
    # dyn.setETolerance(init_t1/100 ) #used to det. stepsize, default: t/1000
    # dyn.run()
    #
    # # initialize parameters for cooling loop
    # InitialParams(rampedParams )
    #
    # # perform simulated annealing
    #
    # protocol.initDynamics(dyn,  
    #    finalTime=0.1, #time to integrate at a given temp.
    #    numSteps=0, # take as many steps as necessary
    #    # eTol_minimum=0.001 # cutoff for auto-TS det.
    # )
    # anneal1.run()

    # mod by anda: second annealing loop, actual annealing  
    # initialize parameters for high temp dynamics.
    # InitialParams(rampedParams )
    # high-temp dynamics setup - only need to specify parameters which
    # differ from initial values in rampedParams
    InitialParams(highTempParams )
    #
    protocol.initDynamics(dyn,  
        initVelocities=1,  
        bathTemp=init_t2,  
        potList=potList,  
        finalTime=50)
    dyn.setETolerance(init_t2/100 ) #used to det. stepsize, default: t/1000
dyn.run()

    # initialize parameters for cooling loop
    InitialParams(rampedParams )
```

210
# perform simulated annealing
#
protocol.initDynamics(dyn,  
    finalTime=0.5, # time to integrate at a given temp.  
    numSteps=0,  # take as many steps as necessary  
    #eTol_minimum=0.001  # cutoff for auto-TS det.  
  )

anneal2.run()
anneal3.run()
#
# torsion angle minimization
#
protocol.initMinimize(dyn, numSteps=5000)
dyn.run()

###
### all atom minimization
###
protocol.initMinimize(minc, potList=potList, numSteps=3000)
minc.run()
#
# perform analysis and write structure
loopInfo.writeStructure(potList, crossTerms)
pass

from simulationTools import StructureLoop  
StructureLoop(numStructures=numberOfStructures,  
    startStructure=startStructure,  
    structLoopAction=calcOneStructure,  
    pdbTemplate=outFilename,  
    genViolationStats=1,  
    averageFilename="average_min.pdb",  
    averageFitSel="not resname ANI and not (name H71 or name H72 or name H73)",  
    averageRefineSteps=15,  
    averageTopFraction=0.1,  
    averagePotList=potList).run()
Script code

.3 Xplor-NIH calculation restraints files and structures

The coordinates of the 10 minimum-energy, violation-free structures can be found at the Protein Databank (PDB) with accession codes 2kh0 (13merHNF, face-down), 2kh1 (13merHNF, face-up), 2kz1 (13merRef) and 2kz2 (13mer2AP). The input files to the structure calculations have been deposited along with the structures.

.4 Lua scripts written for data export from CARA

The following scripts were written for use in CARA only. Their description is given in the header.

language

--- script to filter through peaklist with all graded peaks and compare it
--- to normal peaklist. All peaks that are not integrated are written to
--- a new peaklist for inspection.

--- written by Andre Dallmann April 23 2007

---

--- PREPARATIONS

--- initialize tables for peak information

  t={} t.label = {} t.id = {} t.assx = {} t.assy = {} t.posx = {} t.posy = {} t.ampl = {} t.grade = {} t.vol = {} t.atomtype = {} t.assxlabel = {} t.assylabel = {}

--- get Project

  local ProjectNames = {}
i = 0

212
for a, b in pairs(cara:getProjects()) do
  i = i + 1
  ProjectNames[i] = b:getName()
end

t.ProjectName = dlg.getSymbol("Select Project", "", unpack(ProjectNames))
t.project = cara:getProject(t.ProjectName)

-- get Peaklist with all graded peaks
local PeaklistNames = {}
i = 0
for a, b in pairs(t.project:getPeakLists()) do
  i = i + 1
  PeaklistNames[i] = b:getName()
end
t.PeaklistName = dlg.getSymbol("Select Graded Peaklist", "", unpack(PeaklistNames))
for a, b in pairs(t.project:getPeakLists()) do
  i = i + 1
  if (b:getName() == t.PeerlistName) then
    t.peaklist = t.project:getPeakList(b:getId(a))
  end
end

-- get Peaklist as Reference
local PeaklistNames = {}
i = 0
for a, b in pairs(t.project:getPeakLists()) do
  i = i + 1
  PeaklistNames[i] = b:getName()
end
t.PeaklistName = dlg.getSymbol("Select Reference Peaklist", "", unpack(PeaklistNames))
for a, b in pairs(t.project:getPeakLists()) do
  i = i + 1
  if (b:getName() == t.PeerlistName) then
    t.peaklistref = t.project:getPeakList(b:getId(a))
  end
end

-- Get Output Filename

t.Filename = dlg.getText("Enter the output filename", "", t.ProjectName)

-- open output
outfile = io.output(t.Filename .. ".diff.peaks")

-- write header
outfile:write("#Number of dimensions ..t.peaklist:getDimCount()..\n")
print("#Number of dimensions ..t.peaklist:getDimCount()")
for i = 1, t.peaklist:getDimCount() do
  outfile:write("#INAME ..i.. ..t.peaklist:getAtomType(i)..\n")
  print("#INAME ..i.. ..t.peaklist:getAtomType(i))
end

count = 0

--- Main Body ---

--- Main Body ---

213
Script code

--- read out peak information
for i, j in pairs (t.peaklistref:getPeaks()) do --- cycle through all peaks of refpeaklist
  boolean = 0
  for x, y in pairs (t.peaklist:getPeaks()) do --- cycle through all peaks of graded
    peaklist
    if (j:getAssig() == y:getAssig()) then
      boolean = 1
    end
  end
  if (boolean == 0) then
    t.peak = t.peaklistref:getPeak(i)
    count = count + 1
    t.id[i] = string.format("%9.0f", t.peak:getId())
    t.ass = {t.peak:getAssig()}
    t.assx[i] = string.format("%9.0f", t.ass[1])
    t.assy[i] = string.format("%9.0f", t.ass[2])
    t.label[i] = string.format("%25.25s", t.peak:getLabel())
    t.pos = {t.peak:getPos()}
    t.posx[i] = string.format("%9.3f", t.pos[1])
    t.posy[i] = string.format("%9.3f", t.pos[2])
    t.ampl[i] = string.format("%7.0f", t.peak:getAmp())
    t.vol[i] = string.format("%15.3f", t.peak:getVol())
    outfile:write (count . t.posx[i] . t.posy[i] . "0 U 
      " . t.label[i] . "0\n"
    )
    print (count . t.posx[i] . t.posy[i] . "0 U " . t.label[i])
  end
end --- of first for loop

--------------------------------------------------
End of Main Body
--------------------------------------------------

i = 0
f = nil

--- close oufiles
outfile:close()

--------------------------------------------------
End of Script
--------------------------------------------------

print ("ncheck_intpeaks_byanda is done.")
print ("Have a nice day!")

--- script to export cara peaklist to Sparky
--- written by Andre Dallmann April--23--2007

--------------------------------------------------
214
--- PREPARATIONS ---

--- initialize tables for peak information

t={}  
t.label = {}  
t.id = {}  
t.assx = {}  
t.assy = {}  
t.posx = {}  
t.posy = {}  
t.ampl = {}  
t.grade = {}  
t.vol = {}  
t.atomtype = {}  
t.assxlabel = {}  
t.assylabel = {}

--- get Project
local ProjectNames = {}
i = 0
for a,b in pairs(cara:getProjects()) do
    i = i + 1
    ProjectNames[i] = b:getName()
end

t.ProjectName = dlg.getSymbol("Select Project",**, unpack(ProjectNames))
t.project = cara:getProject(t.ProjectName)

--- get Peaklist with all graded peaks
local PeaklistNames = {}
i = 0
for a,b in pairs(t.project:getPeakLists()) do
    i = i + 1
    PeaklistNames[i] = b:getName()
end

t.PeaklistName = dlg.getSymbol("Select Peaklist",**, unpack(PeaklistNames))
for a,b in pairs(t.project:getPeakLists()) do
    i = i + 1
    if (b:getName()==t.PeaklistName) then
        t.peaklist = t.project:getPeakList(b:getId(a))
    end
end

--- Get Output Filename

t.Filename = dlg.getText("Enter the output filename",**, t.ProjectName)

--- open outfile
outfile = io.output(t.Filename.."_sparky.peaks")

--- write header
```python
outfile: write (string.format("%25.25s","Assignment")...string.format("%9.9s","w1")...string.format("%9.9s","w1")..."\n")
print (string.format("%25.25s","Assignment")...string.format("%9.9s","w1")...string.format("%9.9s","w1")..."\n")
count = 0

--- Main Body ---

--- read out peak information
for i, peak in pairs(t.peaklist:getPeaks()) do — cycle through all peaks
  t.peak = t.peaklist:getPeak(i)
  t.label[i] = string.format("%25.25s", t.peak:getLabel())
  t.ass[i] = { t.peak:getAssig() }
  t.assx[i] = string.format("%9.0f", t.ass[1])
  t.assy[i] = string.format("%9.0f", t.ass[2])
  t.pos = { t.peak:getPos() }
  t.posx[i] = string.format("%9.3f", t.pos[1])
  t.posy[i] = string.format("%9.3f", t.pos[2])
  t.ampl[i] = string.format("%7.0f", t.peak:getAmp())
  t.vol[i] = string.format("%15.3f", t.peak:getVol())
  t.assx = t.project:getSpin(t.assx[i]):getLabel()
  t.assy = t.project:getSpin(t.assy[i]):getLabel()
  resxid = t.project:getSpin(t.assx[i]):getSystem():getId()
  resyid = t.project:getSpin(t.assy[i]):getSystem():getId()
  resx = t.project:getSpin(t.assx[i]):getSystem():getResidue():getType():getShort()
  resy = t.project:getSpin(t.assy[i]):getSystem():getResidue():getType():getShort()
  if (assx==assy) then
    outfile: write(resx...resxid...assx..."-"...assy...t.posx[i]...t.posy[i]..."\n")
    print (resx...resxid...assx..."-"...assy...t.posx[i]...t.posy[i])
  else
    outfile: write(resx...resxid...assx..."-"...resy...resyid...assy...t.posx[i]...t.posy[i]..."\n")
    print (resx...resxid...assx..."-"...resy...resyid...assy...t.posx[i]...t.posy[i])
  end
end -- of second for loop

--- End of Main Body ---

---
i = 0
t = nil
--- close outfiles
outfile: close()
---

--- End of Script ---

print ("\ntestscript_byanda is done."")
print ("Have a nice day!")
```
--- script to filter through peaklist which where exported from homoscope to monoscope
--- produces two files, containing d2o and h2o peaks respectively

--- written by Andre Dallmann April 11 2007

---

PREPARATIONS

--- initialize tables for peak information

local ProjectNames = {}  
i = 0
for a, b in pairs(cara:getProjects()) do
    i = i + 1
    ProjectNames[i] = b:getName()
end

local PeaklistNames = {}  
i = 0
for a, b in pairs(cara:getPeakLists()) do
    i = i + 1
    PeaklistNames[i] = b:getName()
end

--- Get Project

local ProjectName = dlg:getSymbol("Select Project",**, unpack(ProjectNames))
local project = cara:getProject(t.ProjectName)

--- Get Peaklist

local PeaklistName = dlg:getSymbol("Select Peaklist",**, unpack(PeaklistNames))
for a, b in pairs(t.project:getPeakLists()) do
    i = i + 1
    if b:getName() == t.PeaklistName then
        t_peaklist = t.project:getPeakList(b:getId(a))
    end
```
Script code

end

--- Get Output Filename

```
t. Filename = dlg.getText( "Enter the output filename" , "" , t.ProjectName)
```

--- open outfile

```
outfile_h2o = io.output( t. Filename . . "_h2o . peaks" )
outfile_d2o = io.output( t. Filename . . "_d2o . peaks"
```

--- Main Body

--- read out peak information

```
for i , peak in pairs ( t. peaklist : getPeaks( ) ) do -- cycle through all peaks
  t. peak = t. peaklist : getPeak( i )
  t. id [ i ] = string. format ( "%9.0f" , t. peak : getId( ) )
  t. ass = { t. peak : getAssign() }
  t. assx [ i ] = string. format ( "%9.0f" , t. ass [ 1 ] )
  t. label [ i ] = string. format ( "%25.25s" , t. peak : getLabel( ) )
  t. pos = { t. peak : getPos( ) }
  t. posx [ i ] = string. format ( "%9.3f" , t. pos [ 1 ] )
  t. posy [ i ] = string. format ( "%9.3f" , t. pos [ 2 ] )
  t. ampl [ i ] = string. format ( "%7.0f" , t. peak : getAmp( ) )
  t. vol [ i ] = string. format ( "%15.3f" , t. peak : getVol( ) )
end -- of first for loop

--- write header

```
outfile_h2o : write ( "#Number of dimensions " . . t. peaklist : getDimCount( ) . . "\n"
```

```
for i = 1 , t. peaklist : getDimCount( ) do
  outfile_h2o : write ( "#INAME " . . i . . " " . . t. peaklist : getAtomType( i ) . . "\n"
```

```
end

--- initialize counter variables

```
c_h2o = 0
```

```
c_d2o = 0
```

--- write peaklists

```
for x , assx in pairs ( t. assx ) do
  t. assxlabel [ x ] = t. project : getSpin( t. assx [ x ] ) : getLabel( ) -- get Peaklabel
  t. assylabel [ x ] = t. project : getSpin( t. assy [ x ] ) : getLabel( ) -- get Peaklabel
  if (((t. assxlabel [ x ] == "H41") or (t. assxlabel [ x ] == "H42") or (t. assxlabel [ x ] == "H1") or (t. assxlabel [ x ] == "H3")) or (((t. assylabel [ x ] == "H41") or (t. assylabel [ x ] == "H42") or (t. assylabel [ x ] == "H1") or (t. assylabel [ x ] == "H3")))) and (t. assx [ x ] == t. assx [ x ] )
```

```
then
  c_h2o = c_h2o + 1
  outfile_h2o : write ( c_h2o . . t. posx [ x ] . . t. posy [ x ] . . " 0 U
```

```
  t. vol [ x ] . . t. ampl [ x ] . . " 0 " . . t. assx [ x ] . . t. assx [ x ] . . " 0 \n#
```

```
  t. label [ x ] . . "\n")
```

```
-- print ( c_h2o . . t. posx [ x ] . . t. posy [ x ] . . " 0 U " . . t. vol [ x ] . . t. ampl [ x ] . . " 0 " . . t. assx [ x ] . . t. assx [ x ] . . " 0 \n#
```

```
end
```
0''.t.assx[x].t.assy[x]' 0\# ''t.label[x]... h2o')

else
if (((t.assx_label[x]=="H2") and ((t.assy_label[x]=="H2")
or (t.assy_label[x]=="H2'")) or (t.assy_label[x]=="H2''")
or (t.assx_label[x]=="H4") and ((t.assx_label[x]=="H2")
or (t.assx_label[x]=="H2'")) or (t.assx_label[x]=="H2''")
or ((t.assx_label[x]=="H2") and ((t.assx_label[x]=="H2")
or (t.assx_label[x]=="H2'")) or (t.assx_label[x]=="H2''")
or (t.assx_label[x]=="H4") and (t.assy_label[x]=="H2") and
e. leaves H6 connections because of 2AP

i = i + 1 -- dummy statement
else
if (t.assx[x]~=t.assy[x]) then
   c_d2o = c_d2o + 1
   outfile_d2o:write (c_d2o. t.posx[x]. t.posy[x]." 0 U
"...t.vol[x].t.ampl[x]... -- 0''.t.assx[x].t.assy[x]' 0\# ''t.label[x]...\n"

-- print (c_d2o. t.posx[x]. t.posy[x]." 0 U
"...t.vol[x].t.ampl[x]... -- 0''.t.assx[x].t.assy[x]' 0\# ''t.label[x]... d2o')

end
end

end

------------------------------------------------------- End of Main Body -------------------------------------------------------

--- close oufties
outfile_h2o:close()
outfile_d2o:close()

--- End Script ------------------------------------------

-------------------------------------------------------

print ("\nfilterpeaks_byanda is done."
print ("Have a nice day!"

------------------------------------------------------- End filterpeaksbyanda

--- script to filter through peaklist which were exported from homoscope to
--- monoscope
--- produces several peaklists, containing d2o and h2o peaks respectively
--- adjustments need to be made for unnatural nucleobases, peakwidths, and home spectrum for
--- h2o and d2o!!!

--- written by Andre Dallmann May--11--2007

219
-- PREPARATIONS

--- initialize tables for peak information

t={} 
t.label = {} 
t.id = {} 
t.assx = {} 
t.assy = {} 
t.posx = {} 
t.posy = {} 
t.ampl = {} 
t.grade = {} 
t.vol = {} 
t.atomtype = {} 
t.assxlabel = {} 
t.assylabel = {}

--- Get Project
local ProjectNames = {}
i = 0
for a, b in pairs(cara:getProjects()) do 
i = i + 1
    ProjectNames[i] = b:getName()
end

t.ProjectName = dlg.getSymbol('Select Project','' , unpack(ProjectNames))
t.project = cara:getProject(t.ProjectName)

--- Get Peaklist
local PeaklistNames = {}
i = 0
for a, b in pairs(t.project:getPeakLists()) do 
i = i + 1
    PeaklistNames[i] = b:getName()
end

t.PeaeklistName = dlg.getSymbol('Select Peaklist H2O','' , unpack(PeaklistNames))
for a, b in pairs(t.project:getPeakLists()) do 
i = i + 1
    if (b:getName()==t.PeaeklistName) then 
        t.peaklist_h2o = t.project:getPeakList(b:getId(a))
    end
end

t.PeaeklistName = dlg.getSymbol('Select Peaklist D2O','' , unpack(PeaklistNames))
for a, b in pairs(t.project:getPeakLists()) do 
i = i + 1
    if (b:getName()==t.PeaeklistName) then 
        t.peaklist_d2o = t.project:getPeakList(b:getId(a))
    end
end
Lua scripts written for data export from CARA

-- initialize variables
local c = 0
local d = 0
local e = 0
local f = 0
local waminox = 0.027 -- peak width for amino
local wrestx = 0.027 -- peak width for rest in x dim
local wally = 0.027 -- peak width for y dim
local h2o_spectrum = 6 -- home spectrum needs to be adjusted !!!

for i,p in pairs (t.peaklist_h2o:getPeaks()) do -- cycle through all peaks
  t.posx_ref, t.posy_ref = p:getPos() -- get Positions of Reference peak
  t.assx, t.assy = p:getAssig() -- get Assignment of Reference peak
  t.x = t.project:getSpin(t.assx):getLabel() -- get Peaklabel
  t.y = t.project:getSpin(t.assy):getLabel() -- get Peaklabel
  t.HNFx = t.project:getSpin(t.assx):getSystem():getResidue():getId() -- get Residuenumber
  t.HNFy = t.project:getSpin(t.assy):getSystem():getResidue():getId() -- get Residuenumber
  if ((t.x=="H41") and (t.y=="H42") or ((t.x=="H42") and (t.y=="H41"))) and not (t.HNFx==2) or (t.HNFx==12) or (t.HNFy==2) or (t.HNFy==12) or (t.HNFx==14) or (t.HNFx==26) or (t.HNFy==14) or (t.HNFy==26)) then
    -- selects all H41/H42 pairs which are not subject to base pair fraying
    e = e + 1
    if (e==1) then -- for first match establish peaklist
      peaklist3 = spec:createPeakList("1H","1H")
      peaklist3:setName(t.project:getName().."_h41h42_h2o")
      peaklist3:setHome(t.project:getSpectrum(h2o_spectrum))
      t.project:addPeakList(peaklist3)
      peaklist3:getModel(0):setWidth(1, waminox)
      peaklist3:getModel(0):setWidth(2, wally)
      peaklist3:setAttr("WidthX", waminox)
      peaklist3:setAttr("WidthY", wally)
    end
  end
  peak = peaklist3:createPeak(p:getPos()) -- create Peaks in Peaklist
  peak:setAssig(p:getAssig())
  peak:setLabel(p:getLabel())
  for xx,yy in pairs (t.peaklist_h2o:getPeaks()) do -- cycle through all peaks
    t.posx, t.posy = yy:getPos() -- get Positions of peak
    t.assx, t.assy = yy:getAssig() -- get Assignment of peak
    if (((t.posx_ref-t.posx)^2+(t.posy_ref-t.posy)^2)^(1/2)) > (2*wally)) then -- if next peak is more than double peakwidth away, isolated peak
      if ((peak:getAttr("grade")="h") and (peak:getAttr("grade")="c")) then
        end
Script code

```python
peak : setAttr ('grade', 'a')

else if (wally <
    (((((t.posx_ref-t.posx)^2)+(t.posy_ref-t.posy)^2))^(1/2)) and
    (((((t.posx_ref-t.posx)^2)+(t.posy_ref-t.posy)^2))^(1/2)) <
    (2*wally)) then
    if (peak:getAttr('grade')=='c') then
        peak : setAttr ('grade', 'b')
    end
else if (((((t.posx_ref-t.posx)^2)+(t.posy_ref-t.posy)^2))^(1/2)) <
    wally) and (i==xx) then -- very close proximity to next peak,
great error
    peak : setAttr ('grade', 'c')
end
end

-- this procedure is repeated for all peaks of interest
-- different peaklists are required for peaks with different peakwidths
-- for h2o peaks only one half of the diagonal is used
else if ((t.x=='H41')) and ((t.y=='H1') or (t.y=='H3') or (t.y=='H5') or (t.HNFy==7))
    and (t.assx==t.assy) then
    -- selects all relevant H41-peaks
    c = c + 1
    if (c==1) then
        peaklist = spec.createPeakList ('1H', '1H')
        peaklist : setName(t.project : getName() .. '_H41_h2o')
        peaklist : setHome(t.project : get Spectrum(h2o_spectrum))
        t.project : addPeakList (peaklist)
        peaklist : getModel(0) : set Width(1, waminox)
        peaklist : getModel(0) : set Width(2, wally)
        peaklist : setAttr ('WidthX', waminox)
        peaklist : setAttr ('WidthY', wally)
    end
end
peak = peaklist : createPeak(p:getPos())
peak : setAssig(p:getAssig())
peak : setLabel(p:getLabel())

for xx, yy in pairs (t.peaklist_h2o:getPeaks()) do --cycle through all peaks
    t.posx, t.posy = yy:getPos()
    t.assxx, t.assyy = yy:getAssig()
    if (((((t.posx_ref-t.posx)^2)+(t.posy_ref-t.posy)^2))^(1/2)) >
    (2*wally)) then -- if peak next peak is more than double
    peakwidth away, isolated peak
        if (peak:getAttr('grade')=='b') and
            (peak:getAttr('grade')=='c') then
            peak : setAttr ('grade', 'a')
        end
    end
else if (wally <
    (((((t.posx_ref-t.posx)^2)+(t.posy_ref-t.posy)^2))^(1/2)) and
    (((((t.posx_ref-t.posx)^2)+(t.posy_ref-t.posy)^2))^(1/2)) <
    (2*wally)) then
    if (peak:getAttr('grade')=='c') then
        peak : setAttr ('grade', 'b')
    end
else if (((((t.posx_ref-t.posx)^2)+(t.posy_ref-t.posy)^2))^(1/2)) <
    wally) and (i==xx) then -- very close proximity to next peak,
great error
    peak : setAttr ('grade', 'c')
```
.4 Lua scripts written for data export from CARA

```lua
end

elseif ((t.x=="H42") and ((t.y=="H1") or (t.y=="H3") or (t.y=="H5") or (t.HNFy==7)) and (t.assx=t.assy) then
    -- selects all relevant H42 peaks
    f = f + 1
    if (f==1) then
        peaklist4 = spec.createPeakList("1H","1H")
        peaklist4.setName(t.project:getName().."-_H42_h2o")
        peaklist4:setHome(t.project:getSpectrum(h2o_spectrum))
        t.project:addPeakList(peaklist4)
        peaklist4:getModel(0):setWidth(1,waminox)
        peaklist4:getModel(0):setWidth(2,wally)
        peaklist4:setAttr("WidthX",waminox)
        peaklist4:setAttr("WidthY",wally)
    end
    peak = peaklist4:createPeak(p:getPos())
    peak:setAssig(p:getAssig())
    peak:setLabel(p:getLabel())
    for xx,yy in pairs (t.peaklist_h2o:getPeaks()) do -- cycle through all peaks
        t.posx , t.posy = yy:getPos()
        t.assxx , t.assyy = yy:getAssig()
        if (((t.posx_ref-t.posx)^2)+(t.posy_ref-t.posy)^2)(1/2) > (2*wally) then -- if peak next peak is more than double
            peakwidth away, isolated peak
            if (peak:getAttr("grade")~="h") and
                (peak:getAttr("grade")~="c") then
                peak:setAttr("grade","a")
            end
        elseif (wally < (((t.posx_ref-t.posx)^2)+(t.posy_ref-t.posy)^2)(1/2)) < (2*wally) ) then
            if (peak:getAttr("grade")~="c") then
                peak:setAttr("grade","b")
            end
        elseif (((t.posx_ref-t.posx)^2)+(t.posy_ref-t.posy)^2)(1/2) < wally ) then -- very close proximity to next peak,
            great error
            peak:setAttr("grade","c")
        end
    end
    elseif ((t.x=="H1" or t.x=="H3" or t.x=="H2") and (t.y=="H1" or t.y=="H3" or t.y=="H5") or (t.HNFy==7)) and (t.assx=t.assy) then
        -- selects all other exchangable proton peaks plus HNF-crosspeaks
        d = d + 1
        if (d==1) then
            peaklist2 = spec.createPeakList("1H","1H")
            peaklist2:setName(t.project:getName().."-_rest_h2o")
            peaklist2:setHome(t.project:getSpectrum(h2o_spectrum))
            t.project:addPeakList(peaklist2)
            peaklist2:getModel(0):setWidth(1,wrestx)
            peaklist2:getModel(0):setWidth(2,wally)
            peaklist2:setAttr("WidthX",wrestx)
            peaklist2:setAttr("WidthY",wally)
        end
```
script code

```
peak = peaklist2: createPeak(p:getPos())
peak: setAssig(p:getAssig())
peak: setLabel(p:getLabel())
for xx, yy in pairs (t: peaklist_h2o:getPeaks()) do −− cycle through all peaks
    t: posx, t: posy = yy:getPos()
    t: assx, t: assy = yy:getAssig()
    if (((t: posx_ref-t: posx)^2)+(t: posy_ref-t: posy)^2)^((1/2)) >
        (2*wally) then −− if peak next peak is more than double
            peakwidth away, isolated peak
            if (peak: getAttr("grade")="b") and
                (peak: getAttr("grade")="c") then
                peak: setAttr("grade", "a")
            end
        elseif (wally <
            (((t: posx_ref-t: posx)^2)+(t: posy_ref-t: posy)^2)^((1/2)) and
            (((t: posx_ref-t: posx)^2)+(t: posy_ref-t: posy)^2)^((1/2)) <
            (2*wally)) then −− very close proximity to next peak, great error
                peak: setAttr("grade", "c")
            end
        end
    end −− of first for loop

---

function find2x (index) −− function to format the atomlabels
    local Boolean = false
    local Booleanx = false
    local Booleany = false
    for x in string: gfind(t: x, "H[2345]" ) do
        Booleanx=true
    end
    for y in string: gfind(t: y, "H[25]" ) do
        Booleny=true
    end
    if (Booleanx=true) and (Booleny=true) then
        Boolean=true
    else
        Boolean=false
    end
    return Boolean
end

function find2y (index) −− function to filter atoms
    local Boolean = false
    local Booleanx = false
    local Booleny = false
    for x in string: gfind(t: x, "H[2345]" ) do
        Booleanx=true
    end
    for y in string: gfind(t: y, "H[25]" ) do
        Booleny=true
    end
    if (Booleanx=true) and (Booleny=true) then
        Boolean=true
    else
        Boolean=false
    end
    return Boolean
end
```
4 Lua scripts written for data export from CARA

for x in string.gfind(t.x, 'H\[25\]$') do
    Booleanx=true
end
for y in string.gfind(t.y, 'H\[2345\]' \[']*$') do
    Booleany=true
end
if (Booleanx==true) and (Booleany==true) then
    Boolean=true
else
    Boolean=false
end
return Boolean
end

-- initialize variables
local c = 0
local d = 0
local e = 0
local f = 0
local wh2 = 0.040 -- peak width for H2' or H2''
local wrest = 0.025 -- peak width for rest
local d2o_spectrum = 1

for i, p in pairs(t.peaklist_d2o:getPeaks()) do -- cycle through all peaks
    t.posx_ref, t.posy_ref = p:getPos()
    t.assx, t.assy = p:getAssig()
    t.x = t.project:getSpin(t.assx):getLabel() -- get Peaklabel
    t.y = t.project:getSpin(t.assy):getLabel() -- get Peaklabel
    if (t.x=="H41") or (t.x=="H42") or (t.x=="H1") or (t.x=="H3") or
    (t.y=="H41") or (t.y=="H42") or (t.y=="H1") or (t.y=="H3") then
        t.peaklist_d2o:removePeak(p) -- remove all h2o-peaks
    elseif ((t.x=="H2'") or (t.x=="H2''")) and (t.assx==t.assy) and (find2x(i)==false) and
    (find2y(i)==false) then
        local peaklist2 = spec.createPeakList("1H", "1H")
        peaklist2:setName(t.project:getName().."_h2inx_d2o")
        peaklist2:setHome(t.project:getSpectrum(d2o_spectrum))
        t.project:addPeakList(peaklist2)
        peaklist2:getModel(0):setWidth(1, wh2)
        peaklist2:getModel(0):setWidth(2, wrest)
        peaklist2:setAttr("WidthX", wh2)
        peaklist2:setAttr("WidthY", wrest)
        peak = peaklist2:createPeak(p:getPos())
        peak:setAssig(p:getAssig())
        peak:setLabel(p:getLabel())
    end
end

for xx, yy in pairs(t.peaklist_d2o:getPeaks()) do -- cycle through all peaks
    t.posx, t.posy = yy:getPos()
    t.assx, t.assy = yy:getAssig()
if (((t.posx_ref-t.posx)^2)+((t.posy_ref-t.posy)^2))^(1/2) > (2*wh2) then — if peak next peak is more than double peakwidth away, isolated peak
  if (peak:getAttr("grade")=='b' and
    (peak:getAttr("grade")=='c')) then
    peak:setAttr("grade","a")
  end
else if (wh2 < (((t.posx_ref-t.posx)^2)+((t.posy_ref-t.posy)^2))^(1/2) < (2*wh2)) then
  if (peak:getAttr("grade")=='c') then
    peak:setAttr("grade","b")
  end
else if (((t.posx_ref-t.posx)^2)+((t.posy_ref-t.posy)^2))^(1/2) < (2*wh2) and (i==xx) then — very close proximity to next peak, great error
    peak:setAttr("grade","c")
else if (t.ys="H2'" or (t.ys="H2'') and (t.assx==t.assy) and (find2x(i)==false) and
  (find2y(i)==false) then — selects all peaks including H2' to H2'' spins on y-axis (differentiation necessary because of different peakwidth for y and x axis)
  e = e +1
  if (e==1) then
    peaklist3 = spec:createPeakList("1H","1H")
    peaklist3:setName(t.project:getName()..._h2iny_d2o)
    t.project:addPeakList(peaklist3)
    peaklist3:getModel(0):setWidth(1,wrest)
    peaklist3:getModel(0):setWidth(2,wh2)
    peaklist3:setAttr("WidthX",wrest)
    peaklist3:setAttr("WidthY",wh2)
  end
  peak = peaklist3:createPeak(p:getPos())
  peak:setAssig(p:getAssig())
  peak:setLabel(p:getLabel())
  for xx,yy in pairs (t.peaklist_d2o:getPeaks()) do — cycle through all peaks
    t.posx,t.posy = yy:getPos()
    t.assx,t.assyy = yy:getAssig()
    if (((t.posx_ref-t.posx)^2)+((t.posy_ref-t.posy)^2))^(1/2) > (2*wh2) then — if peak next peak is more than double peakwidth away, isolated peak
      if (peak:getAttr("grade")=='b' and
        (peak:getAttr("grade")=='c')) then
        peak:setAttr("grade","a")
      end
    end
  else if (wh2 < (((t.posx_ref-t.posx)^2)+((t.posy_ref-t.posy)^2))^(1/2) < (2*wh2)) then
    if (peak:getAttr("grade")=='c') then
      peak:setAttr("grade","b")
    end
  else if (((t.posx_ref-t.posx)^2)+((t.posy_ref-t.posy)^2))^(1/2) < (2*wh2) then
    if (peak:getAttr("grade")=='c') then
      peak:setAttr("grade","b")
    end
Lua scripts written for data export from CARA

wh2) and (i=='xx') then -- very close proximity to next peak, great error
    peak:setAttr("grade",'c')
end

elseif (t.assx=='t.assy') and (find2x(i)==false) and (find2y(i)==false) then
    -- select all other peaks which are not sugar to H2* or H5*
    f = f +1
    if (f==1) then
        peaklist4 = spec:createPeakList("1H","1H")
        peaklist4:setName(t.project:getName().."_rest_d2o")
        peaklist4:setHome(t.project:getSpectrum(d2o_spectrum))
        t.project:addPeakList(peaklist4)
        peaklist4.setModel(0):setWidth(1,wrest)
        peaklist4.setModel(0):setWidth(2,wrest)
        peaklist4:setAttr("WidthX",wrest)
        peaklist4:setAttr("WidthY",wrest)
    end
    peak = peaklist4:createPeak(p:getPos())
    peak:setAssign(p:getAssign())
    peak:setLabel(p:getLabel())
    for xx,yy in pairs(t.peaklist_d2o:getPeaks()) do -- cycle through all peaks
        t.posx,t.posy = yy:getPos()
        t.assxx,t.assyy = yy:getAssign()
        if (((((t.posx_ref-t.posx)^2)+((t.posy_ref-t.posy)^2))^(1/2)) > (2*wrest)) then
            if (peak:getAttr("grade")~="b") and
                (peak:getAttr("grade")~="c") then
                peak:setAttr("grade","a")
            end
        elseif (wrest<((((t.posx_ref-t.posx)^2)+((t.posy_ref-t.posy)^2))^(1/2)) < (2*wrest)) then
            if (peak:getAttr("grade")="c") then
                peak:setAttr("grade","a")
            end
        elseif (((((t.posx_ref-t.posx)^2)+((t.posy_ref-t.posy)^2))^(1/2)) < (wrest) and (i=='xx') then
            peak:setAttr("grade",'c')
            -- print (p:getLabel()..’..yy:getLabel()..’
            ‘..((t.posx_ref-t.posx)^2)+((t.posy_ref-t.posy)^2))^(1/2))
        end
    end
end
end

--- of first for loop

end

--- End of Main Body

---

--- t.Project:removePeakList(peaklist_d2o)
--- t.Project:removePeakList(peaklist_h2o)
t = nil

--- End of Script

---
print ( "\nfilterpeaks_byanda is done.\n")
print ( 'Have a nice day!' )

--- End filterpeaks_byanda

--- script to generate ppm-file for input to GIF

--- Script

--- PREPARATIONS

--- initialize tables for peak information

$t=$
t.label = {}
t.id = {}
t.assx = {}
t.assy = {}
t.posx = {}
t.posy = {}
t.ampl = {}
t.grade = {}
t.vol = {}
t.atomtype = {}
t.assxlabel = {}
t.assylabel = {}

--- get Project
local ProjectNames = {}
local i = 0
for a, b in pairs(cara:getProjects()) do
  i = i + 1
  ProjectNames[i] = b:getName()
end
t.ProjectName = dlg.getSymbol('Select Project', '', unpack(ProjectNames))
t.project = cara:getProject(t.ProjectName)

--- get Input-Filename
---SparkyList = dlg.getText( 'Enter input filename', '', t.ProjectName.'_sparky_peaks')

--- open input file
---io.open (SparkyList, r)

--- Get Output Filename

--- Write ppm-file

--- End filterpeaks_byanda

--- Script

--- PREPARATIONS

--- initialize tables for peak information

$t=$
t.label = {}
t.id = {}
t.assx = {}
t.assy = {}
t.posx = {}
t.posy = {}
t.ampl = {}
t.grade = {}
t.vol = {}
t.atomtype = {}
t.assxlabel = {}
t.assylabel = {}
A Lua scripts written for data export from CARA

−− open outfile
outfile = io.output( t.Filename )

−− write header
for i, spin in pairs ( t.project:getSpins() ) do
    t.label[i]=spin:getLabel()
    if (t.label[i]=="H5") then
        t.label[i]="HQ5"
    elseif (t.label[i]=="H2") then
        t.label[i]="HQ2"
    end
end
for i, spin in pairs ( t.project:getSpins() ) do
    if (t.label[i]=="H7") then
        outfile:write( "PPM " . spin:getSystem():getResidue():getType():getShort()." 
            " . spin:getSystem():getResidue():getId()." . t.label[i] . " 1
            " . spin:getShift()." 1
        outfile:write( "PPM " . spin:getSystem():getResidue():getType():getShort()." 
            " . spin:getSystem():getResidue():getId()." . t.label[i] . " 2
            " . spin:getShift()." 1
        outfile:write( "PPM " . spin:getSystem():getResidue():getType():getShort()." 
            " . spin:getSystem():getResidue():getId()." . t.label[i] . " 3
            " . spin:getShift()." 1
    else
        outfile:write( "PPM " . spin:getSystem():getResidue():getType():getShort()." 
            " . spin:getSystem():getResidue():getId()." . t.label[i] . " 1
            " . spin:getShift()." 1
        end
end
outfile:close()
t={} print ("done with gifappm")

−− script to remove peak lists, peak lists are picked by their names

−− written by Andre Dallmann
for i, p in pairs ( cara:getProject():getPeakLists() ) do
    temp = string.gsub ( p:getName(), "13merHNF" )
    for y in temp do
        cara:getProject():removePeakList(p)
        print (y)
    end
end
print ("peak lists removed!")

−− writes H shifts from different spins to separate columns in an external file
Script code

--- written by Andre Dallmann

```python
function Format(Number)  --- function to format the chemical shifts
    FormattedNumber = string.format("%7.3f", Number)
    return FormattedNumber
end

function Format2(String)  --- function to format the chemical shifts
    FormattedString = string.format("%7.7s", String)
    return FormattedString
end

--- User editable parameters are below:

--- Spacer between elements of table to write out:
Spacer = " "
Spacer2 = " ",

--- Table of spin labels whose shifts should be written to a column

SpinsInColumns = {}
SpinsInColumns[1] = "H1'
SpinsInColumns[2] = "H2'
SpinsInColumns[3] = "H2''
SpinsInColumns[4] = "H3'
SpinsInColumns[5] = "H4'
SpinsInColumns[6] = "H5'
SpinsInColumns[7] = "H5''
SpinsInColumns[8] = "H1'
SpinsInColumns[9] = "H2'
SpinsInColumns[10] = "H3'
SpinsInColumns[11] = "H41'
SpinsInColumns[12] = "H42'
SpinsInColumns[13] = "H5'
SpinsInColumns[14] = "H6'
SpinsInColumns[15] = "H7'
SpinsInColumns[16] = "H8'
SpinsInColumns[17] = "H4'
SpinsInColumns[18] = "H1''

--- End of user editable section

--- define a table of temporary script variables
t={}

--- Get Parameters from User

--- 1. Get Project:
local projectnames = {}
i=0
for ProjName, Proj in pairs(caragetProjects()) do
    i = i + 1
    projectnames[i] = ProjName
end
if i==1 then
    t.ProjectName = projectnames[i]
```

230
4 Lua scripts written for data export from CARA

else
    t.ProjectName = dlg.getSymbol( "Choose project", "select one", unpack( projectnames ) )
end
if not t.ProjectName then
    error( "No project name defined" )
else
    t.P = cara:getProject( t.ProjectName )
end

-- 2. Get Output filename:  

-- for each line look for each column entry in turn
-- add it to the end of the growing line

Seq = t.P:getSequence()
j = 0
Lines = {}
for ResId, Res in pairs( Seq ) do
    Sys = Res:getSystem()
    if Sys then
        SpinsInSys = Sys:getSpins()
        j = j + 1
        Lines[ j ] = string.format ( "\%7.7s", ResId )
        for k = 1, table.getn( SpinsInColumns ) do
            LabelToFind = SpinsInColumns[ k ]
            MatchingSpin = nil  -- reset to none found
            for SpinId, Spin in pairs( SpinsInSys ) do
                if Spin:getLabel() == LabelToFind then
                    MatchingSpin = Spin
                    end  -- if Spins label matches LabelToFind
            end
            if MatchingSpin then
                FormShift = Format ( MatchingSpin:getShift() )
                Lines[ j ] = Lines[ j ] .. FormShift
            else
                Formzero = Format( "-" )
                Lines[ j ] = Lines[ j ] .. Formzero  -- no shift assigned, leave empty
            end
        end
        for all elements k of SpinsInColumns { try to find a shift for this label }
        end
        end
        -- for all residues in sequence
        --create string "Table" with lines

    end
end

-- 3. Get Label of Spin to write out chemical shifts from
-- I replaced this step with a table , see the top of the script
-- t.Label = dlg.getText( "Enter label of the spins whose chemical shifts you want to write out : ", "Enter Label of spins whose shifts will be written out (e.g. HA): ")


Script code

-- create the first line of table
for m = 1, table:getn( SpinsInColumns ) do
    if m === 1 then
        Labels = string.format("%7.7s", SpinsInColumns[1])
    else
        Formlabels = string.format("%7.7s", SpinsInColumns[m])
       Labels = Labels .. Formlabels
    end
end
res = string.format("%7.7s", "Res")
Header = res .. Labels
for l = 1, table:getn( Lines ) do
    if l === 1 then
        Table = Header .. "\n" .. Lines[1]
    else
        Table = Table .. "\n" .. Lines[1]
    end
end
-- Now write out all lines to a file
file = io.open( t.FileName, "w" )
file:write( Table )
file:flush()
file:close()

print("Wrote out " .. table:getn( Lines ) .. " lines to file " .. t.FileName )
print( "script WriteShiftsInColumns is done" )
t = nil

-- writes 1H shifts from different spins to separate columns in an external file
-- written by Andre Dallmann

function Format( Number ) -- function to format the chemical shifts
    FormattedNumber = string.format("%11.3f", Number )
    return FormattedNumber
end
function Format2( String ) -- function to format the chemical shifts
    FormattedString = string.format("%11.11s", String )
    return FormattedString
end

-- User editable parameters are below: ----------------------------------------

-- Spacer between elements of table to write out:
Spacer = " "
Spacer2 = " 

-- Table of spin labels whose shifts should be written to a column
SpinsInColumns = {}
SpinsInColumns[1] = "Cl"
.4 Lua scripts written for data export from CARA

SpinsInColumns[2] = 'C2'
SpinsInColumns[3] = 'C3'
SpinsInColumns[4] = 'C4'
SpinsInColumns[5] = 'C5'
SpinsInColumns[6] = 'C6'
SpinsInColumns[7] = 'C7'
SpinsInColumns[8] = 'C8'
SpinsInColumns[9] = 'C1'
SpinsInColumns[10] = 'C2'
SpinsInColumns[11] = 'C3'
SpinsInColumns[12] = 'C4'

--- End of user editable section

--- define a table of temporary script variables
t={}  

--- Get Parameters from User

--- 1. Get Project:

local projectnames = {}
i=0
for ProjName, Proj in pairs( cara:getProjects() ) do
  i = i + 1
  projectnames[ i ] = ProjName
end

if i==1 then
  t.ProjectName = projectnames[ i ]
else
  t.ProjectName = dlg:getSymbol( "Choose project", "select one", unpack( projectnames ) )
end

if not t.ProjectName then
  error( "No project name defined" )
else
  t.P = cara:getProject( t.ProjectName )
end

--- 2. Get Output filename:

t.FileName = dlg:getText( "Enter output filename", "output filename", "shiftstable_C.txt" )

--- 3. Get Label of Spin to write out chemical shifts from

--- I replaced this step with a table, see the top of the script
--- t.Label = dlg:getText( "Enter label of the spins whose chemical shifts you want to write out : ", "Enter Label of spins whose shifts will be written out (e.g. HA): " )

--- loop through the sequence and for each residue, create a Line
--- then for each Line look for each column entry in turn
--- add it to the end of the growing line

Seq = t.P:getSequence()
j = 0
Lines = {}
for ResId, Res in pairs( Seq ) do
  Sys = Res:getSystem()
  Lines[j] = Sys
  j = j + 1
end

233
if Sys then — if residue is assigned

    SpinsInSys = Sys:getSpins()
    j = j + 1
    Lines[ j ] = string.format("%11.11s", ResId)

for k = 1, table:getn( SpinsInColumns ) do
    LabelToFind = SpinsInColumns[ k ]
    MatchingSpin = nil — reset to none found

for SpinId, Spin in pairs( SpinsInSys ) do — search for a match to
    LabelToFind
        if Spin:getLabel() == LabelToFind then
            MatchingSpin = Spin
        end — if Spins label matches LabelToFind
    end — for all Spins in System (look for match to this Label)
    if MatchingSpin then
        FormShift = Format( MatchingSpin:getShift() )
        Lines[ j ] = Lines[ j ]..FormShift
    else
        Formzero = Format2( "-" )
        Lines[ j ] = Lines[ j ]..Formzero — no shift assigned, leave
        empty
    end — for all elements k of SpinsInColumns (try to find a shift for this
    label)
end — for all residues in sequence

end — if System is assigned

—create string 'Table' with lines

—create the first line of table
for m = 1, table:getn( SpinsInColumns ) do
    if m == 1 then
        Labels = string.format("%11.11s", SpinsInColumns[ 1 ])
    else
        Formlabels = string.format("%11.11s", SpinsInColumns[ m ])
        Labels = Labels..Formlabels
    end
res = string.format("%11.11s","Res")
Header = res..Labels

for l=1,table:getn( Lines ) do
    if l==1 then
        Table = Header.."\n"..Lines[ 1 ]
    else
        Table = Table.."\n"..Lines[ 1 ]
    end
end

— Now write out all lines to a file
file = io.open( t.FileName, "w" )
file:write( Table )
file:flush()
file:close()
.A Lua scripts written for data export from CARA

print('Wrote out ..table.getn(Lines)..' lines to file '..t.FileName)
print('"script WriteShiftsInColumns is done"

t = nil

--- First part: Script to output all chosen and integrated peaks from one project
--- and combine them in one peaklist.

--- Second part: Choose the best integrated peak among same ones or average over
--- equivalently rated peaks.

--- Third part: Convert peak volumes to distances. Tricky is here the
--- differentiation of d2o and h2o and methyl peaks (all have different
--- reference peaks)!

--- Fourth part: An XPLOR-input file is generated where the distance information
--- and some predefined upper and lower limits (deduced from the maximum
--- deviation of the standard peaks) are used.


---

--- FIRST PART
---

--- Table for all the variables used in the script

--- choosing one project
local ProjectNames = {}local i = 0
for a,b in pairs(cara:getProjects()) do
    i = i + 1
    ProjectNames[i] = b:getName()
end
t.ProjectName=dlg.getSymbol("Select Project",**, unpack( ProjectNames )

--- Get Output Filename

t.Filename = dlg.getText("Enter the output filename", **, t.ProjectName)

--- open output
outfile = io.output( t.Filename..'._all.peaks'

--- Write header to peaklist
local label = string.format('%25.25s', 'Peaklabel')local id = string.format('%9.9s', 'PeakID')local assx = string.format('%9.9s', 'ID(X)')local assy = string.format('%9.9s', 'ID(Y)')local posx = string.format('%9.9s', 'PPM(X)')local posy = string.format('%9.9s', 'PPM(Y)')
local ampl = string.format("%7.7s", "Ampl")
local grade = string.format("%9.9s", "Grade")
local vol = string.format("%15.15s", "Volume")
outfile:write("
ID new... id... label... assx... assy... posx... posy... ampl... vol... grade...
")

--- generate tables for information
local count = 0
local i = 0
t.label = {}
t.id_old = {}
t.ass = {}
t.assx = {}
t.assy = {}
t.pos = {}
t.posx = {}
t.posy = {}
t.ampl = {}
t.grade = {}
t.vol = {}

--- generate list of all peaks graded abc of all peak lists in specified project
for peaklistid, pealkist in pairs(t.project:getPeakLists()) do -- cycle through all peak lists
  for peakid, peak in pairs(t.peaklist:getPeaks()) do -- cycle through all peaks
    t.peak = t.peaklist:getPeak(peakid)
    if ((t.peak:getAttr("grade")=="a") or (t.peak:getAttr("grade")=="b") or
     (t.peak:getAttr("grade")=="c") then -- choose only peaks with grade abc
      i = i + 1 -- this is the index for all the tables, corresponds to new peak id
      t.label[i] = string.format("%25.25s", t.peak:getLabel())
      t.id_old[i] = string.format("%9.9f", t.peak:getId())
      t.ass = {t.peak:getAssig()}
      t.assx[i] = string.format("%9.9f", t.ass[1])
      t.assy[i] = string.format("%9.9f", t.ass[2])
      t.pos = {t.peak:getPos()}
      t.posx[i] = string.format("%9.9f", t.pos[1])
      t.posy[i] = string.format("%9.9f", t.pos[2])
      t.ampl[i] = string.format("%7.7s", t.peak:getAmp())
      t.grade[i] = string.format("%7.7s", t.peak:getAttr("grade")
      t.vol[i] = string.format("%15.15s", t.peak:getVol())
      outfile:write(i..."
      ...	.id_old[i]... t.label[i]... t.assx[i]... t.assy[i]... t.posx[i]... t.posy[i]...
      ...	.ampl[i]... t.vol[i]... t.grade[i]...
    end
end

--- close outfile
outfile:close()
Lua scripts written for data export from CARA

---

PREPARATIONS

---

Preparing combination

```lua
for i, assx in pairs (t.assx) do
  for j, assy in pairs (t.assy) do
    if (((t.assx[i]==t.assx[j]) and (t.assy[i]==t.assy[j])) or
        ((t.assx[i]==t.assy[j]) and (t.assy[i] == t.assx[j]))) and not (j==i) )
      then
        -- select all peaks that have the same assignment (including cross-diagonal peaks)
        counter = counter + 1
        if (t.grade[i]==t.grade[j]) then
          t.vol[i] = string.format ("%15.3f",(t.vol[i] + t.vol[j])) -- average volumes, rest stays
          t.label[j] = nil -- set jth peak to nil
          t.assx[j] = nil
          t.assy[j] = nil
          t.posx[j] = nil
          t.posy[j] = nil
          t.ampl[j] = nil
          t.grade[j] = nil
          t.vol[j] = nil
          t.id_old[j] = nil
```
if ((a==t.grade[i]) and ((t.grade[j]==b) or (t.grade[j]==c))) or
((b==t.grade[i]) and (t.grade[j]==c)) then
    counter = 1
    t.label[j] = nil  -- set jth peak to nil
    t.assx[j] = nil
    t.assy[j] = nil
    t.posx[j] = nil
    t.posy[j] = nil
    t.ampl[j] = nil
    t.grade[j] = nil
    t.vol[j] = nil
    t.id_old[j] = nil
end

if ((a==t.grade[j]) and ((t.grade[i]==b) or (t.grade[i]==b))) or
((b==t.grade[j]) and (t.grade[i]==c)) then
    counter = 1
    t.vol[i] = t.vol[j]  -- transfer volume and grade of better integrated peak (j)
    t.grade[i] = t.grade[j]
    t.label[i] = nil  -- set ith peak to nil
    t.assx[i] = nil
    t.assy[i] = nil
    t.posx[i] = nil
    t.posy[i] = nil
    t.ampl[i] = nil
    t.grade[i] = nil
    t.vol[i] = nil
    t.id_old[i] = nil
end
end  -- if loop

if (counter > 1) then  -- only valid if grades are the same and averaging is needed
    t.vol[i] = string.format("%15.3f", (t.vol[i]/counter))
end

end  -- first for loop

end  -- if loop

--- initialize new tables for the combined peaklist

for i, assx in pairs(t.assx) do  -- generate new table with combined peaks
    t.volnum=tonumber(t.vol[i])
    if (t.volnum>0) then
        t.labelnew = {}  
        t.assxnew = {}  
        t.assynew = {}  
        t.assxlabel = {}  
        t.assylabel = {}  
        t.assxresid = {}  
        t.assyresid = {}  
        t.gradenew= {}  
        t.volnew = {}  
    end
end
Lua scripts written for data export from CARA

```lua
x = x + 1

-- get Peak label

-- get residue id

-- loop to correct for base rectangle sum method error
-- for peaks with grade c or b divide volume by 2 or 1.5 respectively
-- this is a very rough approximation!!!

-- close outfile

-- third part
```

---

239
--- PREPARATIONS ---

## open output

```python
outfile = io.output(t.Filename+'._dist.peaks')
outfile2 = io.output(t.Filename+'._notused.peaks')
```

## initialize variables

```python
local sumcyt = 0
local summet = 0
local sumcytamino = 0
local sumcyth42h5 = 0
local sumcyth41h5 = 0
local countcyt = 0
local countmet = 0
local countcytamino = 0
local countcyth42h5 = 0
local countcyth41h5 = 0
```

---

## Main Body

---

### Setting up Reference Volumes and Distances

---

## sum up reference peaks

```python
for j, assx in pairs {t.assxnew} do
  for y in string.find {t.labelnew[j]}, \H[56]/H[56] \[0-9]\:[C0-9]+\ do -- establish reference for d2o peaks
    countcyt = countcyt + 1
    sumcyt = sumcyt + t.volnew[j]
  end
  for y in string.find {t.labelnew[j]}, \H[67]/H[67] \[0-9]\:[T0-9]+\ do -- establish reference for methyl peaks
    countmet = countmet + 1
    summet = summet + t.volnew[j]
  end
  for y in string.find {t.labelnew[j]}, \H[12]/H[12] \[0-9]\:[C0-9]+\ do -- establish reference for h2o exchangeable peaks
    countcytamino = countcytamino + 1
    sumcytamino = sumcytamino + t.volnew[j]
  end
  for y in string.find {t.labelnew[j]}, \H[21]/H[21] \[0-9]\:[C0-9]+\ do -- establish reference for h2o exchangeable non-exchangeable peaks
    countcyth42h5 = countcyth42h5 + 1
    sumcyth42h5 = sumcyth42h5 + t.volnew[j]
  end
  for y in string.find {t.labelnew[j]}, \H[42]/H[42] \[0-9]\:[C0-9]+\ do -- establish reference for h2o exchangeable non-exchangeable peaks
    countcyth42h5 = countcyth42h5 + 1
    sumcyth42h5 = sumcyth42h5 + t.volnew[j]
  end
  for y in string.find {t.labelnew[j]}, \H[21]/H[21] \[0-9]\:[C0-9]+\ do -- establish reference for h2o exchangeable non-exchangeable with H41 peaks
    countcyth41h5 = countcyth41h5 + 1
  end
```
sumcyth41h5 = sumcyth41h5 + t.volnew[j]
end
for y in string.gfind (t.labelnew[j], "H5/H41 [0–9]:C[0–9]+") do — establish reference
for h2o exchangeable—non-exchangeable with H41 peaks
countcyth41h5 = countcyth41h5 + 1
sumcyth41h5 = sumcyth41h5 + t.volnew[j]
end

refvolcyt = string.format("%.13.3f", sumcyt / countcyt) — average volume of CYT H5–H6
refdistcyt = 2.48 — distance of CYT H5–H6
refvolmet = string.format("%.13.3f", summet / countmet) — average volume of THY H6–H7
refdistmet = 3.09 — distance of THY H6–H7
refvolcytamino = string.format("%.13.3f", sumcytamino / countcytamino) — average volume of
CYT H41–H42
refdistcytamino = 1.70 — distance of CYT H41–H42
refvolcyth42h5 = string.format("%.13.3f", sumcyth42h5 / countcyth42h5) — average volume of
CYT H42–H5
refdistcyth42h5 = 2.40 — distance of CYT H42–H5
refvolcyth41h5 = string.format("%.13.3f", sumcyth41h5 / countcyth41h5) — average volume of
CYT H41–H5
refdistcyth41h5 = 3.62 — distance of CYT H41–H5

——— Prepare standard deviations for references ————

— initialize variables
local stddevsumcyt = 0
local stddevsummet = 0
local stddevsumcytamino = 0
local stddevsumcyth42h5 = 0
local stddevsumcyth41h5 = 0
local maxdev1 = 0
local maxdev2 = 0
local maxdev3 = 0
local maxdev4 = 0
local maxdev5 = 0

for j, assx in pairs (t.assxnew) do
  for y in string.gfind (t.labelnew[j], "H56/H56 [0–9]:C[0–9]+") do
    stddevsumcyt = stddevsumcyt + (t.volnew[j]–refvolcyt)^2 — standard deviation
    dummy1 = math.abs(t.volnew[j]–refvolcyt) — dummy for maximum deviation
    if (dummy1 > maxdev1) then
      maxdev1 = dummy1
    end
  end

  for y in string.gfind (t.labelnew[j], "H67/H67 [0–9]:T[0–9]+") do — establish
  reference for methyl peaks
    stddevsummet = stddevsummet + (t.volnew[j]–refvolmet)^2 — standard deviation
    dummy2 = math.abs(t.volnew[j]–refvolmet) — dummy for maximum deviation
    if (dummy2 > maxdev2) then
      maxdev2 = dummy2
    end
  end

  for y in string.gfind (t.labelnew[j], "H412/H412 [0–9]:C[0–9]+") do — establish
  reference for h2o exchangeable peaks

end

241
Script code

```plaintext
stddevsumcytamino = stddevsumcytamino + (t.volnew[j] - refvolcytamino)^2  -- standard deviation
dummy3 = math.abs(t.volnew[j] - refvolcytamino) -- dummy for maximum deviation
if (dummy3 > maxdev3) then
    maxdev3 = dummy3
end

for y in string.gfind(t.labelnew[j], "H42/H5 [0-9]:C[0-9]+") do -- establish reference for h2o exchangeable-non-exchangeable peaks (appears twice because of selection reasons)
    stddevsumcyth42h5 = stddevsumcyth42h5 + (t.volnew[j] - refvolcyth42h5)^2  -- standard deviation
dummy4 = math.abs(t.volnew[j] - refvolcyth42h5) -- dummy for maximum deviation
    if (dummy4 > maxdev4) then
        maxdev4 = dummy4
    end
end

for y in string.gfind(t.labelnew[j], "H5/H42 [0-9]:C[0-9]+") do -- establish reference for h2o exchangeable-non-exchangeable peaks
    stddevsumcyth42h5 = stddevsumcyth42h5 + (t.volnew[j] - refvolcyth42h5)^2  -- standard deviation
dummy4 = math.abs(t.volnew[j] - refvolcyth42h5) -- dummy for maximum deviation
    if (dummy4 > maxdev4) then
        maxdev4 = dummy4
    end
end

for y in string.gfind(t.labelnew[j], "H41/H5 [0-9]:C[0-9]+") do -- establish reference for h2o exchangeable-non-exchangeable peaks (appears twice because of selection reasons)
    stddevsumcyth41h5 = stddevsumcyth41h5 + (t.volnew[j] - refvolcyth41h5)^2  -- standard deviation
dummy5 = math.abs(t.volnew[j] - refvolcyth41h5) -- dummy for maximum deviation
    if (dummy5 > maxdev5) then
        maxdev5 = dummy5
    end
end

for y in string.gfind(t.labelnew[j], "H5/H41 [0-9]:C[0-9]+") do -- establish reference for h2o exchangeable-non-exchangeable peaks
    stddevsumcyth41h5 = stddevsumcyth41h5 + (t.volnew[j] - refvolcyth41h5)^2  -- standard deviation
dummy5 = math.abs(t.volnew[j] - refvolcyth41h5) -- dummy for maximum deviation
    if (dummy5 > maxdev5) then
        maxdev5 = dummy5
    end
end

-------------------- Calculate standard deviations for references ---------------------

stddevcyt = string.format("%.3f", (stddevsumcyt / countcyt)^(1/2))
stddevmet = string.format("%.3f", (stddevsummet / countmet)^(1/2))
stddevcytamino = string.format("%.3f", (stddevsumcytamino / countcytamino)^(1/2))
stddevcyth42h5 = string.format("%.3f", (stddevsumcyth42h5 / countcyth42h5)^(1/2))
stddevcyth41h5 = string.format("%.3f", (stddevsumcyth41h5 / countcyth41h5)^(1/2))
```

242
--- Prepare maximum deviations in percent for references ---

```lua
if (maxdev1==nil) then
    maxdevcyt = string.format("%.3f", dummy1/refvolcyt)
else
    maxdevcyt = string.format("%.3f", maxdev1/refvolcyt)
end
if (maxdev2==nil) then
    maxdevmet = string.format("%.3f", dummy2/refvolmet)
else
    maxdevmet = string.format("%.3f", maxdev2/refvolmet)
end
if (maxdev3==nil) then
    maxdevcytamino = string.format("%.3f", dummy3/refvolcytamino)
else
    maxdevcytamino = string.format("%.3f", maxdev3/refvolcytamino)
end
if (maxdev4==nil) then
    maxdevcyth42h5 = string.format("%.3f", dummy4/refvolcyth42h5)
else
    maxdevcyth42h5 = string.format("%.3f", maxdev4/refvolcyth42h5)
end
if (maxdev5==nil) then
    maxdevcyth41h5 = string.format("%.3f", dummy5/refvolcyth41h5)
else
    maxdevcyth41h5 = string.format("%.3f", maxdev5/refvolcyth41h5)
end
```

--- Prepare for distance calculation ---

```lua
function f ( String ) -- function to format the atomlabels
    FormattedString = string.format("%.7s", String)
    return FormattedString
end
function f2 ( String ) -- function to format the atomlabels
    FormattedString = string.format("%.9s", String)
    return FormattedString
end
function d2ox ( index ) -- function to find non-exchangeable protons on x axis
    local Boolean=false
    for x in string.gfind(t.assxlabel[index],"H[12345] [']s") do
        Boolean=true
    end
    for y in string.gfind(t.assxlabel[index],"H[2568]") do
        Boolean=true
    end
    return Boolean
end
function d2oy ( index ) -- function to find non-exchangeable protons on x axis
    local Boolean=false
    for x in string.gfind(t.assxlabel[index],"H[12345] [']s") do
        Boolean=true
    end
    for y in string.gfind(t.assxlabel[index],"H[2568]") do
        Boolean=true
    end
```

--- Complete Lua scripts written for data export from CARA ---
function h2ox ( index ) — function to find non-exchangeable protons on x axis
    local Boolean=false
    for x in string.gfind(t.assxlabel[index],"H[13]") do
        Boolean=true
    end
    return Boolean
end

function h2oy ( index ) — function to find non-exchangeable protons on x axis
    local Boolean=false
    for x in string.gfind(t.assylabel[index],"H[13]") do
        Boolean=true
    end
    return Boolean
end

function HNFx ( index ) — function to find non-exchangeable protons on x axis
    local Boolean=false
    local resid=t.project:getSpin(t.assxnew[index]):getSystem():getResidue():getId()
    if resid==7 or resid==20 then
        Boolean=true
    end
    return Boolean
end

function HNFy ( index ) — function to find non-exchangeable protons on x axis
    local Boolean=false
    local resid=t.project:getSpin(t.assylabel[index]):getSystem():getResidue():getId()
    if resid==7 or resid==20 then
        Boolean=true
    end
    return Boolean
end

--- initialize new tables for distance and the lower and upper limit (same)
t.distance = {}
t.limit = {}
local i = nil
local assx = nil

----------------------------- Distance and Limit calculation -----------------------------

--- select atom pairs corresponding to references and calculate distances and
--- limits and write out new peaklist to file

--- limits are calculated by taking the maximum deviation of the corresponding
--- reference peak times the distance

for i, assx in pairs (t.assxlabel) do
    if (((HNFx(i)==true) or (d2ox(i)==true)) and ((HNFy(i)==true) or (d2oy(i)==true))) then
        t.distance[i] = string.format("%7.2f", refdistcyt*(refvolcyt/t.volnew[i]^(1/6))
        if (t.gradenew[i]==a) then --- error bounds scaled by grading of integration
            t.limit[i] = string.format("%7.1f", t.distance[i]*maxdevcyt)
        else
            t.limit[i] = string.format("%7.1f", t.distance[i]*maxdevcyt)
        end
    end
end
.4 Lua scripts written for data export from CARA

```lua
if (t.gradenew[i]==b) then
    t.limit[i] = string.format("%.1f", t.distance[i]*maxdevcyt*1.2)
else
    t.limit[i] = string.format("%.1f", t.distance[i]*maxdevcyt*1.4)
end
```

```lua
ofile:write (f2('d2o: ')...(i)...t.labelnew[i]..assx..t.assylabel[i]..t.gradenew[i]..t.volnew[i]..t.distance[i]..t.limit[i]..'\n')
elseif ((assx=="(H7)") or (t.assylabel[i]=="(H7)")) then
    t.distance[i] = string.format("%.2f", redistmet*(refvolmet/t.volnew[i])^(1/6))
    if (t.gradenew[i]==a) then — error bounds scaled by grading of integration
        t.limit[i] = string.format("%.1f", t.distance[i]*maxdevmet)
    else
        if (t.gradenew[i]==b) then
            t.limit[i] = string.format("%.1f", t.distance[i]*maxdevmet*1.2)
        else
            if (t.gradenew[i]==b) then
                t.limit[i] = string.format("%.1f", t.distance[i]*maxdevmet*1.4)
            end
        end
    end
end
ofile:write (f2('methyl: ')...(i)...t.labelnew[i]..assx..t.assylabel[i]..t.gradenew[i]..t.volnew[i]..t.distance[i]..t.limit[i]..'\n')
elseif (((HNFx(i)==true) or (h2ox(i)==true)) and ((h2oy(i)==true)) or ((t.assylabel[i]=="(H42)") or (h2o2x(i)==true) or (assx=="(H42)") and (HNPy(i)==true)) or (h2oy(i)==true) or (d2o2x(i)==true) or (d2o2y(i)==true)) then
    t.distance[i] = string.format("%.2f", redistcyth42h5*(refvolcyth42h5/t.volnew[i])^(1/6))
    if (t.gradenew[i]==a) then — error bounds scaled by grading of integration
        t.limit[i] = string.format("%.1f", t.distance[i]*maxdevcyth42h5)
    else
        if (t.gradenew[i]==b) then
            t.limit[i] = string.format("%.1f", t.distance[i]*maxdevcyth42h5*1.2)
        else
            t.limit[i] = string.format("%.1f", t.distance[i]*maxdevcyth42h5*1.4)
        end
    end
end
ofile:write (f2('d2o_h2o: ')...(i)...t.labelnew[i]..assx..t.assylabel[i]..t.gradenew[i]..t.volnew[i]..t.distance[i]..t.limit[i]..'\n')
elseif (((HNFx(i)==true) or (h2ox(i)==true)) and (t.assylabel[i]=="(H41)") or ((assx=="(H41)") or (h2ox(i)==true) or (HNPy(i)==true) or (h2oy(i)==true))) then
    t.distance[i] = string.format("%.2f", redistcyth41h5*(refvolcyth41h5/t.volnew[i])^(1/6))
    if (t.gradenew[i]==a) then — error bounds scaled by grading of integration
        t.limit[i] = string.format("%.1f", t.distance[i]*maxdevcyth41h5)
    else
        if (t.gradenew[i]==b) then
```
Script code

t.limit[i] = string.format("%7.1f",
t.distance[i]*maxdevcyth41h5*1.2)
else
t.limit[i] = string.format("%7.1f",
t.distance[i]*maxdevcyth41h5*1.4)
end

end

outfile: write (f2("H41:
')..f(i)..'..t.labelnew[i]..assx..t.assylabel[i]..t.gradenew[i]
..t.volnew[i]..t.distance[i]..t.limit[i]..'\n")
else — write out peaks not used to file!!!
t.distance[i] = string.format("%7.2f",
refdistcyt*(refvolcyt/t.volnew[i])^(1/6))
if (t.gradenew[i]==a) then — error bounds scaled by grading of integration
  t.limit[i] = string.format("%7.1f", t.distance[i]*maxdevcyt)
else
  if (t.gradenew[i]==b) then
    t.limit[i] = string.format("%7.1f",
t.distance[i]*maxdevcyt*1.2)
  else
    t.limit[i] = string.format("%7.1f",
t.distance[i]*maxdevcyt*1.4)
  end
end

outfile2: write (f2("not used (ref on C H5-H6):
')..f(i)..'..t.labelnew[i]..assx..t.assylabel[i]..
t.gradenew[i]..t.volnew[i]..t.distance[i]..t.limit[i]..'\n")
end

----------------------------------- End of Main Body -----------------------------------

----------------------------------- End of THIRD PART -----------------------------------

----------------------------------- End of FOURTH PART -----------------------------------

----------------------------------- PREPARATIONS -----------------------------------

— open outfile
outfile = io.output(t.Filename..'_.reference_peaks')

local i = 0
function f ( String ) -- function to format the atomlabels
    FormattedString = string.format( "%.2f", String )
    return FormattedString
end

outfile: write
("\n-- non-exchangeable proton cross-peaks: CYT H5-H6\nreference_vol ref_dist
maximum_dev(%) n° refvolcyt f(refdistcyt) stddevcyt maxdevcyt\nPeaklabel Volume Dist Dev\n")
for j, assx in pairs (t.assxnew) do
    for y in string.find (t.labelnew[j], "H[56]/H[56] [0-9]/C[0-9]+") do -- reference for
d2o peaks
        outfile: write
        (t.labelnew[j]..t.volnew[j]..t.distance[j]..f(t.distance[j]-refdistcyt).."\n")
    end
end
outfile: write
("\n-- methyl proton cross-peaks: MET H6-H7\nreference_vol ref_dist standard_dev
maximum_dev(%) n° refvolmet f(refdistmet) stddevmet maxdevmet\nPeaklabel Volume Dist Dev\n")
for j, assx in pairs (t.assxnew) do
    for y in string.find (t.labelnew[j], "H[67]/H[67] [0-9]/T[0-9]+") do -- reference for methyl peaks
        outfile: write
        (t.labelnew[j]..t.volnew[j]..t.distance[j]..f(t.distance[j]-refdistmet).."\n")
    end
end
outfile: write
("\n-- exchangeable proton cross-peaks: CYT H41-H42\nreference_vol ref_dist standard_dev
maximum_dev(%) n° refvolcytamino f(refdistcytamino) stddevcytamino maxdevcytamino\nPeaklabel Volume Dist Dev\n")
for j, assx in pairs (t.assxnew) do
    for y in string.find (t.labelnew[j], "H4[12]/H4[12] [0-9]/C[0-9]+") do -- reference for h2o exchangeable peaks
        outfile: write
        (t.labelnew[j]..t.volnew[j]..t.distance[j]..f(t.distance[j]-refdistcytamino).."\n")
    end
end
outfile: write
("\n-- non-exchangeable/exchangeable proton cross-peaks: CYT H42-H5\nreference_vol ref_dist standard_dev
maximum_dev(%) n° refvolcyth42h5 f(refdistcyth42h5) stddevcyth42h5 maxdevcyth42h5\nPeaklabel Volume Dist Dev\n")

247
Script code

for j, assx in pairs (t. assxnew) do
    for y in string.gfind (t.labelnew[j], "H42/H5 [0-9];C[0-9]+") do
        -- reference for h2o exchangeable-non-exchangeable peaks (appears twice because of selection reasons)
        outfile: write
            (t.labelnew[j]..t.volnew[j]..t.distance[j]..f(t.distance[j]-refdistcyth42h5)..."\n")
    end
end

for j, assx in pairs (t. assxnew) do
    for y in string.gfind (t.labelnew[j], "H5/H42 [0-9];C[0-9]+") do
        -- reference for h2o exchangeable-non-exchangeable peaks
        outfile: write
            (t.labelnew[j]..t.volnew[j]..t.distance[j]..f(t.distance[j]-refdistcyth42h5)..."\n")
    end
end

outfile: write
("\n--- End of Main Body ---\n"

--- close outfile
outfile: close ()
i = 0

--- End of FOURTH PART ---

---

--- FIFTH PART ---

---

--- PREPARATIONS ---

248
.4 Lua scripts written for data export from CARA

```lua
-- open outfile
outfile1 = io.output( t.Filename.."_xplor.list" )
outfile2 = io.output( t.Filename.."_xplor_all.list" )
outfile3 = io.output( 'picktbl...t.Filename')
outfile4 = io.output( t.Filename.."_xplor.noe" )
outfile5 = io.output( t.Filename.."_xplor_all.noe" )
outfile6 = io.output( 'picktbl_all...t.Filename')
outfile7 = io.output( t.Filename.."_xplor_longdist.list" )
outfile8 = io.output( t.Filename.."_xplor_longdist.noe" )
outfile9 = io.output( t.Filename.."_xplor_H1H5sug.list" )
outfile10 = io.output( t.Filename.."_xplor_H1H5sug.noe" )
outfile11 = io.output( 'picktbl_longdist...t.Filename')
outfile12 = io.output( 'picktbl_H1H5sug...t.Filename')

function find (index) -- function to find atomlabels
  local Boolean = false
  local Booleanx = false
  local Booleany = false
  for x in string.gfind(t.assxlabel[index],'H[2345]"["*" ]') do
    Booleanx=true
  end
  for y in string.gfind(t.assxlabel[index],'H[2345]"["*" ]') do
    Booleany=true
  end
  if (Booleanx=true) and (Booleany=true) then
    Boolean=true
  else
    Boolean=false
  end
  return Boolean
end

function find_h1 (index) -- function to format the atomlabels
  local Boolean = false
  local Booleanx = false
  local Booleany = false
  for x in string.gfind(t.assxlabel[index],'H[12345]"["*" ]') do
    Booleanx=true
  end
  for y in string.gfind(t.assxlabel[index],'H[12345]"["*" ]') do
    Booleany=true
  end
  if (Booleanx=true) and (Booleany=true) then
    Boolean=true
  else
    Boolean=false
  end
  return Boolean
end

function find_h1h5sug (index) -- function to format the atomlabels
  local Boolean = false
  local Booleanx = false
  local Booleany = false
  for x in string.gfind(t.assxlabel[index],'H[15]"["*" ]') do
    Booleanx=true
  end
end
```

249
for y in string.gfind(t.assylabel[index],"H\[15\] ' \[ ' ∗") do
  Booleany=true
end
if (Booleanx==true) and (Booleany==true) then
  Boolean=true
else
  Boolean=false
end
return Boolean
end

for i,assx in pairs (t.assxlabel) do — iterate over all peaks
  if (t.assxlabel[i]=="H5") then
    t.assxlabel[i]="HQ5"
  elseif (t.assxlabel[i]=="H2") then
    t.assxlabel[i]="HQ2"
  elseif (t.assxlabel[i]=="H7") then
    t.assxlabel[i]="H7#"
  end
end

for i,assy in pairs (t.assylabel) do — iterate over all peaks
  if (t.assylabel[i]=="H5") then
    t.assylabel[i]="HQ5"
  elseif (t.assylabel[i]=="H2") then
    t.assylabel[i]="HQ2"
  elseif (t.assylabel[i]=="H7") then
    t.assylabel[i]="H7#"
  end
end

for i,assx in pairs (t.assxlabel) do — iterate over all peaks
  if (t.assxlabel[i]=="H5") then
    t.assxlabel[i]="HQ5"
  elseif (t.assxlabel[i]=="H2") then
    t.assxlabel[i]="HQ2"
  elseif (t.assxlabel[i]=="H7") then
    t.assxlabel[i]="H7#"
  end
end

for i,assx in pairs (t.assxlabel) do — iterate over all peaks
  if (t.assxlabel[i]=="H5") then
    t.assxlabel[i]="HQ5"
  elseif (t.assxlabel[i]=="H2") then
    t.assxlabel[i]="HQ2"
  elseif (t.assxlabel[i]=="H7") then
    t.assxlabel[i]="H7#"
  end
end
Lua scripts written for data export from CARA

```lua
name"...t.assxlabel[i].." {resid"...t.assysresid[i].." and
name"...t.assxlabel[i].."}...t.distance[i]...t.limit[i]...t.limit[i].."
added!"n"

outfile8: write (t.assxsresid[i]...t.assxlabel[i]...t.assysresid[i]
t.assxlabel[i]...t.labelnew[i]...t.gradenew[i]...t.distance[i].."\n"
outfile1: write ("pick bond (resid"...t.assxsresid[i].." and
name"...t.assxlabel[i].."} (resid"...t.assysresid[i].." and
name"...t.assxlabel[i].."}...t.distance[i]...t.limit[i]...t.limit[i].."
! added!"n"

outfile8: write (t.assxsresid[i]...t.assxlabel[i]...t.assysresid[i]
t.assxlabel[i]...t.labelnew[i]...t.gradenew[i]...t.distance[i].."\n"
elseif (find_h1h5sug(i)==true) then
  outfile2: write ("assign (resid"...t.assxsresid[i].." and
name"...t.assxlabel[i].."} (resid"...t.assysresid[i].." and
name"...t.assxlabel[i].."}...t.distance[i]...t.limit[i]...t.limit[i].."
! added!"n"

outfile5: write (t.assxsresid[i]...t.assxlabel[i]...t.assysresid[i]
t.assxlabel[i]...t.labelnew[i]...t.gradenew[i]...t.distance[i].."\n"
outfile6: write ("pick bond (resid"...t.assxsresid[i].." and
name"...t.assxlabel[i].."} (resid"...t.assysresid[i].." and
name"...t.assxlabel[i].."}...t.distance[i]...t.limit[i]...t.limit[i].."
! added!"n"
end

end

--- close outfile
outfile: close()
outfile2: close()
outfile3: close()
outfile4: close()
outfile5: close()
outfile6: close()
outfile7: close()
outfile8: close()
outfile9: close()
outfile10: close()
```

.4 Lua scripts written for data export from CARA

--- close outfile
outfile: close()
outfile2: close()
outfile3: close()
outfile4: close()
outfile5: close()
outfile6: close()
outfile7: close()
outfile8: close()
outfile9: close()
outfile10: close()
Script code

outfile11: close()
outfile12: close()
i = 0
t = nil

------------------------------- End of FIFTH PART -------------------------------

print ('\ngenerateinput_byanda is done.' )
print ('\nHave a nice day!' )

------------------------------- End generateinput_byanda -------------------------------
In this section a short collection of the most frequently used utility scripts is presented. They are grouped according to their function and a small description is supplied.

**Scripts for converting program inputs/outputs for further use in other programs**

The next two scripts are used to convert a SPARKY resonance table to a CARA atomlist.

```bash
#!/bin/csh
#
# Skript um aus Sparky-resonance-table eine CARA-atomlist zu machen
#
awk 'NR==1 { print "INAME 1 H"; print "INAME 2 H"}; print $2 " $3" 0 U 0.000 0 - 0' $1
```

The next three scripts are all needed for the NOESY back-calculation with the program GIFA (the first to modify the psf-file, the second to generate the ppm-file needed as input and the last to convert the XPLOR-NIH back-calculation output into GIFA-format.

```bash
#!/bin/csh
#
# Skript um aus Sparky-resonance-table eine CARA-atomlist zu machen
#
awk '{NR=1} {gsub (/h2$ /, "HQ2" , $2) ; gsub (/h5$ /, "HQ5" , $7) ; for (i=1;i<=NF;i++) { printf "%5s toupper($i) }} /me/ { printf "%5s%5s%5s
", "H71", "H72", "H73"} {if (NR==1) { printf "$0" "SNF" "SNF"} "$1.shifts" > blu
```

The next three scripts are all needed for the NOESY back-calculation with the program GIFA (the first to modify the psf-file, the second to generate the ppm-file needed as input and the last to convert the XPLOR-NIH back-calculation output into GIFA-format.

```bash
#!/bin/csh
#
# Dateiname des PSF-Files ohne Extension muss uebergeben werden.
# Es werden nur CYT, Thy, Ade, Gua selektiert, wenn andere
# Modifikationen vorhanden entsprechend mit einfuehren.
#
awk '{if (( $(1).[0-9]+) & $(3-/'a-zA-Z'+$) & $(3-/'GUA|ADE|CYT|THY|PUR') } { print "%14 s%5s%5s%5s%6s%16s
", $1, $2, $3, $4, $5, $6, $7, $8; NR=1} else print' $1.psff > $1_spect.psff
```

```bash
BEGIN { getline VAR < "'"$1".'ppm'; split (VAR,a)
```
Script code

```csh
#!/bin/csh

# Script code

if ($4 =~ ^H/) { if ([$4]==$3) { getline VAR < "$1.ppm" ; split (VAR,a) ; print "$1 $4 $5 $6 $7 $8" ; getline VAR < "$2.ppm" ; split (VAR,a) ; for (i=1; i<=211; i++) { if ([$4]==a[i] && $5==a[i] && $6==a[i] && $7==a[i]) { close ("$2.ppm") ; close ("$1.ppm") ; close ("$1.gifa.pdb") ; break } else { getline VAR < "$1.ppm" ; split (VAR,a) ; close ("$1.ppm") ; close ("$1.gifa.pdb") ; close ("$1.gifa.spect") ; break } } ; close ("$1.ppm") ; close ("$1.gifa.pdb") ; close ("$1.gifa.spect") ; break } else print "$1 $4 $5 $6 $7 $8"
```
The next script is used to generate and XPLOR-NIH RDC restraints file from the input to the PALES-program.

```bash
#!/bin/csh
# used to convert pales input to xplor RDC input (with axes2.pdb and axis_500.psf)
awk '{ if (NR==1) { print " ! RDC table " } ; if ( (NR>3)&&(NF==10)) { print " assign ( resid 500 and name OO ) \n ( resid 500 and name Z ) \n ( resid 500 and name X ) \n ( resid "$1" and name "$3" ) \n ( resid "$1" and name "$6" ) " $ (NF-3) "$ (NF-2) "$ (NF-2)$ (n\n")) } \n" > $1.xplor
}
```

The next two scripts are used for generating random data subsets from RDC and NOE restraints files, respectively.

```bash
#!/bin/csh
# script to pick out every fourth ( or fifth etc ) NOE point from input
set VAR='echo $1 '
awk '{ if ( $1 ~/^ assign /) print } ' $1 > bla
awk 'BEGIN {DUMMY=5} { if (NR == DUMMY) {DUMMY = DUMMY+3; print > "$VAR" unused} else print } ' bla > $1.third
rm bla
```

```bash
#!/bin/csh
# script to pick out every fourth ( or fifth etc ) RDC point from PALES input
set VAR='echo $1 '
awk 'BEGIN {DUMMY=6; print "VARS RESID_I RESNAME_I ATOMNAME_I RESID_J RESNAME_J ATOMNAME_J D DD W FORMATT %5d %6s %6s %5d %6s %6s %9.3f %9.3f %.2f " > "$VAR" unused" } { if ( (NR>3)&&(NR == DUMMY) ) {DUMMY = DUMMY+3; print > "$VAR" unused} else print } ' $1 > $1.third
```

The next script is used for generating a single pdb-file from the average and the 10 minimum-energy structures for submission to the PDB databank.

```bash
#!/bin/csh
# script for generating pdb submission file with averaged, minimized structures as model 1
# and the 10 minimum-energy structures as models 2–11. TER cards, chainIDs are inserted
# and ANI residues deleted.

set pdbbs = 'awk '{ if ( $1 ~/pdb /) { printf "%s ",$1 } } ' *_##.pdb.stats'
cat average_min.pdb $pdbbs > bla
awk 'BEGIN { print "MODEL 1"; COUNT1=0;COUNT2=2} { if ( $1=="REMARK" ) print ; if ( (($1="ATOM")&(85<=13))&(84!="ANI")) printf "%5s%6s%6s%6s%6s%6s%6s%6s%6s%6s%6s\n" , $1 , $2 , $3 , $4 , $5 , $6 , $7 , $8 , $9 , $10 , "$ ; next }; if ( (($1=="ATOM")&(85>=14))&(84!="ANI")) { if (COUNT1=0) { print "TER"; printf "%5s%6s%6s%6s%6s%6s%6s%6s%6s%6s%6s\n" , $1 , $2 , $3 , $4 , $5 , $6 , $7 , $8 , $9 , $10 , "$ ; next }); if (COUNT1=1) { printf "%5s%6s%6s%6s%6s%6s%6s%6s%6s%6s%6s\n" , "TER", "ENDMDL", "MODEL",COUNT2=COUNT2+1;COUNT1=0;next}; if
```

255
The next script calls the program XPLOR-VMD with the 10 minimum-energy structures of a calculation for display.

```bash
#!/bin/csh
# script for displaying top 10 structures of xplor-nih python script in avmd
set pdbs = 'awk '{ if ( $1~/pdb/) { printf "%s ", $1;}}' *##.pdb.stats'
vmd-xplor $pdbs
unset pdbs
```

The next script collects the CURVES helical parameter calculation out of the 3DNA output files that were generated for the 10 minimum-energy structures of a XPLOR-NIH calculation and prints them to one file.

```bash
#!/bin/csh
# Einmaliges Skript zum umnummerieren von schon vorhandenen Dateien
set i=1
while ($i<= 10)
  grep -A 17 'Curves' $i/cf_7methods.par >> comp_hel_par.tab
  @ i++
end
unset i
```

The next script selects just a subset of NOEs from an NOE input file, the rest is commented out.

```bash
#!/bin/csh
# uebergabe des .tbl files ohne extension
# awk '{ if ( $1=='assign') { printf "%s
",$0} else else { $1="assign"; printf "%s
",$0}) else
  # $1='\"assign\"' printf "%s
",$0}) $1_2.tbl > bla
awk '$1==81 { printf "%-9s%-10s%-3s%-4s%-5s%-6s%-10s%-3s%-4s%-5s%-12s%-%12s%-%8s%-%8s%-%4s
", $1, $2, $3, $4, $5, $6, $7, $8, $9, $10, $11, $12, $13, $14, $15, $16} bl > $1_3.tbl'
rm bla
#($(66-/[hH] 6) &$(66-/[hH] 6) &$(66-/[hH] 6))
```

256
.5 Utility scripts

Scripts for renaming atoms in files

The script is used to rename certain atoms in an input file to XPLOR-NIH calculations.

```bash
#!/bin/csh
sort $2 > bla
sort $1.list > blal2
awk '{$1="display"; gsub("H5","HQ5",$0); gsub("H2","HQ2",$0); gsub("H7","H7#",$0); print $0"display  \$result"} ' bla > $2
awk '{gsub("H5","HQ5",$0); gsub("H2","HQ2",$0); gsub("H7","H7#",$0); print $0} ' blal2 > $1.tb1
rm blal*
```

The script is used to rename certain atoms in a whole family of calculated structures.

```bash
#!/bin/csh
set i=1
while (i<=100)
  foreach i (*.pdb)
    cp $i blai
    sed ' s/H2/HNF/HNF/g' $i > "blai"$i
    sed ' s/H2/HNF/HNF/g' "blai"$i > "blabla"$i
    sed ' s/H2bb/HNF/HNF/g' "blabla"$i > "13mer_HNF_"$i "pdb"
  end
  awk '{ if ( ($1="REMARK") || ($2==208) || ($2==235) ) { print } else { if ($2==209) {gsub($2,$2+24);DUMMY1=$0; next}; if ($2==210) {gsub($2,$2+24);DUMMY2=$0; next}; if ($2==211) {gsub($2,$2+24);DUMMY3=$0; next}; if ($2==235) {gsub($2,$2-4); print; next} } } } ' blal * > $i
  sed ' s/H2/HNF/HNF/g' "13mer_HNF_"$i "pdb"
  sed ' s/H2/HNF/HNF/g' "blabla"$i > "13mer_HNF_"$i "pdb"
  @ i++
end
rm blai*
```

Scripts for displaying specific information content from files

This script prints out the number of intra-, interresidual and total number of NOEs, as well as the NOEs to the modification site.

```bash
#!/bin/csh
#
# Skript gibt Anzahl der NOE, der intra- und interresidualen und der NOEs zur Modifikation
# (Achtung nur Position 7) aus
awk 'BEGIN {INTER=0;INTRA=0;MOD=0} { if ($1="assign") { DUMMY=0} else { if ($3==$8) { INTRA=INTRA+1} else { if ($3==7) || ($8==7) || ($3==20) || ($8==20) } } } END {TOTAL=INTER+INTRA; print "Total: " TOTAL; print " intraresidual: " INTRA; print " interresidual: " INTER; print " NOE to modification site = " MOD }' $1
```

257
This script prints out the Da and Rh values from all pdb-files in the working directory.

```csh
#!/ bin / csh
# script for displaying Da values of top 10 structures of xplor−nih python script
set pdbs = 'awk '{ i f ( $1 /~pdb/) { printf "%s " , $1 " "}} ' "_##.pdb.stats"

grep "Da: " $pdbs
unset pdbs
```

This script prints out the sorted energies from all pdb-files in the working directory.

```csh
#!/ bin / csh
grep "summary total" * . pdb | awk '{ print $4 " "$5 " "$1 } ' | sort −n
```

This script prints out the sorted energies from just a subset of all pdb-files in the working directory.

```csh
#!/ bin / csh
#
# Skript um die 10 niedrigsten Energien auszugeben.
#
set i =1
grep "total" refine*.pdb | awk '{ print $5 " "$1 " "$2 " "$3 } ' | sort −n | head −50 | awk '{ print $5 " "$1 " "$2 " "$3 } ' | sort −n | head −50 | awk '{ print $3 " "$4 } ' > 50minen
cat 50minen

while ( $i<= 50)
  set VAR='awk 'BEGIN {FS="."} { i f (NR=="$i") printf "%s " , $1 } ' 50minen'
  cp $VAR".pdb "" min10_ "$i".pdb"
  @ i++
end
#rm bla
```

This script generates files which contain the NOE restraints information (which was used as input for the structure calculations) with the corresponding distances in the 10 minimum-energy structures. Thus not only NOE violations but also too tight NOE restraints can be easily identified.

```csh
#!/ bin / csh
#
# Skript um die 10 niedrigsten Energien auszugeben.
# Anschliessend werden die Bindungslaenge der 10 energieniedrigsten Strukturen fuer die
# noe−constraints
# hinter die eingegebenen noe−constraints ausgeschrieben.
# Es muss der Name des noe−constraint−files ohne die extension tbl angegeben werden.
# Es werden mehrere Files ausgegeben die die Auswertung der Strukturen erleichtern
#
set i =1
set VAR='echo "$1"'
```
.5 Utility scripts

grep "enviol" *.pdb | awk '{ print $1" "$2" "$4" "$5 }' | sort -n | awk '{ print $2 " "$3 " "$4 " "$5 }' | head -10

grep "enviol" *.pdb | awk '{ print $1" "$2" "$4" "$5 }' | sort -n | awk '{ print $2 " "$3 " "$4 " "$5 }' | head -10 > blaslabla

awk '{ printf "%20s %5s %5s %5s %5s %5s %5s %5s %5s %5s %5s %5s %5s %5s %5s %5s %5s %5s %5s

" $11 , $12 , " $ (NF−1) , " $NF; if ( $1 ~ / ^ \(! / )

{ printf "%5s \n \n on" } else printf "%5s \n \n on" } $VAR > blaslabla

while ( $i<= 10)

set FILE='awk 'BEGIN {FS="."} { if (NR=="$i") print $1 }' blaslabla ' .noe'

echo $FILE

awk '{ printf "%s; $0; getline < " $FILE"; printf "%8.4f\n."; blasi > bonds.log

awk '{ printf "%s; $0; REF=$3; getline < " $FILE"; DEV=REF-$1; printf "%8.4f\n."; }}'

blasi > noedev.log

@ i++

mv bonds.log blasi

mv noedev.log blasi

awk '{ printf "%s\n

" $0 , " RMSD_AVE: "RMS, "RMSD_NOE: "RMSNOE }' blasi > rmsd.log

rm blaslabla

awk '{ printf "%s\n

" $0 , " RMSD_AVE: "RMS, "RMSD_NOE: "RMSNOE }' blasi > rmsd.log

rm blaslabla

awk '{ printf "%s\n

" $0 , " RMSD_AVE: "RMS, "RMSD_NOE: "RMSNOE }' blasi > rmsd.log

rm blaslabla

awk '{ printf "%s\n

" $0 , " RMSD_AVE: "RMS, "RMSD_NOE: "RMSNOE }' blasi > rmsd.log

rm blaslabla

awk '{ printf "%s\n

" $0 , " RMSD_AVE: "RMS, "RMSD_NOE: "RMSNOE }' blasi > rmsd.log

rm blaslabla
Chemical shifts

In the following the chemical shifts are listed. All proton chemical shifts are referenced to the HOD signal at 4.80 ppm.

.6 Chemical shifts of 13merHNF

The chemical shifts which are unique to HNF and/or the abasic site are not listed in the table and thus are given below in ppm.

HNF (Res 7): H1" 5.32, H4 6.78, C1 109.8, C3 113.7, C4 121.2

Abasic site (Res 20): H1" 4.05
Tabelle 2: $^1H$ chemical shifts of sugar protons of 13merHNF

<table>
<thead>
<tr>
<th>Res</th>
<th>H1'</th>
<th>H2'</th>
<th>H2''</th>
<th>H3'</th>
<th>H4'</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>6.00</td>
<td>2.67</td>
<td>2.79</td>
<td>4.86</td>
<td>4.28</td>
</tr>
<tr>
<td>C2</td>
<td>6.09</td>
<td>2.14</td>
<td>2.54</td>
<td>4.85</td>
<td>4.27</td>
</tr>
<tr>
<td>T3</td>
<td>5.75</td>
<td>2.13</td>
<td>2.45</td>
<td>4.89</td>
<td>4.16</td>
</tr>
<tr>
<td>G4</td>
<td>5.84</td>
<td>2.65</td>
<td>2.69</td>
<td>5.00</td>
<td>4.38</td>
</tr>
<tr>
<td>C5</td>
<td>5.47</td>
<td>1.97</td>
<td>2.31</td>
<td>4.82</td>
<td>4.16</td>
</tr>
<tr>
<td>A6</td>
<td>6.34</td>
<td>2.82</td>
<td>2.78</td>
<td>5.06</td>
<td>4.40</td>
</tr>
<tr>
<td>A7</td>
<td>5.32</td>
<td>2.30</td>
<td>2.38</td>
<td>4.88</td>
<td>4.65</td>
</tr>
<tr>
<td>A8</td>
<td>6.09</td>
<td>2.51</td>
<td>2.81</td>
<td>4.98</td>
<td>4.43</td>
</tr>
<tr>
<td>C9</td>
<td>5.50</td>
<td>1.95</td>
<td>2.34</td>
<td>4.82</td>
<td>4.15</td>
</tr>
<tr>
<td>G10</td>
<td>5.96</td>
<td>2.62</td>
<td>2.78</td>
<td>4.96</td>
<td>4.37</td>
</tr>
<tr>
<td>T11</td>
<td>6.05</td>
<td>2.10</td>
<td>2.47</td>
<td>4.88</td>
<td>4.23</td>
</tr>
<tr>
<td>C12</td>
<td>5.72</td>
<td>2.04</td>
<td>2.38</td>
<td>4.86</td>
<td>4.14</td>
</tr>
<tr>
<td>G13</td>
<td>6.18</td>
<td>2.64</td>
<td>2.40</td>
<td>4.71</td>
<td>4.20</td>
</tr>
<tr>
<td>C14</td>
<td>5.72</td>
<td>1.85</td>
<td>2.36</td>
<td>4.70</td>
<td>4.07</td>
</tr>
<tr>
<td>G15</td>
<td>5.45</td>
<td>2.73</td>
<td>2.80</td>
<td>5.01</td>
<td>4.33</td>
</tr>
<tr>
<td>A16</td>
<td>6.27</td>
<td>2.72</td>
<td>2.93</td>
<td>5.08</td>
<td>4.50</td>
</tr>
<tr>
<td>C17</td>
<td>5.58</td>
<td>1.95</td>
<td>2.32</td>
<td>4.83</td>
<td>4.16</td>
</tr>
<tr>
<td>G18</td>
<td>5.94</td>
<td>2.60</td>
<td>2.76</td>
<td>4.98</td>
<td>4.37</td>
</tr>
<tr>
<td>T19</td>
<td>6.11</td>
<td>2.35</td>
<td>2.53</td>
<td>4.94</td>
<td>4.23</td>
</tr>
<tr>
<td>T20</td>
<td>4.15</td>
<td>2.26</td>
<td>2.27</td>
<td>4.77</td>
<td>4.11</td>
</tr>
<tr>
<td>T21</td>
<td>5.79</td>
<td>2.00</td>
<td>2.45</td>
<td>4.91</td>
<td>4.33</td>
</tr>
<tr>
<td>G22</td>
<td>5.75</td>
<td>2.58</td>
<td>2.66</td>
<td>4.97</td>
<td>4.33</td>
</tr>
<tr>
<td>C23</td>
<td>5.45</td>
<td>1.97</td>
<td>2.31</td>
<td>4.82</td>
<td>4.16</td>
</tr>
<tr>
<td>A24</td>
<td>6.01</td>
<td>2.74</td>
<td>2.90</td>
<td>5.05</td>
<td>4.39</td>
</tr>
<tr>
<td>G25</td>
<td>5.82</td>
<td>2.48</td>
<td>2.66</td>
<td>4.97</td>
<td>4.36</td>
</tr>
<tr>
<td>C26</td>
<td>6.10</td>
<td>2.15</td>
<td>2.22</td>
<td>4.47</td>
<td>4.05</td>
</tr>
</tbody>
</table>
### Tabelle 3: $^1$H chemical shifts of base protons of 13merHNF

<table>
<thead>
<tr>
<th>Res</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H41</th>
<th>H42</th>
<th>H5</th>
<th>H6</th>
<th>H7</th>
<th>H8</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.97</td>
</tr>
<tr>
<td>C2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.32</td>
<td>6.62</td>
<td>5.36</td>
<td>7.53</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T3</td>
<td>-</td>
<td>-</td>
<td>13.96</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.34</td>
<td>1.64</td>
<td>-</td>
</tr>
<tr>
<td>G4</td>
<td>12.70</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.89</td>
</tr>
<tr>
<td>C5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.21</td>
<td>6.40</td>
<td>5.43</td>
<td>7.37</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6</td>
<td>-</td>
<td>7.10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.36</td>
</tr>
<tr>
<td>A7</td>
<td>5.97</td>
<td>-</td>
<td>6.00</td>
<td>-</td>
<td>-</td>
<td>7.04</td>
<td>7.45</td>
<td>-</td>
<td>7.71</td>
</tr>
<tr>
<td>A8</td>
<td>-</td>
<td>7.29</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.99</td>
</tr>
<tr>
<td>C9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.98</td>
<td>6.38</td>
<td>5.09</td>
<td>7.16</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G10</td>
<td>12.73</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.83</td>
</tr>
<tr>
<td>T11</td>
<td>-</td>
<td>-</td>
<td>13.76</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.29</td>
<td>1.41</td>
<td>-</td>
</tr>
<tr>
<td>C12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.63</td>
<td>7.03</td>
<td>5.72</td>
<td>7.51</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.96</td>
</tr>
<tr>
<td>C14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.17</td>
<td>7.01</td>
<td>5.89</td>
<td>7.60</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T15</td>
<td>12.95</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.96</td>
</tr>
<tr>
<td>A16</td>
<td>-</td>
<td>7.90</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.23</td>
</tr>
<tr>
<td>C17</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.09</td>
<td>6.42</td>
<td>5.22</td>
<td>7.19</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G18</td>
<td>12.60</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.84</td>
</tr>
<tr>
<td>T19</td>
<td>-</td>
<td>-</td>
<td>13.21</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.29</td>
<td>1.50</td>
<td>-</td>
</tr>
<tr>
<td>T20</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T21</td>
<td>-</td>
<td>-</td>
<td>12.77</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.29</td>
<td>1.51</td>
<td>-</td>
</tr>
<tr>
<td>G22</td>
<td>12.54</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.83</td>
</tr>
<tr>
<td>C23</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.26</td>
<td>6.35</td>
<td>5.38</td>
<td>7.35</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A24</td>
<td>-</td>
<td>7.68</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.18</td>
</tr>
<tr>
<td>G25</td>
<td>12.93</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.68</td>
</tr>
<tr>
<td>C26</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.14</td>
<td>6.57</td>
<td>5.24</td>
<td>7.33</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
### Chemical shifts

#### Tabelle 4: $^{13}$C chemical shifts of 13merHNF

<table>
<thead>
<tr>
<th>Res</th>
<th>C1'</th>
<th>C3'</th>
<th>C2</th>
<th>C5</th>
<th>C6</th>
<th>C8</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>82.2</td>
<td>76.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.7</td>
</tr>
<tr>
<td>C2</td>
<td>83.9</td>
<td>73.9</td>
<td>-</td>
<td>95.5</td>
<td>140.0</td>
<td>-</td>
</tr>
<tr>
<td>T3</td>
<td>82.8</td>
<td>75.3</td>
<td>-</td>
<td>-</td>
<td>136.4</td>
<td>-</td>
</tr>
<tr>
<td>G4</td>
<td>81.4</td>
<td>76.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.2</td>
</tr>
<tr>
<td>C5</td>
<td>83.3</td>
<td>73.7</td>
<td>-</td>
<td>95.4</td>
<td>139.5</td>
<td>-</td>
</tr>
<tr>
<td>A6</td>
<td>82.3</td>
<td>76.7</td>
<td>151.3</td>
<td>-</td>
<td>-</td>
<td>139.1</td>
</tr>
<tr>
<td>A7</td>
<td>102.0</td>
<td>76.5</td>
<td>118.1</td>
<td>121.9</td>
<td>119.2</td>
<td>-</td>
</tr>
<tr>
<td>A8</td>
<td>82.7</td>
<td>75.2</td>
<td>150.5</td>
<td>-</td>
<td>-</td>
<td>138.3</td>
</tr>
<tr>
<td>C9</td>
<td>83.1</td>
<td>73.8</td>
<td>-</td>
<td>95.1</td>
<td>139.1</td>
<td>-</td>
</tr>
<tr>
<td>G10</td>
<td>81.9</td>
<td>76.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.2</td>
</tr>
<tr>
<td>T11</td>
<td>82.7</td>
<td>75.0</td>
<td>-</td>
<td>-</td>
<td>135.8</td>
<td>-</td>
</tr>
<tr>
<td>C12</td>
<td>85.1</td>
<td>75.2</td>
<td>-</td>
<td>96.0</td>
<td>140.7</td>
<td>-</td>
</tr>
<tr>
<td>G13</td>
<td>81.8</td>
<td>70.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>136.2</td>
</tr>
<tr>
<td>C14</td>
<td>83.7</td>
<td>75.0</td>
<td>-</td>
<td>96.6</td>
<td>140.2</td>
<td>-</td>
</tr>
<tr>
<td>G15</td>
<td>81.2</td>
<td>76.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.9</td>
</tr>
<tr>
<td>A16</td>
<td>82.1</td>
<td>76.6</td>
<td>152.6</td>
<td>-</td>
<td>-</td>
<td>138.1</td>
</tr>
<tr>
<td>C17</td>
<td>83.1</td>
<td>73.7</td>
<td>-</td>
<td>95.4</td>
<td>138.9</td>
<td>-</td>
</tr>
<tr>
<td>G18</td>
<td>82.1</td>
<td>76.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.3</td>
</tr>
<tr>
<td>T19</td>
<td>82.7</td>
<td>73.7</td>
<td>-</td>
<td>-</td>
<td>135.8</td>
<td>-</td>
</tr>
<tr>
<td>T20</td>
<td>67.2</td>
<td>68.8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T21</td>
<td>83.5</td>
<td>75.2</td>
<td>-</td>
<td>-</td>
<td>135.9</td>
<td>-</td>
</tr>
<tr>
<td>G22</td>
<td>81.2</td>
<td>76.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.1</td>
</tr>
<tr>
<td>C23</td>
<td>83.5</td>
<td>73.6</td>
<td>-</td>
<td>95.4</td>
<td>139.9</td>
<td>-</td>
</tr>
<tr>
<td>A24</td>
<td>82.1</td>
<td>76.6</td>
<td>151.4</td>
<td>-</td>
<td>-</td>
<td>138.5</td>
</tr>
<tr>
<td>G25</td>
<td>81.2</td>
<td>76.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>134.4</td>
</tr>
<tr>
<td>C26</td>
<td>84.0</td>
<td>68.6</td>
<td>-</td>
<td>95.1</td>
<td>139.7</td>
<td>-</td>
</tr>
</tbody>
</table>
### Chemical shifts of 13merRefGC

<table>
<thead>
<tr>
<th>Res</th>
<th>H1'</th>
<th>H2'</th>
<th>H2&quot;</th>
<th>H3'</th>
<th>H4'</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>6.02</td>
<td>2.68</td>
<td>2.81</td>
<td>4.87</td>
<td>4.29</td>
</tr>
<tr>
<td>C2</td>
<td>6.11</td>
<td>2.15</td>
<td>2.56</td>
<td>4.87</td>
<td>4.29</td>
</tr>
<tr>
<td>T3</td>
<td>5.77</td>
<td>2.17</td>
<td>2.49</td>
<td>4.91</td>
<td>4.18</td>
</tr>
<tr>
<td>G4</td>
<td>5.88</td>
<td>2.65</td>
<td>2.72</td>
<td>5.01</td>
<td>4.39</td>
</tr>
<tr>
<td>C5</td>
<td>5.48</td>
<td>1.93</td>
<td>2.30</td>
<td>4.82</td>
<td>4.15</td>
</tr>
<tr>
<td>A6</td>
<td>5.94</td>
<td>2.70</td>
<td>2.87</td>
<td>5.04</td>
<td>4.37</td>
</tr>
<tr>
<td>A7</td>
<td>5.46</td>
<td>2.57</td>
<td>2.69</td>
<td>5.00</td>
<td>4.36</td>
</tr>
<tr>
<td>A8</td>
<td>6.18</td>
<td>2.60</td>
<td>2.89</td>
<td>5.02</td>
<td>4.45</td>
</tr>
<tr>
<td>C9</td>
<td>5.58</td>
<td>2.00</td>
<td>2.35</td>
<td>4.81</td>
<td>4.15</td>
</tr>
<tr>
<td>G10</td>
<td>5.97</td>
<td>2.61</td>
<td>2.79</td>
<td>4.94</td>
<td>4.37</td>
</tr>
<tr>
<td>T11</td>
<td>6.06</td>
<td>2.10</td>
<td>2.48</td>
<td>4.89</td>
<td>4.24</td>
</tr>
<tr>
<td>C12</td>
<td>5.73</td>
<td>2.04</td>
<td>2.39</td>
<td>4.86</td>
<td>4.14</td>
</tr>
<tr>
<td>G13</td>
<td>6.19</td>
<td>2.65</td>
<td>2.40</td>
<td>4.71</td>
<td>4.21</td>
</tr>
<tr>
<td>C14</td>
<td>5.74</td>
<td>1.88</td>
<td>2.37</td>
<td>4.71</td>
<td>4.08</td>
</tr>
<tr>
<td>G15</td>
<td>5.47</td>
<td>2.74</td>
<td>2.82</td>
<td>5.02</td>
<td>4.33</td>
</tr>
<tr>
<td>A16</td>
<td>6.28</td>
<td>2.73</td>
<td>2.94</td>
<td>5.08</td>
<td>4.50</td>
</tr>
<tr>
<td>C17</td>
<td>5.60</td>
<td>2.03</td>
<td>2.38</td>
<td>4.83</td>
<td>4.18</td>
</tr>
<tr>
<td>G18</td>
<td>5.97</td>
<td>2.61</td>
<td>2.80</td>
<td>4.94</td>
<td>4.38</td>
</tr>
<tr>
<td>T19</td>
<td>6.04</td>
<td>2.19</td>
<td>2.55</td>
<td>4.89</td>
<td>4.27</td>
</tr>
<tr>
<td>T20</td>
<td>6.00</td>
<td>2.11</td>
<td>2.51</td>
<td>4.81</td>
<td>4.19</td>
</tr>
<tr>
<td>T21</td>
<td>5.71</td>
<td>2.11</td>
<td>2.46</td>
<td>4.88</td>
<td>4.15</td>
</tr>
<tr>
<td>G22</td>
<td>5.85</td>
<td>2.66</td>
<td>2.71</td>
<td>4.99</td>
<td>4.37</td>
</tr>
<tr>
<td>C23</td>
<td>5.49</td>
<td>1.98</td>
<td>2.32</td>
<td>4.83</td>
<td>4.14</td>
</tr>
<tr>
<td>A24</td>
<td>6.01</td>
<td>2.74</td>
<td>2.89</td>
<td>5.05</td>
<td>4.39</td>
</tr>
<tr>
<td>G25</td>
<td>5.83</td>
<td>2.49</td>
<td>2.66</td>
<td>4.97</td>
<td>4.36</td>
</tr>
<tr>
<td>C26</td>
<td>6.13</td>
<td>2.15</td>
<td>2.22</td>
<td>4.47</td>
<td>4.06</td>
</tr>
</tbody>
</table>
### Tabelle 6: $^1$H chemical shifts of base protons of 13merRefGC

<table>
<thead>
<tr>
<th>Res</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H41</th>
<th>H42</th>
<th>H5</th>
<th>H6</th>
<th>H7</th>
<th>H8</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.00</td>
</tr>
<tr>
<td>C2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.33</td>
<td>6.57</td>
<td>5.39</td>
<td>7.56</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T3</td>
<td>-</td>
<td>-</td>
<td>13.92</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.37</td>
<td>1.67</td>
<td>-</td>
</tr>
<tr>
<td>G4</td>
<td>12.92</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.91</td>
</tr>
<tr>
<td>C5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.33</td>
<td>6.64</td>
<td>5.40</td>
<td>7.34</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6</td>
<td>-</td>
<td>7.52</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.15</td>
</tr>
<tr>
<td>A7</td>
<td>12.59</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.69</td>
</tr>
<tr>
<td>A8</td>
<td>-</td>
<td>7.76</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.09</td>
</tr>
<tr>
<td>C9</td>
<td>-</td>
<td>-</td>
<td>8.06</td>
<td>6.40</td>
<td>5.17</td>
<td>7.18</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G10</td>
<td>12.70</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.82</td>
</tr>
<tr>
<td>T11</td>
<td>-</td>
<td>-</td>
<td>13.76</td>
<td>-</td>
<td>-</td>
<td>7.29</td>
<td>1.40</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C12</td>
<td>-</td>
<td>-</td>
<td>8.61</td>
<td>7.00</td>
<td>5.72</td>
<td>7.52</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.97</td>
</tr>
<tr>
<td>C14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.91</td>
<td>7.62</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G15</td>
<td>12.94</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.98</td>
</tr>
<tr>
<td>A16</td>
<td>-</td>
<td>7.89</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.23</td>
</tr>
<tr>
<td>C17</td>
<td>-</td>
<td>-</td>
<td>8.12</td>
<td>6.45</td>
<td>5.24</td>
<td>7.22</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G18</td>
<td>12.74</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.82</td>
</tr>
<tr>
<td>T19</td>
<td>-</td>
<td>-</td>
<td>13.66</td>
<td>-</td>
<td>-</td>
<td>7.27</td>
<td>1.37</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T20</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.33</td>
<td>6.89</td>
<td>5.60</td>
<td>7.59</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T21</td>
<td>-</td>
<td>-</td>
<td>13.79</td>
<td>-</td>
<td>-</td>
<td>7.34</td>
<td>1.67</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G22</td>
<td>12.69</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.90</td>
</tr>
<tr>
<td>C23</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.29</td>
<td>6.56</td>
<td>5.40</td>
<td>7.37</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A24</td>
<td>-</td>
<td>7.64</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.19</td>
</tr>
<tr>
<td>G25</td>
<td>12.72</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.69</td>
</tr>
<tr>
<td>C26</td>
<td>-</td>
<td>-</td>
<td>8.30</td>
<td>6.57</td>
<td>5.32</td>
<td>7.38</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
### Chemical shifts of 13merRef

<table>
<thead>
<tr>
<th>Res</th>
<th>H1’</th>
<th>H2’</th>
<th>H2”</th>
<th>H3’</th>
<th>H4’</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>6.02</td>
<td>2.68</td>
<td>2.81</td>
<td>4.87</td>
<td>4.28</td>
</tr>
<tr>
<td>C2</td>
<td>6.11</td>
<td>2.15</td>
<td>2.55</td>
<td>4.86</td>
<td>4.28</td>
</tr>
<tr>
<td>T3</td>
<td>5.76</td>
<td>2.15</td>
<td>2.47</td>
<td>4.90</td>
<td>4.17</td>
</tr>
<tr>
<td>G4</td>
<td>5.86</td>
<td>2.63</td>
<td>2.70</td>
<td>4.99</td>
<td>4.38</td>
</tr>
<tr>
<td>C5</td>
<td>5.41</td>
<td>1.84</td>
<td>2.22</td>
<td>4.79</td>
<td>4.11</td>
</tr>
<tr>
<td>A6</td>
<td>5.75</td>
<td>2.71</td>
<td>2.84</td>
<td>5.03</td>
<td>4.36</td>
</tr>
<tr>
<td>A7</td>
<td>5.84</td>
<td>2.63</td>
<td>2.84</td>
<td>5.05</td>
<td>4.43</td>
</tr>
<tr>
<td>A8</td>
<td>6.05</td>
<td>2.54</td>
<td>2.83</td>
<td>5.00</td>
<td>4.44</td>
</tr>
<tr>
<td>C9</td>
<td>5.56</td>
<td>1.92</td>
<td>2.32</td>
<td>4.77</td>
<td>4.14</td>
</tr>
<tr>
<td>G10</td>
<td>5.95</td>
<td>2.61</td>
<td>2.78</td>
<td>4.95</td>
<td>4.36</td>
</tr>
<tr>
<td>T11</td>
<td>6.05</td>
<td>2.10</td>
<td>2.47</td>
<td>4.88</td>
<td>4.23</td>
</tr>
<tr>
<td>C12</td>
<td>5.72</td>
<td>2.03</td>
<td>2.38</td>
<td>4.85</td>
<td>4.14</td>
</tr>
<tr>
<td>G13</td>
<td>6.18</td>
<td>2.64</td>
<td>2.40</td>
<td>4.70</td>
<td>4.20</td>
</tr>
<tr>
<td>C14</td>
<td>5.73</td>
<td>1.87</td>
<td>2.37</td>
<td>4.70</td>
<td>4.07</td>
</tr>
<tr>
<td>G15</td>
<td>5.46</td>
<td>2.73</td>
<td>2.80</td>
<td>5.01</td>
<td>4.33</td>
</tr>
<tr>
<td>A16</td>
<td>6.27</td>
<td>2.71</td>
<td>2.92</td>
<td>5.08</td>
<td>4.49</td>
</tr>
<tr>
<td>C17</td>
<td>5.56</td>
<td>1.99</td>
<td>2.36</td>
<td>4.82</td>
<td>4.17</td>
</tr>
<tr>
<td>G18</td>
<td>6.00</td>
<td>2.62</td>
<td>2.82</td>
<td>4.96</td>
<td>4.39</td>
</tr>
<tr>
<td>T19</td>
<td>6.04</td>
<td>2.14</td>
<td>2.61</td>
<td>4.87</td>
<td>4.28</td>
</tr>
<tr>
<td>T20</td>
<td>6.13</td>
<td>2.17</td>
<td>2.62</td>
<td>4.91</td>
<td>4.19</td>
</tr>
<tr>
<td>T21</td>
<td>5.84</td>
<td>2.10</td>
<td>2.48</td>
<td>4.91</td>
<td>4.15</td>
</tr>
<tr>
<td>G22</td>
<td>5.82</td>
<td>2.65</td>
<td>2.67</td>
<td>5.00</td>
<td>4.38</td>
</tr>
<tr>
<td>C23</td>
<td>5.49</td>
<td>1.98</td>
<td>2.32</td>
<td>4.83</td>
<td>4.16</td>
</tr>
<tr>
<td>A24</td>
<td>6.01</td>
<td>2.74</td>
<td>2.89</td>
<td>5.04</td>
<td>4.38</td>
</tr>
<tr>
<td>G25</td>
<td>5.82</td>
<td>2.48</td>
<td>2.66</td>
<td>4.96</td>
<td>4.36</td>
</tr>
<tr>
<td>C26</td>
<td>6.13</td>
<td>2.14</td>
<td>2.20</td>
<td>4.47</td>
<td>4.05</td>
</tr>
</tbody>
</table>
### Tabelle 8: $^1H$ chemical shifts of base protons of 13merRef

<table>
<thead>
<tr>
<th>Res</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H41</th>
<th>H42</th>
<th>H5</th>
<th>H6</th>
<th>H7</th>
<th>H8</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.99</td>
</tr>
<tr>
<td>C2</td>
<td></td>
<td></td>
<td></td>
<td>8.33</td>
<td>6.62</td>
<td>5.39</td>
<td>7.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td></td>
<td>13.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.36</td>
<td>1.66</td>
<td></td>
</tr>
<tr>
<td>G4</td>
<td>12.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.89</td>
</tr>
<tr>
<td>C5</td>
<td></td>
<td></td>
<td></td>
<td>8.34</td>
<td>6.31</td>
<td>5.40</td>
<td>7.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A6</td>
<td>7.17</td>
<td></td>
<td></td>
<td>8.34</td>
<td>7.02</td>
<td>5.72</td>
<td>7.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A7</td>
<td>7.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.11</td>
</tr>
<tr>
<td>A8</td>
<td>7.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.05</td>
</tr>
<tr>
<td>C9</td>
<td></td>
<td></td>
<td></td>
<td>7.91</td>
<td>6.36</td>
<td>5.10</td>
<td>7.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G10</td>
<td>12.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.81</td>
</tr>
<tr>
<td>T11</td>
<td></td>
<td></td>
<td>13.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.27</td>
<td>1.39</td>
</tr>
<tr>
<td>C12</td>
<td></td>
<td></td>
<td>8.62</td>
<td>7.02</td>
<td>5.72</td>
<td>7.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.97</td>
</tr>
<tr>
<td>C14</td>
<td></td>
<td></td>
<td></td>
<td>5.91</td>
<td>7.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G15</td>
<td>12.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.97</td>
</tr>
<tr>
<td>A16</td>
<td>7.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.23</td>
</tr>
<tr>
<td>C17</td>
<td></td>
<td>8.11</td>
<td>6.44</td>
<td>5.23</td>
<td>7.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G18</td>
<td>12.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.83</td>
</tr>
<tr>
<td>T19</td>
<td></td>
<td>13.91</td>
<td>7.25</td>
<td>1.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T20</td>
<td></td>
<td>13.90</td>
<td>7.47</td>
<td>1.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T21</td>
<td></td>
<td>13.70</td>
<td>7.33</td>
<td>1.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G22</td>
<td>12.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.89</td>
</tr>
<tr>
<td>C23</td>
<td>8.32</td>
<td>6.34</td>
<td>5.39</td>
<td>7.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A24</td>
<td>7.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.19</td>
</tr>
<tr>
<td>G25</td>
<td>12.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.69</td>
</tr>
<tr>
<td>C26</td>
<td></td>
<td></td>
<td></td>
<td>5.33</td>
<td>7.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Tabelle 9: $^{13}$C chemical shifts of 13merRef

<table>
<thead>
<tr>
<th>Res</th>
<th>C1'</th>
<th>C3'</th>
<th>C2</th>
<th>C5</th>
<th>C6</th>
<th>C8</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>82.2</td>
<td>75.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.9</td>
</tr>
<tr>
<td>C2</td>
<td>83.9</td>
<td>76.6</td>
<td>-</td>
<td>95.7</td>
<td>140.1</td>
<td>-</td>
</tr>
<tr>
<td>T3</td>
<td>82.8</td>
<td>75.3</td>
<td>-</td>
<td>-</td>
<td>136.6</td>
<td>-</td>
</tr>
<tr>
<td>G4</td>
<td>81.4</td>
<td>76.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.5</td>
</tr>
<tr>
<td>C5</td>
<td>83.3</td>
<td>76.1</td>
<td>-</td>
<td>95.5</td>
<td>139.6</td>
<td>-</td>
</tr>
<tr>
<td>A6</td>
<td>81.7</td>
<td>76.9</td>
<td>151.4</td>
<td>-</td>
<td>-</td>
<td>138.7</td>
</tr>
<tr>
<td>A7</td>
<td>82.2</td>
<td>76.1</td>
<td>151.1</td>
<td>-</td>
<td>-</td>
<td>138.2</td>
</tr>
<tr>
<td>A8</td>
<td>81.9</td>
<td>76.6</td>
<td>152.0</td>
<td>-</td>
<td>-</td>
<td>138.0</td>
</tr>
<tr>
<td>C9</td>
<td>83.1</td>
<td>74.0</td>
<td>-</td>
<td>95.1</td>
<td>139.1</td>
<td>-</td>
</tr>
<tr>
<td>G10</td>
<td>82.0</td>
<td>76.9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.5</td>
</tr>
<tr>
<td>T11</td>
<td>82.7</td>
<td>75.1</td>
<td>-</td>
<td>-</td>
<td>135.8</td>
<td>-</td>
</tr>
<tr>
<td>C12</td>
<td>83.8</td>
<td>76.5</td>
<td>-</td>
<td>96.1</td>
<td>141.0</td>
<td>-</td>
</tr>
<tr>
<td>G13</td>
<td>82.0</td>
<td>70.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.5</td>
</tr>
<tr>
<td>C14</td>
<td>85.2</td>
<td>75.1</td>
<td>-</td>
<td>96.8</td>
<td>140.4</td>
<td>-</td>
</tr>
<tr>
<td>G15</td>
<td>81.3</td>
<td>76.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>136.4</td>
</tr>
<tr>
<td>A16</td>
<td>82.1</td>
<td>76.9</td>
<td>152.7</td>
<td>-</td>
<td>-</td>
<td>138.4</td>
</tr>
<tr>
<td>C17</td>
<td>83.2</td>
<td>74.0</td>
<td>-</td>
<td>95.4</td>
<td>139.2</td>
<td>-</td>
</tr>
<tr>
<td>G18</td>
<td>82.1</td>
<td>75.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.5</td>
</tr>
<tr>
<td>T19</td>
<td>82.7</td>
<td>75.1</td>
<td>-</td>
<td>-</td>
<td>135.8</td>
<td>-</td>
</tr>
<tr>
<td>T20</td>
<td>82.7</td>
<td>74.8</td>
<td>-</td>
<td>-</td>
<td>137.2</td>
<td>-</td>
</tr>
<tr>
<td>T21</td>
<td>81.5</td>
<td>74.7</td>
<td>-</td>
<td>-</td>
<td>136.7</td>
<td>-</td>
</tr>
<tr>
<td>G22</td>
<td>81.3</td>
<td>76.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.5</td>
</tr>
<tr>
<td>C23</td>
<td>83.5</td>
<td>75.3</td>
<td>-</td>
<td>95.5</td>
<td>139.9</td>
<td>-</td>
</tr>
<tr>
<td>A24</td>
<td>82.1</td>
<td>77.0</td>
<td>151.5</td>
<td>-</td>
<td>-</td>
<td>138.8</td>
</tr>
<tr>
<td>G25</td>
<td>81.2</td>
<td>75.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>134.5</td>
</tr>
<tr>
<td>C26</td>
<td>84.0</td>
<td>68.8</td>
<td>-</td>
<td>95.3</td>
<td>140.2</td>
<td>-</td>
</tr>
</tbody>
</table>
## Chemical shifts of 13mer2AP

**Tabella 10: 
$^1$H chemical shifts of sugar protons of 13mer2AP**

<table>
<thead>
<tr>
<th>Res</th>
<th>H1'</th>
<th>H2'</th>
<th>H2''</th>
<th>H3'</th>
<th>H4'</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>6.02</td>
<td>2.68</td>
<td>2.81</td>
<td>4.88</td>
<td>4.30</td>
</tr>
<tr>
<td>C2</td>
<td>6.11</td>
<td>2.15</td>
<td>2.55</td>
<td>4.88</td>
<td>4.29</td>
</tr>
<tr>
<td>T3</td>
<td>5.77</td>
<td>2.15</td>
<td>2.48</td>
<td>4.91</td>
<td>4.17</td>
</tr>
<tr>
<td>G4</td>
<td>5.87</td>
<td>2.66</td>
<td>2.72</td>
<td>5.01</td>
<td>4.39</td>
</tr>
<tr>
<td>C5</td>
<td>5.49</td>
<td>1.92</td>
<td>2.29</td>
<td>4.83</td>
<td>4.15</td>
</tr>
<tr>
<td>A6</td>
<td>5.92</td>
<td>2.70</td>
<td>2.89</td>
<td>5.05</td>
<td>4.36</td>
</tr>
<tr>
<td>X7</td>
<td>5.55</td>
<td>2.57</td>
<td>2.70</td>
<td>5.01</td>
<td>4.38</td>
</tr>
<tr>
<td>A8</td>
<td>6.17</td>
<td>2.60</td>
<td>2.89</td>
<td>5.03</td>
<td>4.47</td>
</tr>
<tr>
<td>C9</td>
<td>5.56</td>
<td>1.97</td>
<td>2.34</td>
<td>4.83</td>
<td>4.14</td>
</tr>
<tr>
<td>G10</td>
<td>5.96</td>
<td>2.63</td>
<td>2.79</td>
<td>4.96</td>
<td>4.38</td>
</tr>
<tr>
<td>T11</td>
<td>6.06</td>
<td>2.11</td>
<td>2.49</td>
<td>4.89</td>
<td>4.24</td>
</tr>
<tr>
<td>C12</td>
<td>5.73</td>
<td>2.05</td>
<td>2.40</td>
<td>4.87</td>
<td>4.15</td>
</tr>
<tr>
<td>G13</td>
<td>6.19</td>
<td>2.66</td>
<td>2.42</td>
<td>4.72</td>
<td>4.22</td>
</tr>
<tr>
<td>C14</td>
<td>5.73</td>
<td>1.87</td>
<td>2.38</td>
<td>4.71</td>
<td>4.09</td>
</tr>
<tr>
<td>G15</td>
<td>5.47</td>
<td>2.74</td>
<td>2.82</td>
<td>5.03</td>
<td>4.34</td>
</tr>
<tr>
<td>A16</td>
<td>6.29</td>
<td>2.74</td>
<td>2.95</td>
<td>5.09</td>
<td>4.51</td>
</tr>
<tr>
<td>C17</td>
<td>5.60</td>
<td>2.03</td>
<td>2.38</td>
<td>4.85</td>
<td>4.19</td>
</tr>
<tr>
<td>G18</td>
<td>6.00</td>
<td>2.66</td>
<td>2.82</td>
<td>4.98</td>
<td>4.40</td>
</tr>
<tr>
<td>T19</td>
<td>6.07</td>
<td>2.14</td>
<td>2.64</td>
<td>4.89</td>
<td>4.27</td>
</tr>
<tr>
<td>T20</td>
<td>6.05</td>
<td>2.21</td>
<td>2.54</td>
<td>4.89</td>
<td>4.21</td>
</tr>
<tr>
<td>T21</td>
<td>5.73</td>
<td>2.14</td>
<td>2.47</td>
<td>4.91</td>
<td>4.15</td>
</tr>
<tr>
<td>G22</td>
<td>5.86</td>
<td>2.64</td>
<td>2.72</td>
<td>5.00</td>
<td>4.38</td>
</tr>
<tr>
<td>C23</td>
<td>5.48</td>
<td>1.97</td>
<td>2.32</td>
<td>4.84</td>
<td>4.15</td>
</tr>
<tr>
<td>A24</td>
<td>6.01</td>
<td>2.75</td>
<td>2.90</td>
<td>5.05</td>
<td>4.39</td>
</tr>
<tr>
<td>G25</td>
<td>5.83</td>
<td>2.49</td>
<td>2.66</td>
<td>4.98</td>
<td>4.37</td>
</tr>
<tr>
<td>C26</td>
<td>6.12</td>
<td>2.16</td>
<td>2.23</td>
<td>4.48</td>
<td>4.06</td>
</tr>
</tbody>
</table>
Tabelle 11: $^1$H chemical shifts of base protons of 13mer2AP

<table>
<thead>
<tr>
<th>Res</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H41</th>
<th>H42</th>
<th>H5</th>
<th>H6</th>
<th>H7</th>
<th>H8</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.99</td>
</tr>
<tr>
<td>C2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.33</td>
<td>6.62</td>
<td>5.38</td>
<td>7.55</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T3</td>
<td>-</td>
<td>-</td>
<td>13.95</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.36</td>
<td>1.68</td>
<td>-</td>
</tr>
<tr>
<td>G4</td>
<td>12.72</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.92</td>
</tr>
<tr>
<td>C5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.30</td>
<td>6.32</td>
<td>5.42</td>
<td>7.34</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6</td>
<td>-</td>
<td>7.33</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.20</td>
</tr>
<tr>
<td>X7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.80</td>
<td>-</td>
<td>7.95</td>
</tr>
<tr>
<td>A8</td>
<td>-</td>
<td>7.72</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.09</td>
</tr>
<tr>
<td>C9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.02</td>
<td>6.40</td>
<td>5.16</td>
<td>7.15</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G10</td>
<td>12.72</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.83</td>
</tr>
<tr>
<td>T11</td>
<td>-</td>
<td>-</td>
<td>13.77</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.29</td>
<td>1.42</td>
<td>-</td>
</tr>
<tr>
<td>C12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.64</td>
<td>7.04</td>
<td>5.73</td>
<td>7.52</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.97</td>
</tr>
<tr>
<td>C14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.90</td>
<td>7.62</td>
<td>-</td>
</tr>
<tr>
<td>G15</td>
<td>12.97</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.98</td>
</tr>
<tr>
<td>A16</td>
<td>-</td>
<td>7.92</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.24</td>
</tr>
<tr>
<td>C17</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.11</td>
<td>6.46</td>
<td>5.25</td>
<td>7.23</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G18</td>
<td>12.72</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.86</td>
</tr>
<tr>
<td>T19</td>
<td>-</td>
<td>-</td>
<td>13.69</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.29</td>
<td>1.41</td>
<td>-</td>
</tr>
<tr>
<td>T20</td>
<td>-</td>
<td>-</td>
<td>13.38</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.46</td>
<td>1.65</td>
<td>-</td>
</tr>
<tr>
<td>T21</td>
<td>-</td>
<td>-</td>
<td>13.57</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.34</td>
<td>1.69</td>
<td>-</td>
</tr>
<tr>
<td>G22</td>
<td>12.66</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.89</td>
</tr>
<tr>
<td>C23</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.34</td>
<td>6.35</td>
<td>5.40</td>
<td>7.36</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A24</td>
<td>-</td>
<td>7.68</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.19</td>
</tr>
<tr>
<td>G25</td>
<td>12.94</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.69</td>
</tr>
<tr>
<td>C26</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.26</td>
<td>7.35</td>
<td>-</td>
</tr>
<tr>
<td>Res</td>
<td>C1'</td>
<td>C3'</td>
<td>C2</td>
<td>C5</td>
<td>C6</td>
<td>C8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td>82.2</td>
<td>73.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>134.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>83.9</td>
<td>76.5</td>
<td>-</td>
<td>95.5</td>
<td>140.1</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>82.8</td>
<td>75.3</td>
<td>-</td>
<td>-</td>
<td>136.6</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G4</td>
<td>81.4</td>
<td>76.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>83.2</td>
<td>76.1</td>
<td>-</td>
<td>95.5</td>
<td>139.7</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A6</td>
<td>81.7</td>
<td>76.8</td>
<td>151.5</td>
<td>-</td>
<td>-</td>
<td>138.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X7</td>
<td>82.2</td>
<td>76.9</td>
<td>-</td>
<td>-</td>
<td>147.6</td>
<td>138.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A8</td>
<td>81.9</td>
<td>76.6</td>
<td>152.0</td>
<td>-</td>
<td>-</td>
<td>138.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C9</td>
<td>83.1</td>
<td>74.0</td>
<td>-</td>
<td>95.1</td>
<td>139.1</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G10</td>
<td>81.9</td>
<td>76.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T11</td>
<td>82.7</td>
<td>75.1</td>
<td>-</td>
<td>-</td>
<td>135.8</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C12</td>
<td>85.1</td>
<td>76.5</td>
<td>-</td>
<td>96.0</td>
<td>141.1</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G13</td>
<td>82.0</td>
<td>70.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C14</td>
<td>83.7</td>
<td>74.9</td>
<td>-</td>
<td>96.8</td>
<td>140.4</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G15</td>
<td>81.2</td>
<td>76.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>136.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A16</td>
<td>82.1</td>
<td>76.9</td>
<td>152.7</td>
<td>-</td>
<td>-</td>
<td>138.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C17</td>
<td>83.2</td>
<td>74.0</td>
<td>-</td>
<td>95.2</td>
<td>139.3</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G18</td>
<td>82.1</td>
<td>75.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T19</td>
<td>82.7</td>
<td>75.1</td>
<td>-</td>
<td>-</td>
<td>135.8</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T20</td>
<td>82.5</td>
<td>74.5</td>
<td>-</td>
<td>-</td>
<td>137.2</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T21</td>
<td>81.4</td>
<td>74.4</td>
<td>-</td>
<td>-</td>
<td>136.7</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G22</td>
<td>81.2</td>
<td>76.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>135.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C23</td>
<td>83.4</td>
<td>74.0</td>
<td>-</td>
<td>95.5</td>
<td>139.9</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A24</td>
<td>82.1</td>
<td>76.1</td>
<td>151.5</td>
<td>-</td>
<td>-</td>
<td>138.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G25</td>
<td>81.2</td>
<td>75.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>134.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C26</td>
<td>84.0</td>
<td>68.6</td>
<td>-</td>
<td>95.4</td>
<td>139.9</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The chemical shift differences are calculated as $X(13\text{merRef}) - X(13\text{mer2AP})$.

**Tabla.13**: $^1H$ chemical shift differences between 13merRef and 13mer2AP. For X7 H2 the difference the CSD between atoms H2 in 13merRef and H6 in 13mer2AP is given.

<table>
<thead>
<tr>
<th>Res</th>
<th>H1'</th>
<th>H2'</th>
<th>H2''</th>
<th>H3'</th>
<th>H4'</th>
<th>H1/H3</th>
<th>H2/H5</th>
<th>H6/H8</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
</tr>
<tr>
<td>C2</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01/0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>T3</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>G4</td>
<td>-0.01</td>
<td>-0.03</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-</td>
<td>-0.03</td>
</tr>
<tr>
<td>C5</td>
<td>-0.08</td>
<td>-0.08</td>
<td>-0.07</td>
<td>-0.04</td>
<td>-0.04</td>
<td>0.03/-0.01</td>
<td>-0.02</td>
<td>-0.04</td>
</tr>
<tr>
<td>A6</td>
<td>-0.17</td>
<td>0.01</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-</td>
<td>-0.16</td>
<td>-0.01</td>
</tr>
<tr>
<td>X7</td>
<td>0.29</td>
<td>0.06</td>
<td>0.14</td>
<td>0.04</td>
<td>0.05</td>
<td>-</td>
<td>-0.67</td>
<td>0.16</td>
</tr>
<tr>
<td>A8</td>
<td>-0.12</td>
<td>-0.06</td>
<td>-0.06</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-</td>
<td>-0.13</td>
<td>-0.04</td>
</tr>
<tr>
<td>C9</td>
<td>0.00</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.06</td>
<td>0.00</td>
<td>-0.03/0.02</td>
<td>-0.06</td>
<td>-0.04</td>
</tr>
<tr>
<td>G10</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.05</td>
<td>-</td>
<td>-0.02</td>
</tr>
<tr>
<td>T11</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.03</td>
<td>-0.02</td>
</tr>
<tr>
<td>C12</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-0.02/-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>G13</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-</td>
<td>-</td>
<td>-0.01</td>
</tr>
<tr>
<td>C14</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>G15</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-</td>
<td>-0.01</td>
</tr>
<tr>
<td>A16</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-</td>
<td>-0.01</td>
<td>-0.02</td>
</tr>
<tr>
<td>C17</td>
<td>-0.04</td>
<td>-0.04</td>
<td>-0.02</td>
<td>-0.03</td>
<td>-0.02</td>
<td>-0.01/-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td>G18</td>
<td>0.00</td>
<td>-0.03</td>
<td>0.00</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.03</td>
<td>-</td>
<td>-0.03</td>
</tr>
<tr>
<td>T19</td>
<td>-0.03</td>
<td>0.00</td>
<td>-0.03</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.21</td>
<td>-0.04</td>
<td>-0.04</td>
</tr>
<tr>
<td>T20</td>
<td>0.09</td>
<td>-0.04</td>
<td>0.09</td>
<td>0.02</td>
<td>-0.03</td>
<td>0.53</td>
<td>-0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>T21</td>
<td>0.11</td>
<td>-0.04</td>
<td>0.01</td>
<td>0.01</td>
<td>0.13</td>
<td>0.00</td>
<td>-</td>
<td>-0.02</td>
</tr>
<tr>
<td>G22</td>
<td>-0.04</td>
<td>0.00</td>
<td>-0.05</td>
<td>0.00</td>
<td>-0.00</td>
<td>-0.05</td>
<td>-</td>
<td>0.00</td>
</tr>
<tr>
<td>C23</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.01</td>
<td>-0.02/-0.02</td>
<td>-0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>A24</td>
<td>0.00</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-</td>
<td>-0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>G25</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.00</td>
<td>-</td>
<td>0.00</td>
</tr>
<tr>
<td>C26</td>
<td>0.01</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.00</td>
<td>-</td>
<td>0.07</td>
<td>0.03</td>
</tr>
</tbody>
</table>
**Helical parameter**

The averaged values for the 10 minimum-energy, violation-free structures for all helical parameters are listed in the Appendix in sections .11 and .12 for 13merRef and 13mer2AP, respectively. Their RMSD is given as the uncertainty.

**.11 Helical parameters for 13merRef**

<table>
<thead>
<tr>
<th>base pair</th>
<th>Shear (Sx)/Å</th>
<th>Stretch (Sy)/Å</th>
<th>Stagger (Sz)/Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1–C26</td>
<td>-0.42 ± 0.01</td>
<td>-0.26 ± 0.00</td>
<td>-0.14 ± 0.04</td>
</tr>
<tr>
<td>C2–G25</td>
<td>0.42 ± 0.02</td>
<td>-0.26 ± 0.01</td>
<td>-0.13 ± 0.05</td>
</tr>
<tr>
<td>T3–A24</td>
<td>-0.07 ± 0.01</td>
<td>-0.26 ± 0.01</td>
<td>0.07 ± 0.04</td>
</tr>
<tr>
<td>G4–C23</td>
<td>-0.37 ± 0.01</td>
<td>-0.28 ± 0.00</td>
<td>0.10 ± 0.06</td>
</tr>
<tr>
<td>C5–G22</td>
<td>0.40 ± 0.01</td>
<td>-0.26 ± 0.01</td>
<td>-0.13 ± 0.05</td>
</tr>
<tr>
<td>A6–T21</td>
<td>0.02 ± 0.02</td>
<td>-0.26 ± 0.01</td>
<td>-0.04 ± 0.03</td>
</tr>
<tr>
<td>A7–T20</td>
<td>0.07 ± 0.01</td>
<td>-0.27 ± 0.01</td>
<td>-0.24 ± 0.02</td>
</tr>
<tr>
<td>A8–T19</td>
<td>0.03 ± 0.02</td>
<td>-0.28 ± 0.01</td>
<td>-0.06 ± 0.04</td>
</tr>
<tr>
<td>C9–G18</td>
<td>0.39 ± 0.01</td>
<td>-0.27 ± 0.00</td>
<td>-0.05 ± 0.07</td>
</tr>
<tr>
<td>G10–C17</td>
<td>-0.40 ± 0.01</td>
<td>-0.26 ± 0.00</td>
<td>0.08 ± 0.03</td>
</tr>
<tr>
<td>T11–A16</td>
<td>-0.07 ± 0.00</td>
<td>-0.27 ± 0.01</td>
<td>0.09 ± 0.07</td>
</tr>
<tr>
<td>C12–G15</td>
<td>0.40 ± 0.01</td>
<td>-0.27 ± 0.00</td>
<td>-0.11 ± 0.05</td>
</tr>
<tr>
<td>G13–C14</td>
<td>-0.40 ± 0.01</td>
<td>-0.26 ± 0.00</td>
<td>-0.03 ± 0.02</td>
</tr>
</tbody>
</table>

average  | 0.00 ± 0.33  | -0.26 ± 0.01   | -0.05 ± 0.11   |
### Helical parameter

**Tabelle 15:** Base pair parameters for 13merRef, rotational

<table>
<thead>
<tr>
<th>base pair</th>
<th>Buckle ((\chi)/^\circ)</th>
<th>Propeller Twist ((\omega)/^\circ)</th>
<th>Opening ((\sigma)/^\circ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1–C26</td>
<td>-1,4±1,5</td>
<td>2,4±0,6</td>
<td>1,7±0,1</td>
</tr>
<tr>
<td>C2–G25</td>
<td>-4,7±1,9</td>
<td>3,8±1,7</td>
<td>1,5±0,2</td>
</tr>
<tr>
<td>T3–A24</td>
<td>-8,0±1,7</td>
<td>-1,8±1,7</td>
<td>-2,9±0,1</td>
</tr>
<tr>
<td>G4–C23</td>
<td>-0,4±0,5</td>
<td>5,4±2,5</td>
<td>1,5±0,1</td>
</tr>
<tr>
<td>C5–G22</td>
<td>5,3±0,5</td>
<td>3,5±1,7</td>
<td>1,6±0,1</td>
</tr>
<tr>
<td>A6–T21</td>
<td>6,3±0,4</td>
<td>-1,8±0,4</td>
<td>-2,9±0,2</td>
</tr>
<tr>
<td>A7–T20</td>
<td>4,9±0,4</td>
<td>-8,2±1,3</td>
<td>-2,5±0,3</td>
</tr>
<tr>
<td>A8–T19</td>
<td>5,1±0,7</td>
<td>-10,6±1,0</td>
<td>-2,7±0,3</td>
</tr>
<tr>
<td>C9–G18</td>
<td>4,3±1,0</td>
<td>-6,0±1,2</td>
<td>1,5±0,1</td>
</tr>
<tr>
<td>G10–C17</td>
<td>-1,2±1,3</td>
<td>1,2±1,8</td>
<td>1,5±0,1</td>
</tr>
<tr>
<td>T11–A16</td>
<td>2,6±2,7</td>
<td>-3,4±1,4</td>
<td>-3,0±0,2</td>
</tr>
<tr>
<td>C12–G15</td>
<td>4,6±1,6</td>
<td>-8,9±0,6</td>
<td>2,0±0,1</td>
</tr>
<tr>
<td>G13–C14</td>
<td>-0,6±0,8</td>
<td>-0,6±0,5</td>
<td>1,4±0,1</td>
</tr>
<tr>
<td><strong>average</strong></td>
<td>1,3±4,4</td>
<td>-1,9±5,2</td>
<td>-0,1±2,2</td>
</tr>
</tbody>
</table>

**Tabelle 16:** Base pair step parameters for 13merRef, translational

<table>
<thead>
<tr>
<th>base pair step</th>
<th>Shift ((S_x)/\text{Å})</th>
<th>Slide ((S_y)/\text{Å})</th>
<th>Rise ((S_z)/\text{Å})</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1–C2</td>
<td>-0,12±0,03</td>
<td>-0,71±0,11</td>
<td>3,40±0,07</td>
</tr>
<tr>
<td>C2–T3</td>
<td>-0,30±0,05</td>
<td>-1,35±0,04</td>
<td>3,38±0,04</td>
</tr>
<tr>
<td>T3–G4</td>
<td>0,44±0,05</td>
<td>-0,57±0,11</td>
<td>2,81±0,06</td>
</tr>
<tr>
<td>G4–C5</td>
<td>0,10±0,05</td>
<td>-0,25±0,10</td>
<td>3,06±0,05</td>
</tr>
<tr>
<td>C5–A6</td>
<td>-0,61±0,11</td>
<td>-0,67±0,06</td>
<td>3,01±0,03</td>
</tr>
<tr>
<td>A6–A7</td>
<td>-0,36±0,03</td>
<td>-0,87±0,02</td>
<td>3,27±0,01</td>
</tr>
<tr>
<td>A7–A8</td>
<td>-0,45±0,07</td>
<td>-0,66±0,06</td>
<td>3,11±0,03</td>
</tr>
<tr>
<td>A8–C9</td>
<td>-0,02±0,08</td>
<td>-0,72±0,08</td>
<td>3,20±0,01</td>
</tr>
<tr>
<td>C9–G10</td>
<td>0,17±0,07</td>
<td>-0,98±0,09</td>
<td>3,21±0,05</td>
</tr>
<tr>
<td>G10–T11</td>
<td>-0,23±0,03</td>
<td>-0,92±0,04</td>
<td>3,14±0,03</td>
</tr>
<tr>
<td>T11–C12</td>
<td>0,34±0,10</td>
<td>-0,28±0,09</td>
<td>3,00±0,06</td>
</tr>
<tr>
<td>C12–G13</td>
<td>0,02±0,04</td>
<td>-0,99±0,05</td>
<td>3,01±0,04</td>
</tr>
<tr>
<td><strong>average</strong></td>
<td>-0,08±0,32</td>
<td>-0,75±0,31</td>
<td>3,13±0,17</td>
</tr>
<tr>
<td>base pair step</td>
<td>Tilt ($\tau$)/°</td>
<td>Roll ($\rho$)/°</td>
<td>Twist ($\Omega$)/°</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>G1–C2</td>
<td>-1.0 ± 0.4</td>
<td>-4.9 ± 1.0</td>
<td>41.1 ± 0.6</td>
</tr>
<tr>
<td>C2–T3</td>
<td>1.6 ± 0.3</td>
<td>-2.6 ± 1.3</td>
<td>30.4 ± 0.3</td>
</tr>
<tr>
<td>T3–G4</td>
<td>-0.6 ± 0.3</td>
<td>3.7 ± 0.9</td>
<td>35.2 ± 0.6</td>
</tr>
<tr>
<td>G4–C5</td>
<td>2.1 ± 0.1</td>
<td>2.1 ± 0.8</td>
<td>40.6 ± 0.6</td>
</tr>
<tr>
<td>C5–A6</td>
<td>-1.1 ± 0.2</td>
<td>4.6 ± 0.4</td>
<td>35.2 ± 0.4</td>
</tr>
<tr>
<td>A6–A7</td>
<td>0.3 ± 0.2</td>
<td>-2.4 ± 0.6</td>
<td>33.4 ± 0.2</td>
</tr>
<tr>
<td>A7–A8</td>
<td>-4.2 ± 0.3</td>
<td>1.3 ± 0.3</td>
<td>36.0 ± 0.5</td>
</tr>
<tr>
<td>A8–C9</td>
<td>-2.7 ± 0.5</td>
<td>-2.8 ± 0.6</td>
<td>37.6 ± 0.6</td>
</tr>
<tr>
<td>C9–G10</td>
<td>-2.3 ± 0.4</td>
<td>0.7 ± 1.2</td>
<td>33.3 ± 0.8</td>
</tr>
<tr>
<td>G10–T11</td>
<td>1.8 ± 0.3</td>
<td>-3.0 ± 0.6</td>
<td>35.3 ± 0.6</td>
</tr>
<tr>
<td>T11–C12</td>
<td>3.3 ± 0.5</td>
<td>2.7 ± 1.4</td>
<td>39.6 ± 0.7</td>
</tr>
<tr>
<td>C12–G13</td>
<td>-1.0 ± 0.5</td>
<td>10.0 ± 1.3</td>
<td>30.6 ± 0.5</td>
</tr>
<tr>
<td>average</td>
<td>-0.3 ± 2.2</td>
<td>0.8 ± 4.2</td>
<td>35.7 ± 3.5</td>
</tr>
</tbody>
</table>
### Helical parameters for 13mer2AP

#### Tabelle 18: Base pair parameters for 13mer2AP, translational

<table>
<thead>
<tr>
<th>base pair</th>
<th>Shear (Sx)/Å</th>
<th>Stretch (Sy)/Å</th>
<th>Stagger (Sz)/Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1–C26</td>
<td>-0.40 ± 0.01</td>
<td>-0.26 ± 0.00</td>
<td>-0.11 ± 0.04</td>
</tr>
<tr>
<td>C2–G25</td>
<td>0.35 ± 0.02</td>
<td>-0.28 ± 0.01</td>
<td>-0.10 ± 0.05</td>
</tr>
<tr>
<td>T3–A24</td>
<td>-0.07 ± 0.00</td>
<td>-0.26 ± 0.00</td>
<td>-0.02 ± 0.05</td>
</tr>
<tr>
<td>G4–C23</td>
<td>-0.37 ± 0.01</td>
<td>-0.29 ± 0.00</td>
<td>0.33 ± 0.03</td>
</tr>
<tr>
<td>C5–G22</td>
<td>0.29 ± 0.03</td>
<td>-0.27 ± 0.01</td>
<td>-0.10 ± 0.05</td>
</tr>
<tr>
<td>A6–T21</td>
<td>0.01 ± 0.02</td>
<td>-0.27 ± 0.00</td>
<td>-0.02 ± 0.04</td>
</tr>
<tr>
<td>2AP7–T20</td>
<td>-0.02 ± 0.00</td>
<td>-0.32 ± 0.00</td>
<td>-0.23 ± 0.03</td>
</tr>
<tr>
<td>A8–T19</td>
<td>0.08 ± 0.01</td>
<td>-0.26 ± 0.00</td>
<td>0.12 ± 0.03</td>
</tr>
<tr>
<td>C9–G18</td>
<td>0.43 ± 0.02</td>
<td>-0.25 ± 0.01</td>
<td>-0.18 ± 0.06</td>
</tr>
<tr>
<td>G10–C17</td>
<td>-0.39 ± 0.02</td>
<td>-0.27 ± 0.00</td>
<td>-0.16 ± 0.03</td>
</tr>
<tr>
<td>T11–A16</td>
<td>-0.07 ± 0.01</td>
<td>-0.26 ± 0.00</td>
<td>-0.13 ± 0.05</td>
</tr>
<tr>
<td>C12–G15</td>
<td>0.41 ± 0.01</td>
<td>-0.27 ± 0.01</td>
<td>-0.17 ± 0.04</td>
</tr>
<tr>
<td>G13–C14</td>
<td>-0.37 ± 0.02</td>
<td>-0.26 ± 0.01</td>
<td>-0.05 ± 0.03</td>
</tr>
<tr>
<td>average</td>
<td>-0.01 ± 0.31</td>
<td>-0.27 ± 0.02</td>
<td>-0.06 ± 0.15</td>
</tr>
</tbody>
</table>

#### Tabelle 19: Base pair parameters for 13mer2AP, rotational

<table>
<thead>
<tr>
<th>base pair</th>
<th>Buckle (χ)/°</th>
<th>Propeller Twist (ω)/°</th>
<th>Opening (σ)/°</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1–C26</td>
<td>-2.2 ± 1.2</td>
<td>2.1 ± 0.4</td>
<td>1.5 ± 0.1</td>
</tr>
<tr>
<td>C2–G25</td>
<td>-3.0 ± 1.4</td>
<td>6.0 ± 2.0</td>
<td>1.0 ± 0.1</td>
</tr>
<tr>
<td>T3–A24</td>
<td>-5.8 ± 1.5</td>
<td>-2.4 ± 1.3</td>
<td>-2.8 ± 0.1</td>
</tr>
<tr>
<td>G4–C23</td>
<td>5.7 ± 0.7</td>
<td>6.9 ± 1.6</td>
<td>1.8 ± 0.2</td>
</tr>
<tr>
<td>C5–G22</td>
<td>8.0 ± 0.6</td>
<td>1.9 ± 1.0</td>
<td>0.9 ± 0.1</td>
</tr>
<tr>
<td>A6–T21</td>
<td>6.4 ± 0.3</td>
<td>-4.9 ± 0.3</td>
<td>-2.8 ± 0.1</td>
</tr>
<tr>
<td>2AP7–T20</td>
<td>2.3 ± 0.3</td>
<td>-1.2 ± 0.3</td>
<td>-5.7 ± 0.1</td>
</tr>
<tr>
<td>A8–T19</td>
<td>5.3 ± 0.7</td>
<td>-2.5 ± 1.1</td>
<td>-3.0 ± 0.1</td>
</tr>
<tr>
<td>C9–G18</td>
<td>8.4 ± 0.7</td>
<td>-3.7 ± 1.1</td>
<td>1.8 ± 0.2</td>
</tr>
<tr>
<td>G10–C17</td>
<td>-0.5 ± 1.3</td>
<td>6.2 ± 1.5</td>
<td>1.3 ± 0.2</td>
</tr>
<tr>
<td>T11–A16</td>
<td>2.8 ± 2.2</td>
<td>0.8 ± 1.0</td>
<td>-2.9 ± 0.1</td>
</tr>
<tr>
<td>C12–G15</td>
<td>2.7 ± 1.4</td>
<td>-6.5 ± 1.1</td>
<td>2.1 ± 0.1</td>
</tr>
<tr>
<td>G13–C14</td>
<td>-2.1 ± 1.5</td>
<td>-1.0 ± 1.0</td>
<td>1.3 ± 0.1</td>
</tr>
<tr>
<td>average</td>
<td>2.1 ± 4.5</td>
<td>0.1 ± 4.3</td>
<td>-0.4 ± 2.6</td>
</tr>
</tbody>
</table>
### Tabelle 20: Base pair step parameters for 13mer2AP, translational

<table>
<thead>
<tr>
<th>base pair step</th>
<th>Shift (Sx)/Å</th>
<th>Slide (Sy)/Å</th>
<th>Rise (Sz)/Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1–C2</td>
<td>-0.03 ± 0.07</td>
<td>-0.44 ± 0.11</td>
<td>3.28 ± 0.05</td>
</tr>
<tr>
<td>C2–T3</td>
<td>-0.04 ± 0.05</td>
<td>-1.25 ± 0.04</td>
<td>3.28 ± 0.05</td>
</tr>
<tr>
<td>T3–G4</td>
<td>0.28 ± 0.02</td>
<td>-0.85 ± 0.04</td>
<td>2.73 ± 0.06</td>
</tr>
<tr>
<td>G4–C5</td>
<td>0.06 ± 0.03</td>
<td>-0.29 ± 0.05</td>
<td>3.16 ± 0.03</td>
</tr>
<tr>
<td>C5–A6</td>
<td>-0.43 ± 0.03</td>
<td>-0.73 ± 0.05</td>
<td>3.07 ± 0.04</td>
</tr>
<tr>
<td>A6–2AP7</td>
<td>-0.58 ± 0.02</td>
<td>-1.07 ± 0.07</td>
<td>3.31 ± 0.02</td>
</tr>
<tr>
<td>2AP7–A8</td>
<td>-0.13 ± 0.03</td>
<td>-0.66 ± 0.04</td>
<td>3.14 ± 0.02</td>
</tr>
<tr>
<td>A8–C9</td>
<td>-0.10 ± 0.05</td>
<td>-0.98 ± 0.05</td>
<td>3.18 ± 0.02</td>
</tr>
<tr>
<td>C9–G10</td>
<td>0.11 ± 0.04</td>
<td>-1.10 ± 0.05</td>
<td>3.29 ± 0.05</td>
</tr>
<tr>
<td>G10–T11</td>
<td>0.06 ± 0.05</td>
<td>-0.87 ± 0.06</td>
<td>3.29 ± 0.03</td>
</tr>
<tr>
<td>T11–C12</td>
<td>0.31 ± 0.07</td>
<td>-0.39 ± 0.13</td>
<td>3.10 ± 0.04</td>
</tr>
<tr>
<td>C12–G13</td>
<td>-0.09 ± 0.09</td>
<td>-0.99 ± 0.08</td>
<td>3.01 ± 0.07</td>
</tr>
<tr>
<td>average</td>
<td>-0.05 ± 0.26</td>
<td>-0.80 ± 0.30</td>
<td>3.15 ± 0.17</td>
</tr>
</tbody>
</table>

### Tabelle 21: Base pair step parameters for 13mer2AP, rotational

<table>
<thead>
<tr>
<th>base pair step</th>
<th>Tilt (τ)/°</th>
<th>Roll (ρ)/°</th>
<th>Twist (Ω)/°</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1–C2</td>
<td>-1.1 ± 0.5</td>
<td>-2.3 ± 1.5</td>
<td>42.3 ± 0.8</td>
</tr>
<tr>
<td>C2–T3</td>
<td>2.3 ± 0.6</td>
<td>-1.1 ± 1.4</td>
<td>31.1 ± 0.4</td>
</tr>
<tr>
<td>T3–G4</td>
<td>-3.0 ± 0.3</td>
<td>-0.5 ± 0.5</td>
<td>34.0 ± 0.5</td>
</tr>
<tr>
<td>G4–C5</td>
<td>4.1 ± 0.3</td>
<td>0.7 ± 0.9</td>
<td>41.3 ± 0.4</td>
</tr>
<tr>
<td>C5–A6</td>
<td>-1.0 ± 0.2</td>
<td>2.5 ± 0.4</td>
<td>35.5 ± 0.5</td>
</tr>
<tr>
<td>A6–2AP7</td>
<td>-2.1 ± 0.1</td>
<td>0.4 ± 0.3</td>
<td>34.0 ± 0.4</td>
</tr>
<tr>
<td>2AP7–A8</td>
<td>-5.3 ± 0.3</td>
<td>-3.3 ± 0.3</td>
<td>37.6 ± 0.3</td>
</tr>
<tr>
<td>A8–C9</td>
<td>-0.3 ± 0.4</td>
<td>-3.0 ± 0.5</td>
<td>34.1 ± 0.5</td>
</tr>
<tr>
<td>C9–G10</td>
<td>-1.5 ± 0.3</td>
<td>3.9 ± 0.9</td>
<td>32.9 ± 0.5</td>
</tr>
<tr>
<td>G10–T11</td>
<td>3.3 ± 0.4</td>
<td>-5.5 ± 0.6</td>
<td>35.8 ± 0.3</td>
</tr>
<tr>
<td>T11–C12</td>
<td>3.0 ± 0.4</td>
<td>4.3 ± 1.6</td>
<td>39.1 ± 1.0</td>
</tr>
<tr>
<td>C12–G13</td>
<td>-0.8 ± 0.3</td>
<td>10.7 ± 1.6</td>
<td>30.6 ± 0.6</td>
</tr>
<tr>
<td>average</td>
<td>-0.2 ± 2.8</td>
<td>0.6 ± 4.3</td>
<td>35.7 ± 3.7</td>
</tr>
</tbody>
</table>
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WC</td>
<td>Watson-Crick</td>
</tr>
<tr>
<td>DNA</td>
<td>DesoxyriboNucleic Acid</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>SNP</td>
<td>Single Nucleotide Polymorphism</td>
</tr>
<tr>
<td>DETEQ</td>
<td>Detection by Electron Transfer controlled Emission Quenching</td>
</tr>
<tr>
<td>FIT</td>
<td>Forced Intercalation probe</td>
</tr>
<tr>
<td>FRET</td>
<td>Fluorescence Resonance Excitation Transfer</td>
</tr>
<tr>
<td>2AP</td>
<td>2-Aminopurin</td>
</tr>
<tr>
<td>HNF</td>
<td>2-Hydroxy-7-nitrofluorene</td>
</tr>
<tr>
<td>ABA</td>
<td>ABAasic site</td>
</tr>
<tr>
<td>A</td>
<td>Adenine</td>
</tr>
<tr>
<td>T</td>
<td>Thymine</td>
</tr>
<tr>
<td>G</td>
<td>Guanine</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>TRIS</td>
<td>Trishydroxymethylaminomethane</td>
</tr>
<tr>
<td>NOE</td>
<td>Nuclear Overhauser Enhancement</td>
</tr>
</tbody>
</table>
Helical parameter

**NOESY** Nuclear Overhauser Enhancement Spectroscopy

**ISPA** Isolated Spin Pair Approximation

**RDC** Residual Dipolar Coupling

**PAS** Principle Axis System

**SA** Simulated Annealing

**MD** Molecular Dynamics

**HPLC** High Pressure Liquid Chromatography

**WATERGATE** Water suppression by GrAdient Tailored Excitation

**DQF-COSY** Double Quantum Filtered COrrrelated Spectroscopy

**TOCSY** TOtal Correlation Spectroscopy

**HMQC** Heteronuclear Multiple Quantum Coherence Spectroscopy

**DFT** Density Functional Theory

**TZVP** Triple Zeta Valence plus Polarization

**MEP** Molecular Electrostatic Potential

**P.E.-COSY** Primitive Exclusive COrrrelated Spectroscopy

**RMSD** Root Mean Square Deviation

**CSD** Chemical Shift Difference
# Index of figures

1.1 DNA duplex sequence with chemical structure of the 2AP-T and A-T base pairs ........................................... 5  
1.2 DNA duplex sequence with chemical structure of the HNF fluorophore and abasic site ........................................... 5  

2.1 Structure and nomenclature of the Watson-Crick base pairs A:T and G:C 8  
2.2 Nomenclature and structure of 2'′-deoxy-β-D-ribose ......................... 9  
2.3 Dihedral angles in the sugar-phosphate backbone of DNA ................ 9  
2.4 Comparison of A- und B-form DNA ........................................ 9  
2.5 Helical parameters for the orientational description of base pairs ...... 10  
2.6 Kinetic scheme of imino proton exchange ................................... 12  
2.7 Scheme of a general inversion recovery pulse sequence .................... 18  
2.8 Pulse shapes and excitation forms .......................................... 19  
2.9 Scheme of a general NOESY pulse sequence ................................ 21  
2.10 The cross-relaxation network of N groups of nuclei ....................... 22  
2.11 Two different approaches to obtain distances from NOE experiments ... 26  
2.12 Steric interaction Pf1-DNA and principle of RDC measurement ........ 29  
2.13 Orientation of the magnetic field and the internuclear vector in the macromolecular frame ........................................ 31  
2.14 Orientation of the internuclear vector in the eigenframe of the alignment tensor and orientational degeneracy of Residual Dipolar Couplings .... 33
# Index of figures

3.1 Imino proton signal intensities for the temperature range 283-328K. in 13mer2AP .......................... 43

4.1 DNA duplex sequence with chemical structure of the 2AP-T and A-T base pairs ............................. 57

4.2 Sequential assignment in the sugar H1' base H6/H8 region of the NOESY-spectrum in D$_2$O for 13merRef and 13mer2AP .......................... 62

4.3 Absolute per-residue chemical shift differences of 13mer2AP to 13merRef .............................. 64

4.4 Structures of 13merRef and 13mer2AP .......................................................... 65

4.5 Overlay of the experimental and back-calculated NOESY-spectra ................................. 66

4.6 Plot of the experimental vs predicted Residual Dipolar Couplings for 13merRef and 13mer2AP .......................... 66

4.7 13merRef and 13mer2AP: Differences in translational helical parameters between base pair partners .......................... 67

4.8 13merRef and 13mer2AP: Differences in rotational helical parameter between base pair partners .......................... 67

4.9 13merRef and 13mer2AP: Differences in translational helical parameter between base pairs .......................... 68

4.10 13merRef and 13mer2AP: Differences in rotational helical parameter between base pairs .......................... 68

4.11 T20 and T3 imino proton intensities for 13mer2AP at 298K .......................... 71

4.12 Saturation transfer experiments in H$_2$O .......................... 72

4.13 Base pair lifetime determination for 13merRef and 13mer2AP .......................... 73

4.14 Melting curves via 13mer2AP absorbance and 2AP fluorescence yield. ............................. 74

4.15 Linear fits for the imino proton exchange times of T3,T11,G15 and G25 vs the inverse base catalyst concentration. .......................... 74

4.16 Comparison of the three central base pairs of the average structures for 13merRef and 13mer2AP .......................... 78

4.17 DNA duplex sequence with chemical structure of the HNF and abasic site ............................. 86
Index of figures

4.18 Sequential assignment in the sugar H1' base H6/H8 region of the NOESY-spectrum in D$_2$O for 13merHNF .................................................. 88
4.19 Comparison of the imino proton region of 13merRef, 13merRefGC and 13merHNF ................................................................. 89
4.20 Absolute per-residue chemical shift differences of 13merHNF to the corresponding native structures .................. 90
4.21 10 minimum-energy, violation-free structures of 13merHNF in face-up and face-down orientation ....................... 91
4.22 Averaged, minimized structures of 13merHNF in face-up and face-down orientation ............................................ 91
4.23 Close-up view of the average structures of 13merHNF in face-up and face-down orientation ........................................... 92
4.24 Overlay of the experimental and back-calculated NOESY-spectra for 13merHNF .................................................... 92
4.25 Plot of the experimental vs predicted Residual Dipolar Couplings for 13merHNF ...................................................... 93
4.26 Absorption changes upon hybridisation, when lowering temperature from 85 to 25°C for 13merHNF .................... 94
4.27 Temperature-dependence of the spectra in Fig. 4.26. ................. 94
# Index of tables

3.1 Experimentally determined Residual Dipolar Couplings for 13merRef  . 48  
3.2 Experimentally determined Residual Dipolar Couplings for 13mer2AP  . 49  
3.3 Overview of structural statistics for 13merRef and 13mer2AP calculations  50  
3.4 Experimentally determined Residual Dipolar Couplings for 13merHNF  . 55  
3.5 Overview of structural statistics for 13merHNF calculations  . . . . . . . 56  
4.1 Selected $^1$H Chemical shift differences (CSD) for 13merRef and 13mer2AP  63  
   .2 $^1$H chemical shifts of sugar protons of 13merHNF  . . . . . . . . . . . 262  
   .3 $^1$H chemical shifts of base protons of 13merHNF  . . . . . . . . . . . 263  
   .4 $^{13}$C chemical shifts of 13merHNF  . . . . . . . . . . . . . . . . . . . . . 264  
   .5 $^1$H chemical shifts of sugar protons of 13merRefGC  . . . . . . . . . . 265  
   .6 $^1$H chemical shifts of base protons of 13merRefGC  . . . . . . . . . . . 266  
   .7 $^1$H chemical shifts of sugar protons of 13merRef  . . . . . . . . . . . 267  
   .8 $^1$H chemical shifts of base protons of 13merRef  . . . . . . . . . . . . . 268  
   .9 $^{13}$C chemical shifts of 13merRef  . . . . . . . . . . . . . . . . . . . . . 269  
   .10 $^1$H chemical shifts of sugar protons of 13mer2AP  . . . . . . . . . . . 270  
   .11 $^1$H chemical shifts of base protons of 13mer2AP  . . . . . . . . . . . . 271  
   .12 $^{13}$C chemical shifts of 13mer2AP  . . . . . . . . . . . . . . . . . . . . . 272  
   .13 $^1$H chemical shift differences between 13merRef and 13mer2AP  . . . . 273  
   .14 Base pair parameters for 13merRef, translational  . . . . . . . . . . . . 275  
   .15 Base pair parameters for 13merRef, rotational  . . . . . . . . . . . . . . . 276  
   .16 Base pair step parameters for 13merRef, translational  . . . . . . . . . . 276  

<table>
<thead>
<tr>
<th>Table No.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Residual Dipolar Couplings for 13merRef</td>
<td>48</td>
</tr>
<tr>
<td>3.2</td>
<td>Residual Dipolar Couplings for 13mer2AP</td>
<td>49</td>
</tr>
<tr>
<td>3.3</td>
<td>Structural statistics for 13merRef and 13mer2AP</td>
<td>50</td>
</tr>
<tr>
<td>3.4</td>
<td>Residual Dipolar Couplings for 13merHNF</td>
<td>55</td>
</tr>
<tr>
<td>3.5</td>
<td>Structural statistics for 13merHNF</td>
<td>56</td>
</tr>
<tr>
<td>4.1</td>
<td>CSD for 13merRef and 13mer2AP</td>
<td>63</td>
</tr>
<tr>
<td>4.2</td>
<td>$^1$H chemical shifts of sugar protons of 13merHNF</td>
<td>262</td>
</tr>
<tr>
<td>4.3</td>
<td>$^1$H chemical shifts of base protons of 13merHNF</td>
<td>263</td>
</tr>
<tr>
<td>4.4</td>
<td>$^{13}$C chemical shifts of 13merHNF</td>
<td>264</td>
</tr>
<tr>
<td>4.5</td>
<td>$^1$H chemical shifts of sugar protons of 13merRefGC</td>
<td>265</td>
</tr>
<tr>
<td>4.6</td>
<td>$^1$H chemical shifts of base protons of 13merRefGC</td>
<td>266</td>
</tr>
<tr>
<td>4.7</td>
<td>$^1$H chemical shifts of sugar protons of 13merRef</td>
<td>267</td>
</tr>
<tr>
<td>4.8</td>
<td>$^1$H chemical shifts of base protons of 13merRef</td>
<td>268</td>
</tr>
<tr>
<td>4.9</td>
<td>$^{13}$C chemical shifts of 13merRef</td>
<td>269</td>
</tr>
<tr>
<td>4.10</td>
<td>$^1$H chemical shifts of sugar protons of 13mer2AP</td>
<td>270</td>
</tr>
<tr>
<td>4.11</td>
<td>$^1$H chemical shifts of base protons of 13mer2AP</td>
<td>271</td>
</tr>
<tr>
<td>4.12</td>
<td>$^{13}$C chemical shifts of 13mer2AP</td>
<td>272</td>
</tr>
<tr>
<td>4.13</td>
<td>$^1$H chemical shift differences between 13merRef and 13mer2AP</td>
<td>273</td>
</tr>
<tr>
<td>4.14</td>
<td>Base pair parameters for 13merRef, translational</td>
<td>275</td>
</tr>
<tr>
<td>4.15</td>
<td>Base pair parameters for 13merRef, rotational</td>
<td>276</td>
</tr>
<tr>
<td>4.16</td>
<td>Base pair step parameters for 13merRef, translational</td>
<td>276</td>
</tr>
</tbody>
</table>
## Index of tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>.17</td>
<td>Base pair step parameters for 13merRef, rotational</td>
<td>277</td>
</tr>
<tr>
<td>.18</td>
<td>Base pair parameters for 13mer2AP, translational</td>
<td>278</td>
</tr>
<tr>
<td>.19</td>
<td>Base pair parameters for 13mer2AP, rotational</td>
<td>278</td>
</tr>
<tr>
<td>.20</td>
<td>Base pair step parameters for 13mer2AP, translational</td>
<td>279</td>
</tr>
<tr>
<td>.21</td>
<td>Base pair step parameters for 13mer2AP, rotational</td>
<td>279</td>
</tr>
</tbody>
</table>
Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den 28.01.2008                                            André Dallmann