Acta Cryst. (1995). D51, 243-244

Crystallization and preliminary X-ray structure solution of *Rhizomucor miehei* aspartic proteinase.

By ZONGCHAO JIA, *Department of Chemistry, University of Saskatchewan, Saskatoon S7N 0W0, Canada*, MARGARET VANDONSELAAR, *Department of Biochemistry, University of Saskatchewan, Saskatoon S7N 0W0, Canada*, PALLE SCHNEIDER, *Department of Enzyme Chemistry, Novo Research Institute, DK-2880 Bagsvaerd, Copenhagen, Denmark* and J. WILSON QUAIL, *Department of Chemistry, University of Saskatchewan, Saskatoon S7N 0W0, Canada*

(*Received* 4 *October* 1993; *accepted* 12 *October* 1994)

Abstract

Rhizomucor miehei aspartic proteinase (M_r = 38 701, 361 residues) has been crystallized in a form suitable for analysis by X-ray diffraction. The flattened rod-shaped crystals were grown from polyethylene glycol 8000 using vapour-diffusion methods. The crystal form is in space group $P2_12_12_1$ [a = 41.67 (2), b = 51.21 (3) and c = 173.3 (2) Å], with Z = 4 and one molecule in the asymmetric unit. Data were collected over the range $0 < h < 14$, $0 < k < 16$ and $0 < l < 62$, resulting in 7032 unique reflections to give 72.1% completeness with a merging R of 0.067 to a resolution limit of 2.8 Å. A molecular-replacement solution of the structure has been obtained using the aspartic proteinase from *Rhizomucor pusillus* as a model. Rigid-body refinement of the model and subsequent refinement using molecular dynamics were performed with *X-PLOR*, leading to a current R factor of 20.1% for 2.8 Å data.

Introduction

The active milk-clotting aspartic proteinase from *Rhizomucor miehei* contains 361 amino acids. It is synthesized as a zymogen containing a 47-amino-acid propeptide from a precursor containing 430 residues (Bech & Foltmann, 1981; Boel *et al.*, 1986; Gray, Hayenga, Cullen, Wilson & Norton, 1986). The structure of this enzyme is of interest to the dairy industry. Mutants may have superior properties for industrial cheese production. So far, it has not been possible to improve on nature with mutants of this protein, but once the structure of this proteinase and detailed features of its active site are known, it may be possible to propose useful mutants.

Experimental

The enzyme is the native form obtained by fermentation of *R. miehei* and is identical to the enzyme which can be obtained commercially as Rennilase (Bech & Foltmann, 1981; Boel *et al.*, 1986). The protein was purified by the following protocol: ion-exchange chromatography on SP-Sephadex C50 in 0.05 *M* citrate buffer, pH 3.2, with 0.1 *M* sodium chloride and elution by a 0.1–0.3 *M* sodium chloride gradient; concentration of the fractions with clotting activity by ultrafiltration using DDS membrane type 800 (M_r cutoff ≃8–10 kDa); exchange of buffer on a Sephadex G50 medium eluted with 0.02 *M* piperazine/HCl buffer, pH 5.3; ion-exchange chromatography on DEAE–Sephadex A50 in 0.02 *M* piperazine/HCl buffer, pH 5.3, and elution by a 0–0.6 *M* sodium chloride gradient; concentration by ultrafiltration using a DDS membrane type 800; gel-filtration on Sephadex G75 superfine using 0.1 *M* acetate buffer, pH 6.0, with 0.1 *M* sodium chloride; concentration by ultrafiltration and continuous washing with distilled water; freeze drying.

The crystals were grown by the hanging-drop vapour-diffusion method. A drop, containing 12 mg ml^{-1} protein, 6% PEG 8000, 30 m*M* citrate buffer, pH 3.6, was equilibrated over a reservoir of 12% PEG 8000, 40 m*M* citrate, pH 3.6. Initially, clusters of small crystals were obtained. Subsequently, a small single crystal from the cluster was placed as a seed in a fresh solution; colourless flattened rod-shaped crystals grew during a period of months. The crystals were mounted in capillary tubes and data were collected on an Enraf–Nonius (Delft, The Netherlands) FAST area-detector X-ray diffractometer using Cu $K\alpha$ radiation from a FR571 rotating anode (λ = 1.5418 Å, 40 kV, 45 mA). The diffraction pattern was indexed using the auto-indexing subroutine of the *MADNES* program (Messerschmidt & Pflugrath, 1987). The program *NRCVAX CREDUC* (Gabe, Le Page, Charland, Lee & White, 1989) and visual checking, confirmed the space group as $P2_12_12_1$ with cell dimensions a = 41.67 (2), b = 51.21 (3) and c = 173.3 (2) Å. This gives a volume V = 370 000 Å^3. If one molecule per asymmetric unit is assumed, each molecule having a molecular mass of 38 701 Da, then the volume-to-mass ratio V_m = 2.4 Å^3 Da^{-1}, and the approximate solvent content is 49%. These values are close to the average values for proteins (Matthews, 1968). The crystal diffracted to at least 2.8 Å resolution. Data were collected on a flattened rod-shaped crystal of dimensions 0.6 × 0.2 × 0.05 mm. The reflection data were evaluated using the program *MADNES* (Messerschmidt & Pflugrath, 1987). Data were collected over the range $0 < h < 14$, $0 < k < 16$ and $0 < l < 62$. Data reduction was carried out using the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994). A total of 16 983 measured reflections were merged to produce 7032 unique reflections (72.1% completeness to 2.8 Å). The R_{merge}, defined as $R_{merge} = \Sigma|I - \langle I\rangle|/\Sigma I$, is 6.7%. A total of 86.3% of all measured reflections had net $I > 2\sigma(I)$. In the resolution shell from 2.99–2.80 Å, 58.4% of the total expected reflections were observed and 68% of these had $I > 2\sigma(I)$.

Endothiapepsin, penicillopepsin, rhizopuspepsin, hexagonal pepsin and chymosin (4APE, 2APP, 2APR, 5PEP, 2CMS) were available from the Protein Data Bank at Brookhaven (PDB, Bernstein *et al.*, 1977). These structures were used as starting models for the molecular-replacement (MR) method. The percentage identity of these models with *R. miehei* protease calculated using the program *FASTA* (Pearson & Lipman, 1988) are as follows: endothiapepsin (28.8%), penicillopepsin (29.4%), rhizopuspepsin (33.5%), pepsin (37.3%), chymosin (30.4%). Initial attempts at a solution using these search models were unsuccessful. However, a recent publication (Aguilar *et al.*, 1993) was very useful in quantifying the best choice of parameters for MR of an aspartic proteinase and made us aware of the existence of the coordinates of the aspartic proteinase from *R. pusillus* (Newman *et al.*, 1993), which also has 361 residues. The program *PC/GENE* (IntelliGenetics, Inc., 1992),

ISSN 0907-4449

using the method of Meyers & Miller (1988) gave an 81.2% identity with *R. miehei* aspartic proteinase for this protein. The coordinates of *R. pusillus* aspartic proteinase (1MPP) were obtained from the PDB (Bernstein *et al.*, 1977), translated to obtain the centre of mass as the origin and used as a search model for MR purposes.

MR calculations were performed using programs from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Structure factors were calculated at 2.8 Å, using atomic coordinates and temperature factors with all original side chains in the PDB coordinates, using *GENSFC* incorporating the *FFT* program for an orthogonal unit cell of the same cell dimensions as *R. miehei* aspartic proteinase and *P*1 symmetry. Normalized structure-factor amplitudes for both the search molecule and the native data were obtained with *ECALC*. Cross-rotation searches were performed with *ALMN*. based on Crowther's *FFT* rotation function (Rossmann, 1972). Maps were calculated using 2.5° steps in each of the Eulerian angles, with Bessel functions of order up to 30. Maps were investigated in the resolution range 10.0–3.0 Å. Searches were conducted over the ranges 10.0–5.0, 10.0–4.0 and 10.0–3.0 Å. A common strongest peak was found in all three ranges at α = 64.0, β = 110.8 and γ = 102.8°. The strongest peak was found for the range 10.0–4.0 Å and integration Patterson radii of 25 Å, with a signal of 6.90σ and 2.68σ difference relative to the first noise peak.

For translation-function calculations the oriented search model was placed in a cell of dimensions identical to the crystal unit cell. A translation function was calculated in space group $P2_12_12_1$ using data from 10.0–3.0 Å by programs *TFSGEN-FFT*, using Crother & Blow's T2 function (Rossmann, 1972) for the translation search, and solution analysis was performed by *MAPSIG*. The strongest peak was found at x = 0.29, y = 0.0 and z = 0.41 (about 13.9σ, and a difference of 2.9σ relative to the first noise peak).

Results and discussion

The translated model was tested and confirmed using the *X-PLOR* (Brünger, 1992; Brünger, Karplus & Petsko, 1989) translation routine. The structure was then refined using *X-PLOR*. For refinement, all non-identical side chains were truncated to either Ala or Gly according to the *R. miehei* aspartic proteinase sequence. The initial R value was 42.8% (8–2.8 Å). After one cycle of *X-PLOR* refinement, using all reflections > 2$\sigma(I)$, the R value was 22.2%. After a session of map fitting using the program *TURBO-FRODO* (Roussel & Cambillau, 1989) installed on a personal IRIS 4D/30TG graphics computer (Silicon Graphics, Inc., USA) and another cycle of *X-PLOR*, using all reflections > 2$\sigma(I)$, the R value was 20.1%. The R value was fairly constant over all reflection shells, and for the resolution range 2.92–2.80 Å was 20.9%. We have used the sequence numbering of Boel *et al.* (1986). There is no electron density for the first five amino-acid residues and they have been left out of the model. The majority of the rest of the main chain of the molecule is in continuous electron density; the active-site residues (Asp38 and Asp237, equivalent to Asp32 and Asp 215 in the pepsin sequence) are well defined and in close proximity. Most of the breaks in the electron density along the main chain involve only one atom and most of these breaks are in the external turns between units of secondary structure or for atoms in glycine units. The poorest regions are the loops from T264 to G267 and D310 to T315. $C\alpha$ coordinates have been deposited with the Protein Data Bank.*

We have observed diffraction spots on an image plate at a synchrotron to 2.3 Å resolution, but we have been unable thus far to collect a full data set to this resolution. Higher resolution data will be collected at a synchrotron facility. With data of higher resolution, the phasing will be extended, the refinement completed and the details of the structure reported.

Support for this work by the Natural Science and Engineering Research Council of Canada in the form of an operating grant (to JWQ) is gratefully acknowledged.

* Atomic coordinates have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1ASI). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0330). A list of deposited data is given at the end of this issue.

References

AGUILAR, C. F., NEWMAN, M. P., SANZ APARICIO, J., COOPER, J. B., TICKLE, I. J. & BLUNDELL, T. L. (1993). *Acta Cryst.* **A49**, 306–315.

BECH, A.-M. & FOLTMANN, B. (1981). *Neth. Milk Dairy J.* **35**, 275–280.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

BOEL, E., RECH, A.-M., RANDRUP, K., DRAEGER, B., FIIL, N. P. & FOLTMANN, B. (1986). *Proteins Struct. Funct. Genet.* **1**, 363–369.

BRÜNGER, A. T. (1992). *X-PLOR Manual.* Version 3.0. Yale Univ., New Haven, CT, USA.

BRÜNGER, A. T., KARPLUS, M. & PETSKO, G. A. (1989). *Acta Cryst.* **A45**, 50–61.

COLLABORATIVE COMPUTATIONAL PROJECT, NUMBER 4. (1994). *Acta Cryst.* **D50**, 760–763.

GABE, E. J, LE PAGE, Y., CHARLAND, J.-P., LEE, F. L. & WHITE, P. S. (1989). *J. Appl. Cryst.* **22**, 384–387.

GRAY, G. L., HAYENGA, K., CULLEN, D., WILSON, L. J. & NORTON, S. (1986). *Gene*, **48**, 41–53.

IntelliGenetics, Inc. (1992). *PC/GENE.* Release 6.7. IntelliGenetics, Inc., 700 East El Camino Real, Mountain View, CA, USA.

MATTHEWS, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

MESSERSCHMIDT, A. & PFLUGRATH, J. W. (1987). *J. Appl. Cryst.* **20**, 306–315.

MEYERS, E.W. & MILLER, W. (1988). *Comput. Appl. Biosci.* **4**, 11–17.

NEWMAN, M. P., WATSON, F., ROYCHOWDHURY, P., JONES, H., BADASSO, M., CLEASBY, A., WOOD, S. P., TICKLE, I. J. & BLUNDELL, T. L. (1993). *J. Mol. Biol.* **230**, 260–283.

PEARSON, W. R. & LIPMAN, D. J. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 2444–2448.

ROSSMANN, M. G. (1972). Editor. *The Molecular Replacement Method*, pp. 1–276. New York: Gordon and Breach.

ROUSSEL, A. & CAMBILLAU, C. (1989). *TURBO-FRODO.* In *Silicon Graphics Geometry Partners Directory*, pp. 77–78. Mountain View, CA: Silicon Graphics.