From DEPARTMENT OF BIOSCIENCES AND NUTRITION
Karolinska Institutet, Stockholm, Sweden

MOLECULAR DYNAMICS SIMULATIONS OF RNA BULGES, LOOPS AND BASE FLIPPING

Boel Nyström Macchion

Stockholm 2008

Published by Karolinska Institutet. Printed by Universitetsservice US-AB
Nanna Svartz väg 4, 171 77 Solna

ISBN 978-91-7357-458-7

ABSTRACT

RNA molecules have several important functions in a cell. They carry the genetic information from DNA to protein, act as catalysts and regulate gene expression. Both DNA and RNA consist of nucleotides, DNA molecules are mainly double stranded while RNA molecules often have single stranded regions that interrupt the double stranded structure. RNA molecules can take a variety of three dimensional structures and the single stranded regions are commonly used as recognition sites and building blocks.

In this thesis the structure and dynamics of different RNA molecules have been studied by molecular dynamics (MD) simulations. The thesis is based on four papers where the RNA structures contain nucleotides that do not form Watson-Crick base pairs. Base flipping is studied by umbrella sampling in two of the papers.

In the first paper terminal loop motifs are studied. Some of the loop motifs are known to be very stable in thermal melting experiments and we have studied if the loops have an intrinsic stability. The loop motifs with purine-purine stacking and hydrogen bonds across the loop are shown to be most capable of retaining the loop structure.

In the second paper an adenosine to inosine (A-to-I) editing site is studied. The A at the editing site is mismatched with a cytidine (C). The editing site (R/G site) is targeted selectively by an editing enzyme from the adenosine deaminases acting on RNA (ADAR) family. This editing site was studied and compared with another A-C mismatch (46 site), a few base pairs away, that is not selectively targeted. The A at the R/G site is shown to be more flexible than the A at the 46 site in the MD simulations. Base flipping of the A at the R/G site and at the 46 site are investigated and the minor groove pathway is found to be preferred over the major groove.

In the third paper the focus is on a part of the spliceosome, the U6 RNA intramolecular stem loop which contains an unpaired uridine (U). In NMR structures this U is in the stack at pH 7.0 and flipped out at pH 5.7. The pathway of this base flipping is studied. We demonstrate that the minor groove pathway is preferred over the major groove pathway, and that protonation on an A adjacent to the flipping U lowers the energy barrier ~3.5 kcal/mol.

In the fourth paper, RNA bulges are studied. The RNA molecules consist of an antisense oligonucleotide forming a complex with a target RNA with an internal bulge loop of different sizes. O2'-methylations are performed on the antisense strand and these methylations are found to influence the minor groove hydration.

LIST OF PUBLICATIONS

I. Nyström B, Nilsson L.
Molecular dynamics study of intrinsic stability in six RNA terminal loop motifs.
Journal of biomolecular structure & dynamics. 2007, 24: 525-535.

II. Hart K, Nyström B, Öhman M, Nilsson L.
Molecular dynamics simulations and free energy calculations of base flipping in dsRNA.
RNA. 2005, 11:609-618.

III. Nyström Macchion B, Nilsson L.
Influence of adenine protonation on uridine bulge dynamics in the U6 RNA intramolecular stem loop.
Submitted 2007.

IV. Nyström Macchion B, Strömberg R, Nilsson L.
Stability and flexibility of RNA complexes containing bulge loops of different sizes.
Submitted 2007.

CONTENTS

LIST OF ABBREVIATIONS

ADAR	Adenosine deaminases acting on ribonucleic acid
AMBER	Assisted Model Building with Energy Refinement
CHARMM	Chemistry at HARvard Molecular Mechanics
COM	Center of mass
DNA	Deoxyribonucleic acid
dsRNA	Double stranded ribonucleic acid
gluR	Glutamate receptor
MD	Molecular dynamics
mRNA	Messenger ribonucleic acid
NMR	Nuclear Magnetic Resonance
OBAN	Oligotide Based Artificial Nuclease
PBC	Periodic Boundary Conditions
PME	Particle Mesh Ewald
PMF	Potential of Mean Force
pre-mRNA	Precursor messenger ribonucleic acid
RNA	Ribonucleic acid
RNase	Ribonuclease
rRNA	Ribosomal ribonucleic acid
siRNA	Small interfering ribonucleic acid
tRNA	Transfer ribonucleic acid
WHAM	Weighted Histogram Analysis Method

1 INTRODUCTION

Living organisms are enormously diverse. Still the cells of a human being use the same mechanisms for basic functions as the cells in a cat or in a tree. Each cell can be compared with a tiny factory where chemical reactions occur and keep the cell alive. Water takes up about 70% and proteins about 20% of the weight of a cell (Alberts 2002). Proteins are the most versatile molecules in living organisms with differing functions like enzymes, receptors and antibodies. It is necessary for cells to have an efficient protein synthesis machinery that fast can deliver new proteins with correct sequence. The subunits of proteins are amino acids which are joined together into one chain which folds into a three dimensional structure. A correct folded protein is necessary for it to function properly.

All cells use deoxyribonucleic acid (DNA) to store genetic information which is passed on from generation to generation. The subunits of DNA are nucleotides and two complementary DNA strands form a double stranded helix. The genetic information in DNA is first copied (transcribed) into a similar molecule, ribonucleic acid (RNA). After transcription the genetic information in the RNA is decoded and translated into a protein at the ribosome. The ribosome is the protein synthesis machine and is a ribozyme (RNA enzyme) which catalyzes the formation of peptide bonds between the amino acids in the protein chain. Not all parts of DNA codes for proteins, different RNA molecules can also be the final products. RNA molecules have different functions in the cell and like proteins they also need to have a certain three dimensional structure to function properly.

RNA molecules have a long history and have been suggested to be important in the origin of life. There are also other functions and aspects of RNA molecules, that can be used in biotechnology and drug design. A few aspects of the importance of RNA molecules are briefly described below.

According to the RNA world hypothesis, origin of life was based on RNA molecules that both store the genetic information and catalyze their own replication so that DNA and proteins were not needed (Gilbert 1986). Later in evolution, proteins became the main molecules that perform catalysis and even later DNA double helices became the molecules that store the genetic information. The RNA world hypothesis is supported by the fact that RNA can act as a ribozyme.

RNA molecules were viewed as the mediators of the genetic information from DNA to protein but are now acknowledged to perform many different functions in a cell, for instance in gene regulation. Fire and Mello were awarded the Nobel prize in physiology and medicine in 2006 for their work on RNA interference (Fire, Xu et al. 1998). Cells use the RNA interference mechanism to regulate gene expression. In RNA interference double stranded RNA molecules are used in a machinery that leads to mRNA degradation which in turn leads to that the protein does not get translated.

In drug design one wants to target a molecule that is associated with disease. Proteins are usually used as targets and the drugs function as inhibitors. RNA molecules can

also be used in drug design. Since the mRNA codes for a protein, the mRNA can be the drug target instead of the protein it codes for. The RNA interference pathway is one way RNA molecules can be used in drug design.

This thesis includes four papers which cover different RNA molecules. The structures I have studied have in common that they contain nucleotides that do not form Watson-Crick base pairs. The parts of RNA molecules that are not double stranded are important for the variety of structures and functions RNA molecules have. In this thesis I first give an introduction to RNA function and structure, followed by RNA in drug design. Next, I describe the methods I have used, the main results from the four papers and finally future perspectives on molecular dynamics simulations and RNA structure.

2 RNA FUNCTION

In the central dogma of molecular biology, DNA→RNA→Protein (Crick 1970), RNA used to be seen as the passive transporter of the genetic information. Today, it is known that RNA is actively participating in gene expression. The best known RNA molecules are those involved in the flow of genetic information but there are also other classes of RNA molecules, like ribozymes and short interfering RNA (siRNA) molecules. Ribozymes are RNA molecules with catalytic activity (Doherty and Doudna 2000) and siRNA:s are double stranded RNA molecules involved in gene regulation (Hannon 2002). This chapter focuses on RNA molecules involved in gene expression and post-transcriptional modifications that are involved in eukaryotic gene expression. Post-transcriptional modifications are important for protein diversity since the nucleotide sequence in the DNA can lead to different versions of a protein.

2.1 RNA IN GENE EXPRESSION

In gene expression, the genetic information that is coded into the nucleotide sequence of the DNA is passed on to the amino acid sequence of the protein or nucleotide sequence of the RNA (Figure 1).

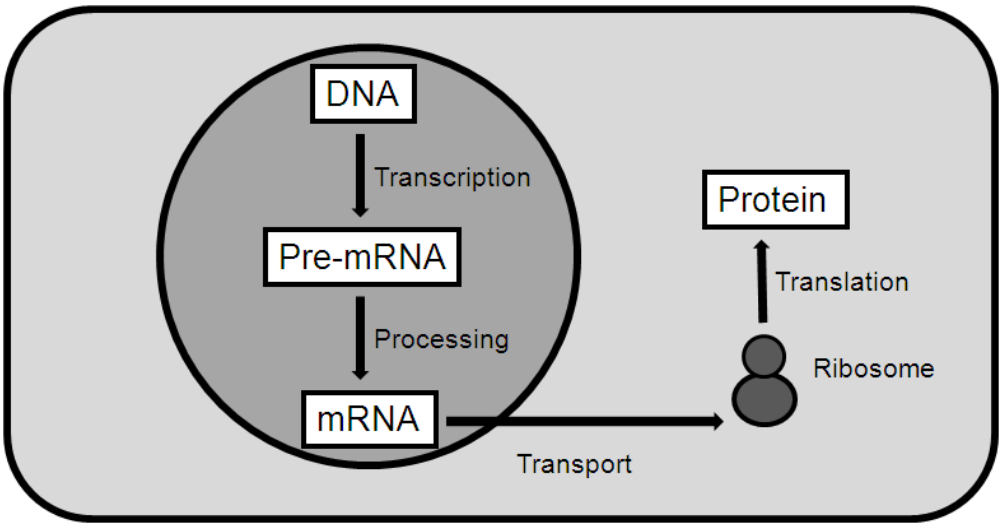

Figure 1. Pathway of genetic information from DNA to protein in an eukaryotic cell.

During transcription the sequence in the DNA is copied into messenger RNA (mRNA) (Stryer 1995). The mRNA carries the genetic code from the DNA to the ribosome where the protein is translated. In contrast to prokaryotic cells which do not have a nucleus, the mRNA in eukaryotic cells needs to be transported from the nucleus to the cytoplasm where the ribosomes are. The genetic information in the mRNA is divided into codons, where each codon consists of three bases. The codon in the mRNA is recognized by the anticodon of a transfer RNA (tRNA). Each transfer RNA (tRNA)

carries a specific amino acid to the site of protein synthesis at the ribosome. The ribosomes consist of ribosomal RNA (rRNA) and proteins. It is rRNA that catalyzes the peptide bond formation in protein synthesis.

2.2 RNA EDITING

Editing is a process where an RNA molecule is altered by insertion, deletion or substitution of nucleotides. If the editing occurs in the mRNA the codons can be changed (Gott and Emeson 2000). The editing can also change start codons, stop codons and splicing sites and is an efficient way to generate protein isoforms from one single mRNA.

2.2.1 A-to-I editing

One common form of RNA editing is the deamination of adenosine (A) to inosine (I), A-to-I editing (Figure 2). During translation, inosine is read as guanosine (G) and if the editing occurs in a coding region the amino acid sequence in the protein is changed and can lead to altered protein function.

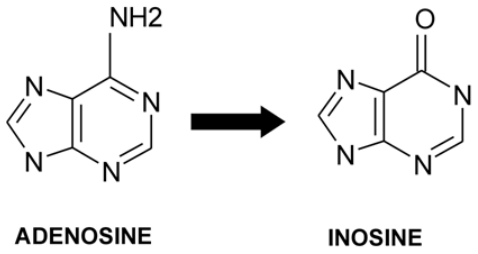

Figure 2. A-to-I editing. Adenosine (A) is converted to inosine (I) during deamination.

The family of adenosine deaminases acting on RNA (ADARs) are enzymes that catalyze A-to-I editing (Bass 2002). ADARs perform two types of editing; hyper-editing and selective editing. The type of editing depends on the structure of the RNA. Hyper-editing occurs when the RNA is long (50 base pairs) and completely double-stranded (Nishikura, Yoo et al. 1991; Polson and Bass 1994). Selective editing can be performed when the double-stranded RNA is interrupted by bulges, loops or mismatches (Bass 2002). The selective editing has been found in pre-mRNA coding for neurotransmitter receptors in the mammalian brain (Sommer, Kohler et al. 1991; Lomeli, Mosbacher et al. 1994; Burns, Chu et al. 1997; Bhalla, Rosenthal et al. 2004). One site that is selectively deaminated is in the mammalian glutamate receptor subunit B (gluR-B) pre-mRNA and this A-to-I editing site is studied in paper II.

2.3 SPLICING

In prokaryotic cells transcription is directly followed by translation, while in eukaryotic cells the transcribed mRNA needs to be processed in several steps before it becomes mature mRNA which can be transported from the nucleus to the cytoplasm where the ribosomes are (Alberts 2002). In eukaryotic cells the transcribed mRNA is called precursor-mRNA (pre-mRNA) and the processing steps are 5'-cap addition, 3'-polyadenylation and splicing. In contrast to prokaryotic DNA which is only coding, eukaryotic DNA contains non-coding regions, exons. In the splicing mechanism of pre-mRNA the non-coding introns are removed and the coding exons are joined together (Figure 3) to form mature mRNA (Brow 2002) .

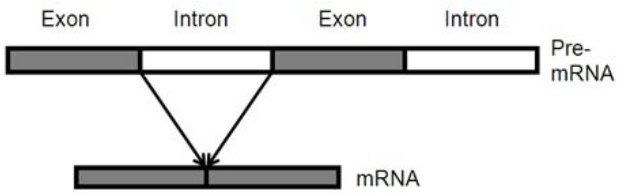

Figure 3. During splicing the coding exons in the pre-mRNA are joined together.

The spliceosome performs the catalysis of splicing. It is a very dynamic ribonucleoprotein complex consisting of five small nuclear RNAs (U1, U2, U4, U5 and U6) and their associated proteins. The spliceosome undergoes several structural rearrangements before it is in its active form. Only parts of it are involved in the actual catalysis. U2 and U6 form a complex that can perform catalysis in a reaction that is similar to the first step of splicing (Valadkhan and Manley 2001). In paper III, the U6 intramolecular stem loop is studied.

Alternative splicing is a mechanism that allows for variations in the resulting mature mRNA by joining together exons in different combinations (Lopez 1998). If several splicing alternatives exist, different proteins can be synthesized from only one piece of DNA. Isoforms of proteins are created by alternative splicing of the pre-mRNA depending on what the cell needs at that moment.

3 RNA STRUCTURE

RNA molecules can have a variety of structures and perform different functions in the cell. Both DNA and RNA are nucleic acids and are built in a similar way, but RNA has a greater variety of three-dimensional structures than DNA. Nucleic acids consist of nucleotide subunits. Despite the difference in three dimensional structures between DNA and RNA both nucleic acids are built up by only four different nucleotides each.

3.1 NUCLEOTIDES

In nucleic acids each nucleotide consists of a base, a sugar and a phosphate group (Figure 4). If the phosphate is removed the nucleotide becomes a nucleoside (Bloomfield, Crothers et al. 2000). The bases which hold the genetic information differ between the different nucleotides while the phosphate and the sugar parts are the same for all nucleotides.

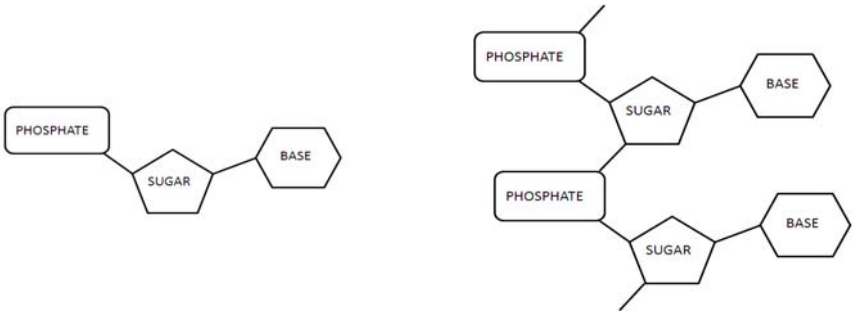

Figure 4. Schematic figures of a nucleotide to the left and two nucleotides in a nucleic acid chain to the right.

In RNA the bases are adenine (A), cytosine (C), guanine (G) and uracil (U). DNA uses thymine (T) instead of U. The bases are divided into purines and pyrimidines depending on whether they consist of one or two aromatic rings. The purines have two rings and are A and G while the pyrimidines have one ring and are C and U/T.

The sugar is ribose in RNA (with –OH on C2') and deoxyribose in DNA (with –H on C2') (Figure 5). The OH-group in RNA is a hydrogen-bonding site and it also makes RNA more susceptible for cleavage.

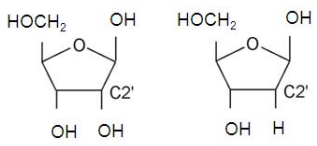

Figure 5. Ribose with –OH on C2' to the left and deoxyribose with –H on C2' to the right.

3.2 DOUBLE HELICAL STRUCTURE AND BASE PAIRING

Watson and Crick described the double helical structure of DNA, with two strands that are joined together by hydrogen bonds between bases of the two strands (Watson and Crick 1953). Since the bases of the two strands form hydrogen bonds, they are positioned on the inside of the double helix while the sugar and phosphates are positioned on the outside (Figure 6). The altering sugars and phosphates form a backbone. The nucleotides are joined together via phosphodiester bonds between the 3´-hydroxyl (-OH) of one nucleotide and the 5'-phosphate (P) of the following nucleotide (Bloomfield, Crothers et al. 2000). The ends of one strand are referred to as the 5'-end and the 3'-end. The two chains in a double helix are in opposite directions. Each base pair consists of one purine and one pyrimidine. The hydrogen bonds that are formed between the bases are of the type N-H···N and N-H···O. The base pairs in double stranded DNA and RNA are called Watson-Crick base pairs and are G-C, A-T (DNA) and A-U (RNA). The G-C base pairs have three hydrogen bonds between the bases while A-T and A-U only have two. In RNA, the G-U wobble pair is also common and adds even more diversity to RNA structure (Varani and McClain 2000). Base pair and wobble pair interactions are not only important for double helical structure, but also in codon-anticodon interactions (Crick 1966).

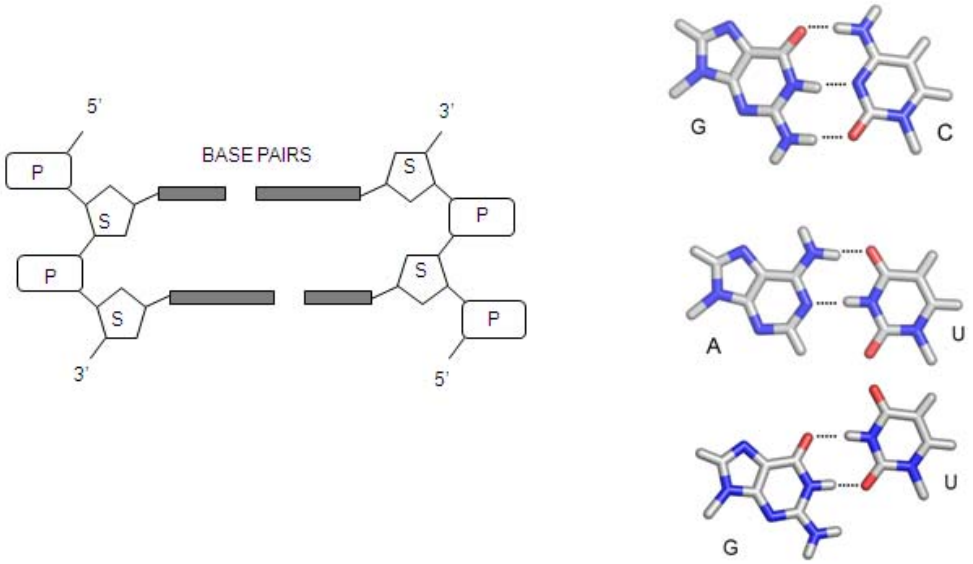

Figure 6. At the left side, a schematic figure of two DNA or RNA strands forming a double strand via base pairs. P is phosphate, S is sugar and the grey rectangles are bases. To the right, G-C, A-U Watson-Crick base pairs and G-U wobble pair.

3.3 BASE STACKING

Apart from base pairing, the bases are also involved in stacking interactions which contribute to the three dimensional nucleic acid structure (Bloomfield, Crothers et al. 2000). The bases are flat and can be arranged so that the aromatic rings are positioned on top of each other (Figure 7). Stacking can be present both between neighboring bases and between bases of different strands like in the figure below. Base stacking involves hydrophobic, electrostatic and dispersion interactions and it is debated which is the most important interaction.

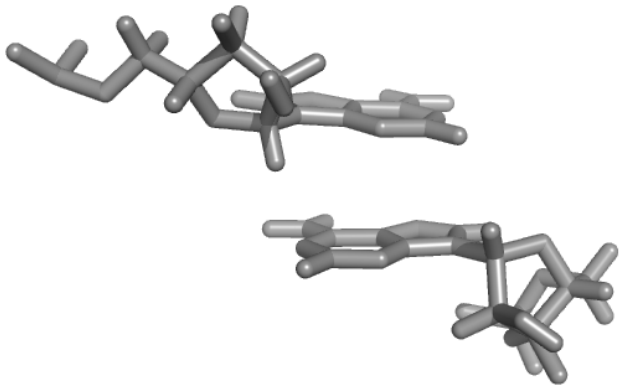

Figure 7. Stacking between two adenines from different RNA strands, cross-strand stacking.

3.4 STRUCTURAL MOTIFS

The structure of DNA is most often double stranded while the double stranded regions of RNA are often interrupted by single stranded regions (Figure 8). There are nucleotide sequences in RNA that often are present where the bases do not form Watson-Crick base pairs and where the phosphodiester backbone has a distinctive fold (Moore 1999; Leontis and Westhof 2003). These sequences are referred to as RNA motifs and they are often involved in tertiary interactions. There are two classes of RNA motifs, according to the topology of the structure, internal loops and hairpin/terminal loops (Moore 1999; Klosterman, Tamura et al. 2002; Leontis and Westhof 2003). In paper I, the intrinsic stability of six terminal loops is studied.

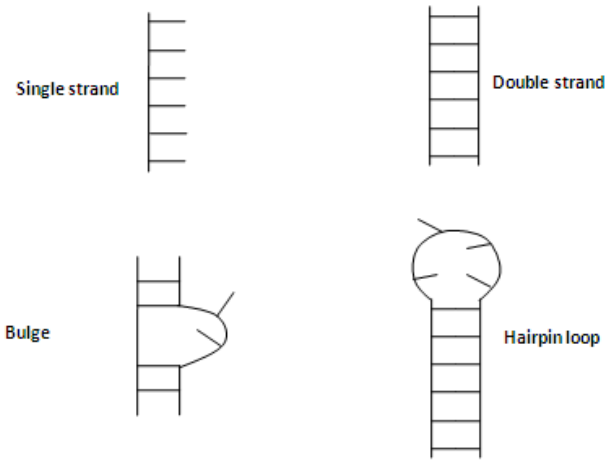

Figure 8. RNA secondary structure.

3.5 BASE FLIPPING

The bases in base pairs are hidden in the helix and cannot serve as recognition sites for interactions like unpaired bases in structural motifs. Base pair opening, where the hydrogen bonds between the bases in the base pair are broken, is common in both DNA and RNA since bases involved in base pairs also need to be recognized. In DNA, base pair opening is required for the strands to be separated during replication and transcription.

While base pair opening is a smaller movement of both bases, base flipping is an event where a base moves out of the helix ~180^{o} and becomes accessible for enzymes. It has been found to occur in systems where enzymes modify or repair individual bases (Roberts and Cheng 1998; Cheng and Roberts 2001). The first crystal structure of such a system to be determined was *Hh*a I methyltransferase, bound to DNA with a flipped out base (Klimasauskas, Kumar et al. 1994). Sequence similarities exist between DNA methyltransferases and ADARs and it has been suggested that base flipping also is involved in ADAR action.

Base pair opening is common in both RNA and DNA and it occurs on a millisecond time scale (Leroy, Charretier et al. 1988; Folta-Stogniew and Russu 1994; Snoussi and Leroy 2001) and for a base to flip out 180^{o} more time is required (Spies and Schowen 2002). The lifetime for these movements is very short and it is difficult to study the pathway for a base to move out of the helix by experiments. Theoretical studies can propose a possible pathway for the base to flip out and also local minima on the way out (Varani and McClain 2000; Giudice, Varnai et al. 2001; Banavali and MacKerell 2002; Giudice, Varnai et al. 2003; Pan and MacKerell 2003). Base flipping is studied in papers II and III.

4 RNA IN DRUG DESIGN

RNA molecules can be used in drug design both as drug targets and as drugs (Eckstein 2007; Melnikova 2007). Instead of targeting a protein associated with a disease, the RNA coding for it can be targeted. Complementary base pairing is an advantage for designing drugs that target mRNA, compared to designing drugs for protein binding sites. Not only mRNA is an RNA drug target, the bacterial ribosome is a drug target for a number of antibiotics and rRNA is involved in drug interactions (Yonath 2005).

There are different classes of RNA drugs such as antisense RNA, siRNA and RNA aptamers. While antisense RNA and siRNA act on mRNA and cause gene silencing, RNA aptamers can bind to various molecular targets, proteins for instance.

There are two approved RNA drugs so far, one RNA aptamer (Macugen) for age-related macular degeneration and one antisense RNA (Vitravene) for AIDS-related cytomegalovirus retinitis (Eckstein 2007; Melnikova 2007). The RNA aptamer targets a protein, the vascular endothelial growth factor, which is involved in the formation of new blood vessels.

In this chapter I will describe the mechanisms of antisense RNA which is used in the oligonucleotide based artificial nuclease (OBAN) systems studied in paper IV.

4.1 ANTISENSE RNA

Antisense oligonucleotides were first used in 1970:s (Stephenson and Zamecnik 1978; Zamecnik and Stephenson 1978). They are single stranded RNA- or DNA-like oligonucleotides, ~18 nucleotides that are complementary to mRNA and can intervene with gene expression by binding to the mRNA (Phillips and Zhang 2000; Eckstein 2007). A protein linked with disease can be prevented from being expressed by introducing synthesized antisense oligonucleotides complementary to the mRNA, also referred to as target RNA.

There are two mechanisms for how the target RNA is prevented from being translated, either by physical blockage at the ribosome by being hybridized with the antisense oligonucleotide or by cleavage of the target RNA by ribonuclease H (RNase H) (Figure 9). Even though the original idea was that the target RNA hybridized with the antisense oligonucleotide and becomes blocked and not translated, it is cleavage by RNase H that is the preferred antisense mechanism.

The antisense oligonucleotide needs to be stable *in vivo* against enzymatic degradation. There are a number of antisense modifications that improve the antisense oligonucleotide by decreasing enzymatic degradation, increasing binding affinity and decreasing toxicity. The early modifications were made to prevent degradation by nucleases and one oxygen atom in the phosphate was exchanged for a sulphur atom (phosphorothioates) (Figure 10a).

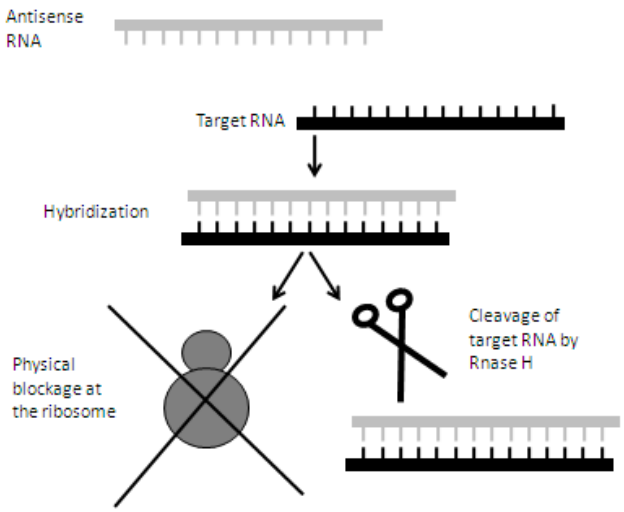

Figure 9. The two mechanisms by which antisense oligonucleotides can stop protein synthesis. The antisense oligonucleotide binds to the mRNA (target RNA) and the mRNA is not translated at the ribosome or is cleaved by RNase H.

The affinity to the target RNA needed to be improved and modifications were made on the sugars of the antisense oligonucleotide, for example O2'-methylations (Figure 10c) (Eckstein 2007). These modifications made the antisense oligonucleotides even more stable against nuclease degradation, but they do not activate RNase H well.

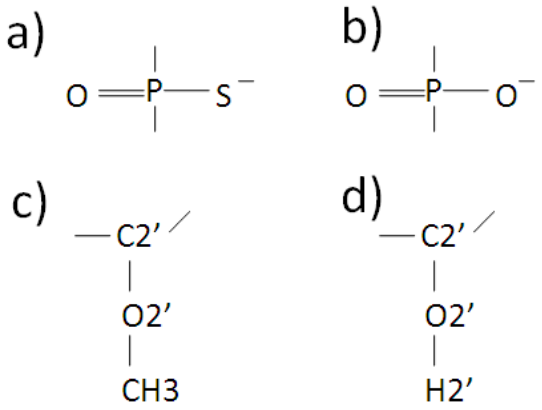

Figure 10. Examples of antisense modifications. In a) one phosphate oxygen atom, as in b), is replaced by a sulphur atom . In c) the H2', as in d), is replaced by a methyl group.

4.1.1 Oligotide Based Artificial Nuclease (OBAN) systems

OBAN systems combine antisense methodology with an artificial cleavage agent (Astrom, Williams et al. 2003). The artificial cleaving agent mimics nucleases and cleave phosphodiester bonds in nucleic acids. In an OBAN system the artificial nuclease is attached to the antisense strand via a linker and the cleavage can be directed to a specific site at the target RNA (Figure 11). The target strand can be made more sensitive for cleavage by the artificial nuclease by introducing a bulge in the target strand when the antisense oligonucleotide and the target RNA hybridize.

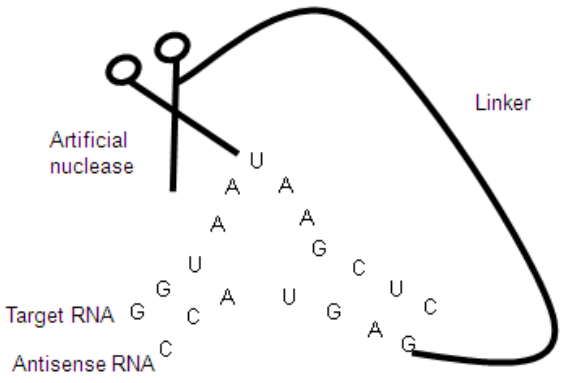

Figure 11. Schematic illustration of an OBAN system where the target RNA contains a bulge and the antisense strand is attached to the artificial cleavage agent via a linker.

5 METHODS

Computer simulations can describe and predict the dynamics of a physical system with known starting conditions. They serve as an important complement to experiments as not every system can be studied by experiments. A computer simulation can be described as an experiment performed in a computer. Molecular dynamics (MD) simulations mimic the behavior of molecules with atomic detail, where the molecular structure is determined by X-ray crystallography, NMR spectroscopy or modeling. The molecule in an MD simulation moves during a period of time and properties of dynamics, structure, interactions etc can be studied. Newton's second law of motion describes the movements of the molecule. In the following chapter the essentials of MD simulations are briefly described. For a comprehensive review please refer to Leach (Leach 2001).

5.1 FORCE FIELDS

A force field sets the rules for how the molecule can move during the MD simulation and is generated by a set of parameters together with a potential energy function:

$$E_{tot} = E_b + E_\theta + E_\phi + E_\omega + E_{vdW} + E_{el} \qquad 1)$$

E_{tot} is the sum of bonded and non-bonded interactions. E_b is the energy for bond length, E_θ for bond angle, E_ϕ for dihedral angle, E_ω for improper dihedral angle, E_{vdW} for van der Waals interactions and E_{ele} for electrostatic interactions (Figure 12).

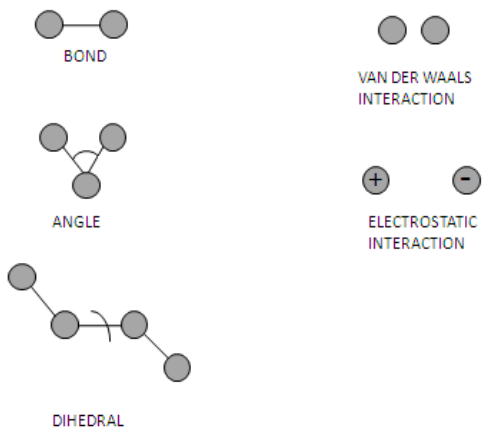

Figure 12. Interactions in the potential energy function in MD simulations.

The parameters used with the potential energy function have been determined using experimental data and quantum mechanical calculations. There are different force fields associated with an MD simulation package like AMBER (Weiner, Kollman et al. 1984) or CHARMM (Brooks, Bruccoleri et al. 1983).

5.2 MOLECULAR DYNAMICS SIMULATIONS

In MD simulations the atoms of a molecule change positions and velocities during time. The force acting on each atom is the negative gradient of the potential energy function. During the simulation a trajectory is created by the integration of Newton's second law of motion. The integration is carried out using a time step, typically 1 fs, and an integration algorithm. The SHAKE (Ryckaert, Ciccotti et al. 1977) algorithm keeps the bond length to hydrogens fixed. The highest frequency motions are removed by using SHAKE and a time step of 2 fs can be used.

Water molecules usually surround a biomolecule in MD simulations. The biomolecule and the water molecules form a finite system with boundaries. By using periodic boundary conditions (PBC) the simulated system is surrounded by image systems in all directions and mimics a larger amount of water molecules (Figure 13). If a particle leaves the system during the simulation it is replaced by an image particle that enters from the opposite side. PBC can be used with different geometrical shapes like a cube or octahedron of water.

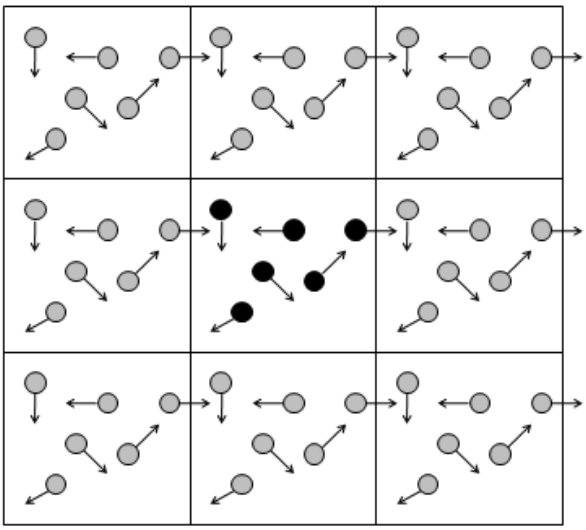

Figure 13. Periodic boundary conditions in two dimenstions. The real system is in the center with black circles.

The most time consuming calculations in MD simulations are those of the long-range electrostatic interactions. These interactions used to be truncated so that only those within a cutoff radius were included to save computational time. The treatment of long-range electrostatic interactions is important in charged molecules like RNA. The atom-based force shift method (Steinbach and Brooks 1994) or the particle mesh Ewald (PME) (Darden, York et al. 1993) summation methods are two ways of dealing with the long-range electrostatic interactions that perform well for nucleic acids (Norberg and Nilsson 2000).

5.3 MD SIMULATIONS OF NUCLEIC ACIDS

The first MD simulation of a biomolecule was performed on the Bovine Pancreatic Trypsin Inhibitor (BPTI) in 1977 (McCammon, Gelin et al. 1977), the first MD simulation of a nucleic acid was of a double-stranded 12 base pair DNA in 1983 (Levitt 1983). The early MD simulations were performed *in vacuo* but now MD simulations of nucleic acids are performed using explicit water molecules around the DNA or RNA molecule. The use of explicit water together with better force fields and better treatment of long-range electrostatic interactions have improved the quality of MD simulations of nucleic acids . Between 1998-2003 a number of reviews on MD simulations on nucleic acids were published (Auffinger and Westhof 1998; Beveridge and McConnell 2000; Cheatham and Kollman 2000; Cheatham and Young 2000; Norberg and Nilsson 2002; Norberg and Nilsson 2003). Today, many different types of RNA have been simulated from an entire tRNA (Auffinger, Louise-May et al. 1999) to an entire ribosome (Sanbonmatsu, Joseph et al. 2005).

5.4 UMBRELLA SAMPLING

Base flipping is an event that occurs on a millisecond time scale, but MD simulations are usually performed on a nanosecond time scale. In MD simulations one can study equilibrium between configurations by studying how often the different configurations exist. There are energetic barriers that make it difficult to study the unfavorable configurations like a flipped out base. Base flipping can still be studied using the umbrella sampling method. It is a method that samples both favorable and unfavorable states. A weighting function is used and becomes larger for configurations that are far away from the equilibrium:

$$w_i(\omega) = k_i(\omega - \omega_i) \qquad 2)$$

where ω, is the reaction coordinate, ω_i is a reference value and k_i is a force constant. The reaction coordinate describes the progression along the reaction pathway and it is usually a geometrical parameter. It is not obvious how to choose the reaction coordinate for base flipping but we used a center of mass (COM) pseudo-dihedral that was used for base flipping before (Figure 14) (Banavali and MacKerell 2002). The choice of reaction coordinate decides the pathway and the resulting free energy.

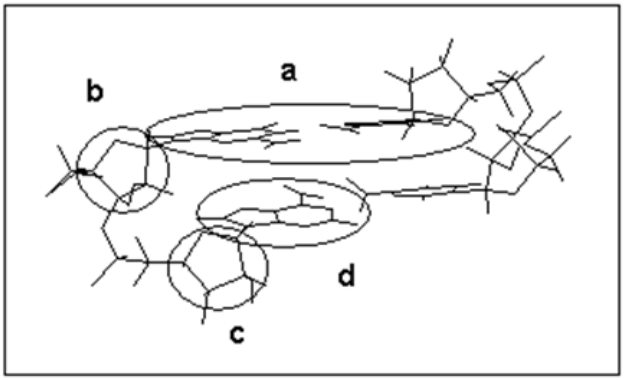

Figure 14. Center of mass dihedral used for base flipping. a) is the base pair 3' to the flipping base, b) is the sugar 3' to the flipping base, c) is the sugar of the flipping base and d) is the flipping base.

The free energy along a reaction coordinate is called the potential of mean force (PMF) and is generated using the Weighted Histogram Analysis Method (WHAM) (Kumar, Bouzida et al. 1992; Boczko and Brooks 1993). Umbrella sampling was used in paper II-III to force a base to move out of the helix 180^{o} through the minor and major groove and the resulting PMF describes how much energy it costs to flip out the base.

6 RESULTS AND CONCLUSIONS

6.1 PAPER I

Molecular dynamics study of intrinsic stability in six RNA terminal loop motifs:
In a terminal loop motif the RNA folds back after the loop and creates a stem, like a tetraloop or a U-turn. There are tetraloop motifs that are common in ribosomal RNA and that have been shown to be very stable in thermal melting experiments (Tuerk, Gauss et al. 1988; Antao, Lai et al. 1991; Antao and Tinoco 1992). Previous work on tetraloop stability has been made in the context of larger structures (Williams and Hall 2000; Li, Ma et al. 2001; Sorin, Engelhardt et al. 2002; Sarzynska, Nilsson et al. 2003; Koplin, Mu et al. 2005; Sorin, Rhee et al. 2005); here we study possible intrinsic stability in six different terminal loop motifs.

The sequences in this study are chosen from the major tetraloop motifs and also one U-turn motif and are $UG_{OM}AA$ (G_{OM}=O2'-methylguanosine), AUCA, GCUUGC, GCAA, AGUU and UUCG (Figure 15). The motifs were simulated at two temperatures (277 K and 300 K) and with adjacent base pairs at 300 K. Three independent 5 ns simulations were performed for all systems.

The degree of unfolding of the loops was measured using a distance (R_{ee}) across the loop. R_{ee} was chosen between two atoms on each side of the loop and was less than 3 Å for all loops in the starting structures. The structures with R_{ee}< 4 Å were defined as folded and the degree of folded structure of the simulation time was calculated. Most of the simulations lead to unfolding of the loop structure according to the R_{ee} distance. Lowering the temperature or adding base pairs to the loop motifs can prevent the loop from unfolding but unfolding still occurs in some of the loop motifs.

The loops that unfold most easily are AUCA and AGUU, both with the fourth base A4 and U4 pointing out of the loop in the starting structures. $UG_{OM}AA$ is found to be the most stable loop because of hydrogen bonding from U1 to A3 and A4 and the three purines stacked upon each other. GCAA, which has a similar structure as $UG_{OM}AA$, also has a hydrogen bond from G1 to A4 and it has two purines stacked upon each other. GCUUGC also show intrinsic stability, because of the two G-C base pairs.

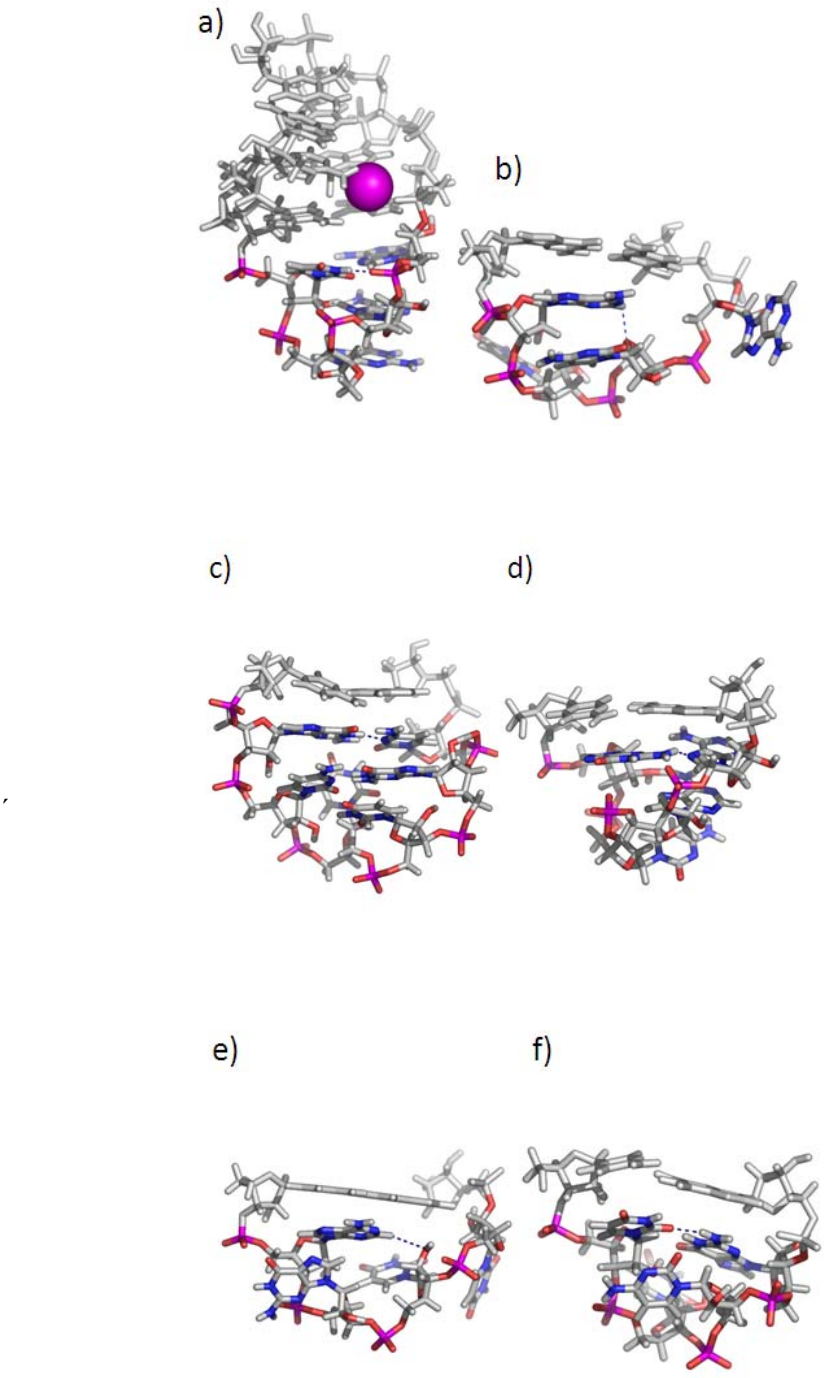

Figure 15. Terminal motifs in color with additional nucleotides, in the long sequences, in white. a) is $UG_{OM}AA$, b) is AUCA, c) is GCUUGC, d) is GCAA, e) is AGUU and f) is UUCG. The dashed lines show the R_{ee} distance across the loop for each structure. The magenta colored sphere in a) indicates the position of a Mg^{2+} ion.

6.2 PAPER II

Molecular dynamics simulations and free energy calculations of base flipping in dsRNA:
Adenosine deaminases acting on RNA (ADARs) are enzymes that catalyze the deamination of adenosines into inosines in RNA that is mainly double stranded. ADAR enzymes act nonselectively in double stranded RNA (dsRNA) and more selectively where the dsRNA in disrupted by loops, mismatches or bulges (Bass 2002).

One site that is selectively deaminated is in the mammalian glutamate receptor subunit B (gluR-B) pre-mRNA (Figure 16). The R/G site (changes arginine to glycine in the codon) of this substrate is efficiently edited by ADAR2 *in vivo* (Higuchi, Maas et al. 2000). The adenosine at the R/G site is mismatched with a cytidine and a few base pairs away there is another AC mismatch (46 site) that is not deaminated *in vivo*.

One of the aims for our simulations was to study the two similar sites and see if there are structural or dynamical differences that direct ADAR2 to the R/G site and not the 46 site. It has been suggested that the adenosine flips out of the helix before deamination. The other aim of the simulations was to study the base flipping at both sites.

In our simulations we have compared the R/G site with the 46 site and mutated the C into a U, to have an A-U base pair, at both sites to investigate the selectivity of ADAR2. Base flipping at both sites was also studied using umbrella sampling. Five sequences were used for the MD simulations; the R/G site (R/G-wt) with three base pairs at each side of AC, the R/G site with the C/U mutation (R/G-mut), the 46 site (46-wt) with three base pairs on each side of AC, the 46 site with the C/U mutation (46-mut) and one longer sequence of 17 base pairs including both the R/G and 46 sites. The long sequence was simulated in three independent runs of 6 ns each. R/G-wt and 46-wt were simulated 10 ns each while R/G-mut and 46-mut were simulated for 5 ns each.

In the MD simulations, the A at the R/G site is more flexible than the A at the 46 site. The nearby GU wobble pair is involved in movements of the A at the R/G site in the short sequence. The free energy calculations show that the minor groove pathway is more favorable than the major groove pathway. There is a local minimum in the minor groove at 50^{o}, where it costs less energy for R/G-wt than 46-wt to be in. The MD simulations of the 17-base pair sequence also show that the A at the R/G-site spends more of the simulation time at 50^{o} than the A at the 46 site.

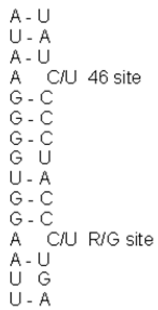

Figure 16. Part of the sequence of glutamate receptor B pre-mRNA.

6.3 PAPER III

Influence of adenine protonation on uridine bulge dynamics in the U6 RNA intramolecular stem loop:
There is an unpaired uridine, U80, in the U6 intramolecular stem loop of the spliceosome (Figure 17) that is stacked at neutral pH and flipped out at lower pH (Reiter, Blad et al. 2004). The uridine bulge in U6 is important for splicing since Mg^{2+} ion binding to U80 is essential for splicing (Yean, Wuenschell et al. 2000). Next to U80 there is a wobble pair, A79-C67, where A79 is protonated at the lower pH. When U80 is out of the stack the conformation of the helix is changed by straightening it up.

MD simulations have been performed on the NMR structures with U80 in the stack and flipped out. The protonation state of A79 has also been changed in the starting structures. Umbrella sampling was used to force U80 out from the stack with and without A79 protonated. These simulations have been made to investigate the dynamical behavior in the region around U80. The pathway of the base flipping cannot be seen in the NMR structures but can be studied with the umbrella sampling.

U80 moves towards the major groove upon protonation of A79 in the unbiased simulations. In the free energy calculations of the U80 base flipping, the energy barrier is approximately 3.5 kcal/mol lower when A79 is protonated and the minor groove pathway is favorable.

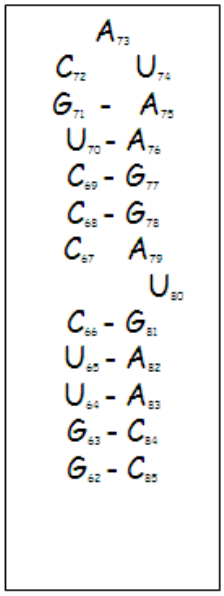

Figure 17. Secondary structure of the U6 intramolecular stem loop.

6.4 PAPER IV

Stability and flexibility of RNA complexes containing bulge loops of different sizes:

RNA bulge loops are used for testing artificial cleavage enzymes in oligonucleotide based artificial nuclease (OBAN) systems (Astrom, Williams et al. 2003). The target strand contains the bulge, which is cleaved, and the antisense strand has a linker with the artificial enzyme attached to it.

There is a bulge loop with structure determined by both X-ray crystallography (Cate, Gooding et al. 1996) and NMR spectroscopy (Luebke, Landry et al. 1997). The sequence from this bulge was used in OBAN-experiments by Roger Strömberg and co-workers and the NMR structures were used for our simulations (Figure 18). In the OBAN experiments, the linker is attached to different positions in the antisense strand in cleavage experiments and different bulge sizes were used (Astrom, Williams et al. 2003). In thermal melting experiments, the stability of the antisense-target complex was tested with different bulge sizes and different modifications in the antisense strand (Sandbrink, Ossipov et al. 2005).

These MD simulations are made to study how different bulge size and O2'-methylations in the antisense strand affect the complex. Bulges with five, four and three nucleotides were simulated.

The bulge introduces flexibility in the double stranded complex and the nucleotides of the bulge are the most flexible within the complexes. The O2'-methylations change the water distribution in the minor groove.

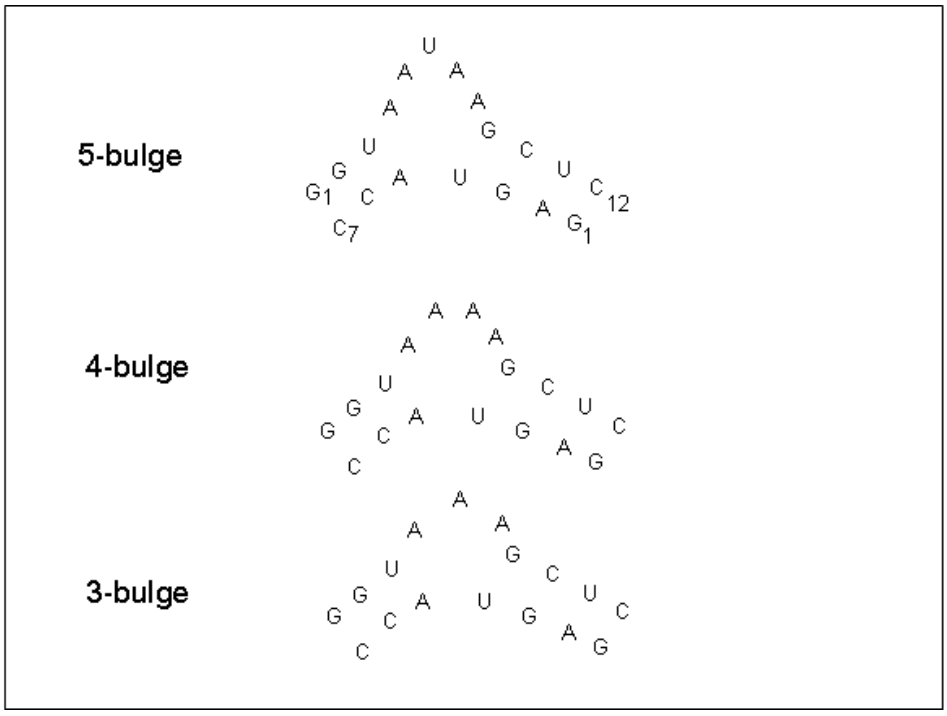

Figure 18. Bulges of different sizes. The bulge with five nucleotides (5-bulge) has the same sequence as in the NMR structures 1ajl and 1ajt.

7 FUTURE PERSPECTIVES

The total number of molecular structures in the Protein Data Bank (PDB) (Berman, Westbrook et al. 2000) was 47, 625 in December 2007. During 2007, 6835 structures were released at PDB and 5887 out of these were solved by X-ray crystallography. MD simulations provide insights on the dynamics of a structure. Molecules in a cell are not static, for example the spliceosome and the ribosome are dynamic machines. A static structure cannot explain the dynamical behavior of a molecule. With the growing number of structures released at PDB there are many questions that just a structure does not provide an answer to. Even though not all structures at PDB are suitable for MD simulations like electron microscopy structures and X-ray structures with low resolution, many structures can be used for MD simulations and their dynamics examined further. A wider range of biological questions can be addressed using MD simulations as the number of PDB structures increase.

MD simulations themselves can be improved in a number of ways like the accuracy, the number of molecules that can be simulated and the speed of the simulations. Force fields are developed continuously to make the simulations more realistic and to make it possible to simulate more types of molecules. MD simulations can be speeded up by faster computers. With greater computational power the simulations can be made longer and/or more simulations can be performed. Longer simulations give better statistics and also a possibility to study events in the cell that occur on a longer time-scale.

RNA molecules are promising in therapeutics since they can be used to prevent gene expression of disease-related proteins. The synthesis of a protein is prevented and not the inhibition of the activity of the protein. The aim of both antisense RNA and siRNA is to destroy mRNA and both types of RNA molecules are potential as RNA drugs. Even with other targets than mRNA molecules, RNA aptamers can be used and bind to a protein and inhibit its activity. RNA-based therapies in clinical trial at the moment (November 2007) act against for instance cancer, asthma and diabetes (Melnikova 2007). Of the presented RNA-based therapies in the review by Melnikova, most are antisense RNA:s and the others are aptamers and siRNA:s. In another recent review (Eckstein 2007) examples of oligonucleotides in clinical trials are presented. Also here, the majority of the oligonucleotides are antisense oligonucleotides. The RNA interference pathway using siRNA is more recently discovered than the antisense mechanism and is already in clinical trials.

RNA molecules are advantageous as drug agents since they target mRNA and can stop any kind of protein from being synthesized and also since base pairing can be used for drug design when the sequence of the mRNA is known. RNA molecules as therapeutic agents are still under development and the difficulties RNA molecules meet are delivery to the cells that are affected by the disease, resistance against enzymatic degradation, binding affinity and the possibility to bind to another part of mRNA to name a few. Despite the difficulties, RNA molecules have a huge potential in therapy of human diseases.

There are only two approved RNA drugs at the moment but investments have been made in the pharmaceutical industry during the past two years to develop RNA-based drugs (Melnikova 2007). Billion-dollar deals have been made between biotechnology and pharmaceutical companies around antisense RNA, RNA interference and aptamers. RNA-based drugs could become the next generation of medicines.

Finally, I believe that both MD simulations and RNA structure will be important for biomedical research in the future.

8 ACKNOWLEDGEMENTS

I would like to thank...

Prof. Lennart Nilsson, for taking me as a student, for your patience and for having a joke ready for every occasion!

The people involved in the papers: Marie Öhman, Samuel Butcher and prof. Roger Strömberg.

The members of the molecular modeling group: Ana, Cristian, Katarina, Olof and Sofia. What would a working day be like without your company and our ENORMOUS lunches at MF?

Past group members: Jan, Jianxin, Johan, Pekka, Peter and Vineet.

All colleagues at CSB, especially Anna-Karin, Caroline, Johan, Katja, Kimberley, Micke, Tobias and Xiaofeng.

Prof. Kurt Berndt and Lotta Hambraeus who made me interested in structural biology and physical chemistry during my undergraduate studies. I learned a lot later while teaching at your courses.

Administrativ och IT-personal, framför allt Erik och Kristina.

Vänner i Stockholm som jag har kul med och som är snälla och låter mig bo hos er: Minna & Abbe, Linda & Christer.

Hanna & Jonathan för roliga utflykter när jag är i Göteborg.

Skåne-kusinerna i Stockholm: Linen & Sara. Det är fantastiskt med familj i närheten.

Alla andra vänner och familjemedlemmar.

Stefan som alltid har god mat redo samt skojarna Ellen & Tilda.

Ina, världens bästa syster!

Mamma & Pappa för allt stöd under den här tiden.

Olivier, le plus grand MERCI du monde!

9 REFERENCES

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. (2002). Molecular biology of the cell. New York, Garland Science.

Antao, V. P., S. Y. Lai, et al. (1991). "A thermodynamic study of unusually stable RNA and DNA hairpins." Nucleic Acids Res **19**(21): 5901-5.

Antao, V. P. and I. Tinoco, Jr. (1992). "Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops." Nucleic Acids Res **20**(4): 819-24.

Astrom, H., N. H. Williams, et al. (2003). "Oligonucleotide based artificial nuclease (OBAN) systems. Bulge size dependence and positioning of catalytic group in cleavage of RNA-bulges." Organic & Biomolecular Chemistry **1**(9): 1461-5.

Auffinger, P., S. Louise-May, et al. (1999). "Molecular dynamics simulations of solvated yeast tRNA(Asp)." Biophys J **76**(1 Pt 1): 50-64.

Auffinger, P. and E. Westhof (1998). "Simulations of the molecular dynamics of nucleic acids." Curr Opin Struct Biol **8**(2): 227-36.

Banavali, N. K. and A. D. MacKerell, Jr. (2002). "Free energy and structural pathways of base flipping in a DNA GCGC containing sequence." J Mol Biol **319**(1): 141-60.

Bass, B. L. (2002). "RNA editing by adenosine deaminases that act on RNA." Annu Rev Biochem **71**: 817-46.

Berman, H. M., J. Westbrook, et al. (2000). "The Protein Data Bank." Nucleic Acids Research **28**(1): 235-242.

Beveridge, D. L. and K. J. McConnell (2000). "Nucleic acids: theory and computer simulation, Y2K." Curr Opin Struct Biol **10**(2): 182-96.

Bhalla, T., J. J. Rosenthal, et al. (2004). "Control of human potassium channel inactivation by editing of a small mRNA hairpin." Nat Struct Mol Biol **11**(10): 950-6.

Bloomfield, V., D. Crothers, et al. (2000). Nucleic acids structures, properties and functions, University Science Books.

Boczko, E. M. and C. L. Brooks (1993). "Constant-Temperature Free-Energy Surfaces for Physical and Chemical Processes." Journal of Physical Chemistry **97**(17): 4509-4513.

Brooks, B. R., R. E. Bruccoleri, et al. (1983). "CHARMM: A program for Macromolecular Energy Minimization and Dynamics Calculations." Journal of Computational Chemistry **4**: 187-217.

Brow, D. A. (2002). "Allosteric cascade of spliceosome activation." Annu Rev Genet **36**: 333-60.

Burns, C. M., H. Chu, et al. (1997). "Regulation of serotonin-2C receptor G-protein coupling by RNA editing." Nature **387**(6630): 303-8.

Cate, J. H., A. R. Gooding, et al. (1996). "Crystal structure of a group I ribozyme domain: principles of RNA packing." Science **273**(5282): 1678-85.

Cheatham, T. E., 3rd and P. A. Kollman (2000). "Molecular dynamics simulation of nucleic acids." Annu Rev Phys Chem **51**: 435-71.

Cheatham, T. E., 3rd and M. A. Young (2000). "Molecular dynamics simulation of nucleic acids: successes, limitations, and promise." Biopolymers **56**(4): 232-56.

Cheng, X. and R. J. Roberts (2001). "AdoMet-dependent methylation, DNA methyltransferases and base flipping." Nucleic Acids Res **29**(18): 3784-95.

Crick, F. (1970). "Central Dogma of Molecular Biology." Nature **227**(5258): 561-&.

Crick, F. H. (1966). "Codon--anticodon pairing: the wobble hypothesis." J Mol Biol **19**(2): 548-55.

Darden, T., D. York, et al. (1993). "Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems." Journal of Chemical Physics **98**(12): 10089-10092.

Doherty, E. A. and J. A. Doudna (2000). "Ribozyme structures and mechanisms." Annual Review of Biochemistry **69**: 597-615.

Eckstein, F. (2007). "The versatility of oligonucleotides as potential therapeutics." Expert Opin Biol Ther **7**(7): 1021-34.
Fire, A., S. Xu, et al. (1998). "Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans." Nature **391**(6669): 806-11.
Folta-Stogniew, E. and I. M. Russu (1994). "Sequence dependence of base-pair opening in a DNA dodecamer containing the CACA/GTGT sequence motif." Biochemistry **33**(36): 11016-24.
Gilbert, W. (1986). "Origin of Life - the Rna World." Nature **319**(6055): 618-618.
Giudice, E., P. Varnai, et al. (2001). "Energetic and conformational aspects of A : T base-pair opening within the DNA double helix." Chemphyschem **2**(11): 673-+.
Giudice, E., P. Varnai, et al. (2003). "Base pair opening within B-DNA: free energy pathways for GC and AT pairs from umbrella sampling simulations." Nucleic Acids Res **31**(5): 1434-43.
Gott, J. M. and R. B. Emeson (2000). "Functions and mechanisms of RNA editing." Annual Review of Genetics **34**: 499-U34.
Hannon, G. J. (2002). "RNA interference." Nature **418**(6894): 244-51.
Higuchi, M., S. Maas, et al. (2000). "Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2." Nature **406**(6791): 78-81.
Klimasauskas, S., S. Kumar, et al. (1994). "HhaI methyltransferase flips its target base out of the DNA helix." Cell **76**(2): 357-69.
Klosterman, P. S., M. Tamura, et al. (2002). "SCOR: a structural classification of RNA database." Nucleic Acids Research **30**: 392-394.
Koplin, J., Y. Mu, et al. (2005). "Structure and dynamics of an RNA tetraloop: a joint molecular dynamics and NMR study." Structure **13**(9): 1255-67.
Kumar, S., D. Bouzida, et al. (1992). "The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules .1. The Method." Journal of Computational Chemistry **13**(8): 1011-1021.
Leach, A. (2001). Molecular Modelling principles and applications. Harlow, England, Prentice Hall.
Leontis, N. B. and E. Westhof (2003). "Analysis of RNA motifs." Curr Opin Struct Biol **13**(3): 300-8.
Leroy, J. L., E. Charretier, et al. (1988). "Evidence from base-pair kinetics for two types of adenine tract structures in solution: their relation to DNA curvature." Biochemistry **27**(25): 8894-8.
Levitt, M. (1983). "Computer simulation of DNA double-helix dynamics." Cold Spring Harb Symp Quant Biol **47 Pt 1**: 251-62.
Li, W., B. Ma, et al. (2001). "Molecular dynamics simulations of the denaturation and refolding of an RNA tetraloop." J Biomol Struct Dyn **19**(3): 381-96.
Lomeli, H., J. Mosbacher, et al. (1994). "Control of kinetic properties of AMPA receptor channels by nuclear RNA editing." Science **266**(5191): 1709-13.
Lopez, A. J. (1998). "Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation." Annu Rev Genet **32**: 279-305.
Luebke, K. J., S. M. Landry, et al. (1997). "Solution conformation of a five-nucleotide RNA bulge loop from a group I intron." Biochemistry **36**(33): 10246-55.
McCammon, J. A., B. R. Gelin, et al. (1977). "Dynamics of folded proteins." Nature **267**(5612): 585-90.
Melnikova, I. (2007). "RNA-based therapies." Nature Reviews Drug Discovery **6**(11): 863-864.
Moore, P. B. (1999). "Structural motifs in RNA." Annu Rev Biochem **68**: 287-300.
Nishikura, K., C. Yoo, et al. (1991). "Substrate specificity of the dsRNA unwinding/modifying activity." Embo J **10**(11): 3523-32.
Norberg, J. and L. Nilsson (2000). "On the truncation of long-range electrostatic interactions in DNA." Biophys J **79**(3): 1537-53.
Norberg, J. and L. Nilsson (2002). "Molecular dynamics applied to nucleic acids." Acc Chem Res **35**(6): 465-72.
Norberg, J. and L. Nilsson (2003). "Advances in biomolecular simulations: methodology and recent applications." Q Rev Biophys **36**(3): 257-306.

Pan, Y. and A. D. MacKerell, Jr. (2003). "Altered structural fluctuations in duplex RNA versus DNA: a conformational switch involving base pair opening." Nucleic Acids Res **31**(24): 7131-40.
Phillips, M. I. and Y. C. Zhang (2000). "Basic principles of using antisense oligonucleotides in vivo." Methods in Enzymology **313**: 46-56.
Polson, A. G. and B. L. Bass (1994). "Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase." Embo J **13**(23): 5701-11.
Reiter, N. J., H. Blad, et al. (2004). "Dynamics in the U6 RNA intramolecular stem-loop: a base flipping conformational change." Biochemistry **43**(43): 13739-47.
Roberts, R. J. and X. D. Cheng (1998). "Base flipping." Annual Review of Biochemistry **67**: 181-+.
Ryckaert, J. P., G. Ciccotti, et al. (1977). "Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes." Journal of Computational Chemistry **23**: 327-341.
Sanbonmatsu, K. Y., S. Joseph, et al. (2005). "Simulating movement of tRNA into the ribosome during decoding." Proceedings of the National Academy of Sciences of the United States of America **102**(44): 15854-15859.
Sandbrink, J., D. Ossipov, et al. (2005). "Investigation of potential RNA bulge stabilizing elements." Journal of Molecular Recognition **18**(4): 318-26.
Sarzynska, J., L. Nilsson, et al. (2003). "Effects of base substitutions in an RNA hairpin from molecular dynamics and free energy simulations." Biophys J **85**(6): 3445-59.
Snoussi, K. and J. L. Leroy (2001). "Imino proton exchange and base-pair kinetics in RNA duplexes." Biochemistry **40**(30): 8898-904.
Sommer, B., M. Kohler, et al. (1991). "RNA editing in brain controls a determinant of ion flow in glutamate-gated channels." Cell **67**(1): 11-9.
Sorin, E. J., M. A. Engelhardt, et al. (2002). "RNA simulations: probing hairpin unfolding and the dynamics of a GNRA tetraloop." J Mol Biol **317**(4): 493-506.
Sorin, E. J., Y. M. Rhee, et al. (2005). "Does water play a structural role in the folding of small nucleic acids?" Biophys J **88**(4): 2516-24.
Spies, M. A. and R. L. Schowen (2002). "The trapping of a spontaneously "flipped-out" base from double helical nucleic acids by host-guest complexation with beta-cyclodextrin: the intrinsic base-flipping rate constant for DNA and RNA." J Am Chem Soc **124**(47): 14049-53.
Steinbach, P. J. and B. R. Brooks (1994). "New Spherical-Cutoff Methods for Long-Range Forces in Macromolecular Simulation." Journal of Computational Chemistry **15**(7): 667-683.
Stephenson, M. L. and P. C. Zamecnik (1978). "Inhibition of Rous sarcoma viral RNA translation by a specific oligodeoxyribonucleotide." Proc Natl Acad Sci U S A **75**(1): 285-8.
Stryer, L. (1995). Biochemistry. New York, W. H. Freeman and Company.
Tuerk, C., P. Gauss, et al. (1988). "CUUCGG hairpins: extraordinarily stable RNA secondary structures associated with various biochemical processes." Proc Natl Acad Sci U S A **85**(5): 1364-8.
Valadkhan, S. and J. L. Manley (2001). "Splicing-related catalysis by protein-free snRNAs." Nature **413**(6857): 701-7.
Varani, G. and W. H. McClain (2000). "The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems." EMBO Rep **1**(1): 18-23.
Watson, J. D. and F. H. Crick (1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid." Nature **171**(4356): 737-8.
Weiner, S. J., P. A. Kollman, et al. (1984). "A New Force-Field for Molecular Mechanical Simulation of Nucleic-Acids and Proteins." Journal of the American Chemical Society **106**(3): 765-784.
Williams, D. J. and K. B. Hall (2000). "Experimental and computational studies of the G[UUCG]C RNA tetraloop." J Mol Biol **297**(5): 1045-61.
Yean, S. L., G. Wuenschell, et al. (2000). "Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome." Nature **408**(6814): 881-4.

Yonath, A. (2005). "Antibiotics targeting ribosomes: resistance, selectivity, synergism and cellular regulation." Annu Rev Biochem **74**: 649-79.
Zamecnik, P. C. and M. L. Stephenson (1978). "Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide." Proc Natl Acad Sci U S A **75**(1): 280-4.