

23.2. PROTEIN–LIGAND INTERACTIONS

edge in forming a pseudo-hydrogen bond between the indole hydrogens and the π -rings of the DNA bases. In *ETSI*, the deformation of the DNA helix resulting from protein intercalation results in the kinking of the helical axis from 45° to about 60° .

Examples of protein intercalation of the DNA helix from the major groove are found in proteins, such as the methyltransferases, that perform chemistry on the bases of the DNA. To perform their enzymatic function, these proteins must extract the target base from the DNA helix and ‘flip’ the base out into the enzyme active site (Cheng, 1995). The resulting void in the DNA is then filled by protein side chains that partially satisfy the hydrogen-bonding and van der Waals interactions that were broken when the target base was flipped. Although there are only a few known structures of DNA–protein complexes with extra-helical bases, base flipping is thought to be a relatively common feature of DNA-modifying enzymes.

23.2.4.2. Single-stranded sequence-nonspecific DNA–protein interactions

There have been a few reports of single-stranded DNA–protein complex structures, all of which involve the sequence-nonspecific recognition of DNA. In the binding of a tetranucleotide to the exonuclease active site of the DNA polymerase I Klenow fragment (Freemont *et al.*, 1988), extensive hydrogen-bonding interactions between the sugar–phosphate backbone and the protein are observed. This provides the most intuitive mechanism for sequence-nonspecific nucleic acid binding, where the protein simply recognizes the phosphate backbone of a single-stranded coil. The protein also appears to form a few hydrophobic interactions with the DNA bases; however, these interactions, which include the partial intercalation between two bases, are thought to be nonspecific.

The structure of replication protein A complexed with single-stranded DNA does not exhibit the intuitive nonspecific mechanism of recognition found in the Klenow fragment (Bochkarev *et al.*, 1997). In this structure, the DNA is extended with its bases splayed out over the surface of the protein. The bases form several pairwise stacking interactions that are interrupted by intercalating protein side chains. Contrary to the sequence-nonspecific nature of recognition, numerous hydrogen bonds are found between the protein and the bases of the DNA strand. These base-dependent contacts require that the protein–DNA interactions must be flexible and plastic in order to accommodate different base sequences.

23.2.4.3. RNA

Although RNA and DNA are chemically similar, RNA presents a much greater variety of shapes and surfaces compared to the relatively simple B-form helix of DNA. Generally single-stranded, RNA often forms secondary structures driven by the base pairing of complementary stretches of sequence within the same strand. The formation of base-paired regions can result in stem loops, bulges and helices which can further assemble into more complicated tertiary structures, such as that observed for transfer RNAs. Protein-mediated recognition of RNA often depends as much on the three-dimensional structure presented by these secondary structures as on the specific identity of the base sequence.

Very little information is currently available on the structural details of protein–RNA interactions (Nagai, 1996). Only a handful of protein–RNA complex structures have been determined. These fall into three basic categories, depending on the secondary structure of the RNA: four tRNA–protein complexes, two stem-loop–protein complexes and a capped single-stranded RNA–protein complex.

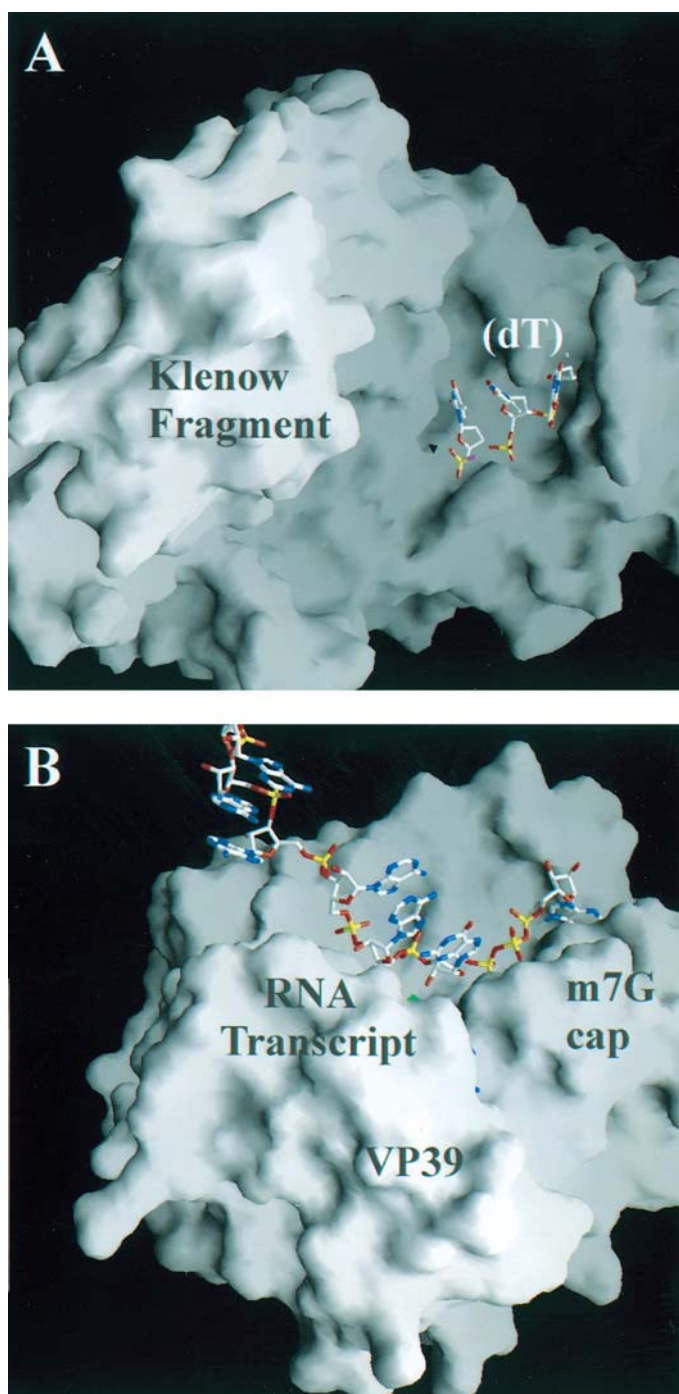


Fig. 23.2.4.4. The sequence-nonspecific recognition of single-stranded nucleic acid. (a) Oligo(dT) bound in the exonuclease active site of DNA polymerase I Klenow fragment (Freemont *et al.*, 1988). (b) A short capped RNA transcript bound to the VP39 RNA methyltransferase (Hodel *et al.*, 1998). Both proteins primarily interact with the backbone of the nucleic acid.

23.2.4.4. Transfer RNA

In the four known structures of tRNA bound to their aminoacyl tRNA synthetases (Cusack *et al.*, 1996a,b; Goldgur *et al.*, 1997; Rould *et al.*, 1991), the effects of RNA's preference for A-form helices on recognition are immediately apparent. The proteins make numerous contacts in the shallow and exposed minor grooves of the RNA helices. This contrasts with the extensive use of the major groove in the recognition of B-form DNA helices. Beyond this generalization, the details of tRNA recognition differ in each