CRYSTAL STRUCTURE ANALYSIS AND REFINEMENT OF TWO VARIANTS OF TRIGONAL TRYPSINOGEN

Trigonal trypsin and PEG (polyethylene glycol) trypsinogen and their comparison with orthorhombic trypsin and trigonal trypsinogen

Wolfram BODE and Robert HUBER

Max-Planck-Institut für Biochemie und Physikalisch-Chemisches Institut der Technischen Universität, D-8033 Martinsried bei München, FRG

Received 14 April 1978

1. Introduction

A large number of crystal structures in the system bovine trypsin, trypsinogen and pancreatic trypsin inhibitor (PTI) has been studied, crystallographically refined and carefully compared (reviewed [2–4]). An interesting finding was, that trypsinogen, the inactive precursor of trypsin is very similar in structure to trypsin in about 85% of the molecule (mean deviation of internal main chain atoms: 0.2 Å), but the rest is entirely different [5,6]. This part (the activation domain) shows no significant electron density in the electron density map of trypsinogen indicating statistical or thermal disorder. This observation may be explained either by at least three to four different conformations of the activation domain or by a temperature factor of more than about 50 Å^2 (corresponding to an r.m.s. displacement of 0.8 Å) as deduced from significance considerations. The ordered part has an average temperature factor around 10 Å^2. Such considerations are based on the fact that density peaks of about one quarter the height of an oxygen atom are of doubtful significance. Such reduction in peak height may be a consequence of statistical disorder or thermal mobility. The real situation in trypsinogen may be a mixture of both types of disorder. Other techniques are required for clarification. In trypsin the activation domain is well defined and rigid. It forms part of the substrate binding area. Trypsin had been analyzed in an orthorhombic modification (OT) (1.5 M ammonium sulphate, pH 5–8 [2,7]), while trypsinogen crystallized in a trigonal space group (2.4 M magnesium sulphate, pH 6.9) (TG) [5,6]. In order to exclude the possibility that the structural differences observed between trypsin and trypsinogen were due to lattice effects we analyzed the trigonal form of trypsin (TT), which is isomorphous to trypsinogen. Trigonal trypsinogen crystals in polyethylene glycol (TGPEG) were analyzed in order to see any influence of the solvent environment on the crystal structure. Trigonal trypsin crystals were obtained from a solution containing 5% trypsin in 0.7 M ammonium sulphate, pH 6.0, by equilibration through vapour diffusion with a 1.5 M ammonium sulphate solution, pH 6.0. The crystallization solutions were seeded with trigonal trypsinogen crystallites soaked in 1.6 M ammonium sulphate. These trypsin crystals were soaked in a solution containing 2.4 M ammonium sulphate, pH 6.9, in the absence of benzamidine. 10 mM ytterbium sulphate was included to detect possible binding of rare earth ions.

Abbreviations: TT, trigonal trypsin; TG, trigonal trypsinogen; TGPEG, trigonal trypsinogen in polyethylene glycol (M_r 2000); OT, orthorhombic trypsin; F_o, observed structure factor amplitudes; F_c, calculated structure factor amplitudes; r.m.s., root mean square

The amino acid sequence numbers refer to the chymotrypsinogen enumeration in [1]

2. Materials and methods

Trigonal trypsin crystals are isomorphous with trigonal trypsinogen. The space group is $P3_121$ and lattice constants are, in trypsin: a, 54.9 Å; c, 109.4 Å; and in trypsinogen: a, 55.1 Å; c, 109.4 Å, respectively. There are very substantial intensity differences between the two crystal forms. Polyethylene glycol trypsinogen crystals were prepared by exchanging the mother liquor of trigonal trypsinogen crystals (2.4 M magnesium sulphate, pH 6.9) for 30% polyethylene glycol (M_r 2000) buffered with 0.07 M Tris/HCl, pH 7.6. 0.1 M $CaCl_2$ was added to saturate the Ca site which was found incompletely occupied in magnesium sulphate trypsinogen crystals [5,6]. The lattice parameters of this form were identical to trypsinogen.

Intensity data of trigonal trypsin crystals were collected from one crystal using rotation photography as described for trypsinogen [5,8]. About 45 000 intensity measurements were merged to yield about 15 000 unique reflexions to 1.8 Å resolution. R_{Merge}, defined as $(\Sigma (<I>_h - I_{hj})^2 / \Sigma(I_{hj})^2)^{1/2}$ was after absorption correction 0.089 [9]. $<I>_h$ is the average intensity of a reflexion h, I_{hj} is the j^{th} measurement of reflexion h. The summation is over all reflexions. Intensity data of polyethylene glycol trypsinogen crystals were collected from one crystal using the same procedures. About 35 000 intensity data were merged to 12 900 reflexions to 1.8 Å resolution. R_{Merge} was 0.065 after absorption correction.

A first difference Fourier map with coefficients $|F_{o,TT}| - |F_{o,TG}|$ and phases calculated from the trypsinogen model had shown high density at the positions of the residues of the activation domain known from the structure of orthorhombic trypsin (see fig.2 in [6]). In addition, it showed a very high peak (5.5 e/Å^3) at the Ca site indicating binding of ytterbium. The starting model for trigonal trypsin refinement was therefore obtained by fitting the internal main chain atomic coordinates of orthorhombic trypsin optimally to the equivalent coordinates of trigonal trypsinogen. The resulting transformation matrix was applied to all atoms of orthorhombic trypsin. The R-value of this model was 0.321. 4 cycles of refinement consisting of phase calculation, Fourier synthesis (with coefficients $2|F_o| - |F_c|$ and $3|F_o| - 2|F_c|$) and real space refinement [10–12] reduced this to 0.262. A difference Fourier map indicated only a few side chain rearrangements. Further cycles intermitted by difference Fourier map inspection brought R to 0.234. 14 499 reflexions were included between 7.8 and 1.8 Å resolutions. Reflexions with a ratio $2||F_o| - |F_c|| / ||F_o| + |F_c||$ above 2.0 were excluded. Individual radii of the polypeptide chain were refined in the last stages but applying a damping scheme. For the localized solvent molecules and the Yb^{3+} occupancies and radii were refined and applied for the structure factor calculation.

Polyethylene glycol trypsinogen crystals were analyzed by a difference Fourier map versus trigonal trypsinogen. This map showed a high peak at the Ca site (1.0 e/Å^3) and some shifts of the liganding groups only. Other density peaks were smaller than 0.3 e/Å^3 indicating virtual identy of the two species. In particular the activation domain showed only residual density below 0.2 e/Å^3. Two automated refinement cycles of the segment including the Ca site and the Ca^{2+} were performed to define the Ca occupancy and site geometry objectively.

3. Results and discussion

3.1. *Trigonal trypsin*

Figure 1 shows several sections through the featureless final difference Fourier map near Asp 194 and Cys 191, which were selected as trypsin and trypsinogen differ substantially here: Asp 194 is entirely differently placed and the activation domain which includes residues A184 to 193 is flexible in trypsinogen, but fixed in trypsin [3,4,6]. Figure 1 is comparable to fig.2 in [6] which represented a difference Fourier map calculated with coefficients $|F_{o,TT}| - |F_{o,TG}|$ and phases of trypsinogen. In contrast, this had shown high residual density at Asp 194 and the segments of the activation domain, as expected.

Orthorhombic trypsin, trigonal trypsin and trigonal trypsinogen were optimally superimposed using internal main chain segments. The mean deviations of main chain atoms (C^{α},C,N,O) for internal segments were: TG–TT, 0.20 Å; TG–OT, 0.18 Å; TT–OT, 0.16 Å; and for all main chain atoms, but excluding residues of the activation domain, not defined in TG: TG–TT, 0.23 Å; TG–OT, 0.22 Å; TT–OT, 0.23 Å. The somewhat larger deviation of external segments compared

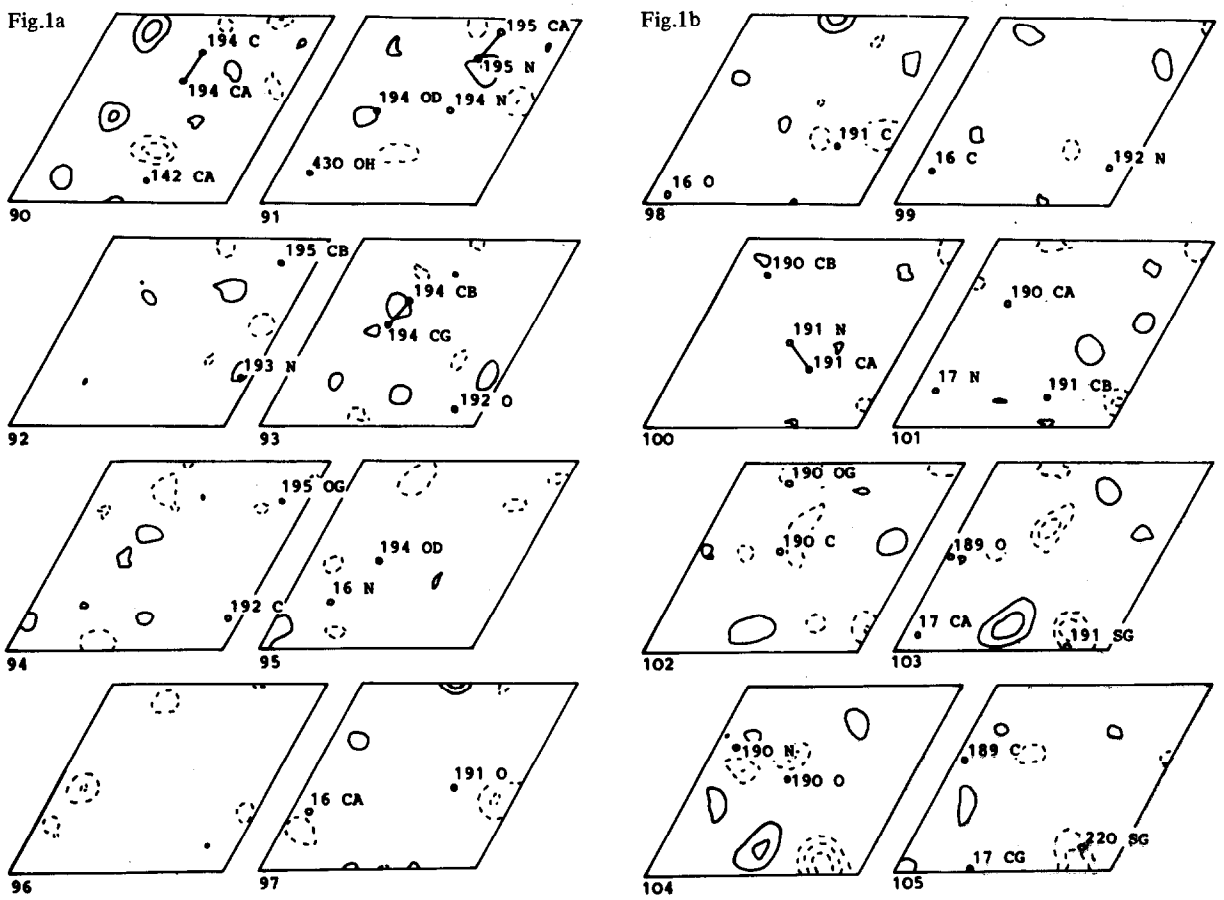

Fig.1a,b. Consecutive sections from 90 to 105 through the final difference Fourier map ($|F_O| - |F_C|$) around Cys 191 and Asp 194. Contours from ± 0.2 e/Å^3 in steps of ± 0.1 e/Å^3.

to internal ones is obvious. The activation domain in trigonal trypsin is insignificantly different from orthorhombic trypsin with a mean deviation of 0.25 Å..

A comparison of the three structures should provide information upon structural changes due to the activation process alone, not influenced by lattice forces (TG and TT) and upon the effect of lattice forces on identical molecules (TT and OT). While this comparison shows a few obviously significant C^{α} deviations, we found no way to correlate these in a reasonable way with the two processes. Isomorphous trigonal trypsin and trypsinogen differ mainly in the activation domain. This implies that the activation process involving a structuring of the activation domain, does not affect the lattice formation. The activation domain is involved in a few lattice contacts in trigonal trypsin. Some atoms of the loop 149–151 touch residue 86 and there is one intermolecular hydrogen bond between 151 N and 86 O. This contact occurs also in orthorhombic trypsin.

The Yb occupation factor in trigonal trypsin is 0.62. Its position deviates from the Ca position in PEG trypsinogen by 0.35 Å, while Ca in PEG trypsinogen deviates from Ca in orthorhombic trypsin (after application of the corresponding transformation matrix) by 0.17 Å. The coordinating residues and atoms are the same. The conformation of the coordinating side chains is very similar. The inner coordination sphere distances are summarized in table 1, and demonstrate the close similarity. The positional

Table 1
Coordination sphere distances to central ligand (Ca or Yb)

Ligand	TG (Ca)	OT (Ca)	TT (Yb)	TGPEG (Ca)
Glu 70 $0^{\epsilon 1}$	2.1	2.4	2.6	2.3
Asn 75 0	2.3	2.4	2.1	2.4
Val 75 0	2.3	2.5	2.4	2.4
Glu 80	2.0	2.4	2.1	2.2
H_2O 711	2.2	2.6	3.0	2.7
H_2O 714	2.1	2.6	2.6	2.4

difference of Ca and Yb is significant and may reflect the difference in charge.

The affinity of trypsin for Yb in solution is $2–4 \times 10^3$ M^{-1} at pH 6.3, 4°C [13]. Under the soaking conditions (10 mM Yb^{3+}) full occupation should be expected. Orthorhombic trypsin had a completely occupied Ca site, although the crystals had been kept in 2.4 M ammonium sulphate where the Ca concentration is 10^{-4} M according to the solubility product. The affinity constant is 3.7×10^3 M^{-1} [14] or 6.2×10^3 M^{-1} [13], 11×10^3 M^{-1} [15] at pH 6.1–6.3, so that only 30–70% occupation would be expected. This anomaly might be caused by a kinetic difficulty of dissociation or by a tightening of binding in the crystalline state. Analogous effects might be responsible for the decreased occupation of Yb in trigonal trypsin.

3.2. *Polyethylene glycol trypsinogen*

Ca, which was found weakly occupied in trypsinogen is fully occupied in polyethylene glycol trypsinogen with 0.1 M $CaCl_2$. While the side chains coordinating Ca in trypsinogen were not well defined, they are well defined in polyethylene glycol trypsinogen. Polyethylene glycol trypsinogen is insignificantly different from orthorhombic trypsin in Ca coordination (table 1). Except for the Ca site trypsinogen and polyethylene glycol trypsinogen are extremely similar. The difference Fourier map (with coefficients $|F_{o,TG}| - |F_{o,TGPEG}|$) shows no significant residual densities anywhere else. This indicates independence of the crystal structure of trypsinogen on the solvent medium, although the variations in salt concentration and dielectric constant are considerable between 2.4 M $MgSO_4$ and 30% polyethylene glycol. We want to emphasize in particular that the activation domain is disordered in both environments. This finding is of relevance as environmental differences appear not to be the cause of the considerable differences in interpreting the trypsinogen model between [6] and [16]. In [16] trypsinogen crystals had been crystallized and analyzed in an aethanol water mixture.

An identical Ca site in trypsin and trypsinogen has now been defined in two different crystal structures. This Ca can be replaced by Yb. In orthorhombic trypsin it can be removed by lowering the pH [2]. We believe it is not justified to cast reasonable doubts on this assignment ([17] but see also [18,19]) instead, one should use it to calibrate spectroscopic distance determinations between the Ca site and the substrate binding site.

References

[1] Hartley, B. S. and Shotton, D. M. (1971) in: The Enzymes (Boyer, P. S. ed) vol. 3, pp. 323–373, Academic Press, New York, London.
[2] Bode, W., Schwager, P. and Huber, R. (1976) in: Proteolysis and Physiological Regulation, Miami Winter Symposia, vol. 11, pp. 43–76, Academic Press, New York, San Francisco, London.
[3] Huber, R. and Bode, W. (1977) in: NMR in Biology (Dwek, R. A., et al. eds) pp. 1–31, Academic Press, New York, San Francisco, London.
[4] Huber, R. and Bode, W. (1978) Acc. Chem. Res. 11, 114–122.
[5] Bode, W., Fehlhammer, H. and Huber, R. (1976) J. Mol. Biol. 106, 325–335.
[6] Fehlhammer, H., Bode, W. and Huber, R. (1977) J. Mol. Biol. 111, 415–438.
[7] Bode, W. and Schwager, P. (1975) J. Mol. Biol. 98, 693–717.
[8] Schwager, P., Bartels, K. and Jones, A. (1975) J. Appl. Cryst. (1975) 8, 275–280.

[9] Schwager, P., Bartels, K. and Huber, R. (1973) Acta Cryst. A29, 291–295.
[10] Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J. and Steigemann, W. (1974) J. Mol. Biol. 89, 73–101.
[11] Deisenhofer, J. and Steigemann, W. (1975) Acta Cryst. B31, 238–250.
[12] Diamond, R. (1971) J. Mol. Biol. 82, 371–391.
[13] Epstein, M., Levitzki, A. and Reuben, J. (1974) Biochemistry 13, 1777–1782.
[14] Delaage, M. and Lazdunski, M. (1967) Biochem. Biophys. Res. Commun. 28, 390–394.
[15] Epstein, M. (1976) Ph. D. Thesis, Weizmann Institute of Science, Rehovot.
[16] Kossiakoff, A. A., Chambers, J. L., Kay, L. M. and Stroud, R. M. (1977) Biochemistry 16, 654–664.
[17] Darnall, D. W., Abbot, F., Gomez, J. and Birnbaum, E. R. (1976) Biochemistry 15, 5017–5023.
[18] Epstein, R. and Reuben, J. (1977) Biochim. Biophys. Acta 481, 164–170.
[19] Dimicoli, J. L. and Bieth, J. (1977) Biochemistry 10, 5532–5537.