Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Strategies for the structural determination of the catalytic domain of *Escherichia coli* cytotoxic necrotizing factor 1

**Lori Buetow,[a] Gilles Flatau,[b] Katy Chiu,[a] Patrice Boquet[b] and Partho Ghosh*[a]**

[a]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093-0314, USA, and [b]INSERM U452, Faculté de Médecine, 28 Avenue de Valombrose, 06107 Nice CEDEX 2, France

Correspondence e-mail: pghosh@ucsd.edu

Received 14 June 2001
Accepted 5 December 2001

Cytotoxic necrotizing factor 1 (CNF1), a 114 kDa toxin produced by certain pathogenic strains of *Escherichia coli*, constitutively activates members of the Rho GTPase family, leading to cytopathic effects. The toxin inhibits GTP turnover by Rho proteins through site-specific deamidation of a Rho glutamine required for GTP hydrolysis. To understand the basis for catalytic activity and target specificity of CNF1, the structure of a catalytically active fragment of CNF1 was sought. Here, strategies that led to successful expression of a soluble 33 kDa active fragment, growth and improvement in the quality of the crystals and determination of phases using a quadruple methionine-substitution mutant of the fragment are presented.

## 1. Introduction

Cytotoxic necrotizing factor 1 (CNF1) is a 114 kDa toxin produced by certain pathogenic strains of *E. coli* (Caprioli *et al.*, 1983). CNF1 has been shown to play a role in the pathogenesis of strains that cause urinary tract infection or neonatal meningitis (Caprioli *et al.*, 1983; Andreu *et al.*, 1997; K. S. Kim, personal communication). CNF1 belongs to a small group of sequence-related toxins that includes *E. coli* cytotoxic necrotizing factor 2 (CNF2) and *Bordetella* dermonecrotic toxin (DNT). In addition, related sequences occur in *Yersinia* spp., although no protein activity has yet been reported. CNF1, CNF2 and DNT constitutively activate members of the Rho GTPase family (RhoA, Rac1 and Cdc42) (Flatau *et al.*, 1997; Schmidt *et al.*, 1997), which are key regulators of the actin cytoskeleton and are implicated in the control of cell adhesion and gene expression (Hall, 1998). CNF1 catalyzes site-specific deamidation of Rho proteins at a glutamine residue (Gln63 in RhoA, Gln61 in both Rac and Cdc42; Flatau *et al.*, 1997; Schmidt *et al.*, 1997) required for GTP hydrolysis. CNF1 and the other toxins are also capable of site-specifically transglutaminating this residue (Schmidt *et al.*, 1998, 1999). Modification of the Rho glutamine greatly reduces its rate of intrinsic or protein-assisted GTP hydrolysis, leading to the persistence of Rho proteins in the activated GTP-bound state. In target cells, constitutive activation of Rho proteins by this group of toxins leads to cytopathic effects (Caprioli *et al.*, 1983).

The C-terminal third of CNF1 is necessary and sufficient to carry out site-specific deamidation of Rho proteins and, when introduced into host cells, to bring about cytopathic effects similar to those caused by the intact toxin (Lemichez *et al.*, 1997; Schmidt *et al.*, 1998). To understand the basis for catalytic activity and target specificity, we have expressed, purified and crystallized a catalytically active region of CNF1 encompassing residues 720–1014 (Buetow *et al.*, 2001). We present here strategies that proved successful in producing soluble fragments of CNF1, in increasing the size and diffraction quality of the initial crystals and in determining phases using a quadruple methionine-substitution mutant.

## 2. Crystallization of the catalytic domain of CNF1

Six fragments of the CNF1 catalytic domain that differ at their ends by ten residues were constructed for expression in *E. coli*. The ends were chosen randomly since information on domain boundaries from sequence or biochemical data was lacking. Overexpression was observed for all six constructs (710–1004, 710–1014, 720–1004, 720–1014, 730–1004, 730–1014), but only the one encoding residues 720–1014 (referred to as CNF1-C) produced soluble protein. Even so, only a fraction (30%) of expressed CNF1-C is soluble (Fig. 1). The soluble fraction of CNF1-C was purified and demonstrated to have activity as a site-specific deamidase of RhoA (Buetow *et al.*, 2001).

Purified CNF1-C exhibited salting-in effects, as evidenced by the solubility limit of CNF1-C increasing from $\sim$6 mg ml$^{-1}$ in the absence

of salt to ~20 mg ml$^{-1}$ in 150 m*M* NaCl. Likewise, precipitating conditions containing less than 150 m*M* ionic strength readily supported the growth of needle-like crystals. Indeed, small needles formed when an equal volume of water was added to concentrated CNF1-C (~15 mg ml$^{-1}$ in 150 m*M* NaCl) (Fig. 2*a*). Approximately one-third of the Jancarik & Kim (1991) factorial screen conditions, which share the feature of being low in ionic strength, yielded small needles in crystallization trials with CNF1-C. The needles were too small for X-ray diffraction data collection and proved refractory to optimization by various means, including macroseeding and microseeding, recrystallization, lysine alkylation of CNF1-C and inclusion of additives.

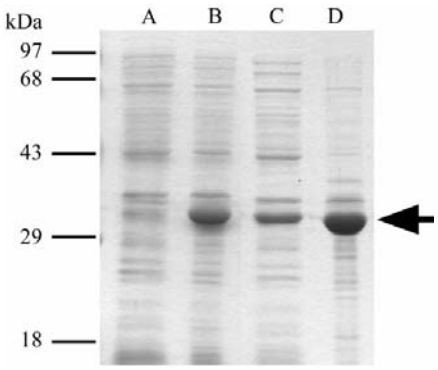


**Figure 1**
SDS–PAGE (12%) of overexpression and solubility of CNF1-C. Lane A, total cell (*E. coli*) protein prior to induction with 0.5 m*M* IPTG; lane B, total cell protein at 5 h post-induction; lane C, soluble fraction of total cell protein; lane D, insoluble fraction of total cell protein. Molecular-weight markers are indicated on the left and CNF1-C is marked by an arrow.

Crystallization by salting-out effects was also evident and yielded very small crystals (Fig. 2*b*). A number of salts including sodium sulfate, sodium potassium phosphate, sodium tartrate, potassium tartrate and sodium citrate, yielded crystals similar in appearance. To increase the size of these crystals, an additive screen using 21 small organic compounds was tried, with the result that only one additive, 1,6-hexanediol, was found to increase crystal size; the effect, however, was slight (Fig. 2*c*). Diffraction from a cluster of these small crystals, which were cryoprotected and mounted in the same loop, was found to extend to 3.3 Å resolution (using a rotating-anode X-ray source). Although the diffraction pattern was weak, it indicated that further efforts at optimization of crystal size were warranted.

Based on the lead with 1,6-hexanediol, a systematic additive screen using aliphatic diols of varying alkyl chain length was tried (Table 1). 1,8-Octanediol was found to decrease the amount of time required for crystallization and to increase the number of crystals, but failed to yield an increase in crystal size. 1,4-Butanediol, however, caused a dramatic change. While both 1,4-butanediol and 1,6-hexanediol yielded crystals of similar size (Fig. 2*d*), crystals grown in the presence of 1,4-butanediol diffracted strongly to a resolution limit of 2.2 Å. Crystal size was also dramatically improved in the 1,4-butanediol condition by using larger volumes of protein (5 µl and equal volume of precipitant) and growing crystals at 310 K (Fig. 2*e*). These crystals were approximately twice as long and three to four times as wide as their non-optimized counterparts and diffracted to a resolution limit of at least 1.76 Å (at 100 K after cryoprotection with Paratone-N). Both volume and temperature changes were essential to improvements in the crystals.

**Table 1**
Properties of crystals grown in high-salt conditions (1.1–1.3 *M* $Na_2SO_4$) using different optimization techniques.

Volume refers to the amount of protein added to a hanging drop. Ppt. refers to precipitation without the formation of crystals. +, ++ and +++++ refer to crystal dimensions of approximately 30 × 5 × 5, 60 × 15 × 15 and 200 × 60 × 60 µm, respectively. nd, not determined.

| Additive | Volume (µl) | Temperature (K) | Crystal size | Resolution limit (Å) |
|---|---|---|---|---|
| — | 1 | 291 | + | nd |
| 1,6-Hexanediol | 1 | 291 | ++ | 3.3 |
| 1,8-Octanediol | 1 | 291 | ++ | nd |
| 1,4-Butanediol | 1 | 291 | ++ | 2.2 |
| 1,4-Butanediol | 5 | 291 | ++ | 2.2 |
| 1,4-Butanediol | 1 | 310 | Ppt. | nd |
| 1,4-Butanediol | 5 | 310 | ++++++ | 1.76 |

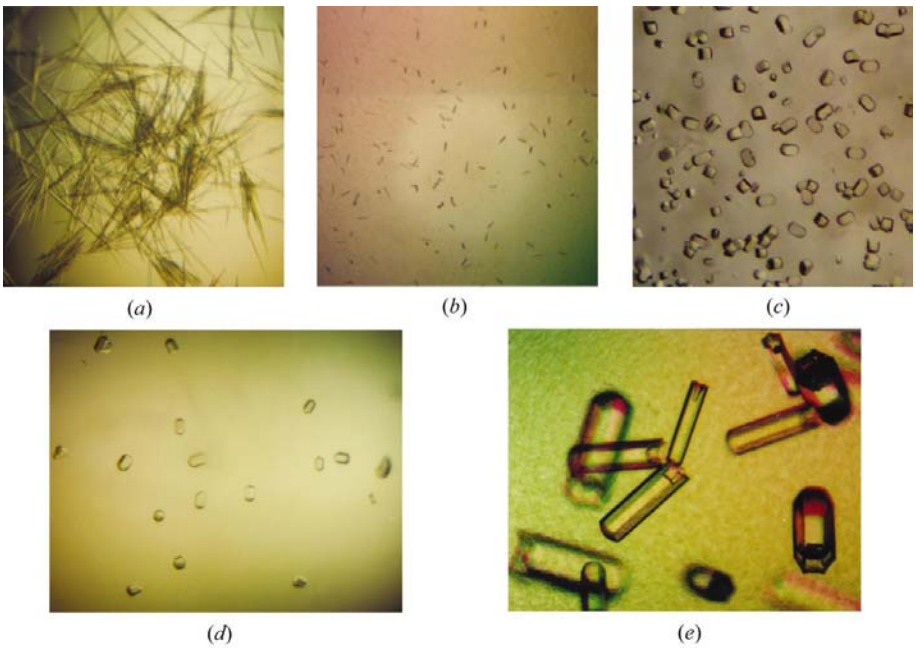


**Figure 2**
Crystallization of CNF1-C. CNF1-C at 15 mg ml$^{-1}$ was crystallized at 291 K using 1 µl of protein and 1 µl of precipitant (except in *e*). (*a*) Salting-in effects: precipitant is water. (*b*) Salting-out effects: precipitant is high salt (1.3 *M* $Na_2SO_4$). (*c*) Optimization of high-salt crystals using 2.5%(*w*/*v*) 1,6-hexanediol. (*d*) Optimization of high-salt crystals using 5%(*v*/*v*) 1,4-butanediol. (*e*) Optimization of high-salt crystals with 5% 1,4-butanediol through changes in temperature (310 K) and crystallization drop volume (5 µl protein and 5 µl precipitant).

The molecular reasons for the dramatic improvements afforded by including 1,4-butanediol in crystallization are not apparent. No molecules of 1,4-butanediol are observed in the crystal structure (Buetow *et al.*, 2001). However, the striking dependence of crystal morphology on the chain length of the aliphatic group suggests a direct role for the diol in stabilizing crystal contacts.

## 3. Obtaining phases

CNF1-C crystals are sensitive to the soaking manipulations that are required for isomorphous replacement techniques. This is evident from the fact that crystals that are soaked as identically as possible give good agreement in diffraction intensities (~3.6% to 1.9 Å resolution), whereas crystals that have been soaked for different amounts of time do not agree (differences of up to 30% to 2.5 Å resolution). Such differences did not arise from alterations in unit-cell para-

meters (which changed by 0.6% at most). Since this sensitivity makes phase determination by replacement techniques using multiple crystals difficult, the multiple-wavelength anomalous dispersion technique using a single crystal was attempted. Selenomethionine was incorporated by standard means, but instead of using wild-type CNF1-C, a quadruple methionine mutant was used to increase anomalous signal beyond that from the one internal methionine in the 295-residue CNF1-C (the initiator Met is proteolytically removed).

Construction of the quadruple methionine mutant began with construction of three double methionine mutants: L769M/F771M, N823M/L824M and L906M/L907M. Sites chosen for mutagenesis were based on three criteria. Firstly, Leu was targeted for mutagenesis because of the documented evolutionary tendency of Met to be substituted by Leu (Dayhoff *et al.*, 1978). Secondly, sites having a second Leu or hydrophobic residue close to the initially identified Leu were preferred (as in L769M/F771M and L906M/L907M) in order to encode two mutations on a single mutagenic DNA primer. Thirdly, homology to DNT was used to identify sites at which substitution by Met may be tolerated (as in N823M/L824M, where Asn823 is predicted to be replaced by Met in DNT; Buetow *et al.*, 2001). This last line of reasoning failed, as N823M/L824M was found to produce insoluble protein. The two other double mutants were at least partially soluble and were combined into a quadruple methionine mutant. The quadruple methionine mutant (L769M/F771M/L906M/L907M) contains a total of five methionines, four of which were ordered and provided phasing information in a MAD experiment. Met769 was not useful for phase determination and structural determination shows that it occupies a surface-exposed position on a flexible loop (Buetow *et al.*, 2001).

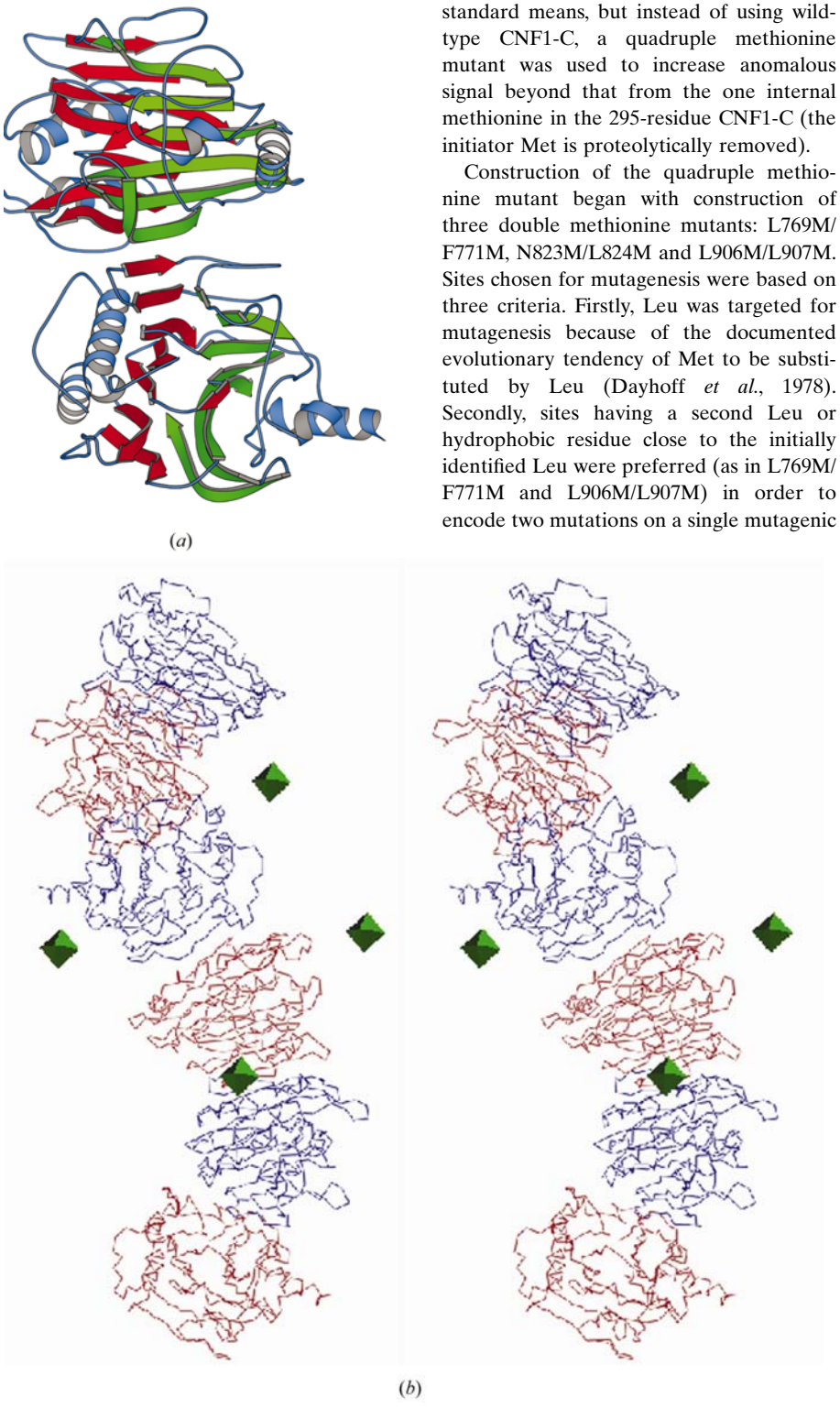


**Figure 3**
Crystal packing of CNF1-C. (*a*) Ribbon diagram of intermolecular $\beta$-sheet formed by symmetry-related molecules. The six-stranded $\beta$-sheet is shown in red, the seven-stranded in blue. (*b*) Stereoview of packing along sixfold-screw ($6_5$) axis (vertical on page). Six CNF-C molecules (alternating in red and blue) are shown. Green diamonds indicate packing interfaces between molecules located along adjacent sixfold screw axes. Fig. 3(*a*) was created using *MOLSCRIPT* (Kraulis, 1991) and Fig. 3(*b*) using *O* (Jones *et al.*, 1991).

## 4. Crystal packing

An interesting packing association is observed in CNF1-C crystals. CNF1-C crystallizes in the space group $P6_5$, with a single protein molecule in the asymmetric unit (conditions as described in Fig. 2*e*). CNF1-C has a central $\beta$-sandwich that is constituted of mixed six- and seven-stranded sheets (Buetow *et al.*, 2001). On one edge of the $\beta$-sandwich, the six-stranded sheet continues by $\beta$-strand hydrogen bonds into the seven-stranded sheet of an adjacent symmetry-related CNF1-C molecule (Fig. 3*a* and 3*b*). On the other edge of the $\beta$-sandwich, the seven-stranded sheet continues by $\beta$-strand hydrogen bonds into the six-stranded sheet of a second adjacent symmetry-related CNF1-C molecule. A large number of hydrogen bonds in addition to $\beta$-sheet hydrogen bonds are found at these interfaces. CNF1-C molecules that run along the sixfold screw axis are linked by the intermolecular 13-stranded $\beta$-sheet. The diamonds in Fig. 3(*b*) show how the sixfold screw axes pack against one another to give rise to the three-dimensional crystal.

A similar 13-stranded intermolecular $\beta$-sheet is also observed in an unrelated crystal form of CNF1-C obtained in low salt with PEG as a precipitant [15%(*w*/*v*) PEG 4K, 50 m*M* Tris pH 8.5, 100 m*M* sodium acetate, CNF1-C at $\sim$5 mg ml$^{-1}$]. In the PEG condition, CNF1-C crystallizes in the space group $P4_3$ (data not shown), and the intermolecular 13-stranded $\beta$-sheet links CNF1-C molecules along the fourfold screw axis. The crystals obtained in PEG are smaller and diffract to lower resolution ($\sim$2.9 Å) than those obtained at high salt. CNF1-C appears to have a propensity for forming crystallographic associations

through intermolecular $\beta$-sheets and the exact geometry of this packing (*e.g.* fourfold or sixfold screw) appears to be modulated by the precipitant (PEG or high salt).

## 5. Conclusions

In summary, the keys to obtaining crystals that diffract to atomic resolution in this case were inspection of diffraction properties of crystals, even if judged to be too small to yield single-crystal data, and systematic screening of compounds chemically related to an additive identified in a random screen. Phase determination by MAD depended on site-directed substitution of four residues; the substitutions were carried out incrementally in pairs. These strategies enabled structural determination of CNF1-C and provided essential information for devising strategies to inhibit this cytotoxin.

We thank Lawrence Knight and Ninh Nguyen for excellent technical assistance, Henry Bellamy of SSRL for help in MAD data collection and members of the laboratory for technical assistance and discussion. We thank Alan Hall for his gift of the RhoA-expressing vector. This work is based upon research conducted at the SSRL, which is funded by the Department of Energy (BES, BER) and the National Institutes of Health (NCRR, NIGMS). LB was supported by a GANN fellowship. PG is a W. M. Keck Distinguished Young Scholar.

## References

Andreu, A., Stapleton, A. E., Fennell, C., Lockman, H. A., Xercavins, M., Fernandez, F. & Stamm, W. E. (1997). *J. Infect. Dis.* **176**, 464–469.

Buetow, L., Flatau, G., Chiu, K., Boquet, P. & Ghosh, P. (2001). *Nature Struct. Biol.* **8**, 584–588.

Caprioli, A., Falbo, V., Roda, L. G., Ruggeri, F. M. & Zona, C. (1983). *Infect. Immun.* **39**, 1300–1306.

Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. (1978). *Atlas of Protein Sequence and Structure*, edited by M. O. Dayhoff, pp. 345–358. Washington DC: National Biomedical Research Foundation.

Flatau, G., Lemichez, E., Gauthier, M., Chardin, P., Paris, S., Florentini, C. & Boquet, P. (1997). *Nature (London)*, **387**, 729–733.

Hall, A. (1998). *Science*, **279**, 509–514.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kraulis, P. (1991). *J. Appl. Cryst.* **24**, 946–950.

Lemichez, E., Flatau, G., Bruzzone, M., Boquet, P. & Gauthier, M. (1997). *Mol. Microbiol.* **24**, 1061–1070.

Schmidt, G., Goehring, U. M., Schirmer, J., Lerm, M. & Aktories, K. (1999). *J. Biol. Chem.* **274**, 31875–31881.

Schmidt, G., Sehr, P., Wilm, M., Selzer, J., Mann, M. & Aktories, K. (1997). *Nature (London)* **387**, 725–729.

Schmidt, G., Selzer, J., Lerm, M. & Aktories, K. (1998). *J. Biol. Chem.* **273**, 13669–13674.