glycosylated glycoproteins can be mutated to other amino acids (Horan *et al.*, 1998) to avoid or reduce glycosylation.

Other post-translational modifications occur less frequently. Prenylation and N-myristoylation can occur at the C- and N-termini, respectively. Expression in *E. coli*, often using truncated versions of target proteins, is a common remedy (Pai *et al.*, 1990).

4.3.8. Stabilization of protein targets

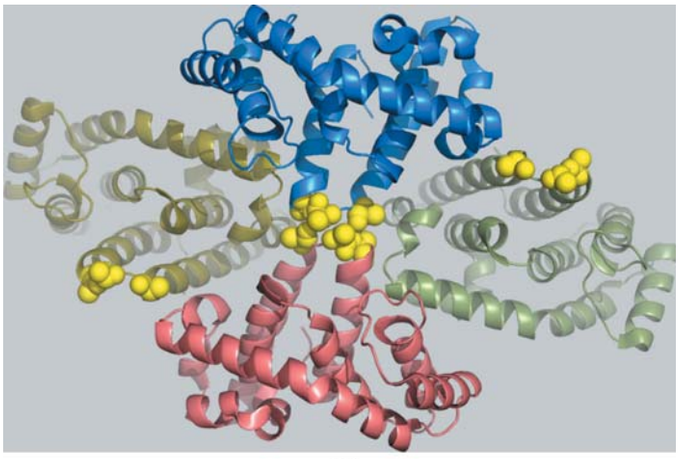
There is currently no clear consensus regarding a possible correlation between the thermostability of a protein and its propensity to form crystals. It is often assumed on the basis of somewhat anecdotal evidence that thermostable proteins are more readily crystallizable and therefore if a specific target protein is recalcitrant to crystallization then a homologue from a thermophilic organism should instead be used. In some cases, low thermostability may correlate with the presence of unstructured loops or termini and consequently the construct-optimization strategies, as described above, are likely to yield a more crystallizable variant with a concomitantly increased stability. For example, a study of MAPKAP kinase 2 showed that truncated variants with increased thermal stability also showed higher crystallization propensity (Malawski *et al.*, 2006). However, it is uncertain whether the relationship is causal or serendipitous. A recent analysis of large-scale data from a structural genomics project showed that when partly or fully unfolded proteins and hyperstable proteins (with melting temperatures T_m of greater than 90 °C) are excluded from comparisons, thermostability *per se* does not correlate with propensity for crystallization (Price *et al.*, 2009). Consequently, it appears that in a general case of a well behaving protein, attempts to increase thermostability by site-directed mutagenesis may not necessarily yield variants with enhanced crystallization properties, although when the prospective crystallization target is inherently unstable, engineering more stable variants may be helpful. In fact, this strategy has been successfully used for membrane proteins, which are often unstable in detergent environments. The first structure of a recombinant G-protein-coupled receptor (GPCR), *i.e.* bovine rhodopsin in complex with 11-*cis* retinal, was obtained using a thermostable variant with an engineered disulfide bond (Standfuss *et al.*, 2007). In the more recent case of the turkey adrenergic β2 receptor, 318 variants were screened and six mutations were identified that increased thermostability. A variant containing all six mutations had an apparent T_m that was 21 °C higher than that of the native protein in dodecylmaltoside (DDM), was more stable in short-chain detergents and was successfully crystallized (Warne *et al.*, 2008, 2009). The effort required for such vast screening is substantial, but it appears that within protein families (such as GPCRs) the pattern of mutations enhancing thermostability is preserved, thus making it possible to transfer the mutations from one family member to another (Serrano-Vega & Tate, 2009).

Finally, it should be noted that protein thermostability is strongly dependent on solution parameters such as the ionic strength. In the case of ribonuclease SA, the melting temperature (T_m) increased from ∼40 to ∼60 °C on transfer from pure 50 m*M* diglycine buffer to 0.9 *M* ammonium sulfate (Trevino *et al.*, 2007). Similarly, binding small molecules, either physiological or non-physiological ligands, typically promote stability (Matulis *et al.*, 2005). High-throughput screening methods have been developed to aid in screening for conditions and ligands that enhance stability (Vedadi *et al.*, 2006; Mezzasalma *et al.*, 2007) and in a general case this strategy appears to hold better promise than attempts to engineer higher stability through mutagenesis.

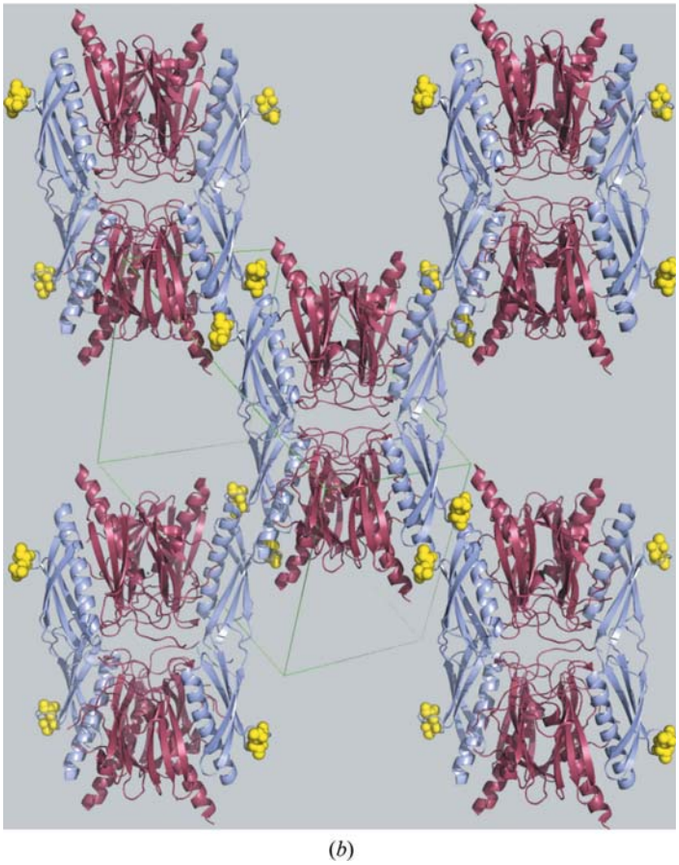
4.3.9. Surface-entropy reduction (SER)

For well ordered intrinsically stable proteins that show none of the problems addressed above, the propensity of the molecules to associate together and form crystals mediated by weak but specific interactions is ultimately defined by the physical chemistry and topology of the molecular surface. As already pointed out, large flexible amino acids on the surface, such as Lys, Glu and Gln, constitute an impediment to intermolecular interactions and consequently to protein crystallization (Cieślik & Derewenda, 2009; Price *et al.*, 2009). In fact, it has been suggested that these residues and the 'entropy shield' that they form play a role in protein evolution which, given the high average concentration of proteins in cells, disfavours protein–protein interactions unless they are biologically functional (Doye, 2004). Thus, an intuitively obvious way to generate crystallizable variants is to replace selected large and surface-exposed residues with smaller residues such as alanine. This crystal-engineering strategy based on the surface-entropy reduction (SER) concept was extensively tested using as a model system the globular domain of the human Rho-specific guanine nucleotide-dissociation inhibitor (RhoGDI), which is recalcitrant to crystallization in its wild-type form owing to a high content of Lys and Glu residues, which constitute more than 20% of the sequence (Longenecker, Garrard *et al.*, 2001; Mateja *et al.*, 2002; Derewenda, 2004; Cooper *et al.*, 2007). These experiments established that in order to be most effective the SER strategy requires simultaneous mutations of clusters of two to three solvent-exposed high-entropy amino acids, typically Lys, Glu or Gln, located in close sequence proximity. These amino acids are replaced with alanine, although threonine and tyrosine, which is known to make a positive contribution at protein–protein interfaces such as antibody–antigen complexes (Fellouse *et al.*, 2006), can also be used (Cooper *et al.*, 2007). Engineered low-surface-entropy variants of RhoGDI produced new and unique crystal forms, many with superior diffraction quality when compared with the wild-type protein. Importantly, in the vast majority of these crystals the mutated surface patches mediated crystal contacts, suggesting that SER engineering directly drives crystallization in a rational fashion by creating suitable crystal-contact-forming interfaces. The general utility of the method was further established by the crystallization of several novel protein targets found to be recalcitrant to crystallization in their wild-type form (Longenecker, Lewis *et al.*, 2001; Derewenda *et al.*, 2004; Devedjiev *et al.*, 2004; Janda *et al.*, 2004).

SER is quickly becoming a method of choice for engineering crystallizable variants of both individual proteins and protein–protein complexes. To date (December 2009), there have already been more than 100 depositions made to the Protein Data Bank (Berman *et al.*, 2007) based on diffraction studies of crystals generated by SER and corresponding to 47 novel structures, seven novel protein–protein complexes, several studies of proteins in complexes with drug leads aimed at rational drug development and two membrane proteins. The current list of crystal structures obtained using SER crystals includes a number of cases of exceptionally high biological interest. For example, the EscJ protein from enteropathogenic *E. coli*, the oligomerization of which initiates assembly of the type III bacterial secretion system, was crystallized with three entropy-reducing mutations (E62A, K63A and E64A) forming a key contact in the crystal structure (Yip *et al.*, 2005). Likewise, the HIV CcmK4 capsid



(*a*)



(*b*)

Figure 4.3.9.1
Two examples of proteins crystallized by the surface-entropy reduction (SER) method. (*a*) The RGSL domain of the PDZRhoGEF nucleotide-exchange factor (PDB code 1htj; Longenecker, Lewis *et al.*, 2001); the yellow spheres show the alanines introduced by mutagenesis, which mediate an isologous crystal contact across a crystallographic twofold axis. (*b*) The crystal structure of EpsI complexed with EpsJ (PDB code 2ret; Yanez *et al.*, 2008); the EpsI protein (pale blue) contains two surface mutations, shown by yellow spheres, which mediate heterologous crystal contacts.

protein only crystallized after an entropy-reducing mutation (E104Y) was introduced into the protein (Pornillos *et al.*, 2009). The protein–protein complexes solved to date underscore the utility of the method, which extends beyond individual proteins because high-entropy patches occur outside complex interfaces. For example, the complex of c-Src and its inactivator Csk was crystallized using a variant of Csk carrying K361A and K362A mutations (Levinson *et al.*, 2008). Similarly, it was possible to crystallize the complex of two pseudopilins EpsI and EpsJ from the type 2 secretion system of *Vibrio vulnificus* when a variant of EpsI carrying two mutations (E128T and K129T) was used (Yanez *et al.*, 2008; Fig. 4.3.9.1).

An interesting variation of the SER method was used in the investigation of the RACK1 protein, which was crystallized as an in-line fusion with an MBP variant carrying D82A, K83A and K239A mutations (Ullah *et al.*, 2008). This is the first example of the application of the surface-entropy reduction strategy to a carrier protein and not the crystallization target itself.

The SER strategy is attractive not only because of its efficacy but also because of its simplicity: once an expression construct for a target protein is available several rounds of mutagenesis can easily create variants with systematically enhanced crystallizability. To assist in the design of crystallizable variants, a server has been developed that uses the amino-acid sequence of the target to identify suitable mutation sites (Goldschmidt *et al.*, 2007).

4.3.10. Improvement of crystal quality

In most cases, protein engineering is used as a tool of last resort to obtain variants for proteins for which no crystals can be grown using the wild-type form. However, it may sometimes be necessary to obtain a new, different crystal form even when the wild-type protein does crystallize. Such a need may arise, for example, in drug-design investigations, where high-resolution structures are particularly critical for evaluation of the interactions between lead compounds and the target protein and may not always be possible using wild-type crystals. A novel crystal form may also be necessary if the wild-type crystals contain the target protein in an orientation in which the active site is obscured by crystal contacts, making it impossible to soak in drug lead compounds and screen small-molecule libraries by high-throughput crystallography (Blundell & Patel, 2004).

One possible strategy for obtaining a new crystal form is to modify the existing crystal contacts by replacing some of the participating amino acids. While this approach occasionally leads to improvement of the X-ray data resolution (Liu *et al.*, 2007; Mizutani *et al.*, 2008), modification of crystal contacts is typically counterproductive as it abolishes the propensity of the target to crystallize in one form but does not necessarily induce another (Charron *et al.*, 2002). A more successful strategy is to generate a novel crystal form by engineering new crystal contacts through SER. For example, a novel crystal form of the insulin-like growth factor 1 receptor kinase domain, a putative drug target, was obtained using a double mutant (E1067A and E1069A); the new form diffracted to 1.5 Å resolution, whereas the wild-type crystals only diffracted to 2.7 Å resolution (Munshi *et al.*, 2003). In the case of the catalytic domain of activated factor XI, a key enzyme in the blood coagulation cascade and another potential drug target, a single K437A mutation allowed the preparation of a crystal form that diffracted to 2.0 Å resolution (Jin *et al.*, 2005). Entropy-reducing mutations were also key in the preparation of a crystal form of HIV-1 reverse transcriptase for structure-based drug design that diffracted to 1.8 Å resolution, in contrast to the typical 2.5–3.0 Å range observed for the wild-type protein crystals (Bauman *et al.*, 2008).

4.3.11. Conclusions

Protein engineering has become a routine tool that is used to generate crystallizable macromolecules and their complexes. While some approaches may only apply to very specific targets, a