Acta Cryst. (1982). B38, 1462–1472

On the Disordered Activation Domain in Trypsinogen: Chemical Labelling and Low-Temperature Crystallography

BY JOACHIM WALTER, WOLFGANG STEIGEMANN AND TEJ PAL SINGH*

Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, Federal Republic of Germany

HANS BARTUNIK

European Molecular Biology Laboratory (EMBL), Notkestrasse 85, D-2000 Hamburg 52, Federal Republic of Germany

AND WOLFRAM BODE AND ROBERT HUBER

Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, Federal Republic of Germany

(*Received* 26 *August* 1981; *accepted* 21 *January* 1982)

Abstract

In trypsinogen a substantial part of the molecule, the activation domain, is disordered. The disulphide Cys 191–Cys 220 in the flexible activation domain of trypsinogen can be selectively reduced and mercurated. The crystal structure of such modified trypsinogen has been analysed in its free form and complexed with PTI and the dipeptide Ile-Val. In trypsinogen the Hg atom which is covalently inserted between the two S atoms is crystallographically undetectable. In the trypsinogen PTI complex the Hg atom is well defined. Crystallographic refinement of the complex has been performed at 2·1 Å resolution. The R value is 0·20. These observations indicate thermal or static disorder of the activation domain characterized by a B factor of more than about 200 Å^2. Trypsinogen crystals were crystallographically analysed at 173 and 103 K using synchrotron radiation, the crystal structures refined and the molecular structures and B factors compared with the room-temperature data. The overall B factor is reduced by 4·5 Å^2 at 173 K and by 4·6 Å^2 at 103 K. A number of additional ordered water and methanol molecules are observed. The order of the activation domain is not increased detectably at low temperature except at the N-terminal region. The residues at the boundaries of the flexible segments show no substantial temperature-dependent decrease in B factor, except near the N-terminus. These observations are consistent with predominantly static disorder of the activation domain.

1. Introduction

A large number of crystal structures in the system trypsin, trypsinogen and pancreatic trypsin inhibitor [PTI (Trasylol)]† have been studied, crystallographically refined and compared (for a review see Huber & Bode, 1978). These studies have indicated a close similarity between trypsinogen and trypsin except in the activation domain which is flexible in the inactive proenzyme (Fehlhammer, Bode & Huber, 1977) and ordered in the active enzyme (Bode & Schwager, 1975). The activation domain contains four segments tightly interdigitating in trypsin: N-terminus to Gly 19, Gly 142 to Pro 152, Gly 184 to Gly 193, and Gly 216 to Asn 223. Trypsinogen is activated by proteolytic cleavage of the N-terminal hexapeptide, but it may also be converted to a trypsin-like state upon strong ligand binding (Bode, 1979). Its complex with PTI and the dipeptide Ile-Val is extremely similar to the trypsin PTI complex (Bode, Schwager & Huber, 1978; Perkins & Wuethrich, 1980). The studies described in the following aim at defining more clearly the disorder of the activation domain. Thermal motion and conformational heterogeneity will produce similar effects on the electron density distribution, either by smearing it out over a large volume or by distributing it on several unresolved sites.

The activation domain contains a disulphide bridge, Cys 191–Cys 220, which is disordered in trypsinogen. It can be selectively reduced and mercurated. The crystal structure analysis of such heavy-atom-labelled trypsinogen and its complex with PTI and the dipeptide Ile-Val extends the level of significance of the crystallographic analysis beyond the observations on the unmodified compounds, but cannot discriminate between static and dynamic disorder. This may be achieved by low-temperature studies as the dynamic component only is influenced. We had performed a preliminary analysis of trypsinogen at 213 K and found

* Present address: Department of Physics, University of Indore, Khandwa Road, Indore-452001 (MP), India.

† Trasylol is a registered trademark of Bayer, Leverkusen.

0567-7408/82/051462-11$01.00

substantial changes of the crystal properties, mosaic spread and diffuse scattering, indicative of a phase change and annealing of crystal imperfections. This was probably due to the extremely slow cooling procedure applied in this experiment. But we found no indications of ordering of the activation domain. However, the intensity data set was incomplete (Singh, Bode & Huber, 1980). We therefore undertook the analysis at 173 and 103 K on the basis of complete and more accurate data sets obtained with synchrotron radiation. In this experiment the crystals were rapidly cooled in a special device mounted on the rotation camera (Bartunik & Schubert, 1982).

2. Experimental

Mercuration

Trypsinogen was selectively reduced at cystine 191–220 following the procedure of Light and colleagues (Knights & Light, 1976). It is briefly reported here, as a few modifications were necessary for the subsequent reaction with mercury dichloride: 25 mg trypsinogen (Merck, Darmstadt) and 10 mg benzamidine hydrochloride (Merck) were dissolved in 0·5 ml water saturated with N_2 and brought to pH 8 with a small amount of concentrated alkaline phosphate buffer. The solution was cooled in an ice bath. N_2 was blown over the solution and 0·5 ml 0·2 *M* $NaBH_4$ (Merck) solution in water added. The solution was gently stirred for 1 h. After addition of anti-foam (Bender & Hobein, München), excess borohydride was quenched by adding acetone in an amount equivalent to $NaBH_4$. The solution was brought to pH 6·5 with a small amount of concentrated acidic phosphate buffer. 20 µl 0·1 *M* $HgCl_2$ solution was added with stirring and after 15 min another 20 µl. The reaction solution was then dialysed against 10^{-3} *M* HCl for 48 h and chromatographed on Sephadex G25 (Pharmacia, Sweden) equilibrated with 10^{-3} *M* HCl. This material was lyophilized and used for the crystallization experiments.

The amount of Hg bound was determined by complex formation with dithizone (1,5-diphenylthiocarbazone) (Holzbecher, Divis, Kral, Sucha & Vlacil, 1976).

Protein concentration was determined from $A^{1\%}_{280\,nm}$: 15·3. This preparation contained 1·2 mol Hg per mol trypsinogen. This product will be abbreviated TgHg in the following.

Crystallization

Crystallization of TgHg from $MgSO_4$ solutions followed the procedure optimized for trypsinogen (Bode, Fehlhammer & Huber, 1976).

To 20 µl of a 10 mg ml^{-1} solution of TgHg, PTI was added at a molar ratio of 4/1. Crystallization was achieved by adding 5 µl of 1·5 *M* $MgSO_4$ solution, pH 7, and concentrating by vapour diffusion against 1·5 *M* $MgSO_4$ solution. The crystals were transferred to a 2·4 *M* $MgSO_4$ solution for the X-ray experiments.

The complex of TgHg with PTI and the dipeptide Ile-Val (abbreviated TgHg-PTI-IV) was crystallized following the procedure for the trypsinogen complex (Bode, Schwager & Huber, 1978). Crystallization is achieved under conditions identical to those described for TgHg except that an equimolar amount of PTI and 0·6 µl of a solution containing 3·5 mg ml^{-1} Ile-Val dipeptide is added. These crystals were transferred to a 2·4 *M* $MgSO_4$ solution containing 1 m*M* Ile-Val.

Data collection of TgHg and TgHg-PTI-IV

Intensity data were collected on a rotation camera with a focus-to-crystal distance of 30 cm and crystal-to-film distance of 7·5 cm. Graphite-monochromatized Cu $K\alpha$ radiation was used. The rotation angle between successive photographs was 2 to 3°. Still photographs between rotation photographs were used for refinement of the crystal alignment (Schwager, Bartels & Jones, 1975). Scanning on an automatic densitometer, evaluation and processing of data were performed with the programs of Schwager & Bartels (1977). For TgHg 8481 measurements above a 2·5σ significance level were obtained from one crystal, yielding 5562 unique reflections to a maximum resolution of 2·3 Å. In the TgHg-PTI-IV case 37 553 observations above the 2·5σ significance level were collected from two crystals. These data were corrected for absorption effects (Schwager, Bartels & Huber, 1973). 14 327 unique reflections were obtained up to 2·1 Å resolution. The R_{merge} values, defined as $R = \sum_{hj} |(\langle I_h \rangle - I_{hj})/N_h \langle I_h \rangle|$, are 0·063 for TgHg and 0·089 for its complex with PTI and Ile-Val.

Low-temperature analysis

The crystal. Bovine trypsinogen crystals were transferred from 2·4 *M* $MgSO_4$ (pH = 6·9) *via* Ficoll (Pharmacia, Sweden) to a 70% methanol–30% water mixture buffered with 0·1 *M* Na cacodylate, 0·1 *M* $CaCl_2$ at pH = 7·0 at room temperature. The freezing point of such a methanol–water mixture is reported to be 156 K (Hui Bon Hoa & Douzou, 1973) or 188 K (Douzou, 1977).

The low-temperature data. The trypsinogen crystals were cooled in a device developed for an Arndt–Wonacott camera (Bartunik & Schubert, 1982). The device is mounted on a normal goniometer head. It consists of two concentric chambers with thin (8 µm) Mylar walls (Fig. 1). The inner chamber receives a flow

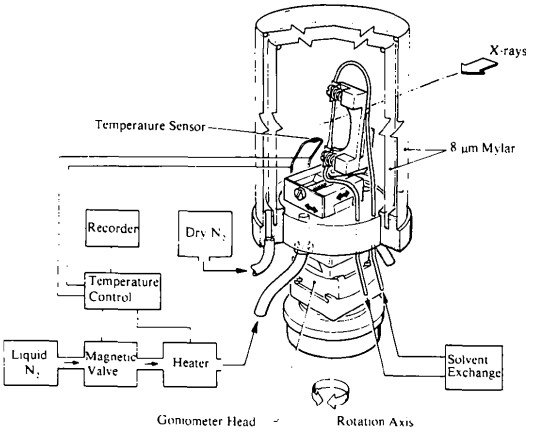

Fig. 1. Schematic drawing of the cooling device.

of cold N_2 gas, whose temperature is controlled *via* a Si-diode sensor and a heating coil. Temperature gradients in the environment of the crystal are negligible. Warm dry N_2 circulates through the outer chamber in order to prevent the walls from icing.

After initial alignment one crystal was cooled to 173 K in two steps: first to 243 K within 5 min, then to 173 K within another 7 min. The protocol for the experiment at 103 K was: initial alignment of the crystal, cooling to 173 K within 40 min, to 153 K within 5 min, and finally to 103 K within another 7 min. After each cooling step slight realignment was needed in order to correct for movements of the crystal caused by the temperature change. The temperature variation during the data collection was ±0·5 K at 173 K and ±2 K at 103 K. Remarkably the solvent phase of the crystal did not freeze in the last experiment, although the temperature was about 60 K below the reported freezing point. External solvent in the capillary froze and gave rise to Debye–Scherrer rings of the ice crystals.

Data collection with synchrotron radiation. The low-temperature data were collected on the instrument X11 at the EMBL outstation, Hamburg, using synchrotron radiation from the storage ring DORIS operated at a particle energy of 3·3 GeV and a maximum current of about 75 mA. X11 has double-focusing optics with quartz mirrors for vertical and a bent triangular Ge(111) monochromator for horizontal focusing. Film data were collected with a modified Arndt–Wonacott camera working under computer control on the basis of incident-beam monitoring for automatic compensation of effects of time-variable incident intensity from the synchrotron radiation source (Bartunik, Clout & Robrahn, 1981).

A wavelength λ of 1·069 Å with a resolution of $\Delta\lambda/\lambda = 3 \times 10^{-3}$ was employed. In order to obtain a broad enough dynamical range in the film response at this short wavelength, steel (NIROSTA) foils of a thickness of 25 μm were used as attenuator foils between films. The absorption factors were determined to be 1·4 for the OSRAY T4 film, and 5·8 for the attenuator foil. The fraction of polarization of the radiation in the plane of the synchrotron f_h was estimated to be 0·98. The program used for film evaluation (Schwager & Bartels, 1977) was modified accordingly. Recent computations of this polarization correction after the Ge(111) monochromator gave an f_h of about 0·92 (Bartels, 1982). The effect of this difference on the observed structure factors is negligible.

34 340 (30 429) reflections were collected from one crystal at 173 K (103 K) and merged to yield 15 890 (15 524) independent reflections. This corresponds to 72·4% (71·0%) of all significant ($>2{\cdot}5\sigma$) reflections to 1·7 Å resolution. In the resolution range 1·73 to 1·70 Å there are 386 (588) unique significant reflections corresponding to 35·6% (54·9%) of the possible reflections. The crystal size in both experiments was about 0·5 mm. A collimator with a limiting aperture of 0·3 mm was used. The exposure time was typically 5 min per degree of rotation. The collection of one data set took in total about 3–4 h. The crystal did not move during this time, as indicated by the change of only 0·15° (0·28°) in the mis-setting angles as determined by the film-evaluation program.

Data processing. By minimization of the mean shifts between observed and calculated positions of the reflections, the wavelength and cell constants were refined: $\lambda = 1{\cdot}0688$ Å, $a = b = 55{\cdot}4$, $c = 108{\cdot}5$ Å at 173 K and $a = b = 55{\cdot}5$, $c = 107{\cdot}8$ Å at 103 K. After the measurements had been made at 173 K the warmed crystal showed cell constants of $a = b = 55{\cdot}3$, $c = 109{\cdot}3$ Å. These values are close to the room-temperature data and indicate reversibility of the lattice-constant change. This observation was confirmed by another experiment: a trypsinogen crystal (70% methanol–30% water), mounted in a capillary (diameter 0·7 mm), was put into liquid N_2 ($T = 77$ K) for about 4 h. Comparison of precession photographs taken before and after cooling showed no change in lattice constants.

R_{merge} of the data was 0·062 at 173 K and 0·078 at 103 K.

3. Results and discussion

Mercury trypsinogen in native and complexed form

A difference Fourier map between mercurated trypsinogen and native trypsinogen in 50% (v/v) methanol–water mixture was calculated using phases from the refinement of the latter structure (see later and Singh, Bode & Huber, 1980). No significant density was found in the vicinity of the disulphide bridge Cys 191–Cys 220. The largest features in this map are near

His 57 (a peak of 6σ) and at the Ca binding site (a minimum of -7σ). These were also observed in the difference Fourier map between trypsinogen in methanol–water mixture and trypsinogen in $MgSO_4$ (Singh, Bode & Huber, 1980). The refinement showed that the Ca site is fully occupied in methanol–water and partially occupied in $MgSO_4$. His 57 is better localized in $MgSO_4$ and has a lower B factor. The B factor is defined as $B = 8\pi^2\langle x^2\rangle$, where $\langle x^2\rangle$ is the mean-square displacement of the atoms along the normal to the reflection planes (Debye, 1914). B values are indicators of dynamic and static disorder causing a decrease in scattering factor by the temperature factor $\exp[-B(\sin^2\theta/\lambda^2)]$.

The absence of electron density of the Hg atom in the difference map must be explained by a distribution of the density over a large volume. A decision between thermal mobility or microheterogeneity cannot be made from this experiment. A B factor of at least 200 Å^2 must be assumed for the Hg atom, corresponding to an isotropic r.m.s. deviation from the equilibrium location of about 1·6 Å. An Hg atom with a B value of 200 Å^2 placed at an arbitrary position in the crystal structure is almost invisible in a difference Fourier map calculated with this atom.

As reference for the TgHg-PTI-IV complex the refined crystal structure of the complex with native trypsinogen (Bode, Schwager & Huber, 1978) has been taken. This crystal structure had been further improved with the use of Deisenhofer's version of the Jack & Levitt (1978) refinement procedure combining crystallographic and energy refinement. Energy parameters for bond lengths, bond angles, torsion angles and non-bonded interactions were those recommended by Levitt (1974). The R value calculated from the corresponding model was 0·22 for the reflections between 6·8 and 2·1 Å resolution (with a rejection ratio $RR = 2|F_o - F_c|/|F_o + F_c|$ of 1·2). A difference Fourier map between the mercurated and the native complex using phases from this refinement shows only significant density changes in the region of the 191–220 cystine. Sections in the vicinity of this disulphide are shown in Fig. 2. The positive peak has a height of 14 e.s.d.'s (calculated from the fluctuation of the difference map in the whole molecule); it represents the bound Hg atom. The negative density (−6 e.s.d.'s) is located at the disulphide. Binding of Hg forces the S atoms to adopt new positions. The conformation of the polypeptide main chain around 191–220 is almost unchanged as indicated by the absence of significant difference density between the native and the mercurated complexes. The visibility of the Hg atom is in congruence with the observation that the activation domain in trypsinogen is rigidified upon forming a complex with PTI and the dipeptide Ile-Val (Bode, Schwager & Huber, 1978; Huber & Bode, 1978). Disorder of the Hg atom in native trypsinogen on the other hand emphasizes the flexibility of the activation domain. Apart from the implications for flexibility in the trypsinogen system by study of chemical labelling the stereochemistry of Hg bonding in the 191–220 cystine is of considerable interest. Approximate tetrahedral coordination has been found in dichloro(L-cysteine)-mercury(II) and approximate linear S–Hg–S bonding in (L-cysteinato)(L-cysteine)mercury(II) chloride hemihydrate (Taylor & Carty, 1977). Initial model building showed compatibility with a tetrahedral conformation. Little change in the positions of main-chain atoms is

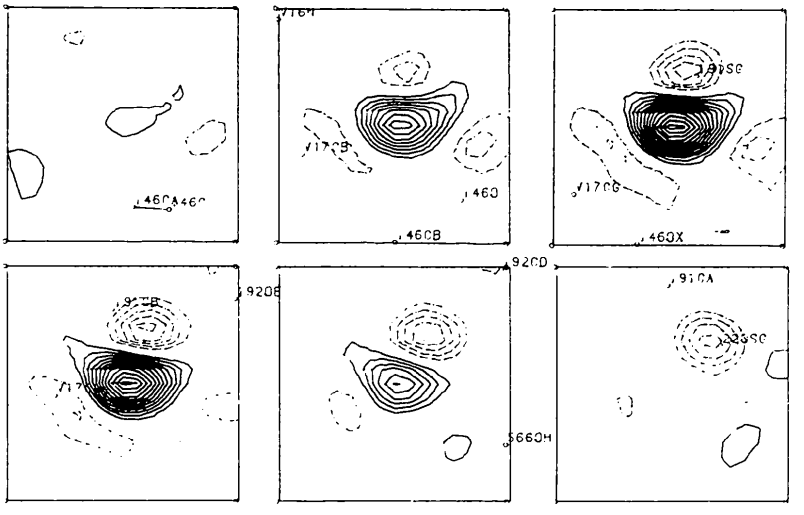

Fig. 2. Difference Fourier map between the mercurated trypsinogen complexed with PTI and the dipeptide Ile-Val and the unmercurated complex. Phases of the latter have been used. The vicinity of the disulphide 191–220 is shown in a number of consecutive sections. Positive contours are fully drawn, negative contours dashed. Contour lines are drawn at levels of $+1{\cdot}5\sigma$ in increments of 1σ. The superimposed model represents the mercurated complex TG-PTI-IV around the disulphide Cys 191–Cys 220.

required as a consequence of Hg insertion. Bi-coordination with a straight S–Hg–S bond angle seems unlikely since the S atoms cannot adopt positions 4·5–5 Å apart without affecting the main chain considerably. It is, however, obvious from the elongated shape of the electron density (Fig. 2), that the interpretation as a single Hg site is inadequate. The Hg atom must adopt a number of positions equally distant from the S atoms, which are approximately located on the arc of a circle. The angular range is limited by the maximally allowed distortion of the C_β–S_γ–Hg bond angle. As an approximation of the manifold possibilities the density was modelled by three Hg sites with $\frac{1}{3}$ occupancy each, one located at the centre of the peak and two 1·8 Å apart in a symmetrical arrangement. Holding these parameters constant, this model was tested with Rossmann's heavy-atom refinement procedure modified by Ten Eyck (private communication) and us, using as reference ('native') compound the unmercurated complex with the refined phases. The *B* factors were refined to values of 49, 85 and 83 Å^2 for the central and side atoms, respectively. The negative density at the S locations was modelled by negative scatterers of −16e (S) with their coordinates fixed. *B* factors of 80 and 120 Å^2 were obtained, indicating the close proximity of the S positions in the mercurated complex compared to those in the unmercurated complex. A residual map calculated for this model did not contain remarkable density.

Further crystallographic refinement was performed (Jack & Levitt, 1978) obeying stereochemical considerations. For this purpose a special Hg-cystine group was constructed which had three independent positions for the C_β, S_γ and Hg atoms. The latter were artificially interconnected at fixed distances of 1·8 Å from each other to avoid uncontrolled merging in the continuous density of the Hg atom. The distance Hg–S was set to 2·33 Å as found in the linear conformation, since no other ligands of the Hg could be interpreted in the map. The angle S–Hg–S was set to 110° with a flexibility slightly higher than the bond angles in the main chain. After two cycles of coordinate refinement of the whole molecule and one cycle of temperature-factor refinement with separate averaging for main and side chains (except in the residues Cys 191 and Cys 220) the *R* value dropped from 0·217 to 0·200 (RR = 1·2). An $F_o - F_c$ Fourier map calculated with the resulting coordinates was relatively clean.

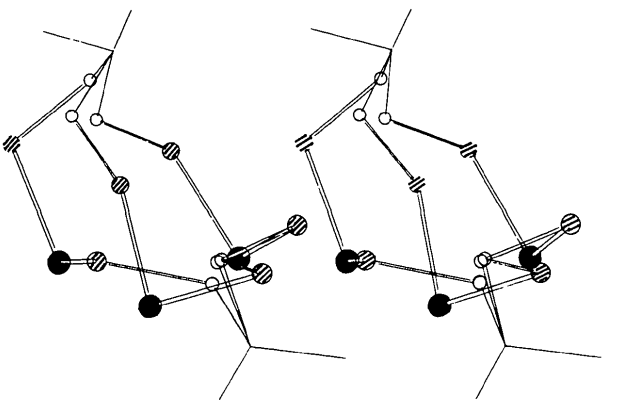

Fig. 3. Stereopicture of the Hg-cystine group in mercurated trypsinogen complexed with PTI and the dipeptide Ile-Val. Three alternative positions of the Hg-cystine are shown. The main-chain atoms (except C_β) of the residues Cys 191 and Cys 220 are interconnected by thin lines. C_β ○, S_γ ⊘, Hg ●.

Fig. 3 shows a stereopicture of the mercurated cystine 191–220 of complexed trypsinogen with three alternative locations. The atomic coordinates of the Hg-cystine fragment have been deposited.*† The S–Hg–S bond angle is 105° on average with a standard deviation of 5°. The bond angles C_α–C_β–S_γ and C–S–Hg have average values of 116 and 106° with standard deviations of 7 and 5°, respectively. These values do not deviate considerably from the values observed for tetrahedral coordination (Taylor & Carty, 1977). However, no other ligands are found within a coordination sphere of 3 Å. As has been noted before, the main chain is affected only slightly by the bonding of the Hg atom. The largest shifts are between 0·2 and 0·3 Å which are observed only at the N and C_α positions of the residues 191 and 220. Apart from these the structure of the mercurated complex does not differ from the unmercurated one. The r.m.s. shift between the two coordinate sets is 0·05 Å.

Refinement of the low-temperature structures

For refining the structure at 173 K, we started with the room-temperature-model coordinate set containing 93 ordered water molecules (Singh, Bode & Huber, 1980) and a *B* value of 15·0 Å^2 for all atoms. The first structure factor calculation gave an *R* of 35%. Several cycles of refinement using Deisenhofer's version of the Jack & Levitt refinement program finally reduced *R* to 18·7% (RR = 1·2) in a resolution range of 1·7 to 6·5 Å. The diagonal-matrix least-squares method was also used to refine isotropic temperature factors. At

* The atomic coordinates of the Hg-cystine fragment and details of the course of the trypsinogen refinement at 173 K have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 36651 (4 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

† Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 2PTN, 3PTN, 2TGA, 1TGC, 1TGT, 2TGT, 2TPI, R2PTNSF, R3PTNSF, R2TGASF, R1TGCSF, R1TGTSF, R2TGTSF and R2TPISF), and are available in machine-readable form from the Protein Data Bank at Brookhaven or one of the affiliated centres at Cambridge, Melbourne or Osaka. The data have also been deposited with the British Library Lending Division as Supplementary Publication No. SUP 37004 (14 microfiche). Free copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

several stages of the refinement the calculated difference electron density maps were inspected. Several density peaks could be interpreted as localized solvent molecules. At a graphics display system (Jones, 1978) some side chains, lying outside the electron density, were corrected. Also, some undefined atoms could be moved into electron density and were activated. From the individual atomic B factors two mean B values were computed for each residue: one for the four main-chain atoms including C_β, and the other for the side-chain atoms without C_β. For estimation of the accessible surface area of the residues a computer program, based on the method of Lee & Richards (1971), was used.

The refined model coordinates of trypsinogen at 173 K served as a starting point for the refinement of the structure at 103 K. The refinement was carried out in the same way as for the 173 K structure. The R value dropped from 30 to 20·9% (resolution range 1·7 to 6·5 Å, RR = 1·2). Details of the comprehensive course of refinement of the low-temperature trypsinogen crystal structures have been deposited.*

Refinement of the room-temperature structures

The crystal structures of trypsinogen in a 50% methanol–50% water mixture (Singh, Bode & Huber, 1980), trypsinogen in 2·4 M $MgSO_4$ (Bode, Fehlhammer & Huber, 1976; Bode & Huber, 1978), trypsin in a trigonal space group (Fehlhammer, Bode & Huber, 1977) and trypsin in an orthorhombic space group (Bode & Schwager, 1975) had been refined previously mainly by applying Diamond's (1971, 1974) real-space refinement. These four structures have been further refined applying the same procedures as for the low-temperature data. The amount of data, resolution range, final R factors and B (overall) values are summarized in Table 1.

* See first deposition footnote.

In the difference electron density map of trypsinogen in methanol–water two high negative peaks (about -6σ) occurred at the carbonyl O atoms of Gly 216 and Try 141. Gly 216 and Try 141 are residues at the boundary of the activation domain. After deactivating these atoms the residual density vanished.

Owing to positive density peaks in the difference map of trigonal trypsin, two atoms, Ser 86 O_γ and Thr 125 $O_{\gamma 1}$, were activated. The highest residual density peak (about $+8\sigma$) is found in the vicinity of the Yb^{3+} ion, which replaces the Ca^{2+} ion.

In the difference map of the orthorhombic modification of trypsin a few positive residual density peaks (between 5 and 8σ) occurred. At the graphics display system 26 undefined side-chain atoms could be moved into density and were activated. Also, 10 additional water molecules were included in the refinement procedure.

The crystal structure at 173 *and* 103 K

The structure of the trypsinogen molecule at 173 K as well as at 103 K differs only slightly from the structure at room temperature. The r.m.s. distance of the main-chain atoms (including C_β) after optimal superposition of the molecules is 0·19 Å at 173 K and 0·22 Å at 103 K. The centre of gravity of the molecule is shifted by 0·01 Å (0·03 Å) in the x, 0·50 Å (0·73 Å) in the y and 0·21 Å (0·41 Å) in the z directions at 173 K (103 K), reflecting the anisotropic thermal expansion also indicated by the lattice-constant changes described before. In the final difference electron density maps (calculated in the resolution range 1·7 to 7·0 Å; contours from 2σ, in steps of σ) the regions where the flexible segments N-terminus to Gly 19, Gly 142 to Pro 152, Gly 184 to Gly 193, and Gly 216 to Asn 223 must be located, were carefully inspected. Scattered residual density peaks (about 5σ, compared to 8σ for a fully localized water molecule) are observed, but no course of the polypeptide chain is detectable even at 103 K.

Table 1. *Details of the refined structures of trypsinogen and trypsin*

Crystal	Trypsinogen	Trypsinogen	Trigonal trypsin	Orthorhombic trypsin	TGHG-PTI-IV	Trypsinogen at 173 K	Trypsinogen at 103 K
Solvent	50% methanol–50% water	2·4 M $MgSO_4$	2·4 M Ammonium sulphate	2·4 M Ammonium sulphate	2·4 M $MgSO_4$	70% methanol–30% water	70% methanol–30% water
Maximum resolution	1·8 Å	1·8 Å	1·7 Å	1·55 Å	2·1 Å	1·7 Å	1·7 Å
Number of significant unique data	12 067 (65·2%)	13 952 (75·4%)	15 425 (70·4%)	23 560 (72·8%)	14 327 (61·6%)	15 890 (72·4%)	15 524 (71·0%)
Number of significant unique data in the innermost resolution shell	182 (20·1%) 1·83 to 1·8 Å	327 (36·1%) 1·83 to 1·8 Å	190 (17·3%) 1·73 to 1·7 Å	296 (18·5%) 1·58 to 1·55 Å	310 (13·2%) 2·14 to 2·1 Å	386 (35·6%) 1·73 to 1·7 Å	588 (54·9%) 1·73 to 1·7 Å
B (overall)	16·1 Å^2	12·4 Å^2	14·9 Å^2	14·5 Å^2	21·0 Å^2	11·6 Å^2	11·5 Å^2
R factors							
(RR = 1·2)	18·0%	19·7%	19·8%	19·3%	20·0%	18·7%	20·9%
(RR = 2·0)	18·6	20·9	20·6	19·9	20·7	19·3	21·9
Resolution range	1·8 to 6·5 Å	1·8 to 6·5 Å	1·7 to 6·5 Å	1·55 to 6·5 Å	2·1 to 6·5 Å	1·7 to 6·5 Å	1·7 to 6·5 Å

Since atomic scatterers with high B factors contribute mostly to the low-resolution data, we calculated a $2F_o - F_c$ Fourier map for the 103 K structure using SIM-weighted phases (Hendrickson & Lattman, 1970) in the resolution range 3·5 to 20·0 Å. This map, also, did not allow chain segments of the activation domain to be traced. In trial calculations parts of the flexible chain segments (Gly 18 to Ile 16, Ser 217 and Asn 223 to Ala 221) were activated to interpret some continuous density. After four cycles of energy and crystallographic refinement and one cycle of B-factor refinement, the R value increased by 0·5%. A SIM-weighted $2F_o - F_c$ Fourier map was calculated and inspected. The density observed did not increase and is of the same order of magnitude as in the solvent regions (apart from the residues Gly 18 and Val 17, which are at the boundary of the flexible N-terminus). Following the main chain to the N-terminus the last residue showing observable, but already very weak, density ($B = 47·3$ Å^2) in the Fourier map at room temperature is Gly 19 (Fig. 4*a*). At $T = 173$ K, also, Gly 18 shows defined electron density (Fig. 4*b*) and a B factor of 37·7 Å^2. By cooling to 103 K the observed electron density of Gly 18 increases and the B value is reduced to 28·3 Å^2. Furthermore, there is additional density for part of residue Val 17 (Fig. 4*c*). The location of Gly 18 and Val 17 (C,C_α) deviates considerably between trypsinogen and the isomorphous trigonal trypsin (Bode, Fehlhammer & Huber, 1976), as Fig. 4(*d*) shows. In trypsin Gly 18 is at the mouth of the pocket which accommodates the N-terminus Ile 16-Val 17. The position of Gly 18 and Val 17 (C,C_α) in trypsinogen found at 173 and 103 K points away from the pocket towards the surface of the molecule (close to Leu 158). The carbonyl O of Gly 18 forms a hydrogen bond to water molecule 470 OH. The positional

Table 2. *Side-chain atoms additionally activated during refinement of the low-temperature structures and the according side-chain B factors at room temperature* (*RT*), 173 K *and* 103 K

Residue	Activated atoms at 173 K	Activated atoms at 103 K	B (RT) (Å^2)	B (173 K) (Å^2)	B (103 K) (Å^2)
Lys 60	N_ζ	–	19·4	20·1	11·6
Asn 79	$N_{\delta 2}$, $O_{\delta 1}$	–	29·8	29·0	24·7
Gln 135	$N_{\varepsilon 2}$, $O_{\varepsilon 1}$	–	33·2	36·8	31·2
Ser 178	O_γ	–	–	38·5	20·6
Ser 236	O_γ	–	–	35·7	31·9
Gln 240	C_γ, C_δ, $O_{\varepsilon 1}$, $N_{\varepsilon 2}$	–	–	32·9	36·8
Lys 109	–	C_ε	28·9	27·2	23·5
Ser 110	–	O_γ	–	–	31·5
Ser 113	–	O_γ	–	–	33·9
Ser 116	–	O_γ	–	–	33·9
Arg 117	–	C_δ	23·6	11·0	9·8
Lys 159	–	N_ζ	39·7	26·5	31·0
Ser 122	–	O_γ	–	–	22·4
Ser 170	–	O_γ	–	–	31·4
Asp 153	–	$O_{\delta 2}$	2·5	4·7	12·4
Asn 165	–	C_γ, $O_{\delta 1}$, $N_{\delta 2}$	–	–	43·4

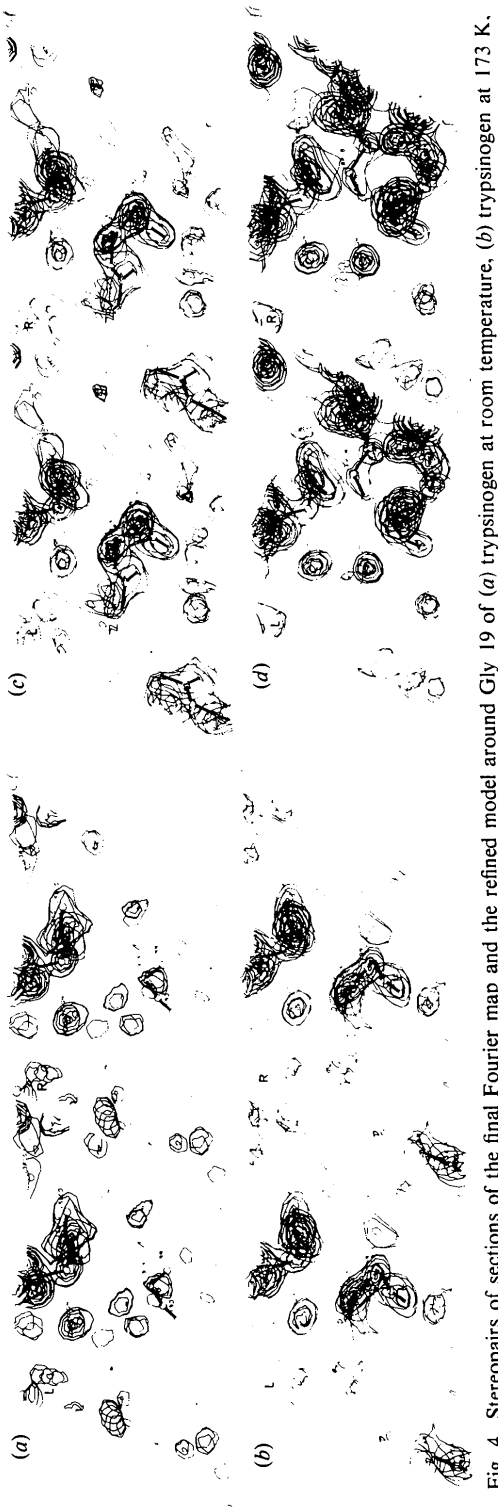

Fig. 4. Stereopairs of sections of the final Fourier map and the refined model around Gly 19 of (*a*) trypsinogen at room temperature, (*b*) trypsinogen at 173 K, (*c*) trypsinogen at 103 K, and (*d*) trigonal trypsin at room temperature. Contours are drawn above the significance level σ in steps of σ.

differences in trypsinogen compared to trypsin after optimal superposition of the molecules are, for Gly 18: 3·9 (C_α), 2·7 (O), 2·3 (C) and 6·2 Å (N) and, for Val 17: 7·3 (C) and 10·3 Å (C_α).

Apart from residue Gly 18 some side-chain atoms, which were 'unobserved' at room temperature, appeared at 173 and 103 K. In Table 2 these additional atoms and the corresponding side-chain B factors are listed. The B factors of these additional activated atoms are generally high, so that the average side-chain B factors rise and, therefore, do not compare well with the room-temperature values.

The largest positional changes at low temperature are observed for the side-chain atoms of Thr 26 ($O_{\gamma 1}$ 2·61/2·53; $C_{\gamma 2}$ 2·26/2·35), Lys 87 (C_γ 0·91/1·42; C_δ 1·20/1·24), Lys 156 (C_δ 2·02/2·01; C_ε 1·08/1·15), Gln 50 ($O_{\varepsilon 1}$ 0·74/1·12; $N_{\varepsilon 2}$ 0·67/1·19) and Val 154 ($C_{\gamma 2}$ 0·91/1·05). (In parentheses positional differences in Å at 173/103 K compared to room temperature are listed.)

In the difference map at 173 K an elongated density is observed for Ser 110 O_γ. Ser 110 O_γ is undefined at room temperature. This feature can be interpreted as two differently occupied alternative conformations of the Ser O_γ. The two locations of the O_γ are separated by about 2 Å. In the difference map at 103 K no significant residual density is detectable at the lower occupied site of the O_γ. The same situation holds for Ser 120 O_γ.

We found 31 (46) water molecules which become ordered at 173 K (103 K) in addition to the solvent localized at room temperature. Indications for nine of these are already observable in the final difference map of the room-temperature structure. 15 (24) are hydrogen bonded to other solvent molecules, forming water chains consisting of up to ten molecules. Two water molecules at room temperature are replaced by methanol molecules at low temperatures. One of these forms hydrogen bonds to Ser 45 O_γ and Gly 196 O, the other to a water molecule. The evidence for methanol molecules is based on the close proximity of the independently refined and fully occupied CH_3 and OH moieties with distances of 1·44 and 1·46 Å respectively. The increased order of most of the solvent molecules at low temperatures is indicated by the decrease in their B factors. The mean reduction in B of the 93 water molecules, which are ordered at room temperature, is 6·8 Å^2 (10·2 Å^2) at 173 K (103 K). In contrast to the protein, for which the overall B factor does not decrease further between 173 and 103 K (see later), the solvent B factors are reduced on the average by 3·4 Å^2.

The thermal parameters of trypsinogen

Two B factors were assigned to each residue by averaging main-chain and side-chain B values, respectively. In order to obtain an estimate of the standard deviation σ of these B factors we compared the B factors of trypsinogen in 50% methanol–50% water with those of trypsinogen in 2·4 M $MgSO_4$. We compared only those 43 residues which are not accessible to solvent. This comparison is problematic as the B factors are on the average lower by 3·2 Å^2 in $MgSO_4$ than in methanol–water (see later). After scaling the two sets of temperature parameters, the r.m.s. deviation σ was 2·45 Å^2. The r.m.s. deviation for the difference of B factors ΔB is therefore estimated as 3·5 Å^2.

The data of trypsinogen in 2·4 M $MgSO_4$ were absorption corrected; those of trypsinogen in 50% methanol–50% water and the low-temperature data were not. Dependence of reflection intensities on the Bragg scattering angle due to absorption effects would influence the temperature factor. This angular dependence of the absorption factor was estimated for a 50% methanol–50% water mixture as well as for a 70% methanol–30% water mixture of cylindrical and spherical geometry (*International Tables for X-ray Crystallography*, 1959) and found to be negligible. Although the crystal shape is neither cylindrical nor spherical, the error in B due to absorption effects in the methanol–water mixtures is expected to be small.

The overall B factor of trypsinogen in methanol–water is reduced from 16·1 Å^2 at room temperature to 11·6 Å^2 by cooling to 173 K. Cooling to 103 K causes no further change in overall B (11·5 Å^2). This may indicate that thermal motion is largely frozen at 173 K. An alternative interpretation may be that the freezing of the solvent on the surface of the crystal at 103 K causes some disorder which compensates a further decrease in the thermal component of B.

The distribution of the main-chain B factors *vs* residue number (Fig. 5) shows a similar course for all three temperatures. The B factors of the undefined residues are set to 60 Å^2. As was shown in test calculations, the residual electron density of an O atom with $B > 50$ Å^2 vanishes in the difference map. Besides the flexible region and its neighbouring residues the highest peak in the B-factor distribution corresponds to the residues Ser 61, Gly 62 and Ile 63, which form a loop exposed to the solvent. This loop does not form hydrogen bonds to the rest of the molecule except between Ile 63 O and Ala 85 N. In general, exposed residues, *e.g.* Asn 97, Ser 116, Ser 130, Asn 165, Ala 171, have high B factors. The changes in the B factors induced by cooling (Fig. 6*a*,*b*) show a strong reduction of the high B values near residue Gly 19. This fact as well as the additionally observable residue Gly 18 at 173 and 103 K indicates that the large r.m.s. displacements of the flexible N-terminus are predominantly due to thermal motion. In contrast, approaching the other three flexible parts of the molecule, there is no tendency for increased reduction

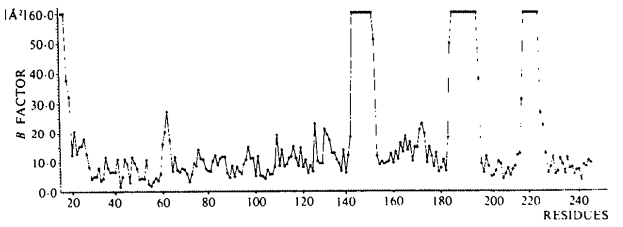

Fig. 5. Main-chain B factors of trypsinogen at 173 K *vs* residue number.

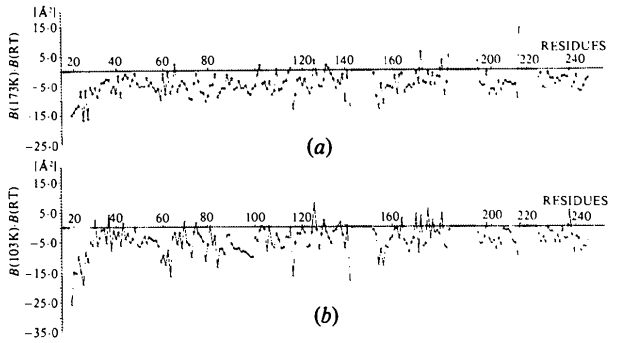

Fig. 6. Difference of the main-chain B factors of trypsinogen between (*a*) 173 K [B(173 K)] and room temperature [B(RT)], and (*b*) 103 K [B(103 K)] and room temperature [B(RT)] *vs* residue number.

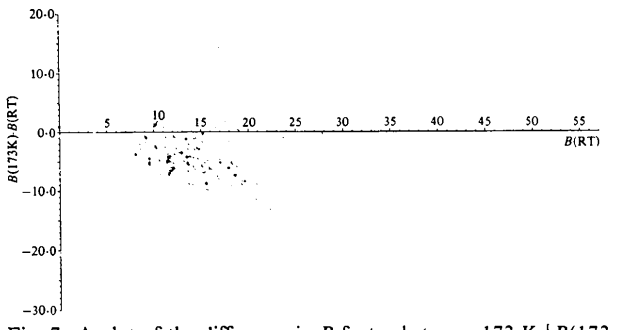

Fig. 7. A plot of the difference in B factor between 173 K [B(173 K)] and room temperature [B(RT)] *vs* the room-temperature B factors [B(RT)] of the main-chain groups.

in B. It should be recalled that these segments are linked to the rest of the molecule twice, while the N-terminus is fixed only once. Also the defined electron density at the boundaries of these flexible segments is not further extended at low temperatures. An abruptly starting, hinge-like motion could explain these features. It seems more likely, however, that the polypeptide chain segments of the activation domain adopt a large number of different conformations or substates. Between 173 and 103 K the changes in the main-chain B factors are distributed around 0·0 Å^2 with a standard deviation of 3·4 Å^2. This value is close to the estimated σ of ΔB. Only for a few main-chain groups is the reduction in B between 173 and 103 K large ($>2\sigma$): for Gly 18 and Gly 19 (near the N-terminus), Gly 23, Lys 60, Gly 62, Ile 63, and Gly 216 (the carbonyl O, activated during refinement of the 173 K structure, becomes better defined at 103 K). As shown in Fig. 7, reduction in B is correlated with B for the main-chain atoms. This holds also for the side-chain groups. This is in contrast to observations in myoglobin crystals, where large B values are essentially temperature independent (Frauenfelder, Petsko & Tsernoglou, 1979).

Although for the overwhelming majority of the residues the B factors decrease from room temperature to 173 K (103 K) (Figs. 6, 7), there are two (seven) main-chain and five (eight) side-chain groups (neglecting those with additionally activated atoms at 173 and 103 K) with positive ΔB greater than the estimated standard deviation of 3·5 Å^2. An explanation could be the existence of conformational substates (Frauenfelder, Petsko & Tsernoglou, 1979). We did not refine occupancies. Occupancies and temperature factors are correlated in protein crystal structure refinement due to the limited resolution of the data. A reduction of the occupancy at low temperature would be expressed by an artificial increase of B. No significant electron density was detectable near those residues so that many (only partially occupied) substates must be postulated in this case. Another explanation would be that the potentials are flattened by configurational changes of neighbouring residues, permitting larger displacements at the low temperatures.

Considering the side-chain B factors, 37 out of the 41 residues with $B > 25{\cdot}0$ Å^2 at room temperature are polar groups – 18 of them Ser O_γ atoms. Almost all are located on the surface and interact with solvent molecules. The B factors of most of these polar side chains are decreased considerably at low temperatures. At 173 K (103 K) there are only 26 (22) of these groups with $B > 25{\cdot}0$ Å^2. This may be a consequence of the increased order of the solvent shell at low temperatures.

Between 173 and 103 K the B factors of a few of the side-chain groups are reduced considerably, showing that even below 173 K a large component of disorder is due to thermal motion for these residues: Thr 26 (−16·5 Å^2), Val 154 (−12·6 Å^2) and Lys 156 (−16·1 Å^2) (these side-chain groups are located differently compared to room temperature), Ser 178 (−17·9 Å^2) (O_γ was activated during refinement of the 173 K structure), Leu 67 (−17·5 Å^2) and Leu 99 (−12·8 Å^2).

The side-chain B factors tend to increase with distance from the molecular centre as well as with an increase of the accessible surface area of each residue. The influence of accessible surface area and distance from the molecular centre on the main-chain B values is weaker than for the side-chain groups. In both cases the characteristic B-factor distributions do not change with temperature.

A comparison of the B factors of trypsinogen in 50% methanol–50% water and in 2·4 M $MgSO_4$ stresses the influence of the solvent on molecular ordering. The overall B factor of trypsinogen in $MgSO_4$ is 12·4 Å^2, which is 3·7 Å^2 lower than in 50% methanol–50% water. Besides this reduction in B, the distributions of the B factors *vs* residue number have the same characteristics in both solvents, suggesting that in the methanol–water mixture an, essentially constant, additional contribution of disorder is involved. As far as dynamic disorder is concerned, the higher viscosity η of 2·4 M $MgSO_4$ (η = 5·8 mPa s) compared to that of 50% methanol–50% water (η = 1·8 mPa s) may hinder rigid-body motion as well as intramolecular mobility. This could possibly explain the lower overall B value of trypsinogen in $MgSO_4$. But, of course, we are not able to discriminate the influence of viscosity from other solvent properties (such as dielectric constant) affecting the solvent–protein interaction.

Comparison with the slow-cooling experiment

In the experiment described above the crystals were rapidly cooled compared to a previous study where we had cooled trypsinogen crystals to 213 K over a period of several days and had taken diffraction photographs at intermediate temperatures (Singh, Bode & Huber, 1980). In this experiment substantial changes in crysta' properties were noticed, indicative of an annealing of crystal imperfections. Such observations were not made in the experiment described above. The annealing process can probably take place only under conditions of very slow gradual cooling. We emphasize, however, that the experiments at the synchrotron were not well suited to comparing diffuse scattering or mosaic spread as the data sets at 173 and 103 K have been collected with different crystals.

Comparison of the thermal parameters of trypsin in the trigonal and orthorhombic space groups

In both modifications of trypsin the overall B factor is about the same, 14·9 and 14·5 Å^2 respectively. It should be recalled that the trigonal as well as the orthorhombic crystals were soaked in the same solvent (2·4 M ammonium sulphate), differing in pH (6·9 and 8·0 respectively); in the case of trigonal trypsin 10 mM ytterbium sulphate was included. Also, the distribution of the B factors *vs* residue number has the same characteristics (Fig. 8), indicating that this distribution represents a characteristic property of the molecule. This feature agrees with the result found for hen and human lysozyme (Artymiuk, Blake, Grace, Oatley, Phillips & Sternberg, 1979).

A detailed description of the refinement of orthorhombic trypsin is in preparation.

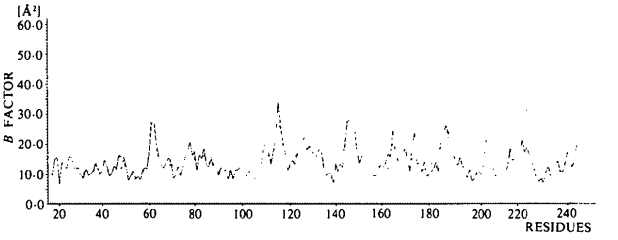

Fig. 8. Main-chain B factors of (*a*) orthorhombic trypsin (full line), and (*b*) trigonal trypsin (broken line) *vs* residue number.

4. Conclusion

In the disordered activation domain of trypsinogen a covalently inserted Hg atom is crystallographically undetectable. This implies thermal or static disorder characterized by a B factor of at least 200 Å^2. Reducing thermal disorder by cooling to 103 K does not increase the order of the flexible activation domain. At the boundaries of these segments no increased reduction of the B factors is observed, nor is the defined electron density further extended. An abruptly starting, hinge-like motion could explain these features. It is more likely, however, that the polypeptide-chain segments of the activation domain adopt a large number of different conformations or substates. In contrast to the activation domain, flexibility of the N-terminus is decreased substantially at low temperatures and is predominantly due to thermal vibrations. Also, the temperature behaviour of the B factors near the N-terminus indicates that collective internal motions are still active even at 103 K.

References

ARTYMIUK, P. J., BLAKE, C. C. F., GRACE, D. E. P., OATLEY, S. J., PHILLIPS, D. C. & STERNBERG, M. J. E. (1979). *Nature (London)*, **280**, 563–568.

BARTELS, K. (1982). To be published.

BARTUNIK, H. D., CLOUT, P. N. & ROBRAHN, B. (1981). *J. Appl. Cryst.* **14**, 134–136.

BARTUNIK, H. D. & SCHUBERT, P. (1982). *J. Appl. Cryst.* **15**. Submitted.

BODE, W. (1979). *J. Mol. Biol.* **127**, 357–374.

BODE, W., FEHLHAMMER, H. & HUBER, R. (1976). *J. Mol. Biol.* **106**, 325–335.

BODE, W. & HUBER, R. (1978). *FEBS Lett.* **90**, 265–269.

BODE, W. & SCHWAGER, P. (1975). *J. Mol. Biol.* **98**, 693–717.

BODE, W., SCHWAGER, P. & HUBER, R. (1978). *J. Mol. Biol.* **118**, 99–112.

DEBYE, P. (1914). *Ann. Phys. (Paris)*, **43**, 49–95.

DIAMOND, R. (1971). *Acta Cryst.* A**27**, 436–452.

DIAMOND, R. (1974). *J. Mol. Biol.* **82**, 371–391.

DOUZOU, P. (1977). *Cryobiochemistry, an Introduction*. London: Academic Press.

FEHLHAMMER, H., BODE, W. & HUBER, R. (1977). *J. Mol. Biol.* **111**, 415–438.

FRAUENFELDER, H., PETSKO, G. A. & TSERNOGLOU, D. (1979). *Nature (London)*, **280**, 558–563.

HENDRICKSON, W. A. & LATTMAN, E. E. (1970). *Acta Cryst.* B**26**, 136–143.

HOLZBECHER, Z., DIVIS, L., KRAL, M., SUCHA, L. & VLACIL, F. (1976). *Handbook of Organic Reagents in Inorganic Analysis*. New York, London, Sydney, Toronto: John Wiley.

HUBER, R. & BODE, W. (1978). *Acc. Chem. Res.* **11**, 114–122.

HUI BON HOA, G. & DOUZOU, P. (1973). *J. Biol. Chem.* **248**, 4649–4654.

International Tables for X-ray Crystallography (1959). Vol. II. Birmingham: Kynoch Press.

JACK, A. T. & LEVITT, M. (1978). *Acta Cryst.* A**34**, 931–935.

JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.

KNIGHTS, R. J. & LIGHT, A. (1976). *J. Biol. Chem.* **251**, 222–228.

LEE, B. & RICHARDS, F. M. (1971). *J. Mol. Biol.* **55**, 379–409.

LEVITT, M. (1974). *J. Mol. Biol.* **82**, 393–420.

PERKINS, S. J. & WUETHRICH, K. (1980). *J. Mol. Biol.* **138**, 43–64.

SCHWAGER, P. & BARTELS, K. (1977). *The Rotation Method in Crystallography*, edited by U. W. ARNDT & A. J. WONACOTT, ch. 10. Amsterdam, New York, Oxford: North-Holland.

SCHWAGER, P., BARTELS, K. & HUBER, R. (1973). *Acta Cryst.* A**29**, 291–295.

SCHWAGER, P., BARTELS, K. & JONES, T. A. (1975). *J. Appl. Cryst.* **8**, 275–280.

SINGH, T. J., BODE, W. & HUBER, R. (1980). *Acta Cryst.* B**36**, 621–627.

TAYLOR, N. J. & CARTY, A. J. (1977). *J. Am. Chem. Soc.* **99**, 6143–6145.