# THE SINGLE CALCIUM-BINDING SITE OF CRYSTALLINE BOVINE β-TRYPSIN

W. BODE and P. SCHWAGER

*Max-Planck-Institut für Biochemie, Abteilung Strukturforschung II und Physikalisch-Chemisches Institut der Technischen Universität, München, 8033 Martinsried bei München, West Germany*

Received 16 May 1975

## 1. Introduction

It is well known that calcium ions accelerate the conversion of trypsinogen to trypsin, prevent the formation of inert proteins during activation and protect trypsin against autolysis. It has been shown that trypsinogen has two different binding sites for calcium ions while trypsin has only one [3,4]. The low affinity calcium binding site ($pK_{Ca} \approx 1.8$ [3]) is found only in the zymogen and has been assigned to the two aspartyl residues 13 and 14 neighbouring the important Lys-15–Ileu-16 bond that is split during activation [4,5]. The second site, with a much higher affintiy for calcium ($pK_{Ca} \approx 3.4$ [3,4]) and related ions, is common to both trypsinogen and trypsin and is presumably also shared by the chymotrypsins [4]. Abita and Lazdunski [4] have shown that chemical modification of the carboxylic side chains of trypsinogen and chymotrypsinogen destroys this binding site. Epstein et al. [6] determined by titration that the calcium ligands of porcine trypsin should have p$K$ values of about 5.2. They therefore concluded that some glutamic and aspartic carboxylic side chains are presumably involved in this calcium binding. In order to determine this site crystallographically, Stroud et al. [7] have tried without sucess to diffuse neodynium ions into trypsin crystals. In their recently published 2.7 Å resolution structure of inhibited bovine trypsin, Stroud et al. could not present any evidence for the calcium-binding site [8,9]. During the course of the constrained crystallographic refinement of the crystal structure of benzamidine-inhibited bovine β-trypsin, we have detected a single site in the trypsin molecule which by all available criteria is occupied by a strongly bound calcium ion. In this paper we will only describe the trypsin structure around the calciumbinding site and discuss possible consequences of this finding on the bovine trypsin sequence data. A detailed description of the biochemical and physical methods applied and of the refined protein structure will be presented in a subsequent paper [10].

Amino acids of trypsin are identified by the residue number of the homologous amino acid in chymotrypsin [1].

The nomenclature recommended by IUPAC-IUB (1970) is used in this paper [2].

## 2. Materials and methods

Orthorhombic crystals of benzamidine-inhibited trypsin, space group $P2_12_12_1$ with cell constants a = 54.89 Å, b = 58.52 Å and c = 67.63 Å were grown from solutions of bovine β-trypsin, ammonium sulfate, benzamidine and small amounts of calcium ions using a vapour diffusion technique. Intensity data to about 1.8 Å resolution were collected photographically [11] from one single crystal, immersed in 2.4 M ammonium sulfate solution and adjusted with small amounts of phosphate buffer to pH 7.0. The merged data set contained about 83% of the expected unique reflexions to 1.8 Å resolution. Given the refined atomic coordinates as found in the crystal structure of bovine trypsin-pancreatic trypsin inhibitor complex [12] the trypsin molecules were located within the trypsin crystals, applying Patterson search techniques [13,14]. Starting with these orientation and translation parameters, the molecular structure of bovine trypsin was fitted to the observed intensity data by a constrained crystallo-

graphic refinement procedure. The technique has already been applied successfully to the crystal structures of the pancreactic trypsin inhibitor [15], the trypsin–pancreatic trypsin inhibitor complex [12] and the Bence-Jones protein Rei [16]. The refinement procedure consists of cycles of structure factor calculation, Fourier-synthesis and real space refinement [15,17]. After several automatic cycles, difference Fourier - syntheses are calculated to detect and to correct gross errors in the model. The refinement of the trypsin structure has progressed to a crystallographic R value of 0.23 ($R = \Sigma||F_o|-|F_c||/\Sigma|F_o|$) summing over 16600 unique reflexions from 6.8 to 1.8 Å resolution ($|F_o|$ and $|F_c|$ are observed and calculated structure factor amplitudes respectively). The accuracy of the atomic positions is estimated to be 0.1 Å for the trypsin atoms and better for the calcium [10].

The calcium content of the β-trypsin used for crystallization and of the trypsin crystals has been determined using an atomic absorption spectrophotometer, kindly provided by Beckman Instr., München.

## 3. Results and discussion

After several rounds of refinement cycles, many positive residual density peaks were found in the difference Fourier map, all of which could, with one exception, be satisfactorily interpreted as external and internal bound solvent molecules. After inclusion of oxygen atoms in the subsequent calculation, a large positive density peak remained in the difference Fourier map only at this particular site. The height of this residual peak corresponds to an atom with 16 to 20 electrons assuming 100% occupancy. The location of the side is in the external loop containing residues 68–80 (fig.1). Considering the ions present in the crystallization solutions, we replaced the original water molecule at this site with a calcium ion and found satisfactory agreement with the electron density map (fig.2). The calcium content of the trypsin crystals used was determined by atomic absorption spectroscopy. The result was almost exactly one calcium ion per trypsin molecule. The calcium and the surrouding water molecules were also observed in the crystal structure of the trypsin–trypsin inhibitor complex [12]. It was, however, mistaken for a water molecule, presumably due to incomplete occupancy (Huber, private communication). Complex crystals are kept in phosphate buffer where the solubility of calcium ions is about 10 times lower than in ammonium sulfate used in this work.

The final difference map near the calcium binding loop is without interpretable features. The highest residual density is on the external side chain of Asn-79 indicating some mobility of this group. This density is 0.4 e/Å$^3$ compared to about 2.0 e/Å$^3$ of a representative carbonyl oxygen atom and to 4.9 e/Å$^3$ for the calcium site in the Fourier map. The highest residual density at main or internal side chain atoms is 0.25 e/Å$^3$.

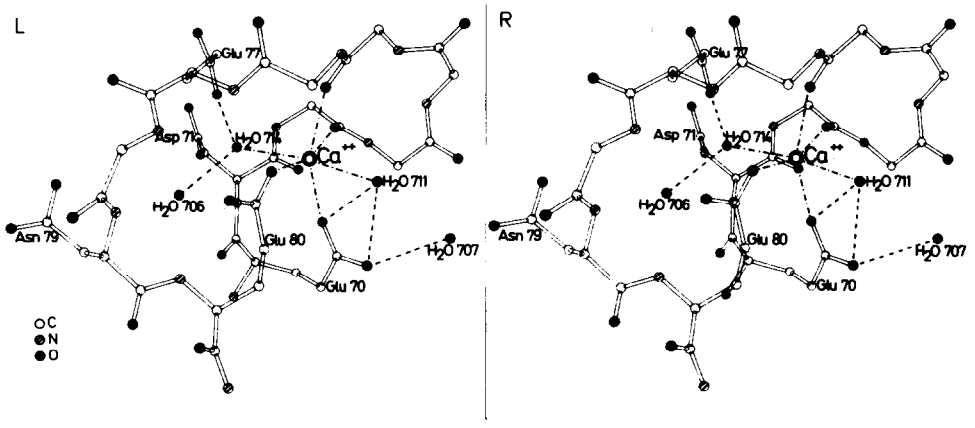


Fig.1. Stereo pair of the calcium-binding loop including the calcium and internal water molecules.

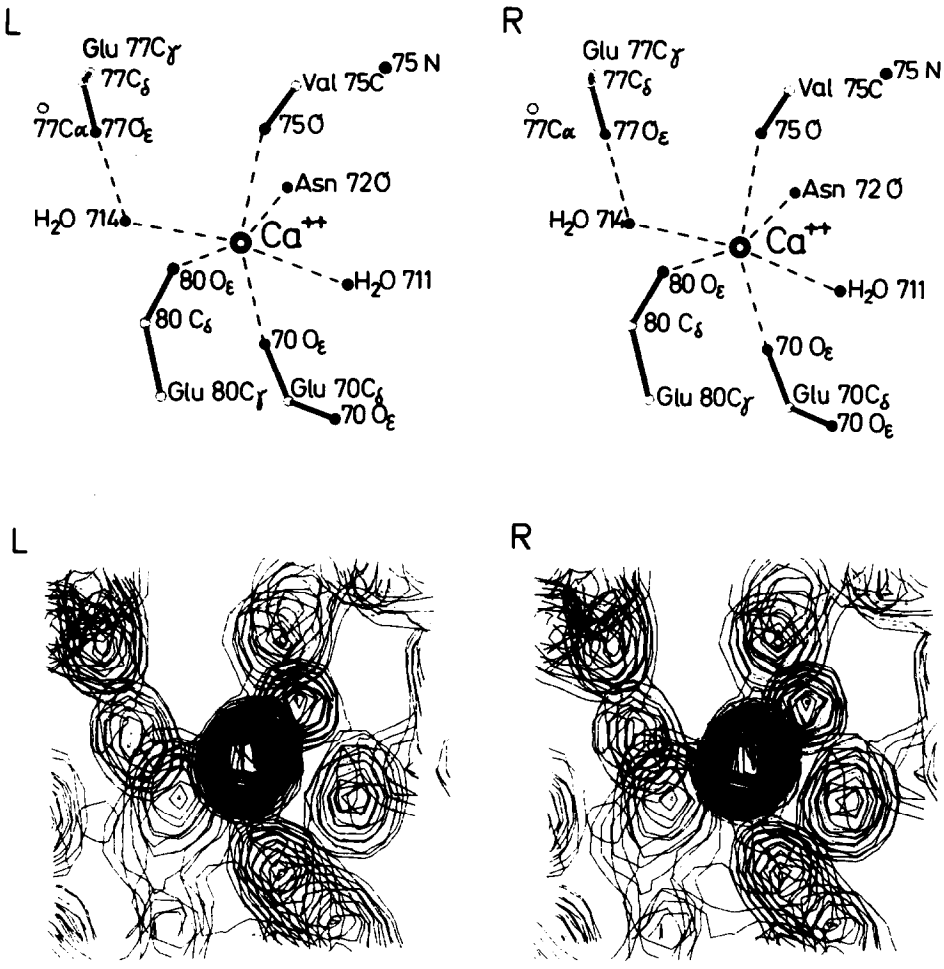


Fig.2. Stereo pair of the calcium co-ordination sphere. (a) Refined model with labelled atoms. (b) Corresponding refined electron density (1.8 A resolution; Contour levels in steps of 0.3e/A$^3$, starting from 0.3e/A$^3$).

As can be seen in fig.2, the calcium ion is complexed by six different ligands positioned at the edges of an almost regular octahedron around the calcium and each about 2.5 Å distant from it (table 1). Two vicinal ligands are represented by the carbonyl oxygens of Asn-72 and Val-75. Two further ligands opposing each other in the octahedron are water molecules labelled 711 and 714 (see fig.2a). $H_2O$ 714 is moreover close (2.3 Å) to one of the two oxygen atoms of the carboxylic group of Glu-77, presumably forming with it a very stable hydrogen bond. The second water ligand, $H_2O$ 711, is 3.1 Å from the side chain of residue 70. The residual two coordinating ligands are

Table 1
Calcium ion co-ordination sphere

| Ligand | | Distance from calcium (Å) |
|---|---|---|
| Gl(n) | 70 NOε | 2.4 |
| Asn | 72 O | 2.4 |
| Val | 75 O | 2.5 |
| Gl(n) | 80 NOε | 2.4 |
| $H_2O$ | 711 O | 2.6 |
| $H_2O$ | 714 O | 2.6 |

Table 2
Optimally aligned amino acid sequences (residues 69 to 80) of different sources

| Species | Residue number 69 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 | 80 | 81 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Chymotrypsinogen A, Cow | Gly | Glu | Phe | Asp | Gln | Gly | Ser | Ser | Ser | Glu | Lys | Ile | Gln |
| Chymotrypsinogen B, Cow | Gly | Glu | Phe | Asp | Gln | Gly | Leu | Glu | Thr | Glu | Asp | Thr | Gln |
| Elastase, Pig | Gly | Glu | His | Asn | Leu | Asn | Gln | Asn | Asn | Gly | Thr | Glu | Gln |
| Trypsinogen, Pig | Gly | Glu | His | Asn | Ile | Asp | Val | Leu | Glu | Glu | Asn | Glu | Gln |
| Trypsinogen, Dogfish | Gly | Glu | His | Asp | Ile | Ser | Ala | Asn | Glu | Gly | Asp | Glu | Thr |
| Trypsinogen, Lungfish | Gly | Glu | His | Asn | Ile | Glu | Val | Asn | Glu | ? | ? | ? | ? |
| Trypsinogen, Cow | Gly | Gln | Asp | Asn | Ile | Asn | Val | Val | Glu | Gly | Asn | Gln | Gln |

the side chains of residues 70 and 80, both being glutamines according to the sequence data for bovine trypsinogen commonly accepted at present [1,18–20].

Based on this primary structure (see table 2), the calcium ion makes no direct contact to an anionic group, so that the positive charge on the calcium ion could only be partly dissippated through the hydrogen bonding network via $H_2O$ 714 to the carboxylate group of Glu-77 and partly compensated by the negative dipole charges of both the carbonyl and water oxygen ligands. But the titration and modification data mentioned above indicate the direct involvement of carboxylic groups in the calcium binding of trypsin. Contrary to the d-transition metal ions, the alkaline earth metal ions normally form complexes primarily with anionic oxygen donor ligands [21]. This is for example the case in the only two calcium-binding proteins that are solved at atomic resolution, viz. thermolysin [22,23] and calciumbinding parvalbumin [24] where the four and two calcium-binding sites resp. are co-ordinated by one to three carboxylic groups.

This raises some doubt on the identification of residues 70 and/or 80 as carboxamides. It is interesting to reinspect the sequence work on bovine trypsinogen. Originally in 1964 Walsh and Neurath [25] had identified both residues 70 and 80 as glutamic acids. In 1966 Hartley et al. [26] agreed with this assignment in a correcting note. At the same time, Mikeš et al. [27] publishing their complete bovine trypsinogen sequence data, assigned residues 70 and 80 to be glutamic amides due to careful electrophoretic determinations. This assignment has obviously been accepted by the other scientists working in this field. In table 2 the amino acid sequences around residues 70 to 80 of some trypsinogen, elastase and chymotrypsinogen species are presented optimally aligned as described by De Haën et al. [20] where the original literature may be found.

As can be seen, residue 70, one of the co-ordination ligands, is in all cases a glutamic acid with the exception of bovine trypsinogen. At position 80 the bovine sequence differs from the sequences to two other known species in being an amide. Both chymotrypsinogens sequences are congruent at residue 70 with the trypsinogens, but differ strongly at position 77 and 80. This is reflected in the large differences in tertiary structure between α-chymotrypsin and trypsin (huber et al. [12] and table 2). Neglecting these large differences, Stroud et al. [7] and Abita et al. [4] proposed the residues Asp-71, Glu-77 and Asp-153 as possible calcium-binding ligands by looking for equally positioned glutamic and aspartic residues in the published sequences of bovine chymotripsinogen and trypsinogen. We see now that Asp-71 and Asp-153 are oriented towards the solution in bovine trypsin. This fits well with the fact that in the porcine trypsinogen sequence, which has been published recently by Hermodson et al. [28], Asp-71 and Asp-153 of the bovine species are replaced by the nonanionic groups His and Ser resp.. According to the binding study of Epstein et al. [6], however, porcine trypsinogen binds calcium as well as the bovine species. We therefore suggest, that the character of residues 70 and 80 in the bovine sequence should be reinvestigated and perhaps revised.

Our data suggest that the calcium, being more than 21 Å from the active site Ser-195 Oγ, does not directly influence the proteolytic activity [29]. Knowing the calcium position, however, it is still difficult to ex-

plain the well known stabilizing effect of the calcium ion on trypsin, namely inhibiting its autodigestion and increasing its thermal stability, especially in the case of the single split α-trypsin [29]. Nevertheless, it seems reasonable to assume that the calcium makes the loop 68–80 more rigid. It prevents the interaction of the charged carboxyl groups with the solute and may stabilize other neighbouring regions including the loops containing the α-splitting site (Lys-145–Ser-146) and another preferred splitting site between Arg-117 and Val-118. Epstein et al. [6], studying the fluorescence enhancement of terbium during binding to porcine trypsin, come to the conclusion that tryptophan residue(s) should be within a very short distance from the calcium site. A possible candidate would be Trp-141 with an average distance from the calcium of about 8.5 Å.

## Acknowledgements

The financial assistance of SFB 51 is acknowledged. We thank Dr R. Huber for frequent and helpful discussions, and H. Fehlhammer, Dr P. M. Colman, R. Reinhardt, Dr J. Deisenhofer and Dr W. Steigemann for valuable support. We thank Mrs Dulnowitch and Beckman Instr., München, for the performance of the calcium determination by atomic absorption spectroscopy.

## References

[1] Hartley, B. S. and Shotton, D. M. (1971) The Enzymes (Boyer, P. D., ed.), Vol.3, p. 323–373, Academic Press, New York and London.
[2] IUPAC–IUB Commission on Biochemical Nomenclature (1970) Biochemistry 9, 3471–3479.
[3] Delaage, M. and Lazdunski, M. (1967) Biochem. Biophys. Res. Commun. 28, 390–394.
[4] Abita, J. P. and Lazdunski, M. (1969) Biochem. Biophys. Res. Commun. 35, 707–712.
[5] Radhakrishnan, T. M., Walsh, K. A. and Neurath, H. (1969) Biochemistry 8, 4020–4027.
[6] Epstein, M., Levitzki, A. and Reuben, J. (1974) Biochemistry 13, 1777–1782.
[7] Stroud, R. M., Kay, L. M. and Dickerson, R. E. (1971) Cold Spring Harbor Symp. Quant. Biol. 36, 125–140
[8] Stroud, R. M., Kay, L. M. and Dickerson, R. E. (1974) J. Mol. Biol. 83, 185–208.
[9] Krieger, M., Kay, L. M. and Stroud, R. M. (1974) J. Mol. Biol. 83, 209–230.
[10] Bode, W., Fehlhammer, H. and Schwager, P. (1975), manuscript in preparation.
[11] Schwager, P., Bartels, K. S. and Jones, A. T. (1975) J. Appl. Cryst. 8, 275–280.
[12] Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J. and Steigemann, W. (1974) J. Mol. Biol. 89, 73–101.
[13] Crowther, N. A. and Blow, D. M. (1967) Acta Crystallogr. 23, 544–548.
[14] Huber, R. (1969) Crystallographic Computing Proc. (1969) (F. R. Ahmed, ed.) Munsgaard.
[15] Deisenhofer, J. and Steigemann, W. (1975) Acta Cryst. B 31, 238–250.
[16] Huber, R. and Steigemann, W. (1974) FEBS Lett. 48, 235–237.
[17] Diamond, R. (1974) J. Mol. Biol. 82, 371–391.
[18] Dayhoff, M. O. (1972) Atlas of Protein Sequence and Structure, Vol.5, Nation. Biomed. Res. Found. Washington, D. C.
[19] Dayhoff, M. O. (1973) Atlas of Protein Sequence and Structure, Vol.5, supplement I, Nation. Biomed. Res. Found., Washington, D. C.
[20] De Haën, C., Neurath, H. and Teller, D. C. (1975) J. Mol. Biol. 92, 225–259.
[21] Darnall, D. W. and Birnbaum, E. R. (1970) J. Biol. Chem. 245, 6484–6488.
[22] Colman, P. M., Jansonius, J. N. and Matthews, B. W. (1972) J. Mol. Biol. 70, 701–724.
[23] Matthews, B. W. and Weaver, L. H. (1974) Biochemistry 13, 1719–1725.
[24] Moews, P. C. and Kretsinger, R. H. (1975) J. Mol. Biol. 91, 201–228.
[25] Walsh, K. A. and Neurath, H. (1964) Proc. Natl. Acad. Sci. U.S. 52, 884–889.
[26] Hartley, B. S. and Kauffman, D. L. (1966) Biochem. J. 101, 229–231.
[27] Mikeš, O., Holcyšovský, V., Tomášek, V. and Šorm, F. (1966) Biochem. Biophys. Res. Commun 24, 346–352.
[28] Hermodson, M. A., Ericsson, L. H., Neurath, H. and Walsh, K. A. (1973) Biochemistry 12, 3146–3153.
[29] Sipos, T. and Merkel, J. R. (1970) Biochemistry 9, 2766–2775.