Protein Engineering vol.6 no.4 pp.341–348, 1993

The sequence and X-ray structure of the trypsin from *Fusarium oxysporum*

W.R.Rypniewski[2], S.Hastrup[1], Ch.Betzel, M.Dauter, Z.Dauter, G.Papendorf, S.Branner[1] and K.S.Wilson

European Molecular Biology Laboratory (EMBL), c/o DESY, Notkestrasse 85, D-2000 Hamburg 52, Germany and [1]Novo Nordisk, Novo Alle, DK-2880 Bagsvaerd, Denmark

[2]To whom correspondence should be addressed

The trypsin from *Fusarium oxysporum* is equally homologous to trypsins from *Streptomyces griseus, Streptomyces erythraeus* and to bovine trypsin. A DFP (diisopropylfluorophosphate) inhibited form of the enzyme has been crystallized from 1.4 M Na_2SO_4, buffered with citrate at pH 5.0–5.5. The crystals belong to space group $P2_1$ with cell parameters $a=33.43$ Å, $b=67.65$ Å, $c=39.85$ Å and $\beta=107.6°$. There is one protein molecule in the asymmetric unit. X-ray diffraction data to a resolution of 1.8 Å were collected on film using synchrotron radiation. The structure was solved by molecular replacement using models of bovine and *S.griseus* trypsins and refined to an *R*-factor of 0.141. The overall fold is similar to other trypsins, with some insertions and deletions. There is no evidence of the divalent cation binding sites seen in other trypsins. The covalently bound inhibitor molecule is clearly visible.

Key words: Fusarium oxysporum/serine proteinase/trypsin/X-ray structure

Introduction

Trypsins are characterized by specifically cleaving the peptide bond on the C-terminal side of lysine or arginine (Kraut, 1977; Steitz and Shulman, 1982). *Fusarium oxysporum* trypsin shows a reversed Arg/Lys specificity compared to bovine trypsin. The *F.oxysporum* trypsin is more Arg active than Lys active, while bovine trypsin shows greater activity towards Lys peptides than towards Arg peptides. The pH activity profile shows a broad activity optimum in the range pH 8–11. The temperature optimum is approximately 40°C at pH 9.5 (NOVO-Nordisk, unpublished). The enzyme consists of 224 amino acid residues and has a mol. wt of 22 190. Sequence alignment reveals an equally close homology to bovine trypsin (Walsh and Neurath, 1964) and to bacterial trypsins. When the first bacterial trypsin was discovered in *Streptomyces griseus*, its close similarity to bovine trypsin became an argument in a dispute about the evolutionary relationship between trypsins and trypsin-like proteins. Hartley (1970, 1979) argued for a gene transfer from higher organisms to the bacterium while Hewett-Emmett *et al.* (1981) succeeded in constructing an evolutionary tree with the bacterial enzyme at the root. Based on structural comparisons, Reed and James (1988) suggested that tryptic specificity imposes stricter structural requirements than does the specificity of related enzymes. Recently, another bacterial trypsin, from *Streptomyces erythraeus* (Miyamoto *et al.*, 1979; Yamane *et al.*, 1991), was investigated and found to have similar properties to that from *S.griseus*.

A considerable number of high resolution crystal structures of serine proteinases or their complexes with various inhibitors have been determined in recent years (Sawyer *et al.*, 1978; Blevins and Tulinsky, 1985; Delbaere and Brayer, 1985; Tsukada and Blow, 1985; Wang *et al.*, 1985; Fujinaga *et al.*, 1987; Meyer *et al.*, 1988; Reed and James, 1988; Navia *et al.*, 1989; Earnest *et al.*, 1991; Yamane *et al.*, 1991). This has made it possible to investigate the catalytic mechanism and to make comparative studies between the different structures. However, high resolution structures of trypsin have been limited to bovine trypsin (Chambers and Stroud, 1977, 1979) and the bacterial trypsin from *S.griseus* (Reed and James, 1988). The bacterial trypsin from *S.erythraeus* has been determined at 2.7 Å resolution (Yamane *et al.*, 1991). The recent results, as well as the results reported here, make it possible to reassess the relationship between the microbial and mammalian trypsins.

Here we report the sequence and the refined crystal structure of trypsin from the fungus *F.oxysporum* at 1.8 Å resolution.

Materials and methods

mRNA preparation

Mycelium from *F.oxysporum* was crushed under liquid nitrogen, resuspended in 5 M guanidinrhodanide, separated on a CsCl–sarcosyl gradient, washed with ethanol and fractionated on an oligo-dT column (Chirgwin *et al.*, 1978, Truelsen *et al.*, 1979). The quality of the mRNA was tested in an *in vitro* translation assay using reticulocyte lysate and [^{35}S]methionine. An SDS–PAGE analysis showed bands from below 20 to over 100 kDa.

Preparation and screening of a cDNA library

A cDNA library of the mRNA was constructed following the procedure of Okayama and Berg (1982). After transformation and plating ~15 000 colonies were replicated onto Whatman 540 filter paper. A mixed probe based on the N-terminal of the enzyme was synthesized (TM42-52/C), as shown in Figure 1. The probe, split on the isoleucine codons, was divided into three pools of 32 and three sets of filters were screened. After washing at 42°C and reisolation of possible candidates three positive clones, all found with the ATC–isoleucine probe, were identified.

Identification and analysis of the gene

Plasmids were prepared from the three clones—two turned out to be identical and the third was ~100 bp shorter. The longer

```
N-terminus  :  I V G G T S A G D F P F I V S A N G

                  C   C       C   A
Probe       :  G A T T T T C C N T T T A T C G T
                                        T
```

Fig. 1. The N-terminal sequence of the *F.oxysporum* trypsin and the probe used for screening the cDNA library.

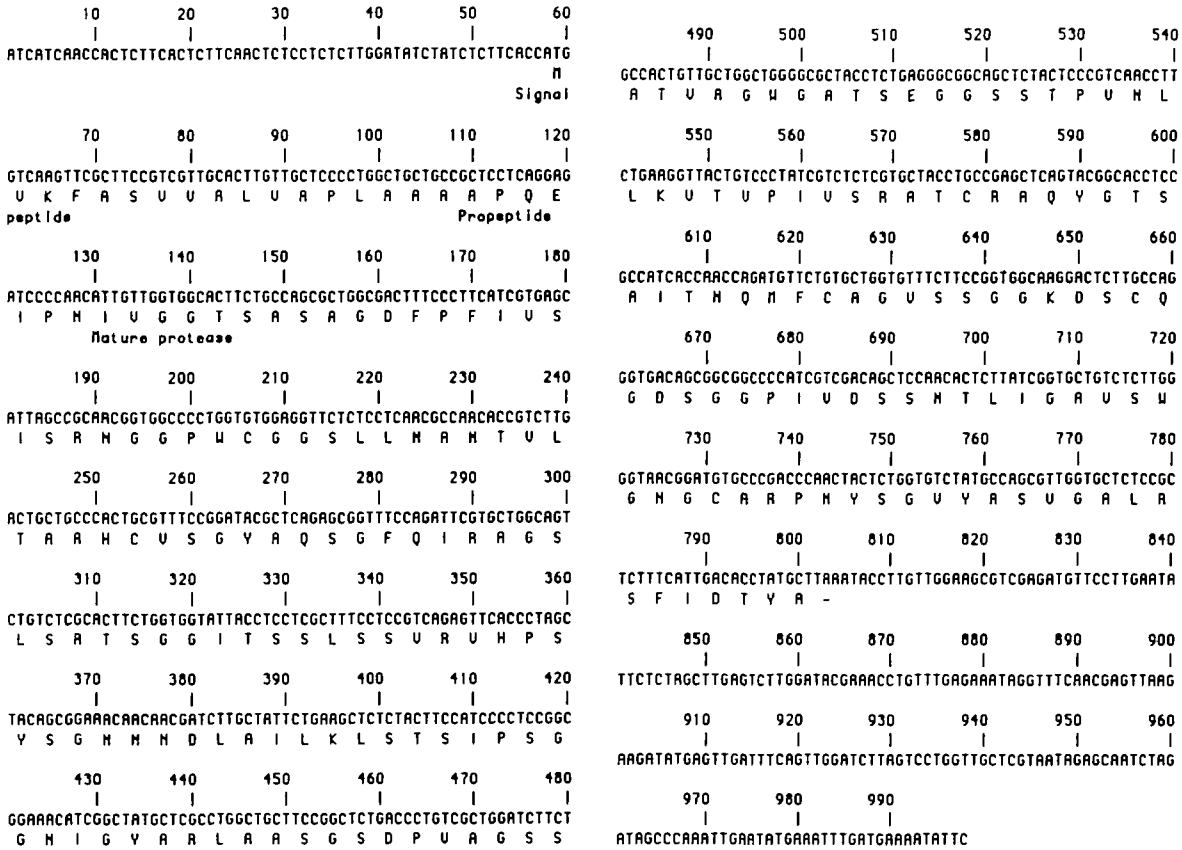

Fig. 2. The sequence of *F oxysporum* preprotrypsin cDNA and the deduced amino acid sequence.

clones were sequenced on both strands using the method of Maxam and Gilbert (1980). Analysis of the sequence (Figure 2) revealed a gene coding for a trypsin-like proteinase with a 17 amino acid signal peptide according to the rules of von Heijne (1986) and a propeptide of seven amino acids (Ala−Pro−Gln−Glu−Ile−Pro−Asn) between the signal peptidase cleavage site and the known N-terminus of the enzyme. The role of this hypothetical propeptide is unknown but proteinases are normally made as inactive zymogens.

Expression of the gene product

The gene was inserted into a suitable vector and transformed into *Aspergillus oryzae* as described by Christensen *et al.* (1988). When grown in YPD medium a protein appeared in the supernatant, which was slightly larger than the trypsin isolated from *F.oxysporum* supernatants. There was no activity as measured by hydrolysis of L-benzoyl-arginoyl-pNA (a trypsin substrate). After addition of an aspartyl protease isolated from *F.oxysporum* supernatants, the protein from the transformed *A.oryzae* was the same size as the standard on an SDS gel and the sample showed tryptic activity (data not shown). This indicated that the propeptide as deduced from the DNA sequence really existed and that its function was to keep the enzyme in an inactive form.

Protein purification

The *F.oxysporum* trypsin was separated from two other *F.oxysporum* proteinases by 0−1 M/0−25% sodium chloride/isopropanol gradient elution at pH 7.0 from a bacitracin−silicapolyol affinity column (Mortensen *et al.*, 1989). The two other proteinases separated from the trypsin by this procedure were a subtilisin-like enzyme and an aspartic proteinase with specificity required for the activation of the *F.oxysporum* protrypsin (Asn−Ile cleavage). The other purification steps were traditional and included ion exchange chromatography on Sepharose CL-6B (prior to the bacitracin separation), gel filtration and concentration by ultrafiltration. The specific absorption for a 0.1% solution of the purified *F.oxysporum* trypsin at 280 nm was determined to be 1.32. The amino acid composition of the purified trypsin was in agreement with the composition deduced from the DNA sequence.

Crystallization

Crystals of *F.oxysporum* trypsin were grown in the presence of DFP (diisopropylfluorophosphate) by hanging drop/vapour diffusion. The wells contained 1.4 M Na_2SO_4, 0.1 M sodium citrate, pH 5.0−5.5. The droplets initially contained 12.5 mg/ml protein and DFP at 5:1 molar ratio, 0.7 M Na_2SO_4, 50 mM citrate, pH 5.0−5.5. The crystals belong to space group $P2_1$ with cell parameters $a=33.43$ Å, $b=67.65$ Å, $c=39.85$ Å and $\beta=107.6°$. The asymmetric unit contained one protein molecule and was estimated to contain 36% solvent by volume (Matthews, 1968).

Data collection and processing

X-ray diffraction data were collected on CEA film to 1.8 Å resolution using synchrotron radiation on the EMBL beamline X11 at the DORIS storage ring, Hamburg. Films were digitized on an Optronics densitometer. The integrated intensities were measured using the program MOSFLM (Leslie *et al.*, 1986). The merging *R*-factor ($R_{merge}=\Sigma|I_i-<I>|/\Sigma<I>$, where I_i is an individual intensity measurement and $<I>$ is the mean intensity for this reflection; the summation is over all the data) was 0.047 for 14 026 independent reflections.

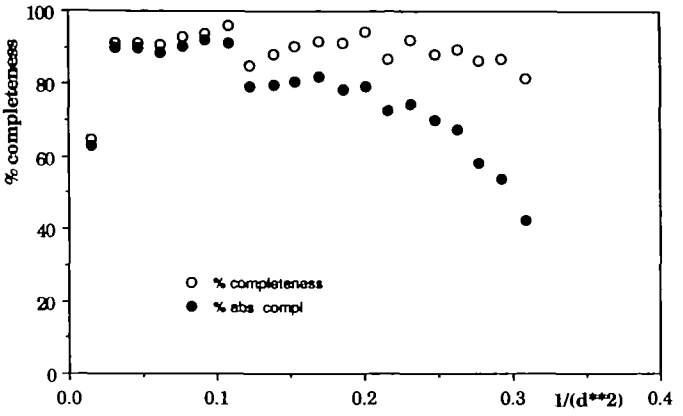

Fig. 3. Completeness of X-ray data as a function resolution is indicated by open circles. Filled circles indicate reflections stronger than three standard deviations as a percentage of all possible reflections.

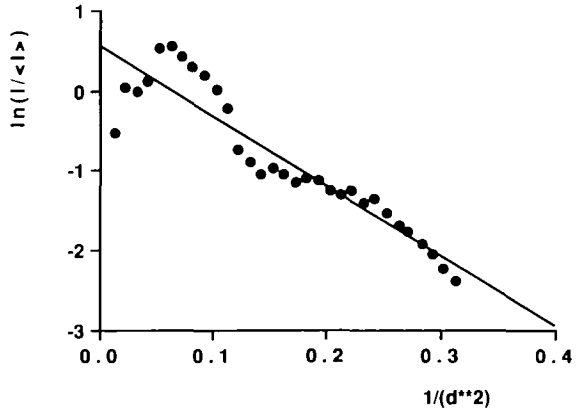

Fig. 4. Wilson plot for the 1.8 Å data.

Completeness of the data was 89.1% to 1.8 Å. Figure 3 shows the completeness of the data as a function of resolution and reflections with intensities greater than three standard deviations as a percentage of all possible reflections. In the highest resolution range the accuracy of the measurements falls off considerably. A Wilson plot (Wilson, 1942) for the data is shown in Figure 4. The fit to the theoretical straight line is satisfactory with the expected deviations for the protein in the low resolution range. The overall temperature factor estimated from the plot is 10.2 Å^2.

Molecular replacement and restrained least-squares refinement

The structure was solved by molecular replacement. Although the protein is homologous to other serine proteinases, in particular to bovine trypsin and trypsin from *S.griseus*, there are extensive regions without any apparent homology or with only a weak homology and any insertions and deletions can easily be missed. This was a reason for some initial difficulties in refining the structure based on a molecular replacement solution using bovine trypsin alone as the starting model. The following procedure was used eventually. A model was prepared based on bovine trypsin (Chambers and Stroud, 1977, 1979) code 4PTP, from the Protein Brookhaven Data Bank (Bernstein *et al.*, 1977; Abola *et al.*, 1987). The side chains that differed were cut down to alanine. In addition, the regions where insertions or deletions were observed or suspected from sequence alignment were deleted. The initial position of the molecule was obtained by molecular

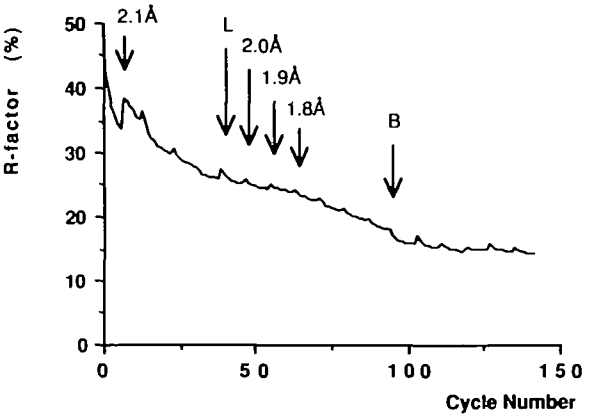

Fig. 5. The R-factor ($\Sigma||F_o|-|F_c||/\Sigma|F_o|$) plotted as a function of the cycle number of the refinement. Manual rebuilds of the model can be seen as kinks in the graph. The successive stages of extending the resolution of the data in refinement are shown. L indicates inclusion of low resolution data, below 6 Å. B indicates the inclusion of bulk solvent in the refinement.

Table 1. Weighting parameters for the least-squares refinement and the final standard deviations

Parameter	Weight	Standard deviation	Number of parameters
Distances (Å)			
Bond lengths (1−2 neighbours)	0.020	0.017	1612
Bond angles (1−3 neighbours)	0.040	0.045	2192
Dihedral angles (1−4 neighbours)	0.050	0.051	578
Planar groups	0.020	0.015	275
Chiral volumes	0.200	0.226	248
Non-bonded contacts (Å)			
Single torsion	0.500	0.181	532
Multiple torsion	0.500	0.180	76
Torsion angles (°)			
Peptide plane (ψ)	5.0	2.7	232
Staggered (aliphatic χ_1)	15.0	16.0	230
Orthonormal (aromatic χ_2)	20.0	19.3	17
Thermal factors (Å^2)			
Main chain bond	2.0	2.13	1013
Main chain angle	3.0	3.11	1216
Side chain bond	3.0	3.92	597
Side chain angle	4.5	5.33	972

replacement and was further refined by varying the position in small increments and recalculating the R-factor. The resulting model was subjected to constrained least-squares refinement using the program CORELS (Sussman *et al.*, 1977; Herzberg and Sussman, 1983) and then to several cycles of stereochemically restrained Hendrickson−Konnert least-squares refinement (Konnert, 1976; Konnert and Hendrickson, 1980). After this initial refinement the R-factor was 0.35 for data between 6 and 2.1 Å.

Although only bovine trypsin gave a clear solution to the molecular replacement procedure, a model based on the atomic coordinates of *S.griseus* trypsin (Reed and James, 1988), PDB code 1SGT (Berstein *et al.*, 1977; Abola *et al.*, 1987), with appropriate deletions, was aligned in a similar position to that initially obtained for bovine trypsin. The model was then refined in the same way as described above. Phases from the two models

were combined (Reed, 1986), using the phase probability coefficients (Hendrickson and Lattman, 1970) and the resulting map was interpreted. The rationale of using two starting models and then combining the information was to make use of the fact that the homologous regions complemented each other to some extent and that systematic errors typical of molecular replacement would cancel out considerably. The model resulting from interpreting the 'combined' map was 77% complete and gave an *R*-factor of 0.296 between 6 and 2.1 Å after 10 cycles of least-squares refinement. Further refinement proceeded in the usual manner with rounds of least-squares minimization followed by manual rebuilding of the model based on inspection of $2F_o$-F_c and F_o-F_c difference maps. Gradually, the residues missing initially from the model became visible in the difference maps and were added to the model. The resolution was extended to include all data. Ordered solvent molecules were included in the model. From refinement cycle 95 onwards contribution from bulk solvent was approximated in the refinement as a constant density in the solvent space. The method used was as described by Phillips (1980) and Fermi *et al.* (1984). The final value for the solvent

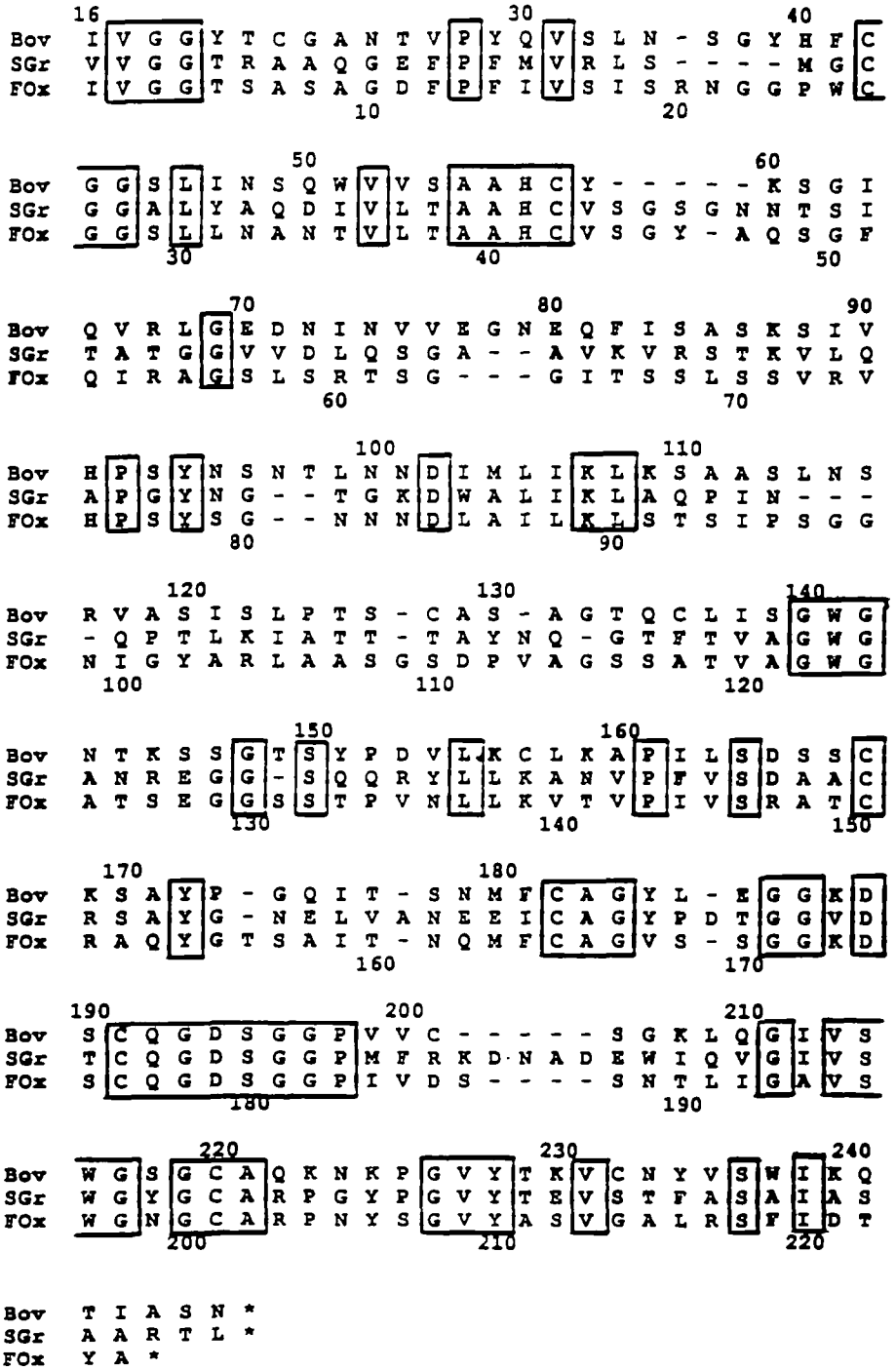

Fig. 6. The amino acid sequence alignment based on the superposed crystal structures of trypsins from *Bos taurus* (Bov) (Walsh and Neurath, 1964), *S.griseus* (SGr) (Olafson *et al.*, 1975) and *F.oxysporum* (FOx). The upper numerals show the sequence number of bovine chymotrypsinogen while the lower numerals indicate the sequence number of *F.oxysporum* trypsin. The conserved residues are indicated by boxes.

density ϱ_{solv} was 0.39 e/Å^3 and for the smoothing temperature factor B_{solv} it was 52 Å^2. Figure 5 summarizes the progress of the refinement. Table 1 lists the weighting parameters used in the refinement and the final standard deviations.

Results and discussion

The comparison of sequences based on superposition of the crystal structures of *F.oxysporum* trypsin with bovine and *S.griseus* trypsins (Figure 6) shows a homology of 43 and 44% respectively. The sequence alignment for trypsins from *F.oxysporum* and *S.erythraeus* (atomic coordinates are not available) gave a homology of 45%. The homology between the bovine and bacterial enzymes is 35%, while between the two bacterial proteins it is 41% (Figure 7). This simple comparison puts *F.oxysporum* somewhere between the two classes of organisms. This is easily understood in terms of the standard evolutionary model. No interspecies gene transfer mechanism is necessary to explain the homology between trypsins. The fact that other, related, serine proteinases frequently show less homology, even within the same organism, than trypsins between distantly related species, can be considered as a separate issue. Having different specificities, the other serine proteinases are in reality different proteins and are subject to different selection pressures or they may have arisen by different mechanisms, possibly via a 'silent gene' stage (Hartley, 1966, 1974).

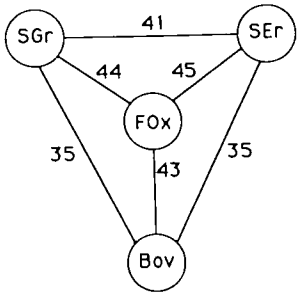

Fig. 7. A diagram illustrating the homology between the trypsins from *F.oxysporum* (FOx), *S.griseus* (SGr), *S.erythreaeus* (SEr) and bovine trypsin (Bov). In the case of FOx, SGr and Bov the homology was calculated based on superposed crystal structures. The homology between SEr and the other enzymes was based on optimal sequence alignment.

After a total of 142 cycles of refinement and 15 rounds of manual adjustment, the *R*-factor is 0.141 for all data to 1.8 Å. The Ramachandran plot (Ramakrishman and Ramachandran, 1965) is shown in Figure 8. The dihedral angles for nearly all the non-glycine residues lie within the permitted regions of the plot. Glycines, marked by squares, show a much greater conformational freedom. The mean coordinate error can be estimated from the distribution of the *R*-factor as a function of the resolution (Luzzati, 1952), as shown in Figure 9. The straight lines correspond to a series of mean coordinate errors derived from Luzzati's theory. The plot indicates a mean error in the coordinates ~0.15 Å. This is generally an overestimate of the true errors as the theory ignores errors in the diffraction data. The deviation at low resolution arises from the relatively poor model for the solvent structure.

The protein model is complete with almost all the protein atoms lying in a density greater than one standard deviation in the $2F_o$-F_c map. A representative region is shown in Figure 10. The mean temperature factor for all atoms was 17.8 Å^2. The mean

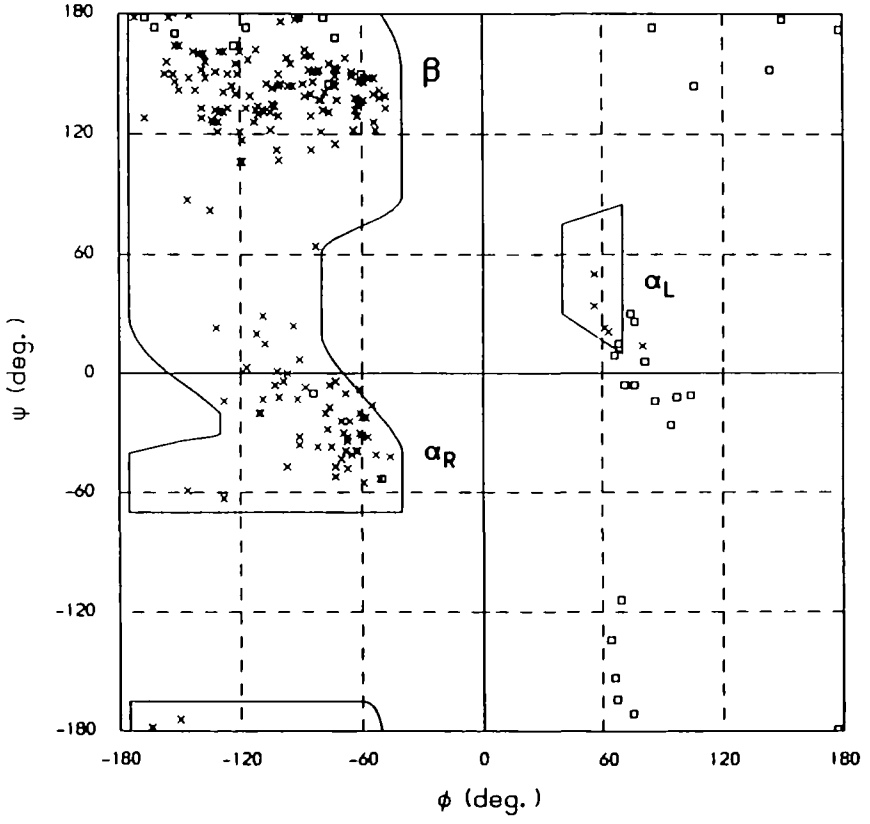

Fig. 8. The Ramachandran plot of the (ϕ,ψ) angles for the refined model of *F.oxysporum* trypsin. The permitted regions of conformational space for all residues other than glycine are shown as bounded parts. The regions for right-handed α-helices, β-pleated sheets and left-handed α-helices are indicated by α_R, β and α_L respectively. The angles for glycines are shown by squares, for all other residues as crosses.

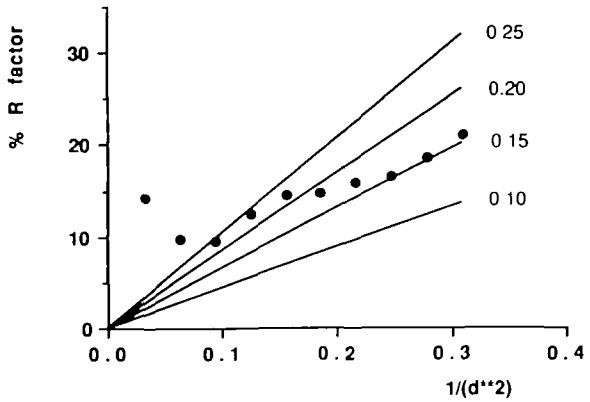

Fig. 9. Plot of the *R*-factor as a function of resolution. The straight lines show the theoretical dependence of the *R*-factor on resolution for the coordinate errors shown on the right (Luzzati, 1952).

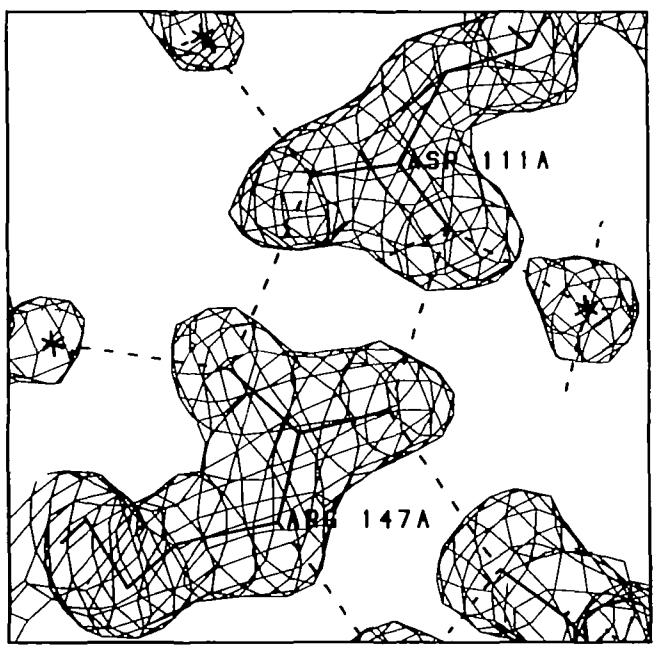

Fig. 10. Representative region of the final $2F_o$-F_c map showing a salt bridge between Arg147 and Asp111. Possible hydrogen bonds are drawn as dashed lines. This example shows a close to optimal utilization of hydrogen bonding potential.

values were 13.3, 15.7, 14.34 and 34.2 Å^2 for the main chain, side chains, all protein atoms and water molecules respectively. Standard deviations of the mean temperature factor per residue were 6.2 and 9.2 Å^2 for the main chain and side chains respectively. Only one residue (Ser170) has a mean *B* value for the main chain (32.3 Å^2) more than three standard deviations above the overall mean. Two side chains (Asn205 and Tyr223) also have mean temperature factors more than three standard deviations above the overall mean for the side chains (44.6 and 43.8 Å^2 respectively). Five side chains, Ser6 Ser19, Ser49, Ser170 and Asn177, have two alternative conformations, each with an occupancy of approximately 50%. Three of them (Ser6, Ser19 and Ser49) make clear hydrogen bonds with either protein atoms or water molecules. The region around Ser170 is less clearly defined. The Asn177 side chain is on the surface of the molecule with no clearly defined solvent. Several other side chains

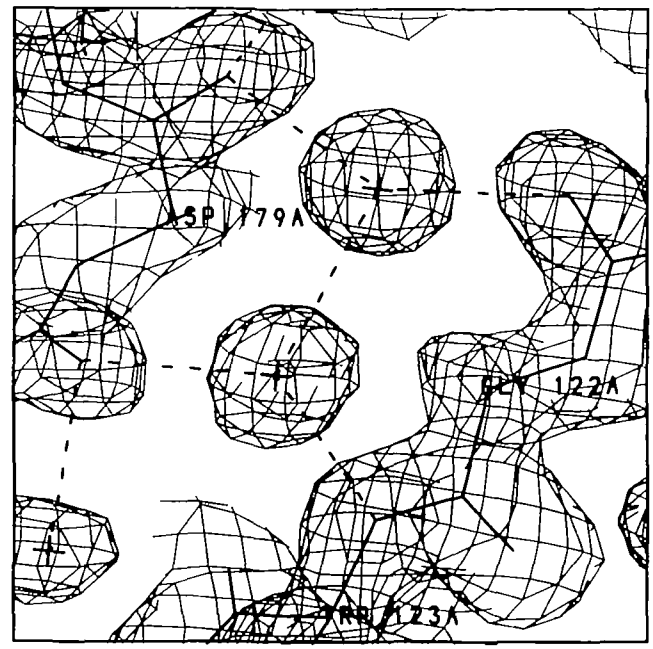

Fig. 11. Two internal water molecules. Most internal water molecules are very well defined with temperature factors similar to the temperature factors of protein atoms.

(Asn189, Asn205, Tyr223), show substantial disorder, reflected in their high temperature factors. Two of the nine arginines (Arg73 and Arg104), both lying on the surface of the molecule, are also poorly defined. The bound DIP (diisopropylphosphoryl) group is clearly visible attached to Ser180 and is included in the model. In the inhibitor molecule one isopropyl group has two alternative conformations. There is no density for the other isopropyl group (cf. Krieger *et al.*, 1974). A total of 329 solvent molecules have been included in the model. The water structure in various trypsins was recently analysed (Finer-Moore *et al.*, 1992) and the internal waters were found to be well conserved. Similarly, in the *F.oxysporum* trypsin 20 out of the 28 observed internal waters are conserved, as compared to bovine and *S.griseus* trypsins. Figure 11 shows an example of two tightly bound internal water molecules. Three of the internal waters that are not conserved fill the space that in bovine trypsin is occupied by the side chain of His40 and in *S.griseus* trypsin by the side chain of Arg32. Some surface water molecules are also conserved between the different structures.

There is no evidence of any divalent cation binding sites. The regions corresponding to the calcium binding sites in the other structures were compared to *F.oxysporum* trypsin. Figure 12 shows the comparison of the region where calcium binds in *S.griseus* trypsin. The site is clearly vacant in the bovine enzyme but in the *F.oxysporum* trypsin the space is filled by the side chain of Arg147, with one of the nitrogens occupying the position of calcium and making a hydrogen bond with the carbonyl oxygen of Met163. In addition, the arginine forms a salt bridge with Asp111 (Figure 10), further stabilizing the local conformation. The site where calcium binds in bovine trypsin is filled in *F.oxysporum* trypsin by the side chain Ser57 and the carbonyl oxygen of the main chain of Ser62. The corresponding region in *S.griseus* trypsin, which also does not contain calcium, is harder to compare owing to its considerably different main chain conformation.'

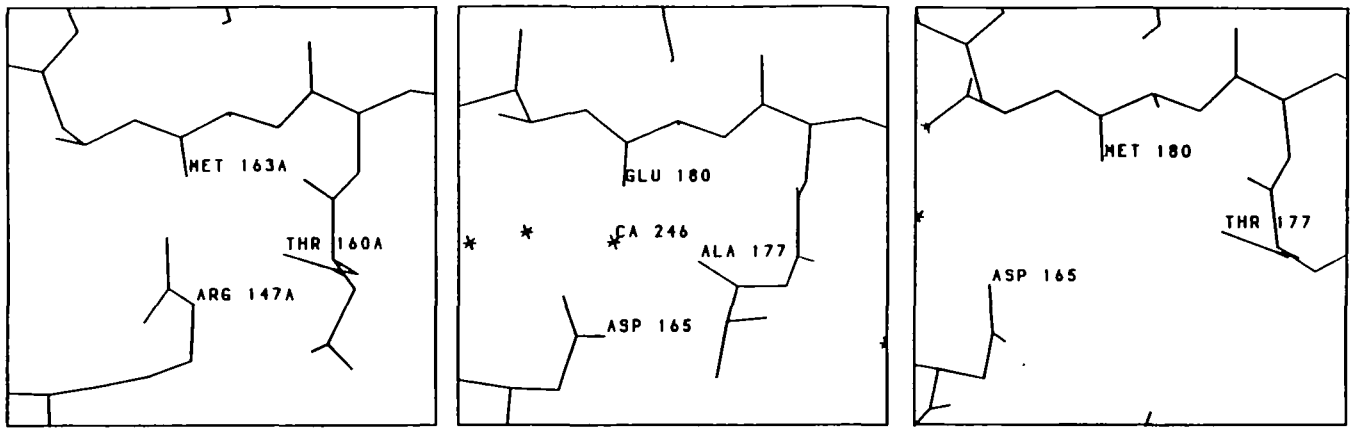

Fig. 12. Comparison of the calcium binding site in *S.griseus* trypsin (middle) with the corresponding regions in *F.oxysporum* trypsin (left) and bovine trypsin (right). See text for discussion.

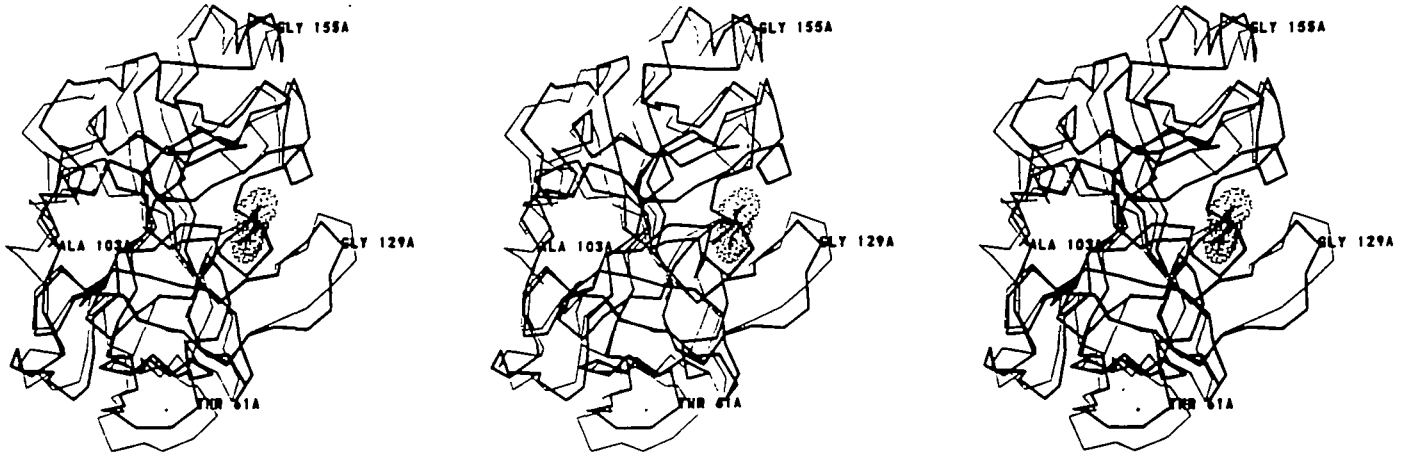

Fig. 13. Superposition of the refined model of *F.oxysporum* trypsin (thick lines) and bovine trypsin (thin lines). The molecules were superposed using the conserved residues. Only α-carbons are shown. A particularly large difference can be seen in the loop of the calcium binding site in bovine trypsin. The inhibitor molecule, bound in the active site of the *F.oxysporum* trypsin, is represented by the dotted regions.

Comparison of *F.oxysporum* trypsin with other trypsins shows that while the structures are very similar near the active and binding sites and in the core of the protein, there are considerable differences in other regions, particularly in the loops on the surface of the molecule. Figure 13 shows the superposition of α-carbons of the final model of *F.oxysporum* trypsin and bovine trypsin.

Acknowledgement

M.D. was supported by a grant from NOVO-Nordisk.

References

Abola,E.E., Bernstein,F.C., Bryant,S.H., Koetzle,T.F. and Weng,J. (1987) In Allen,F.H., Bergerhoff,G. and Sievers,R. (eds), *Crystallographic Databases—Information Content, Software Systems, Scientific Applications*. Data Commision of the International Union of Crystallography, Bonn/Cambridge/Chester, pp. 107–132.

Bernstein,F.C., Koetzle,T.F., Williams,G.J.B., Meyer,E.F.,Jr, Brice,M.D., Rodgers,J.R., Kennard,O., Shomanouchi,T. and Tasumi,M. (1977) *J. Mol. Biol.*, **112**, 535–542.

Blevins,R.A. and Tulinsky,A. (1985) *J. Biol. Chem.*, **260**, 4264–4275.

Chambers,J.L. and Stroud,R.M. (1977) *Acta Crystallogr.*, **B33**, 1824–1837.

Chambers,J.L. and Stroud,R.M. (1979) *Acta Crystallogr.*, **B35**, 1861–1874.

Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1978) *Biochemistry*, **18**, 5294–5299.

Christensen,T., Woeldike,H., Boel,E., Mortensen,S.B., Hjortshoej,K., Thim,L. and Hansen,M.T. (1988) *Biotechnology*, **6**, 1419–1422.

Delbaere,L.T.J. and Brayer,G.D. (1985) *J. Mol. Biol.*, **183**, 89–103.

Earnest,T., Fauman,E., Craic,C.S. and Stroud,R. (1991) *Proteins*, **10**, 171–187.

Fermi,G., Perutz,M.F., Shaanan,B. and Fourme,R. (1984) *J. Mol. Biol.*, **174**, 159–174.

Finer-Moore,J., Kossiakoff,A.A., Hurley,J.H., Earnest,T. and Stroud,R.M. (1992) *Proteins*, **12**, 203–222.

Fujinaga,M., Sielecki,A.R., Read,R.J., Ardelt,W., Laskowski,M.Jr and James,M.N.G. (1987) *J. Mol. Biol.*, **195**, 397–418.

Hartley,B.S. (1966) *Adv. Sci., Lond.*, May, 47–54.

Hartley,B.S. (1970) *Phil. Trans. R. Soc. Lond.*, **B257**, 77–87.

Hartley,B.S. (1974) In Carlile,M.J. and Skehel,J.J. (eds), *Evolution in the Microbial World* (Soc. Gen. Microbiol. Symp. no. 24). Cambridge University Press, pp. 151–182.

Hartley,B.S. (1979) *Proc. R. Soc. Lond.*, **B205**, 443–452.

Hendrickson,W.A. and Lattman,E.E. (1970) *Acta Crystallogr.*, **B26**, 136.

Herzberg,O. and Sussman,J.L. (1983) *J. Appl. Cryst.*, **16**, 144–150.

Hewett-Emmett,D., Czelusniak,J. and Goodman,M. (1981) Ann. NY Acad. Sci., **370**, 511–527.

Kornert,J.H. (1976) *Acta Crystallogr.*, **A32**, 614–617.

Kornert,J.H. and Hendrickson,W.A. (1980) *Acta Crystallogr.*, **A36**, 344–350.

Kraut,J. (1977) *Ann Rev. Biochem.*, **46**, 331–358.

Krieger,M., Kay,L.M. and Stroud,R.M. (1974) *J. Mol. Biol.*, **83**, 209–230.

Leslie,A.G.W., Brick,P. and Wonacott,A.J. (1986) *CCP4 Newslett.*, **18**, 33–39.

Luzzati,V. (1952) *Acta Crystallogr.*, **5**, 802–810.

Matthews,B.W. (1968) *J. Mol. Biol.*, **33**, 491–498.

Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499.

Meyer,E., Cole,G. and Radhakrishnan,R. (1988) *Acta Crystallogr.*, **B44**, 26–38.

Miyamoto,K., Matsuo,H. and Narita,K. (1979) In Dai 30 Kai Tanpakushitsu

Kouzou Touronkai Kouen Youshishuu (in Japanese), 77–80.
Mortensen,S.B., Thim,L., Christensen,T., Woeldike,H., Hjortshoej,K. and Hansen,M.T. (1989) *J. Chromatogr.*, **476**, 227–233.
Navia,M.A., McKeever,B.M., Springer,J.P., Lin,T.-Y., Williams,H.R., Fluder,E.M., Dorn,C.P. and Hoogsteen,K. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 7–11
Okayama,H. and Berg,P. (1982) *Mol. Cell. Biol.*, **2**, 161–170.
Olafson,R.W., Jurasek,L., Carpenter,M.R. and Smille,L.B. (1975) *Biochemistry*, **14**, 1168–1177.
Phillips,S.E.V. (1980) *J. Mol. Biol.*, **142**, 531–554.
Ramakrishnan,C. and Ramachandran,G.N. (1965) *Biophys. J.*, **5**, 909–933.
Reed,R.J. (1986) *Acta Crystallogr.*, **A42**, 140–149.
Reed,R.J and James,M.N.G. (1988) *J. Mol. Biol.*, **200**, 523–551.
Sawyer,L., Shoton,D.M., Campbell,J.W., Wendell,P.L., Muirhead,H., Watson,H.C., Diamond,R. and Ladner,R.C. (1978) *J. Mol. Biol*, **118**, 137–208.
Steitz,T.A. and Shulman,R.G. (1982) *Ann. Rev. Biochem. Biophys.*, **11**, 419–444.
Sussman,J.L., Holbrook,R.S., Church,G.M. and Kim,S.-H. (1977) *Acta Crystallogr.*, **A33**, 800–804.
Truelsen,E., Gausing,K., Jochimsen,B., Jurgensen,P. and Marcker,K.A. (1979) *Nucleic Acids Res.*, **6**, 3061–3072.
Tsukada,H. and Blow,D.M. (1985) *J. Mol. Biol.*, **184**, 703–711.
von Heijne,G. (1986) *Nucleic Acids Res.*, **14**, 4683–4690.
Walsh,K.A. and Neurath,H. (1964) *Proc. Natl Acad Sci. USA*, **52**, 884.
Wang,D. Bode,W. and Huber,R. (1985) *J. Mol. Biol.*, **185**, 595–624.
Wilson,A.J.C. (1942) *Nature*, **150**, 151–152.
Yamane,T., Kobuke,M., Tsutsui,H., Toida,T., Suzuki,A., Ashida,T. Kawata,Y. and Sakiyama,F. (1991) *J. Biochemistry*, **110**, 945–950.

Received October 30, 1992; revised March 2, 1993; accepted March 10, 1993