*Acta Cryst.* (1994). D**50**, 469-471

# Crystallizing Proteins – a Rational Approach?

BY ALLAN D'ARCY

*Departments of Pharmaceutical Research – New Technologies, F. Hoffmann–La Roche Ltd, CH-4002 Basel, Switzerland*

(*Received* 29 *November* 1993; *accepted* 20 *December* 1993)

## Abstract

The advances in recombinant DNA technology in recent years have had a dramatic effect on the area of protein crystallization. Large amounts of pure protein produced in various expression systems have made it possible to conduct experiments that would have been impossible with material from natural sources. With many more laboratories becoming involved in crystallizing proteins a great deal of new information has been generated on techniques to eliminate the so called 'bottleneck of crystallization' in determining a three-dimensional protein structure. More and more new and interesting proteins are being submitted to this laboratory for crystallization. Certain criteria may be set before crystallization trials are started, such as solubility, purity and aggregation tendencies. The introduction of robots now facilitates the screening of crystallization conditions. In cases where no crystals have been obtained after initial screening it can now be decided which possible modifications can be made to the protein itself to improve the chances of obtaining crystals.

Table 1. *Light-scattering results*

| Proteins measured | Size distribution | Crystals grown |
|---|---|---|
| Total 66 | → | 41 (100%) |
| | Narrow unimodal = 44 | 34 (83%) |
| | Broad unimodal = 10 | 6 (15%) |
| | Multimodal = 12 | 1 (2%) |

## Strategy for crystallization

Fig. 1 is a flow diagram illustrating the various steps towards crystallizing a protein and the options available to us when problems are encountered. After determining that a protein is soluble, SDS-PAGE (Laemmli, 1971) is used to determine the purity of the protein. We generally consider that a major band on a Coomassie blue stained gel is sufficient for initial screening. This does not exclude the possibility that subsequent purification may be necessary to improve crystal quality. The protein sample is then examined using laser light-scattering

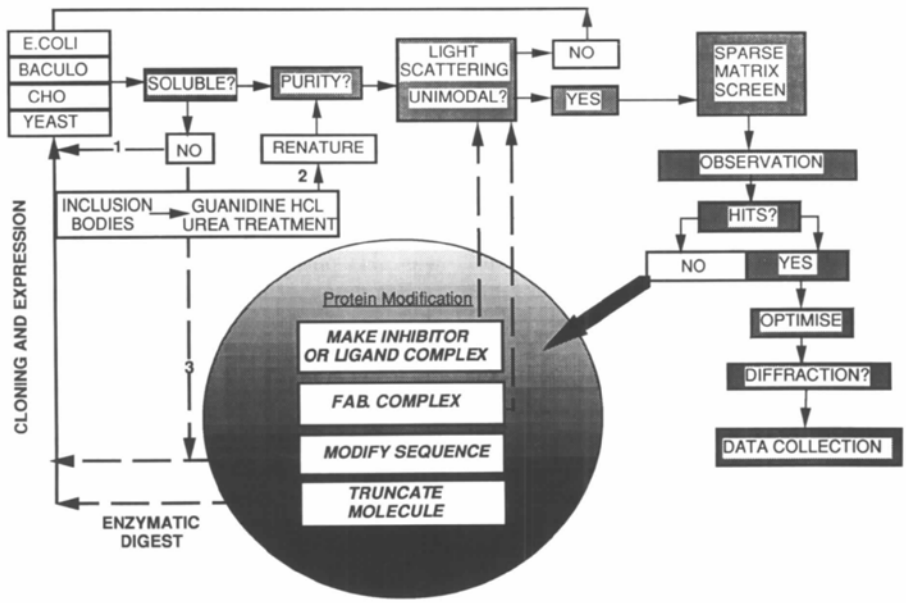


Fig. 1. Crystallization strategy.

techniques as previously described by Zulauf & D'Arcy (1992). This method allows us to detect the presence of aggregates in the protein solution which may be detrimental to crystal nucleation or growth. In a number of cases these aggregates can be eliminated by centrifugation in an airfuge at 100 000*g* for 30 min, or simply changing the buffer conditions. We now use the Dp801 molecular size detector (Protein Solutions) to perform these routine measurements. Table 1 summarizes the results of light-scattering measurements made on over 60 different proteins.

Having established that the protein is unimodal, we then proceed to a screening procedure using a sparse matrix screen based on the original described by Jancarik & Kim (1991); ammonium sulfate, phosphate and Jeffamine 6000 solutions have been included in our screen. If only small amounts of protein are available we can now conduct this screening using the microbatch method with a Douglas Instruments robot (Chayen, Shaw Stuart & Blow, 1992). The disadvantage of this method at present is that no vapour diffusion takes place and in many cases the protein cannot be brought to sufficient saturation to cause nucleation. For this reason we have designed a new crystallization dish for use with this robot system. Douglas Instruments have made some plates for testing in order to optimize the design before a final version is produced.

## Problem proteins

Unfortunately, not all proteins that are produced can crystallize readily and we need to identify the possible problems and try to find a reasonable rationale for solving them. One of the most common problems that we have encountered with recombinant proteins produced in *E. coli* is the formation of inclusion bodies containing large amounts of insoluble protein. In these cases, as shown in Fig. 1, we have three possible solutions to the problem to obtain the desired protein in a soluble form. The classical approach is to solubilize the inclusion bodies in guanidine hydrochloride or urea and slowly re-equilibrate the buffer to physiological conditions, in the hope that the protein will assume a correct and active fold. This approach has proved reliable for some proteins but unsuccessful with others. Sequence modifications using site-directed mutagenesis may improve the solubility of a particular protein or introduce more favourable crystal contacts as previously reported by McElroy, Sisson, Schoetlin, Aust & Villafranca (1992). We have employed this method to obtain crystals of the trimethoprim-resistant dihydrofolate reductase enzyme from *S. aureus* which formed inclusion bodies when expressed in *E. coli* and could not be solubilized using classical methods of denaturation and renaturation. A number of single and double mutations were made to surface residues of the enzyme and cell breakage supernatants screened on SDS-PAGE (Laemmli, 1971) to determine solubility. A double mutation of Asn148Glu and Asn130 to Asp proved to be the most promising option and this mutant has now been crystallized.

When neither of these methods are successful we may have to change to a eucaryotic expression system. This proved to be the case in our laboratory with the soluble $\gamma$-interferon receptor which was expressed both in *E. coli* and *Baculovirus* infected insect cells, but only the protein produced in the *Baculovirus* system proved suitable for crystallization (Fountoulakis, Juranville, Maris, Ozmen & Garotta, 1990; Fountoulakis *et al.*, 1991). In cases where a particular protein fulfils all the criteria for crystallization and still does not produce crystals after extensive screening we should consider what modifications we can make to the protein before continuing crystallization trials. Our first and often the simplest option is to form a complex of our protein either with an inhibitor or a ligand; examples have been reported by Knighton, Nguyen, Taylor & Sowadski (1991) and Badasso, Sibanda, Dealwis & Wood (1992). For many of the proteins we study in our laboratory inhibitors are already available and have proved useful for co-crystallization. One example is that of the catalytic domain of human recombinant fibroblast collagenase which, despite being both highly pure and narrow unimodal, proved extremely difficult to crystallize. The binding of a specific inhibitor to this protein drastically changed its crystallization properties and seven 'hits' were observed in the sparse matrix screen, in some of which crystals were already suitable for X-ray analysis.

A similar situation exists with soluble receptors as in most cases the ligands have already been studied and are available in sufficient quantities. We have attempted to crystallize the soluble extracellular domains of the p55 TNF receptor (D'Arcy *et al.*, 1993) and the IFN-$\gamma$ receptor. In both cases either no crystals or only non-diffracting crystals grew. Purified complexes of both receptors and their respective ligands gave crystals in the initial screening which could be optimized for data collection. More recently we have been successful in crystallizing the IL-5$\alpha$ receptor-ligand complex, which was also purified as a preformed complex.

For many of the proteins prepared for crystallization at Roche monoclonal antibodies are available. This provides us with another tool in cases where the native protein proves difficult to crystallize, and examples have been reported by Ruf *et al.* (1992). Our attempts to crystallize the soluble P75 TNF receptor alone and as a complex with the ligand have proved unsuccessful so far. As monoclonal antibodies to this molecule were available we decided to make Fab fragments and try to crystallize the Fab-receptor complex. A first screening produced crystals which could later be grown reproducibly using seeding techniques; although the diffraction limit of these crystals extends to only ~4 Å, syn-

Table 2. *Crystallization conditions*

| Protein | Concentration (mg ml $^{-1}$) | Reservoir mixed with drop 1:1 | Comments |
|---|---|---|---|
| *S. Aureus* DHFR TMP sensitive | 7.5 | 25% PEG 4000<br>200 m*M* ammonium acetate<br>100 m*M* cacodylate pH 5.5 | Soluble enzyme |
| *S. Aureus* DHFR TMP resistant | 6 | 25% PEG 4000<br>100 m*M* $MgCl_2$<br>100 m*M* cacodylate pH 5.5 | Soluble enzyme produced by mutating surface residues |
| Human recombinant fibroblast collagenase | 13 | 25% PEG 4000<br>100 m*M* LiCl<br>100 m*M* cacodylate pH 6.5 | Inhibitor essential for crystallization |
| P55 TNF receptor complex | 20 | 15% PEG 4000<br>400 m*M* $MgCl_2$<br>100 m*M* cacodylate pH 5.5<br>2% $\beta$-octylglucoside | Complex with ligand, detergent essential |
| $\gamma$ interferon–receptor complex | 12 | 14% PEG 8000<br>100 m*M* $MgCl_2$<br>100 m*M* tris pH 8.0<br>2.85% $\beta$-octylglucoside | Complex with ligand, detergent essential, grown under oil (batch) |
| IL-5–receptor complex | 24 | 12% PEG 8000<br>200 m*M* $Li_2SO_4$<br>100 m*M* cacodylate 6.0 | Complex with ligand |
| Fab–P75 TNF receptor complex | 6 | 10% PEG 4000<br>50 m*M* ammonium sulfate<br>50 m*M* Hepes pH 7.5 | Fab complex necessary for crystallization |

chrotron measurements may permit data collection to higher resolution.

A further possibility for obtaining crystals of proteins that are perhaps too flexible or multidomain domain structures is to produce a truncated molecule that retains biological activity. This may be done by enzymatic digestion or by directly cloning and expressing specific domains of the molecule. We are now in the process of sequencing fragments from proteolytic digests of another receptor molecule which we have not been able to crystallize using the methods described. The conditions used for growing the crystals described are summarized in Table 2.

## Concluding remarks

With the many different techniques now available for crystallizing biological macromolecules we should continue to see an increase in the number of proteins crystallized and subsequent structures solved. The combination and application of traditional methods and those available to us through the developments in molecular biology will be instrumental if we are to remove the so-called 'bottleneck of crystallization' from the process of determining protein crystal structures.

## References


BADASSO, B. L., SIBANDA, J. B., DEALWIS, C. G. & WOOD, S. P. (1992). *J. Cryst. Growth*, **122**, 393-399.

CHAYEN, N. E., SHAW STUART, P. D. & BLOW, D. (1992). *J. Cryst. Growth*, **122**, 176-180.

D'ARCY, A., BANNER, D. W., JANES, W., WINKLER, F. W., LOETSCHER, H., SCHÖNFELD, H. J., ZULAUF, M., GENTZ, R. & LESSLAUER, W. (1993). *J. Mol. Biol.* **229**, 555-557.

FOUNTOULAKIS, M., JURANVILLE, J. F., MARIS, A., OZMEN, L. & GAROTTA, G. (1990). *J. Biol. Chem* **265**(22), 13268-13275.

FOUNTOULAKIS, M., SCHLAEGER, E. J., GENTZ, R., JURANVILLE, J. F., MANNABERG, M., OZMEN, L. & GAROTTA, G. (1991). *Eur. J. Biochem.* **198**, 441-450.

JANCARIK, J. & KIM, S. H. (1991). *J. Appl. Cryst.* **24**, 409-411.

KNIGHTON, D. R., NGUYEN, H. U., TAYLOR, S. & SOWADSKI, J. (1991). *J. Mol. Biol.* **220**, 217-220.

LAEMMLI, U. K. (1971). *Nature (London)*, **227** 680-685.

MCELROY, H. E., SISSON, G. W., SCHOETLIN, W. E., AUST, R. M. & VILLAFRANCA, J. E. (1992). *J. Cryst. Growth*, **122**, 265-272.

RUF, W., STURA, E., LAPOLLA, R., SYED, R., EDGINGTON, T. S. & WILSON, I. A. (1992). *J. Cryst. Growth*, **122**, 253-264.

ZULAUF, M. & D'ARCY, A. (1992). *J. Cryst. Growth*, **122**, 102-106.