23. STRUCTURAL ANALYSIS AND CLASSIFICATION

23.3.4.6. Single-stranded sequence-nonspecific RNA–protein interactions

There is a single example of a single-stranded RNA–protein complex which is sequence-nonspecific. The structure of the vaccinia RNA methyltransferase VP39 bound to a 5' m7G-capped RNA hexamer reveals a mechanism of nonspecific recognition reminiscent of the Klenow fragment–DNA tetramer complex (Hodel et al., 1998). The RNA forms two short single-stranded helices of three bases each. The first of these helices binds in the active site of VP39 solely through hydrogen bonds between the protein and the ribose–phosphate backbone. The bases of the RNA strand stack together as trimers, but do not form any interactions with the protein (Fig. 23.3.4.4). Like the Klenow–DNA complex, this observation suggests an intuitive mechanism for sequence-nonspecific nucleic acid binding, where the single-stranded RNA forms short transient helices driven by intramolecular stacking interactions. The protein then recognizes and stabilizes the helical backbone conformation formed by this transient stacking without interacting with the bases themselves.

23.3.4.7. The recognition of alkylated bases

The complex of VP39 with capped RNA also illustrates a final example of the diversity of protein–ligand interactions in the specific recognition of the 7-methylguanosine cap. When guanosine is methylated at the N7 position, a positive charge is introduced to the π-ring system of the base. Eukaryotic cells utilize the methylation of a guanosine base at the N7 position as a tag or cap for the 5' end of messenger RNA. The m7G5'ppp mRNA cap is specifically recognized in the splicing of the first intron in nascent transcripts, in the transport of mRNA through the nuclear envelope and in the translation of the message by the ribosome (Varani, 1997). Two structures of specific m7G binding proteins are now known: VP39 and the ribosomal cap-binding protein IF-4E (Hodel et al., 1997; Marcotrigiano et al., 1997). Each structure offers clues as to how the proteins can discriminate between the charged methylated m7G base and the unmodified guanosine base. The m7G base is stacked between aromatic protein side chains and hydrogen bonded to acidic protein residues (Fig. 23.3.4.5). One long-held hypothesis is that IF-4E, with dual tryptophan residues, binds specifically to the positively charged form of the base through a charge-transfer complex (Ueda, Iyo, Doi, Inoue & Ishida, 1991). The formation of a charge-transfer complex is evident in small-molecule studies and spectroscopic studies with IF-4E (Ueda, Iyo, Doi, Inoue, Ishida et al., 1991). However, VP39 performs the same discrimination with the much less electronegative phenylalanine and tyrosine side chains (Hodel et al., 1997). So far, no charge-transfer complex has been observed in VP39.

The recognition of charged methylated bases is important not only in mRNA processing, but also in the repair and recognition of DNA damaged by alkylating carcinogens. The mechanism by which the charged m7G base is recognized is probably similar to how other positively charged bases, such as 3-methyladenosine, O2-methylcytosine and O2-methylthymidine, are recognized. In fact, the E. coli DNA repair enzyme, AlkA, will catalyse the glycosylation of all of these bases (Lindahl, 1982). The structure of AlkA is known, but only in the absence of a substrate (Labahn et al., 1996). In this structure, a number of solvent-exposed tryptophan residues are found at the putative active site. This observation suggests that AlkA may recognize positively charged bases through an aromatic ‘sandwich’, much like that found in IF-4E and VP39.