centrifugation and SDS–PAGE analysis were used to determine which mutant proteins were sufficiently soluble to appear in the supernatant. The initial application of this method to 30 mutants resulted in one, Phe185 → Lys, which was soluble and which was subsequently crystallized and its structure determined (Dyda *et al.*, 1994). The protein formed a dimer, and the mutated residue was observed at the periphery of the dimer interface where the introduced lysine formed a hydrogen bond with a backbone atom of the second subunit, an interaction not possible for the unmutated protein. The position of the mutation was remote from the active site, and the physiological relevance of the observed dimer interaction was later confirmed by studies on an avian retroviral integrase (Bujacz *et al.*, 1995).

In further mutational work, it was observed that the HIV-1 integrase core-domain mutant suffered from an inability to bind to Mg^{2+} in the crystal, despite the evidence that Mg^{2+} or Mn^{2+} is needed for activity. The original crystallization took place using cacodylate as a buffer and also had dithiothreitol present in the crystallization medium. Under these conditions, cacodylate can react with –SH groups, and there were two cysteines in the structure that were clearly bonded to arsenic atoms. To avoid this problem, attempts were made to crystallize in the absence of cacodylate. These were successful only when a second mutation, designed to improve solubility, was introduced, Trp131 → Glu (Jenkins *et al.*, 1995; Goldgur *et al.*, 1998). The use of this mutant led to crystals that had the desired property of binding to Mg^{2+} and, in addition, revealed the conformation of a flexible loop that had not been previously defined.

4.3.3. Use of fusion proteins

Fusion proteins have been frequently used in a variety of applications (reviewed by Nilsson *et al.*, 1992), such as preventing proteolysis, changing solubility and increasing stability. They have also been used – although less frequently – for crystallization. The disadvantage in the context of crystallography is that the length and flexibility of the linker chain often introduce mobility of one protein domain relative to the other, which can impede, rather than enhance, crystallization.

Donahue *et al.* (1994) were able to determine the three-dimensional structure of the 14 residues representing the platelet integrin recognition segment of the fibrinogen γ chain by constructing a fusion protein with lysozyme, which was then crystallized from ammonium sulfate. Kuge *et al.* (1997) successfully obtained crystals of a fusion protein consisting of glutathione S-transferase (GST) and the DNA-binding domain (residues 16–115) of the DNA replication-related element-binding factor, DREF, under crystallization conditions similar to those used for GST alone.

In many cases, a fusion protein is made to aid in the isolation and purification of the target protein, and the intervening linker is engineered to contain a proteolytically susceptible sequence. However, subsequent cleavage to separate the two proteins can introduce the possibility of accidental proteolysis elsewhere in the protein. This was observed with a fusion protein between thioredoxin and VanH, a D-lactate dehydrogenase, where attempts to remove the carrier resulted in non-specific proteolysis and VanH inactivation (Stoll *et al.*, 1998). Fortunately, cleavage was unnecessary, and conditions were identified under which the authors were able to crystallize the intact fusion protein.

A novel approach to crystallizing membrane proteins is provided by the fusion protein in which cytochrome b_{562} was inserted into a central cytoplasmic loop of the lactose permease from *Escherichia coli* (Privé *et al.*, 1994). Although crystals have not yet been reported, the cytochrome attachment provides increased solubility together with the ability to use the red colour to assay the progress of crystallization trials.

4.3.4. Mutations to accelerate crystallization

A common problem encountered in crystallization is that certain crystals appear late and grow slowly. Sometimes, the slow appearance of crystals is the result of proteolytic processing, but often the reasons are not apparent. There are several examples where protein engineering has resulted in an increase in the rate of crystallization.

Heinz & Matthews (1994) explored the crystallization of T4 phage lysozyme using a strategy based on their understanding of the structure of the enzyme and its crystallization properties. The crystallization of the wild-type protein required the presence of β-mercaptoethanol (BME), an additive which could not be replaced with dithiothreitol. It had also been observed that the oxidized form of BME, hydroxyethyl disulfide, was trapped in the dimer interface between two lysozyme molecules (Bell *et al.*, 1991). It was hypothesized that dimer formation might be the rate-limiting step in crystallization, so dimerization was enhanced by cross-linking two monomers by disulfide-bridge formation. Applying rules developed for constructing S–S bridges, they selected Asn68 → Cys and Ala93 → Cys. In the presence of oxidized BME, the rate of crystallization of these mutant proteins was substantially increased, with crystals reaching full size in two days, in contrast to two weeks for the unmutated protein. Furthermore, they were able to crystallize a previously uncrystallizable mutant. Unexpectedly, however, the dimer formed in this way was lacking in activity, despite the selection of mutation sites on the opposite side of the molecule to the active site.

Mittl *et al.* (1994) wanted to improve the resolution of their crystals of glutathione reductase. From the 3 Å map, they could see a hole in the crystal packing where two molecules within 6 Å of each other just missed forming a crystal contact; they filled this hole by mutating Ala90 → Tyr and Ala86 → His. This designed double mutant did not improve the resolution, but did increase the rate of crystallization 40-fold, *i.e.*, initial crystals were observed within 1.5 h *versus* 60 h for the wild-type enzyme.

4.3.5. Mutations to improve diffraction quality

Another commonly encountered situation is that crystals can be obtained, but they diffract poorly. There are many examples where investigators have applied protein engineering in an effort to overcome this problem.

Proteolytic trimming is one possible approach to improving diffraction quality. For example, Zhang *et al.* (1997) attempted to crystallize a homodimer of the C2 domain of adenylyl cyclase. The initial crystals diffracted poorly (to 3.8 Å), so the effects of limited proteolysis with chymotrypsin, trypsin, GluC and LysC were investigated. A stable cleavage product was observed with GluC, approximately 4 kDa smaller than the full-length protein, but in order to avoid minor products formed during GluC proteolysis, the cleavage site was re-engineered as a thrombin site. Since there was already an atypical thrombin site seven residues from this site, proteolysis resulted in a smaller protein than expected; nevertheless, this modified protein crystallized readily and diffracted to 2.2 Å.

The importance of applying a variety of strategies to improve crystal quality is exemplified by the work of Oubridge *et al.* (1995), in which initial attempts to crystallize wild-type U1A complexed with RNA hairpins resulted in cubic crystals diffracting to 7–8 Å. By mutating surface residues, changing the N-terminal sequence to reduce heterogeneity and varying the sequence of the RNA hairpin, a new crystal form which diffracted to 1.7 Å was ultimately crystallized. However, in order to achieve this result, many variants were constructed and examined. For the protein, mutations were introduced which it was believed (incorrectly) would affect the crystal packing, and which were selected based on the observed similarity of space group and cell dimensions between crystals of

references