Acta Cryst. (1994). D**50**, 380–384

Enzymatic Deglycosylation as a Tool for Crystallization of Mammalian Binding Proteins

BY HEATHER M. BAKER,* CATHERINE L. DAY, GILLIAN E. NORRIS AND EDWARD N. BAKER

Department of Chemistry and Biochemistry, Massey University, Palmerston North, New Zealand

(*Received* 15 *November* 1993; *accepted* 30 *November* 1993)

Abstract

Enzymatic deglycosylation has been used in attempts to crystallize several glycoproteins with the aim of overcoming the problems resulting from heterogeneity and flexibility of the attached glycan chains. An endoglycosidase preparation from *Flavobacterium meningosepticum*, comprising the enzymes endo F and PNGase-F, was used in experiments on the mammalian binding proteins lactoferrin and haemopexin. Significant differences were found in the susceptibility of different proteins to deglycosylation. For human lactoferrin (Lf) and its recombinant N-terminal half-molecule (Lf_N), deglycosylation was rapid and complete, and was essential for obtaining high-quality crystals of both apo-Lf and Lf_N; for bovine Lf, however, complete deglycosylation did not occur. Similarly, for rabbit haemopexin the carbohydrate chain on the C-terminal domain was easily removed, but the three chains on the N-terminal domain proved more resistant and their removal led to some fragmentation of the protein. Nevertheless, this approach provided the only means of crystallizing the C-terminal domain and is likely to be useful for other glycoproteins.

Introduction

Many proteins of eukaryotic origin are glycosylated, and such glycoproteins are frequently difficult to crystallize. Two main factors are likely to contribute to this difficulty, *i.e.* chemical heterogeneity and conformational flexibility in the carbohydrate moiety. Since the latter is attached to the protein surface this variability can lead to a lack of specificity in intermolecular interactions and inhibit the formation of an ordered crystal lattice.

These effects are illustrated by lactoferrin, an iron-binding glycoprotein of the transferrin family (Baker, Rumball & Anderson, 1987), which can be isolated from milk and other body fluids. Human lactoferrin has two glycosylation sites, one in each half of the bilobal molecule. Chemical analysis of the carbohydrate structure, however, shows considerable heterogeneity, with at least three different chemical structures present in a given preparation (Spik, Strecker, Fournet, Bouquelet & Montreuil, 1982). The chains have a common core, but differ towards their termini with respect to branching, sugar residues and sialation. In the crystal structure analysis of the diferric form of human lactoferrin little or no electron density can be seen for the carbohydrate, which is apparently disordered in the solvent spaces of the crystal (Anderson, Baker, Norris, Rice & Baker, 1989). Heterogeneity of a different type occurs for the recombinant N-terminal half-molecule of human lactoferrin. When this protein, which has a single glycosylation site, is expressed in a eukaryotic cell-culture system (Day, Stowell, Baker & Tweedie, 1992) the resultant preparations contain mostly glycosylated protein, but with a small amount of non-glycosylated material.

A number of strategies have been suggested (Stura, Nemerow & Wilson, 1992) for overcoming these problems of heterogeneity and flexibility in glycoproteins. We have chosen to use an endoglycosidase preparation obtained from *Flavobacterium meningosepticum* (Elder & Alexander, 1982). This preparation contains a mixture of enzymatic activities, including three β-*N*-acetylglucosidase F (endo F) activities (F_1, F_2 and F_3) as well as peptide-*N*-glycosidase F (PNGase F) (Plummer, Elder, Alexander, Phelan & Tarentino, 1984; Trimble & Tarentino, 1991). It gives essentially complete removal of Asn-linked carbohydrate chains, with cleavage occurring either at the Asn–oligosaccharide bond (PNGase F) or at the glycosidic linkages between the two core *N*-acetylglucosamine residues (endo F). Here we describe our experiences, using this preparation in attempts to deglycosylate and crystallize two glycoproteins, lactoferrin and haemopexin. Lactoferrin is lightly glycosylated (~ 6% carbohydrate by weight), the human protein having two glycosylation sites and the bovine protein four. Haemopexin, which is a serum protein whose function it is to bind and transport haem (Morgan & Smith, 1984), is more heavily glycosylated (~ 20% by weight) and has four glycosylation sites.

* To whom correspondence should be addressed (by air mail).

Materials and methods

Human lactoferrin was prepared from fresh colostrum as described previously (Norris, Baker & Baker, 1989). Bovine lactoferrin was prepared by a similar procedure (Norris, Anderson, Baker, Baker, Gartner, Ward & Rumball, 1986) but with a final step incorporating preparative isoelectric focusing (M. Haridas, unpublished work). The recombinant N-terminal half-molecule of human lactoferrin (Lf_N) was obtained by expression of its cloned cDNA in baby hamster kidney (BHK) cells and purified as described (Day, Stowell, Baker & Tweedie, 1992). Rabbit haemopexin, and its separate N-terminal and C-terminal domains, was prepared from fresh rabbit serum (Morgan & Smith, 1984); the two domains were separated by ion-exchange chromatography after digestion of apohaemopexin with plasmin.

For isolation of the endoglycosidase preparation, cultures of *Flavobacterium meningosepticum* were obtained from the National Health Institute of New Zealand (NHI 847). Isolation followed the procedures of Elder & Alexander (1982), with the following modifications. After removal of the cells from 2 l of culture medium by centrifugation, ethylenediaminetetraacetic acid (EDTA) was added to a concentration of 50 m*M* to inhibit proteolytic activity. Ammonium sulfate was then added, to 95% saturation, the suspension allowed to equilibrate for at least 2 h, then centrifuged (13 800*g*, 40 min) and the pellet suspended in 100 ml of 0.025 *M* Tris/50 m*M* EDTA/50% saturated ammonium sulfate, adjusted to pH 7.2 with HCl. This suspension was stirred for 1 h before being centrifuged (17 300*g*, 20 min) and the pellet then resuspended in 20 ml of 0.025 *M* Tris–HCl, pH 7.2, containing 0.15 *M* NaCl and 5 m*M* EDTA. Any insoluble material was removed by centrifugation (17 300*g*, 20 min) and the dark brown supernatant applied to an Ultrogel AcA_{44} (IBF, France) column. Three main active peaks were obtained; these were combined, concentrated to between 0.2 and 0.5 ml, and glycerol added to 50%(*v*/*v*) for storage at 253 K. For comparative purposes, pure PNGase F was purchased from New England Biolabs.

Deglycosylation of the glycoprotein samples was generally carried out in a buffer comprising 0.1 *M* sodium phosphate, 50 m*M* EDTA, 0.5% Nonidet P-40, pH 6.0–6.5. Protein samples of 25–50 mg were equilibrated in this buffer solution at protein concentrations of 5–10 mg ml^{-1}, and incubated with the concentrated endoglycosidase preparation; typically 10–20 μl of the latter were used per 50 mg of protein. Deglycosylation both with and without Nonidet in the buffer was tried, but EDTA was always included because the endoglycosidase preparation reportedly contains some proteolytic impurity (Elder & Alexander, 1982). Deglycosylation was carried out at room temperature, with the progress of the reaction being monitored by taking samples at regular intervals and observing the reduction in molecular weight on sodium dodecyl sulfate (SDS) gels. The presence of carbohydrate could also be detected by staining the gels with thymol–H_2SO_4 or Fuchsin basic (Gandar, 1985).

Following the reaction, the deglycosylated protein was separated from the incubation mixture by gel filtration. The solution was first dialyzed against 0.05 *M* Tris–HCl, 5 m*M* EDTA, 0.15 *M* NaCl, pH 7.5 before passage down an Ultrogel AcA_{44} (IBF) column (1.0 × 40.0 cm) equilibrated with the same buffer. This also generally removed the Nonidet P-40 from the protein. Further purification was by ion-exchange chromatography if necessary.

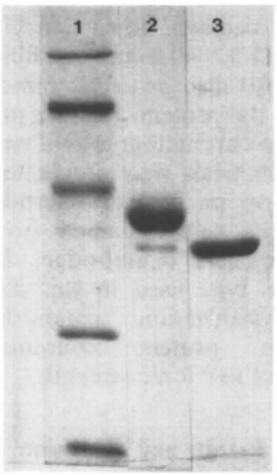
Crystallization trials were conducted using either the hanging-drop method or microdialysis, in the latter case using cells obtained from Cambridge Repetition Engineers (Cambridge, England). For proteins of the type used in this study the most successful crystallization conditions generally required high protein concentrations (20–40 mg ml^{-1}) and low ionic strength.

Results and discussion

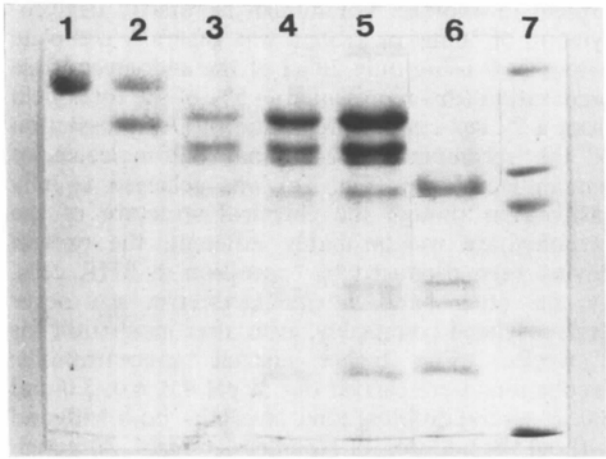
The ease of deglycosylation varied greatly from one protein to another. For human lactoferrin deglycosylation of 50 mg of protein was generally complete within 24 h using only 20 μl of the endoglycosidase preparation (this represented ~ 5% of the total yield from a 2 l cell-culture preparation). Deglycosylation of the recombinant N-terminal half-molecule of human lactoferrin (Fig. 1*a*) was achieved equally easily even though the chemical structure of the carbohydrate was probably different, the protein having been produced by expression in BHK cells. On the other hand, bovine lactoferrin was never deglycosylated completely, even after incubation for 2 weeks, using higher enzyme concentrations. Incubations were carried out at pH 4.0, 6.0, 7.0 and 8.0, at several different ionic strengths, both with and without the addition of Nonidet detergent. Although some carbohydrate was clearly removed, as shown by the reduction in molecular weight on SDS gels, staining for carbohydrate showed that deglycosylation was incomplete. The difference between human and bovine proteins is probably attributable to the accessibility of the glycosylation sites. Both sites on human lactoferrin, Asn137 and Asn478, are highly exposed, on external loops on the molecular surface. For bovine lactoferrin, however, a subsequent crystal structure analysis of the glycosylated protein has shown that at least one of the glycosylation sites, Asn545, is located in a fairly deep cleft

in the protein structure. The carbohydrate chain extending from this site has quite well defined electron density (Haridas, Anderson & Baker, unpublished work) and it seems likely that the carbohydrate at this site is protected by the surrounding protein structure from endoglycosidase attack.

In the case of haemopexin the N-terminal and C-terminal half-molecules showed striking differences in their susceptibility to deglycosylation. The C-terminal domain, which comprises residues 215–435, has only one glycosylation site, at Asn218, very close to the N-terminus of the half-molecule; the carbohydrate at this site is completely removed within 1 h, presumably because of its high accessibility. The N-terminal half-molecule, residues 1–214, has, however, three glycosylation sites, at Asn9, Asn41 and Asn163 (Morgan, Tatum, Muster, Alam & Smith, 1993) and the time course of deglycosylation (Fig. 1*b*) suggests that these vary widely in their susceptibility to attack. One chain (probably that at Asn9) is removed after 24 h, a second after 3–4 d of incubation, and deglycosylation is finally complete after ~ 6 d. The gels of Fig. 1(*b*) also show a more serious problem, however, in that some fragmentation of the polypeptide chain also occurs; this is shown by the presence of some lower-molecular-weight bands. Even the use of a cocktail of proteinase inhibitors (EDTA, phenylmethylsulfonyl fluoride, iodoacetate), or the use of commercially available pure PNGase F (New England Biolabs), failed to prevent this fragmentation, the origin of which is not known.



(*a*)



(*b*)

Fig. 1. SDS gels used to monitor glycosylation. (*a*) The recombinant N-terminal half-molecule of human lactoferrin, Lf_N. Lane 1, molecular weight markers; lane 2, Lf_N as obtained from BHK cells, showing the minor contaminant of non-glycosylated material; lane 3, deglycosylated Lf_N. (*b*) Time course of deglycosylation of N-terminal domain of rabbit haemopexin. Lane 1, native N-terminal domain; lane 2, after 2 h, showing partial removal of one glycan chain; lane 3, after 6 h, showing complete removal of one chain; lane 4, after 24 h, showing partial removal of the remaining two; lane 5, after 3 d, showing some fragmentation of the protein has also occurred; lane 6, after 14 d, showing complete deglycosylation but more fragmentation; lane 7, molecular weight markers.

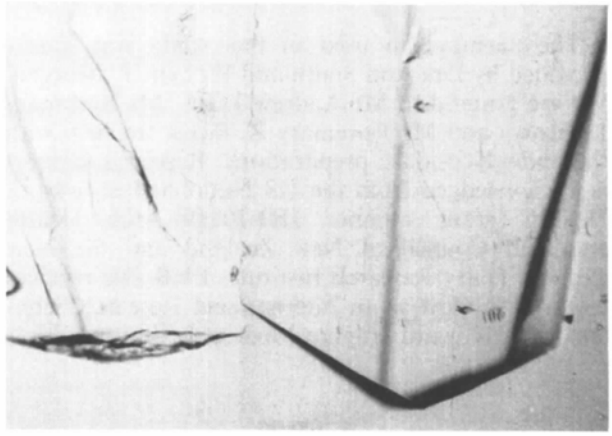
Commercial samples of PNGase F were generally less effective than the endoglycosidase preparation described here. Substantially longer incubation times were required; for human lactoferrin, for example, 50 000 units of PNGase F had failed to remove all the carbohydrate from a 50 mg sample even after 10 d incubation, compared with less than 24 h for complete deglycosylation using the laboratory sample. The difference may be attributable to the advantages of (i) using a freshly prepared enzyme sample, and (ii) using the mixture of four enzyme activities, endo F_1, F_2 and F_3, and PNGase F. At pH 6.0–6.5 all have substantial activity (Plummer, Elder, Alexander, Phelan & Tarentino, 1984), whereas at pH 4.0 most of the activity derives from the endo F enzymes and at pH 8.0 most from PNGase F. We therefore found a pH of 6.0–6.5 to be generally optimal.

A key factor in the most successful deglycosylations was the ability to perform the deglycosylation rapidly. For this reason the use of enzyme samples which were fresh and of high activity, and of concentrated solutions and high enzyme-to-protein ratios, was advantageous.

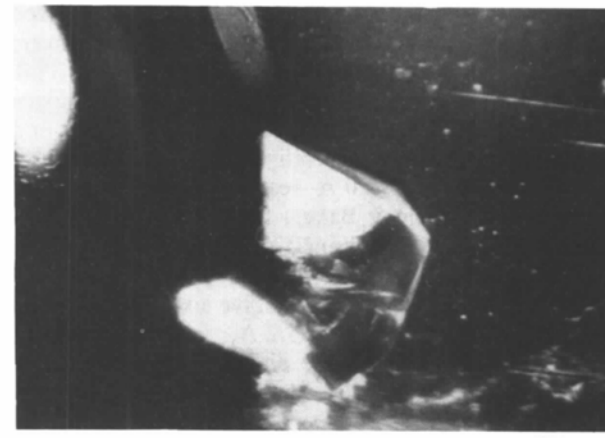
In several cases deglycosylation was essential for the preparation of high-quality crystals, some of which are shown in Fig. 2. Although crystals of the diferric form of native (glycosylated) human lactoferrin can be easily obtained (Baker & Rumball, 1977) and diffract to nearly 2.0 Å resolution, the apoprotein, which undergoes a substantial conformational change on iron removal and is conformationally more flexible (Anderson, Baker, Norris, Rumball & Baker, 1990) gives only poor crystals. These crystals are unstable, with a high solvent

content and a large unit cell ($a = 220.0$, $b = 115.0$, $c = 76.5$ Å, $\alpha = \beta = \gamma = 90°$, space group $P2_12_12_1$) and diffract to only 4.0 Å resolution. The deglycosylated apo-protein, however, gave good quality crystals ($a = 152.1$, $b = 94.0$, $c = 56.0$ Å, $\alpha = \beta = \gamma = 90°$, space group $P2_12_12_1$) which diffracted to at least 2.0 Å resolution (Norris, Baker & Baker, 1989) and have been used for the successful X-ray structure analysis (Anderson, Baker, Norris, Rumball & Baker, 1990).

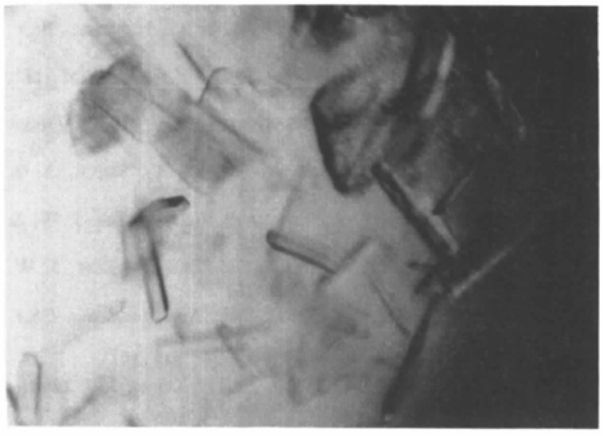
In the case of the recombinant N-terminal half-molecule of human lactoferrin, crystals of the ferric



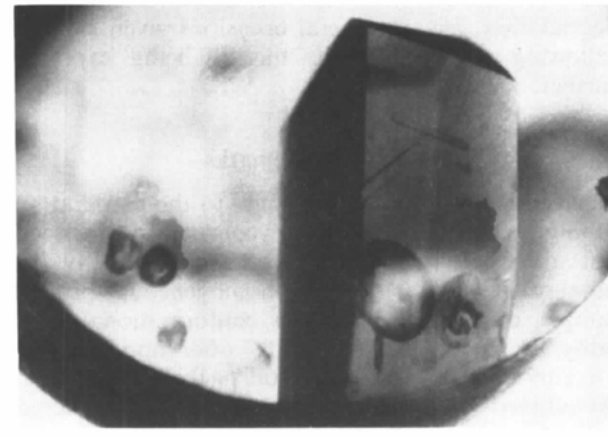
(*a*)



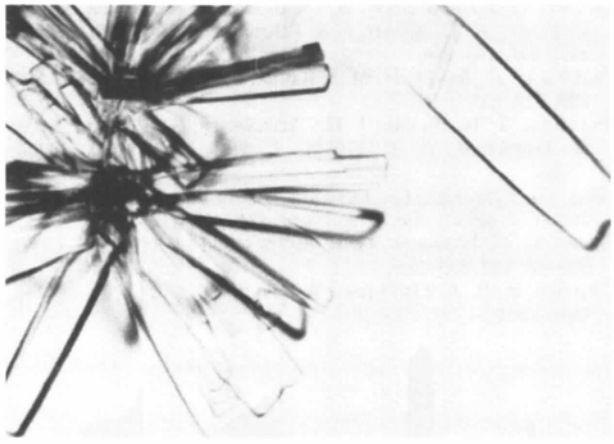
(*b*)



(*c*)



(*d*)



(*e*)

Fig. 2. Crystals of (*a*) native human apo-lactoferrin, (*b*) deglycosylated human apo-lactoferrin, (*c*) ferric form of native Lf_N, (*d*) ferric form of deglycosylated Lf_N and (*e*) deglycosylated C-terminal domain of rabbit haemopexin.

form of the glycosylated protein could be obtained. Preparations were usually contaminated with a small proportion of non-glycosylated material (Fig. 1*a*), which was difficult to remove, however, and the crystals, prepared by dialysis at low ionic strength (0.05 *M* Tris, 12% isopropanol, pH 7.8), were not easily reproducible and were both temperature and radiation sensitive. Deglycosylation gave a homogeneous preparation, which could be crystallized under very similar conditions. The crystals were isomorphous with the glycosylated form ($a = 133.0$, $b = 58.3$, $c = 58.3$ Å, $\alpha = \gamma = 90°$, $\beta = 115.7°$, space group $C2$) but were obtained much more easily and reproducibly, could be stabilized readily and diffracted to better than 2.0 Å resolution (Day, Norris, Anderson, Tweedie & Baker, 1992). Finally, no crystals have yet been obtained of any form of glycosylated haemopexin. The deglycosylated C-terminal half-molecule does, however, give excellent crystals ($a = 41.0$, $b = 65.0$, $c = 85.2$ Å, $\alpha = \beta = \gamma = 90°$, space group $P2_12_12_1$); these diffract to at least 2.4 Å resolution, are stable and reproducible, and are currently the subject of an X-ray structure analysis (Baker, Norris, Morgan, Smith & Baker, 1993). Even the N-terminal half-molecule, which is prone to degradation, has on several occasions given crystals following deglycosylation; this is being explored further.

Concluding remarks

Deglycosylation can offer a route to the preparation of high-quality crystals of glycoproteins which do not otherwise crystallize readily. This may be because it gives a more homogeneous protein sample, or because it reduces conformational flexibility. It may, however, simply offer opportunities for alternative modes of crystal packing because of the altered molecular surface, or lead to altered crystallization conditions because of changes to the protein solubility.

Deglycosylation is not always easily achieved, its ease varying greatly from one protein to another. This may be more of a factor when using endoglycosidases, as in the present study, since the site of attack is close to the protein surface. Other approaches are possible, however, including removal of the terminal sialic acid residues with neuraminidase, or using a 'cocktail' of exoglycosidases to cut back the glycan chain progressively.

Finally, thought must be given to the biological effects of deglycosylation. In the case of lactoferrin, deglycosylation has no effect on its properties of iron binding and release, its stability to temperature or proteolysis, or its conformation (as judged by circular dichroism spectroscopy and X-ray analysis). For other proteins this is not necessarily the case and the relative structural and functional properties of glycosylated and deglycosylated proteins should always be considered.

The haemopexin used in this study was kindly provided by Drs Ann Smith and William T. Morgan. We are grateful to Mr Andrew Flaus, Ms Rosemary L. Brown and Ms Rosemary R. Baker for help with the endoglycosidase preparations. Research support is acknowledged from the US National Institutes of Health (grant number HD-20859), the Health Research Council of New Zealand and the New Zealand Dairy Research Institute. ENB also receives research support as an International Research Scholar of the Howard Hughes Medical Institute.

References

ANDERSON, B. F., BAKER, H. M., NORRIS, G. E., RICE, D. W. & BAKER, E. N. (1989). *J. Mol. Biol.* **209**, 711–734.

ANDERSON, B. F., BAKER, H. M., NORRIS, G. E., RUMBALL, S. V. & BAKER, E. N. (1990). *Nature (London)*, **344**, 784–787.

BAKER, E. N. & RUMBALL, S. V. (1977). *J. Mol. Biol.* **111**, 207–210.

BAKER, E. N., RUMBALL, S. V. & ANDERSON, B. F. (1987). *Trends Biochem. Sci.* **12**, 350–353.

BAKER, H. M., NORRIS, G. E., MORGAN, W. T., SMITH, A. & BAKER, E. N. (1993). *J. Mol. Biol.* **229**, 251–252.

DAY, C. A., NORRIS, G. E., ANDERSON, B. F., TWEEDIE, J. W. & BAKER, E. N. (1992). *J. Mol. Biol.* **228**, 973–974.

DAY, C. L., STOWELL, K. M., BAKER, E. N. & TWEEDIE, J. W. (1992). *J. Biol. Chem.* **267**, 13857–13862.

ELDER, J. H. & ALEXANDER, S. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 4540–4544.

GANDAR, J. E. (1985). *Methods Enzymol.* **104**, 447–451.

MORGAN, W. T. & SMITH, A. (1984). *J. Biol. Chem.* **259**, 12001–12006.

MORGAN, W. T., TATUM, F., MUSTER, P., ALAM, J. & SMITH, A. (1993). *J. Biol. Chem.* **268**, 6256–6262.

NORRIS, G. E., ANDERSON, B. F., BAKER, E. N., BAKER, H. M., GARTNER, A. L., WARD, J. & RUMBALL, S. V. (1986). *J. Mol. Biol.* **191**, 143–146.

NORRIS, G. E., BAKER, H. M. & BAKER, E. N. (1989). *J. Mol. Biol.* **209**, 329–331.

PLUMMER, T. H., ELDER, J. H., ALEXANDER, S., PHELAN, A. W. & TARENTINO, A. L. (1984). *J. Biol. Chem.* **259**, 10700–10704.

SPIK, G., STRECKER, G., FOURNET, B., BOUQUELET, S. & MONTREUIL, J. (1982). *Eur. J. Biochem.* **121**, 413–419.

STURA, E. A., NEMEROW, G. R. & WILSON, I. A. (1992). *J. Cryst. Growth*, **122**, 273–285.

TRIMBLE, R. B. & TARENTINO, A. L. (1991). *J. Biol. Chem.* **266**, 1646–1651.