4. Introduction

The central aim of molecular biology can be seen as to explain the workings of living systems at the level of interaction of individual molecules. Central to our understanding of life is the observation that genetic information is stored in DNA molecules and used by a mostly protein based machinery to synthesize a set of different proteins which in turn give rise to all other biological phenomena. It is obvious that the interactions between proteins and DNA are essential for the proper functioning of living systems.

Many DNA binding proteins do not bind to DNA in general, but rather recognize specific DNA sequences. In this way proteins can be localized to certain positions on the DNA molecule. One example are the interactions of transcription factors with their cognate sequence motifs in the promoter regions of genes. It is therefore of great importance to establish the sequence specificity of the interaction of a protein with DNA. For example, knowledge of the preferred binding sequence of a transcription factor enables the prediction of genes that might be influenced by this transcription factor. In the past, there used to be no straightforward way to establish the sequence specificity of a protein, except to synthesize individual DNA oligomers and test them individually for binding. However, due to the sheer number of possible different sequences, such an approach is limited in practice to cases were there is some indication as to what the actual binding sequence might be. In recent years the newly developed methodology of *in vitro* selection has changed this situation completely.

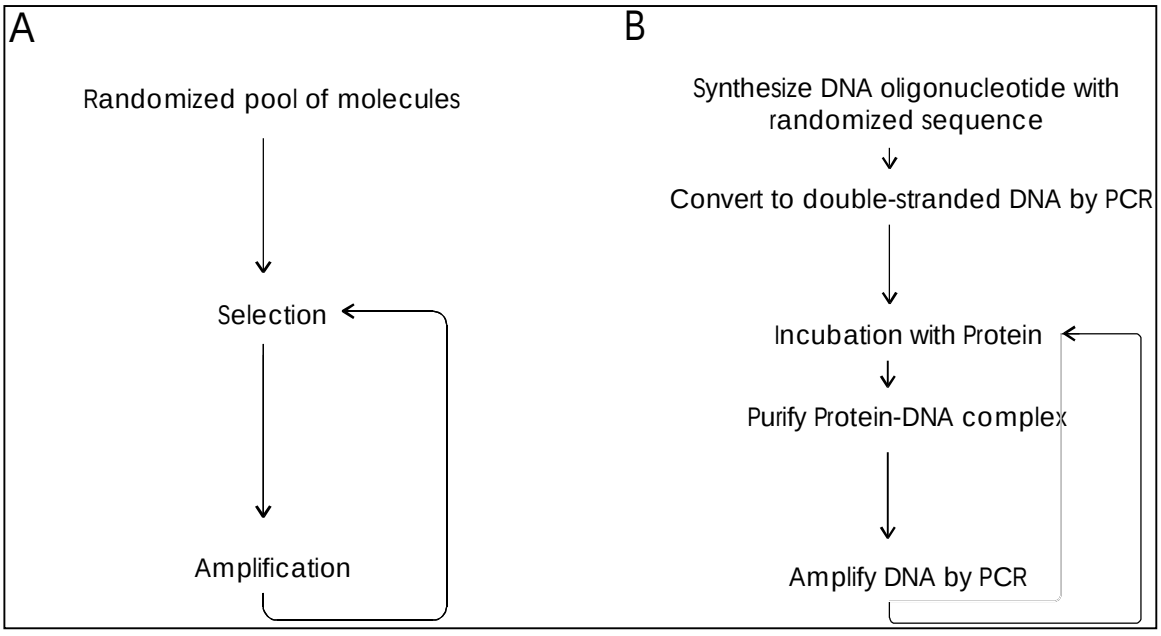

Figure 4.1 Principle of *in vitro* selection

4.1. *In vitro* selection

The typical goal of an *in vitro* selection experiment is the generation of a set of RNA or DNA sequences able to perform a certain function, such as to bind to a ligand or catalyze a chemical reaction (Bartel and Szostak 1993). The strategy used to achieve this goal as exemplified by a DNA binding protein is depicted in Figure 4.1. Panel B. A pool of randomized sequences is prepared and incubated with protein followed by purification of the protein:DNA complex. The selected DNA molecules are then amplified by PCR and subjected to further rounds of selection. While the selection of DNA ligands is the topic of this work, the *in vitro* selection procedure can be understood in a more general way (Panel A). Examples are the selection of intrinsically bent DNA (Beutel and Gold 1992), or RNA molecules with a catalytic function such as RNAzyme Ligases (Bartel and Szostak 1993). Central to the success of *in vitro* selection is the possibility to use several cycles of selection and amplification. It allows for successful selection despite limited discrimination between molecules with and without the desired characteristics. After each round of selection the pool is more and more enriched for those molecules with the desired characteristics. This characteristic behavior of an *in vitro* selection is the basis for the term SELEX (Systematic Evolution of Ligands by Exponential enrichment) (Tuerk and Gold 1990) and the selected molecules have been called ´aptamers´ (Ellington and Szostak 1990). Since the ligands have to be present in excess over the selecting molecules, the selected ligands have to be amplified. Therefore the experiment consists of cycles of selection and amplification reactions. In practice it turns out that the generation of the initial pool of ligands and the subsequent selection step are relatively easy to achieve. The amplification step, however, can often not be achieved. Amplification of a pool of different molecules requires that each molecule must in some way contain the information needed for its replication. Nucleic acids are one class of molecules that do fulfill this criterion. Here, enzymes like DNA- or RNA-dependent polymerases can be used to amplify the molecules. Accordingly pools of nucleic acids have been most commonly used for *in vitro* selection methods. Nevertheless, recent years have seen the adaptation of *in vitro* selection to pools of oligopeptides and libraries of chemical compounds (reviewed in Methods in Enzymology Volume 267 (1996)). The term combinatorial chemistry is therefore increasingly used to describe these methods.

Turning to the specific case of *in vitro* selection of a ligand, the principle scheme can be as follows : A pool of all possible ligands is prepared and then incubated with the molecule of interest. Conditions are chosen such that the ligand pool is in large excess over the molecules of interest. The molecule:ligand complexes are then separated from the unbound ligands and the bound ligands are afterwards dissociated from the molecule of interest. By applying the procedural steps of first generating a pool of ligands and then selecting those which can bind to the molecule of interest, a population of ligands is obtained which binds with higher affinity to the molecule than did the original pool. After the first round of selection the resultant pool is most likely not made up of high affinity ligands, but low affinity ligands still dominate. Presumably, the few high affinity ligands had to compete with an excess of low affinity ligands for the few binding sites. In order to enrich for high affinity ligands the selection step has to be repeated. After each successive selection cycle the high affinity ligands make up a greater proportion of the population.

In the case of the *in vitro* selection of a DNA ligand the initial pool of ligands can come from either of two sources. On the one hand genomic DNA, usually cut into smaller fragments, can be used. More commonly, a population of DNA oligonucleotides is synthesized chemically. The molecules are usually made up of a central block of randomized nucleotides encompassed on both sides by blocks of fixed nucleotide sequence. These flanking sequences allow for the hybridization of oligonucleotide primers which enable the amplification of the oligonucleotide pool by PCR. *In vitro* selection experiments of DNA ligands for proteins have been performed with both single-stranded and double-stranded DNA molecules. Whereas the former case has mostly been used to develop DNA ligands for diagnostic purposes, the latter is typically used to analyze biological relevant interactions. Single-stranded DNA can be prepared either by using a protocol for asymmetric PCR in which one primer is in excess, or by using a biotinylated primer and adsorption of one DNA strand on immobilized streptavidin agarose. Several different methods have been used to separate the DNA:protein complexes from the unbound DNA. The protein molecules can be immobilized on a macromolecular matrix such as agarose beads, resulting in a resin for affinity chromatography. With proteins as binding partners immobilization is often achieved by the use of a fusion protein where the protein of interested is fused to a tag. Glutathione-synthetase, which binds to immobilized glutathione, is one such tag. Another example are runs of six histidine amino-acids,

which can bind to immobilized metal-chelators. Immunoprecipitation has also been used to recover the protein:DNA complexes (Pollock & Treisman 1990). In other experiments, the protein:DNA complex has been isolated by filtration of the binding reaction through a nitrocellulose membrane (Thiesen and Bach 1990). Whereas the unbound DNA molecules can pass freely through the membrane, those bound to protein are retained on the membrane. A third approach has been the use of non-denaturing polyacrylamide gels in which the lower electrophoretic mobility of the protein:DNA complexes is used to separate these complexes from unbound DNA (Blackwell and Weintraub 1990).

In experiments that use double-stranded DNA as the ligand, the goal has mostly been to determine the sequence preference of DNA-binding proteins, especially those of transcription factors. Underlying is the assumption that binding of these factors to DNA is solely determined by the DNA sequence. However, this may be too simple a model to completely explain the biology of chromosomal DNA. Factors like bending and supercoiling do appear to be important for the biological function of DNA.

4.2. Supercoiling

o viv
The structure of DNA is normally represente1.73 TD.1(al(a)-0.83 TD.1)5.1(A)0.9(te1.7o)-0.9(r)5.9
yots,
The(str)5.9(u)-0.9(ct)10.7(ur)5.9(eTD.1)5tfef(i)5.1(s)2.9(sese)9.1(n)-0.9TD.yndermed byblo f or r
e
e rco-8.9n TD43.9(e(r)5.2eo-8.9d) -113 kwnde-8.9nscrstruct)-10of as D
ling.8 results( )10(f)10145(r4181om((tp.8e tkw)40istling.8 o(f)10145
uA(tTD.145(s)321 TD.145(s)321)10e(q)945(TD.145vs)1321au(et tso)

keep on twisting the ends relative to each other, the rubber band suddenly starts to form coils. The coils are characterized by the fact that the axis of the rubber band crosses over itself in space. Similar structures have been seen in electron micrographs of supercoiled DNA (Griffith 1976). The number of turns of this so called superhelix is called ´writhe´ and denoted with the letter W. For closed circular DNA it can be shown that the sum of T and W is constant as long as no covalent bond along the molecule is broken, which would allow one of the strands of the helix to rotate. The sum of T and W is known as the linking number L (L = T + W). Since only L is constant it follows that a circle can adopt different structures which meet this requirement, each differing in the number of supercoils and accordingly in the number of turns of the helix. These different structures are known as topoisomers.

In order to quantify the degree of supercoiling the term superhelical density σ has been introduced. It is defined as the average number of superhelical turns per helical turn of DNA. :

$$\sigma \equiv (L - L_0) / L_0;$$

where L is the linking number of the molecule and L_0 is the linking number of the molecule in its relaxed state. The relaxed state of DNA is the conformation which DNA adopts when its ends can rotate freely.

Most DNA in living systems, including chromosomal and plasmid DNA in bacteria as well as DNA in human cells, is negatively supercoiled. This means that the helix is underwound. The level of supercoiling is regulated by Topoisomerases and Gyrases, therefore the actual level varies in different biological states. Current methods used to measure the superhelical density *in vivo* are not very precise and depending on the method used different results have been obtained. For *E. coli*, values between $\sigma = 0.025$ and $\sigma = 0.05$ have been published (Sinden 1997).

The importance of DNA supercoiling can be inferred from extend of regulation that applies to supercoiling. In both prokaryotic and eukaryotic cells several enzymes exist that regulate the degree of supercoiling. Type I topoisomerases interrupt the sugar-phosphate backbone of only one DNA strand during their catalytic cycle. Type II topoisomerases break both strands and pass double-stranded DNA through the cut. The latter reaction is important in the decatenation of DNA following replication.

While most topoisomerases only relax existing topological strain, *E.coli* DNA gyrase can actively introduce negative supercoils. The other source of negative supercoils are the processes of DNA transcription and replication. Mutations that destroy the activity of either DNA gyrase or topoisomerse I can be lethal in *E.coli* (Sinden 1997). The experimentally shown importance of both DNA gyrase and topoisomerase I lead to a model of bacterial regulation of supercoiling in which gyrase introduces negative supercoiling, wherase toposiomerase I relaxes it. The balance of these two activities controls the precise level of supercoiling.

DNA supercoiling exerts biological effects at several levels. Several pronounced changes of DNA conformation are stabilized by supercoiling. For example Z-DNA, cruciform structures and intramolecular triplex DNA are all stabilized by negative supercoiling. At a different level the distortion of DNA present in the B-DNA conformation introduced by supercoiling can have effects on gene expression. The underwinding of the DNA helix inherent in negatively supercoiled DNA can promote the local melting of the DNA double helix. The local melting of the DNA double helix is a prerequisite for the RNA polymerase to start the synthesis of RNA. It appears that different promoters have been ´tuned´ to work best at different levels of DNA supercoiling (Sinden 1997). One example of regulation of gene expression by the level of supercoiling are the genes for enzymes that regulate the degree of supercoiling themselves : DNA gyrase and DNA Topoisomerase I (Menzel and Gellert 1983; Tse-Dinh and Beran 1988).

4.3. Z-DNA

Based on the known physicochemical and biological data on DNA, and in particular the fiber diffraction studies of Rosalind Franklin, James Watson and Francis Crick in 1953 developed their famous model of the DNA structure (Watson & Crick 1953). The strength of the model was not only its consistence with all known data, but it also suggested a simple way in which genetic information could be stored, read and duplicated. Thus, their model soon became accepted as most likely representing the structure of DNA. A definite proof, however, had to wait for development of techniques to synthesize sufficient amounts of DNA oligomers, allowing for the crystallization of DNA and subsequent determination of the DNA structure using X-

ray diffraction data at high resolution. When this was achieved in the late 70´s, it turned out that the first DNA-structure solved did not resemble the model of Watson and Crick (Wang et al. 1979). As opposed to the Watson-Crick model, the DNA did not adopt a right handed double helix, but rather a left handed double helix. Furthermore the DNA structure was characterized by the glycosidic bonds of purine nucleotides adopting the syn conformation, as opposed to the anti conformation found in the Watson and Crick model. Another difference was the distinct zick-zack of the sugar-phosphate backbone, which prompted the term Z-DNA to be used for this DNA conformation. The model described by Watson and Crick is known as B-DNA and a comparison of important structural parameters is given in Table 4.1.

	B-DNA	**Z-DNA**
Diameter of the Helix [nm]	2,0	1,8
Orientation of the Helix	right-handed	left-handed
Basepairs / turn	10,5	12
Length / Helical turn [nm]	3,4	4,46
Conformation of the Desoxyribose	C3´ endo	C2´ endo
Conformation of the glycosidic bond	anti	purines : syn pyrimidines : anti
Minimal distance between phosphate groups [nm]	1,17	0,77

Table 4.1 Characteristic structural parameters.000 e0 0lNA a-6.2(1 T-50640. e0 0lN-6.)12.A 4T4 1 Tf12 0 0 1356.8

Figure 4.2 Schematic drawing of the B-DNA and Z-DNA structures (Adapted from Wang et al. 1979)

methylation of cytosine at C5 and the bromination of both cytosine at C5 (Behe & Felsenfeld 1981) and guanosine at C8 (Möller et al.1984). Methylation appears to fill a void in the structure that is not easily solvated, whereas bromination of guanosine prevents the nucleotide from adopting the anti conformation (Rich et al. 1984). A third way in which Z-DNA can be stabilized is torsional strain. Due to its left-handed helix, Z-DNA can relax the torsional strain of negatively supercoiled DNA (Peck et al. 1982).

From the time of its discovery speculations were put forward as to whether Z-DNA has a biological function. Sequences of alternating purines and pyrimidines of sufficient length can be found in the genomes of many organisms and stabilization could occur by the negative supercoiling present in most genomic DNA. Furthermore, methylation of cytosine is a modification found *in vivo* and has been implicated in the control of gene expression (Doerfler 1983). One early indication that Z-DNA might indeed exist *in vivo* was the discovery that the serum of patients suffering from the disease lupus erythematodes can contain antibodies directed against Z-DNA (Lafer et al. 1981, Lafer et al. 1983). Further evidence came from studies in which plasmids containing Z-DNA forming sequences were introduced into bacteria. The presence of Z-DNA *in vivo* was then inferred from the protection of these sequences from chemical modification by psoralen (Sinden & Kochel 1987) or the EcoR I methylase (Jaworski et al. 1987). Another approach to probe for Z-DNA formation *in vivo* has been chemical modification of B-Z junctions with osmium tetroxide (Rahmouni & Wells 1989). While these studies provided good evidence for the existence of the Z-DNA conformation in plasmids introduced into bacteria, different techniques were needed to detect the presence of Z-DNA in the genome of eukaryotic cells. Soon after antibodies directed against Z-DNA had been discovered, they were used to probe eukaryotic cells using immunologic techniques. Clear patterns of staining were

observed among others with the polytene chromosomes of *Drosophila melanogaster* (Nordheim et al. 1981). These results, however, could not be repeated in all laboratories and are nowadays interpreted as artifacts (Hill 1983). The largest obstacle for detection of Z-DNA, was to introduce the antibody into the nucleus, without at the same time destroying the chromatin structure. This problem was solved by the use of permeabilized nuclei embedded in agarose microbeads (Wittig et al. 1989). After cells were stabilized in agarose beads, the cell membrane was permeabilized by mild lysis, and antibodies was able to diffuse into the nucleus. The antibodies were then covalently crosslinked to the DNA by exposing the sample to light from a strong UV-laser. The antibody now served as tag for the Z-DNA site and allowed the DNA molecule containing the Z-DNA site to be purified and analyzed. Using this method it was shown that transcription is associated with Z-DNA formation (Wittig et al. 1991). In the case of the human c-myc gene, the formation of Z-DNA was mapped to discrete regions of the gene (Wittig et al. 1992). Furthermore, Z-DNA could only be detected in the c-myc gene while the gene was transcribed. This type of analysis has been extended to the human corticotropin-releasing hormone gene and the human ß-globin gene cluster (Wölfl et al. 1996; Müller et al. 1996).

While the results mentioned above have established that Z-DNA is not just an *in vitro* artifact, it remained a mystery as to whether Z-DNA has a biological function and what it might be. Given the fact that Z-DNA is stabilized by negative supercoiling, and in turn can relax the torsional strain exerted on a stretch of DNA, Z-DNA has been implicated to play a role in a number of biological processes which influence the superhelicity of DNA. According to the two domain model (Liu & Wang 1987) RNA transcription results in the generation of positive supercoiling in front of the RNA polymerase and negative supercoiling behind the RNA polymerase. Based on this model several modes of actions can be envisioned for Z-DNA. First of all, it seems plausible that the degree of supercoiling behind the RNA polymerase could be reduced by DNA adopting the Z-DNA conformation. This in turn should help the polymerase, since excessive levels of negative supercoiling probably inhibit the polymerase. Z-DNA would remove transient high levels of negative supercoiling until the Topoisomerases can restore the natural level of supercoiling. An indication that this process might be important comes from an analysis of human sequences likely to form Z-DNA. It was found that Z-DNA sites are more common at or slightly upstream of the transcriptional start site than elsewhere (Schroth et al. 1992). A more

elaborate model has been proposed, in which the negative supercoiling induced by a passing RNA polymerase leads to the formation of Z-DNA, which in turn stops another RNA polymerase from passing. Thereby RNA polymerases would be kept at a distance from each other, which might be important so as to prevent their RNA transcripts from interfering with each other during further steps of RNA processing (Rich 1994). The ability of Z-DNA to stop transcription by a RNA polymerase has been demonstrated *in vitro* (van de Sande et al. 1982), however, whether this is true *in vivo* is not as clear (Sinden 1996). Other ways in which Z-DNA could interact with RNA transcription could be that the RNA polymerase or transcription factors are prevented from binding to a promoter due to the presence of Z-DNA. Alternatively, a protein could bind specifically to Z-DNA and promote or inhibit transcription. Yet another possibility exists. There are numerous examples of protein binding sites in promoters whose spatial relationship to each other is important. If a sequence which can adopt the Z-DNA conformation, is present between such two binding sites, it could serve to change the topology of the promoter such that the two proteins can no longer interact (Esposito & Sinden 1987). While this is an interesting hypothesis, hard data has been difficult to come by. Insertions of Z-DNA sequences into the promoters of different genes have led to inconclusive results, ranging from inhibition of transcription to no observable effect and even enhancement of transcription (Sinden 1996).

Genetic recombination is another process for which a role of Z-DNA has been postulated. Hotspots for recombination in eukaryotic cells often correlate with $(GT)_n$ or $(CA)_n$ tracts (Sinden 1994). When sequences of $(GT)_n$ and $(GC)_n$ were cloned into a plasmid and inserted into *E. coli*, multimerization (Murphy and Stringer 1986) as well as deletion of dinucleotides have been observed (Klysik et al. 1982).

While many different biological functions of Z-DNA have been proposed, experimental proof for any of them is still outstanding. One of the most promising approaches to elucidate the biological function of Z-DNA has been the search for Z-DNA binding proteins. Numerous reports of Z-DNA binding proteins exist in the literature. However, most of them must nowadays be treated with a great deal of skepticism. Most studies used affinity chromatography of cellular extracts on immobilized DNA stabilized in the Z-DNA conformation by bromination. In most papers, only the retention of proteins on such columns was reported and only a few

of the isolated proteins were cloned. Over the years a number of reports have appeared that identified putative Z-DNA binding proteins (Table 4.2).

Organism	Protein	Remarks	Citation
E. coli	Rho		Lafer et al 1988
E. coli	RecA	Shown not to be Z-DNA specific (Krishna et al. 1991)	Blaho & Wells 1987
E. coli	ParC (Topoisomerase IV)		Lafer et al. 1988
S. cerevisiae	Zuotin	Also claimed to be tRNA binding (Wilhelm et al. 1994)	Zhang et al. 1992
Bull testis	HMG1 / HMG2	Shown not to be Z-DNA specific (Rohmer et al. 1990)	Gut et al. 1987
Drosophila melanogaster	Topoisomerase II		Glikin et al. 1991 Arndt-Jovin et al. 1993
Homo sapiens	Topoisomerase II		Bechert et al. 1994
Gallus gallus	ADAR1		Herbert et al. 1995
Homo sapiens	ADAR1		Herbert et al. 1998

Table 4.2 Putative Z-DNA binding proteins

There are several reasons why the initial claims of Z-DNA binding proteins could often not be substantiated. First and foremost, it has to be noted that the circumstance that a protein binds to a Z-DNA affinity column is not sufficient proof that it actually recognizes the Z-DNA conformation. There are many reasons why a non Z-DNA binding protein could be retained on such a column, including electrostatic interactions and interactions with DNA ends. It has been observed too, that the binding of a number of the isolated proteins could be competed with phospholipids, indicating that they are not DNA binding proteins at all (Krishna et al. 1990). To date there are few proteins that are generally believed to bind to Z-DNA and none for which binding to Z-DNA has been demonstrated *in vivo*. Probably the best candidate for a genuine Z-DNA binding protein is the ADAR1 enzyme.

4.3. ADAR1

Double-stranded RNA-specific adenosine deaminases (ADARs) are a group of closely related enzymes that bind to double-stranded RNA and catalyze the oxidative deamination of adenosines. This activity has also become known as RNA editing. If the site of deamination lies within an exon, the resulting inosine is read by the translation apparatus as if it were a guanosine base, which can result in a different aminoacid being inserted into the protein. Alternatively, RNA editing may result in the

elimination of a stop codon. Originally observed as a dsRNA unwinding enzyme, the first ADAR protein isolated (ADAR1) has also been referred to as dsRAD and DRADA. The editing activity of the ADARs is employed in two different biological settings. First of all, they have been implicated in the biased hypermutations

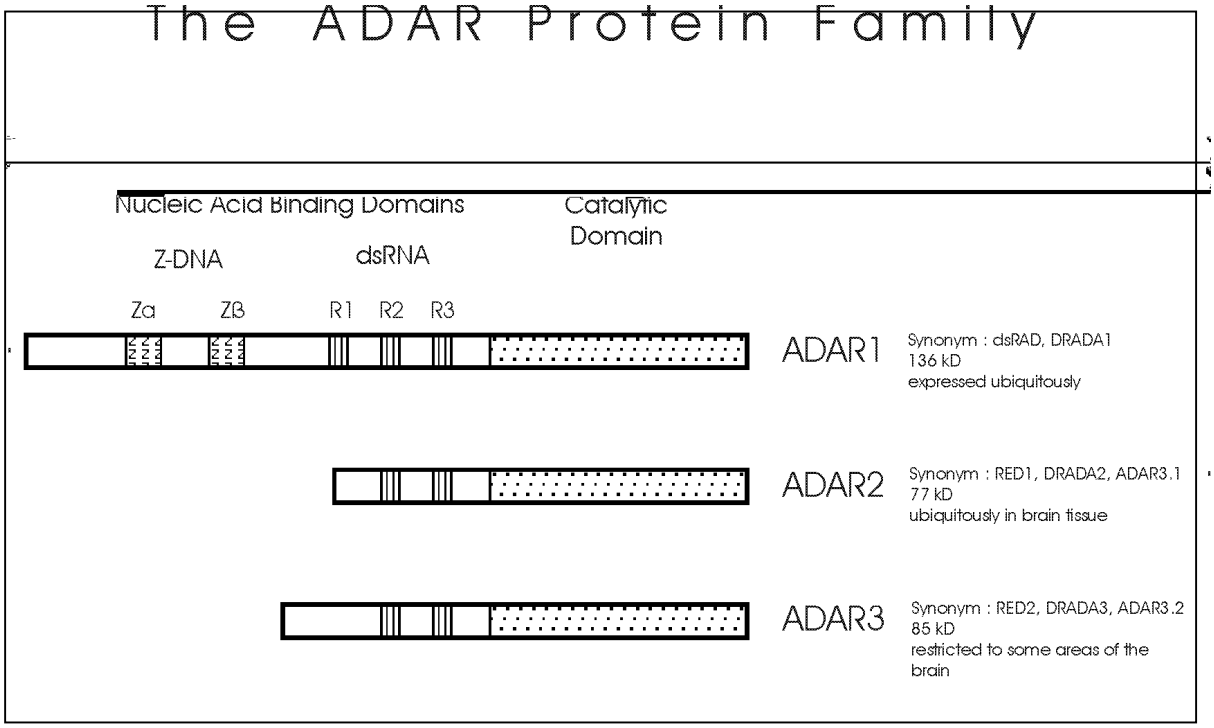

Figure 4.3 Domain structure of the ADAR proteins

observed in measles virus and other negative-stranded RNA virus genomes during lytic and persistent infections (Cattaneo and Billeter 1992). Consistent with an anti viral function is the observation that expression of the ADAR1 enzyme can be stimulated by interferons (Patterson and Samuel 1995). Furthermore, an inosine dependent RNAse has been discovered (Scadden and Smith 1997). This suggests a model whereby viral double stranded RNA is edited by ADAR enzymes and the inosines introduced into RNA serve as a tag for destruction of viral RNA by an inosine dependent RNAse. A different mode of action has been discovered in the case of the GluR receptor. Here, an intronic sequence of the pre-mRNA folds back on an exonic sequence and specific adenosines in the double stranded region are desaminated. The inosine bases created by editing, result in a changed aminoacid sequence of the receptor, which in turn influences its electrophysiologic properties (Sommer et al. 1991). A similar mechanism has been observed in the case of the serotonin receptor (Burns et al. 1997). It appears that in both cases RNA editing is just another method of producing different polypeptide chains from the same gene and thereby generate receptors with different characteristics. A third example of site

specific editing activity is the Hepatitis Delta Virus (HDV) where editing destroys a stop codon (Casey and Gerin 1995).

The general structure of the ADAR proteins is shown in Figure 4.3 . All ADAR proteins have a central region with two or three double stranded RNA binding motifs denoted R1, R2 and R3 . The catalytic domain is located in the C-terminal domain. The ADAR1 protein has an extended N-terminal domain which contains two Z-DNA binding domains Zα and Zß. Z-DNA binding activity had originally been observed in extracts from chicken erythrocytes (Herbert et al. 1993) and was later assigned to the ADAR1 protein (Herbert et al. 1995). By producing recombinant versions of human ADAR1 protein, the Z-DNA binding activity was then mapped to a domain of app. 80 aminoacids, which was termed the Zα-domain. Sequence analysis has revealed a second domain with similar aminoacid sequence, which was named the Zß-domain. Using a DNA probe stabilized in the Z-DNA conformation by bromination, the dissociation constant for the Zα:DNA complex was determined to be 4 nM. A fusionprotein of the Zα-domain and the catalytic domain of Fok I endonuclease was shown to cut plasmid-DNA close to sequences of alternating purines and pyrimidines (Kim et al. 1997). Final proof that the Zα-domain does indeed bind to the Z-DNA conformation will probably be supplied in the near future, since co-crystallization of the Zα-peptide with a DNA molecule has been achieved and analysis by x-ray diffraction indicates that the DNA is present in the Z-DNA conformation (T. Schwartz; personal communication).

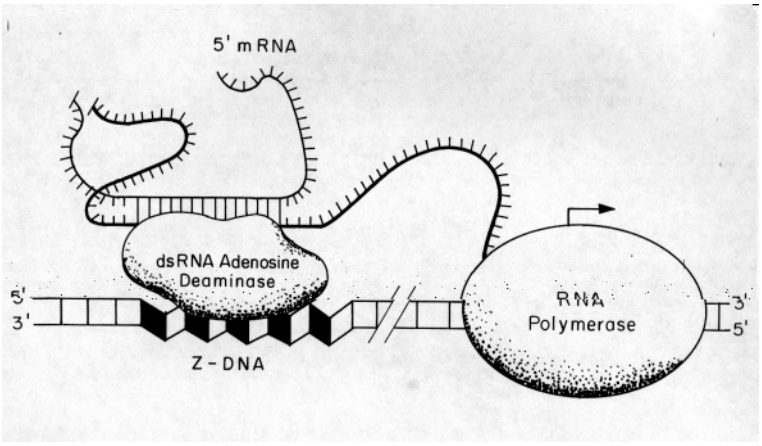

Figure 4.4 (Herbert and Rich 1996)

It is puzzling, why an enzyme involved in post-transcriptional processing of pre-mRNAs should also interact with DNA. A model has been put forward, which tries to provide an explanation (Herbert and Rich 1996). In all known examples, editing of mRNAs by ADAR proteins is dependent on intronic sequences base pairing with exonic sequences. Accordingly, editing has to take place before splicing occurs. Therefore it would make sense, if the editing enzyme is in some way directed to

places of active RNA transcription. As mentioned above, transcription is able to induce the Z-DNA conformation by the negative supercoiling generated in the wake of RNA polymerase. Accordingly, an enzyme which binds to Z-DNA would be located on an actively transcribed gene (Herbert and Rich 1996). A cartoon depicting this situation is shown in Figure 4.4 . Only a few cases of sequence specific editing are known and the search for further examples has been difficult and frustrating (G.Hoppe; personal communication). If the proposed role for Z-DNA in locating the editing enzyme is correct, then it should be possible to obtain a list of potentially edited genes by simply looking for sequences with a high propensity to adopt the Z-DNA conformation. However, in the known cases of site specific editing no DNA sequences likely to adopt the Z-DNA conformation have been found (A. Herbert, personal communication). Hence binding of the ADAR1 enzyme to DNA may not be necessary for editing. Alternatively, one of the two other ADAR enzymes could be responsible for editing of these mRNAs. Even if Z-DNA binding is a targeting mechanism for the ADAR1 enzyme, another problem remains. Analysis of known DNA sequences reveals a large number of sites with the potential to adopt the Z-DNA conformation. It is unlikely, however, that so many editing sites exist. In essence, by binding to Z-DNA the ADAR1 enzyme would be located to many genes which do not contain editing sites. One possible explanation for this paradox would be if the ADAR1 protein does not bind to every Z-DNA site, but has some preference for certain Z-DNA sequences.

The idea that the Zα domain may target the ADAR1 protein to certain genes was one of the reasons why I decided to further characterize the sequence dependency of the binding of ADAR1 to Z-DNA. At the same time the interaction of a Z-DNA specific protein with DNA provided an ideal experimental system for development of novel *in vitro* selection techniques that reflect topological parameters.