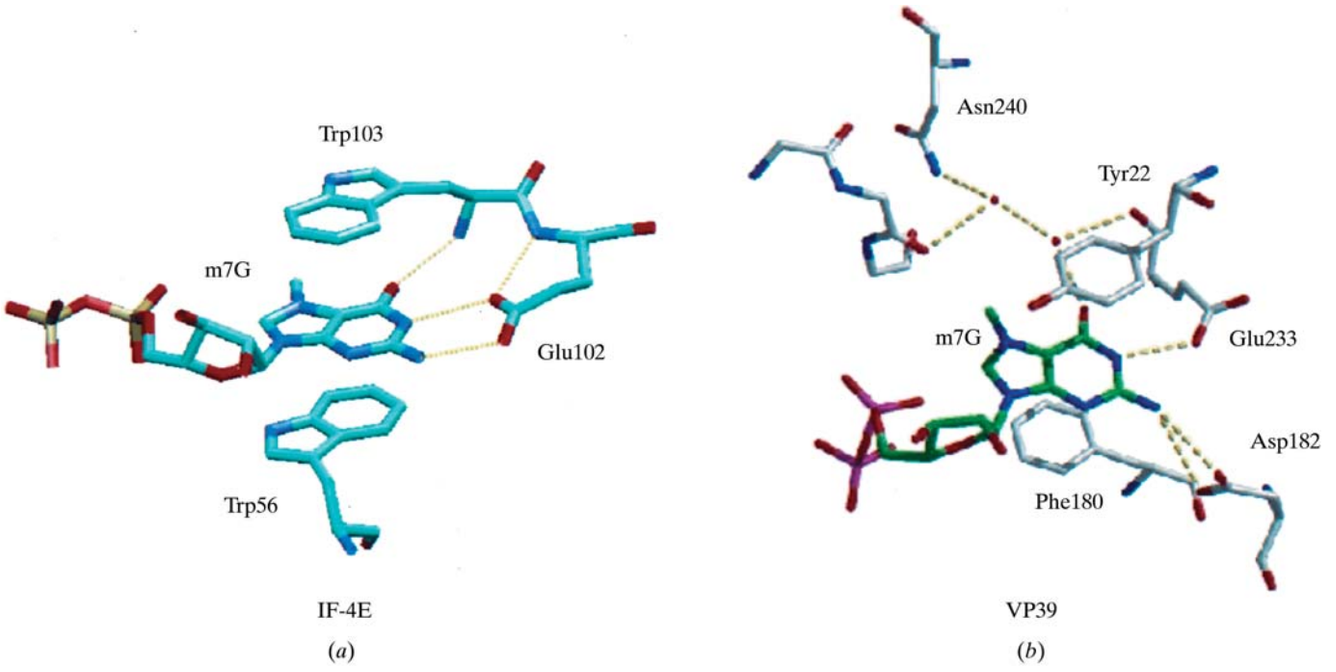

Fig. 23.2.4.5. The specific recognition of the messenger RNA 7-methylguanosine cap. (*a*) The residues contacting the m^7G base in the cap-binding protein, IF-4E (Marcotrigiano *et al.*, 1997). (*b*) The residues interacting with the cap in the vaccinia RNA methyltransferase VP39 (Hodel *et al.*, 1997). Both proteins bind to the charged, methylated base by stacking aromatic amino acids on both sides of the base.

specific case. Comparison of the protein-bound tRNA to the structure of free tRNA reveals that the proteins tend to distort the RNA conformation and partially unwind the helices near the anti-codon loop. In one case, namely the structure of glutamyl-tRNA synthetase (Rould *et al.*, 1991), the final base pair near the acceptor stem of the tRNA is broken, and the CCA acceptor makes a dramatic hairpin turn into the enzyme active site.

23.2.4.5. *Stem loops*

One fascinating observation in viewing the structures of RNA-binding proteins, even in the absence of RNA, is that aside from the tRNA-binding synthetases, they all appear to have evolved from or towards a very similar general fold (Burd & Dreyfuss, 1994). This fold, exemplified by the RNP domain found in numerous RNA-binding proteins, consists of a β-sheet surrounded on one side by α-helices and solvent-exposed on the opposing face. This general folding architecture is found in RNP domains, ribosome proteins, K-homologous domains (KH), double-stranded RNA-binding domains and cold shock proteins. Although each of these subsets of RNA-binding domains has a different topology and most probably bind to RNA with different surfaces, they all appear to have this alpha–beta–solvent architecture.

Two proteins with this architecture have been co-crystallized with their specific RNA stem-loop ligands (Nagai *et al.*, 1995; van den Worm *et al.*, 1998). In both cases, the loop of the RNA binds to the open face of the β-sheet where solvent-exposed aromatic amino-acid side chains stack with the extrahelical bases of the RNA. Unpaired bases from the RNA also form numerous specific hydrogen bonds with protein side chains and polar backbone groups, imparting sequence specificity in the interaction. These structures suggest that the flat, open face of a β-sheet provides a good surface for RNA binding, where the extra-helical bases can make extensive and specific contacts with the protein.

23.2.4.6. *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 $5'm^7G$-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. 23.2.4.4). 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.

23.2.4.7. *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 $m^7G(5')$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 m^7G 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

references