Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Extraction of functional motion in trypsin crystal structures

Andrea Schmidt* and Victor S. Lamzin

European Molecular Biology Laboratory (EMBL), Hamburg Outstation, c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany

Correspondence e-mail: andrea@embl-hamburg.de

Received 6 April 2005
Accepted 25 May 2005

PDB References:
F. oxysporum trypsin, room temperature, 1xvm, r1xvmsf; pH 6, 1xvo, r1xvosf.

The analysis of anisotropic atomic displacement parameters for the direct extraction of functionally relevant motion from X-ray crystal structures of *Fusarium oxysporum* trypsin is presented. Several atomic resolution structures complexed with inhibitors or substrates and determined at different pH values and temperatures were investigated. The analysis revealed a breathing-like molecular motion conserved across trypsin structures from two organisms and three different crystal forms. Directional motion was observed suggesting a change of the width of the substrate-binding cleft and a change in the length of the specificity pocket. The differences in direction of motion across the structures are dependent on the mode of substrate or inhibitor binding and the chemical environment around the active-site residues. Together with the occurrence of multiple-residue conformers, they reflect spatial rearrangement throughout the deacylation pathway.

1. Introduction

A protein molecule is never rigid and its functional motion is preserved to a considerable extent in the crystalline state. High-resolution X-ray crystal structures can provide information on the directional motion of individual atoms in the molecule in the form of anisotropic atomic displacement parameters (anisotropic ADPs; Dunitz, Schomaker *et al.*, 1988). This prompted us to extract the dynamic properties of *Fusarium oxysporum* trypsin (EC 3.4.21.4) from its crystal structures. Trypsin is a very flexible enzyme that displays an induced-fit mode of substrate binding (Schmidt *et al.*, 2003). It forms stable complexes with a variety of substrates and inhibitors (Perona, Craik *et al.*, 1993; Wilmouth *et al.*, 2001) and yields crystals that diffract to atomic resolution. Trypsin is a serine protease that cleaves oligopeptides at the C-terminal side of arginines or lysines in two steps, termed 'acylation' (cleavage and covalent attachment of the N-terminal part of the substrate to a catalytic serine) and 'deacylation' (hydrolysis of the serine–substrate bond and product release).

Many approaches have been tried to collect snapshots of catalytic reactions from crystal structures of biological macromolecules. Although applications of classic time-resolved X-ray crystallography are limited, a number of spectacular results have been achieved, *e.g.* on the characterization of ligand-binding intermediates of myoglobin by application of both monochromatic and Laue X-ray techniques (Schlichting *et al.*, 1990; Schlichting & Chu, 2000). It was shown that catalysis does occur within a protein crystal and that the crystal structure can indeed be used to map such dynamic processes. Another elegant way of retrieving information on dynamics was shown for the crystal structure of

Table 1
Summary of the structures under study.

Values in parentheses are for the outer shell.

	Substrate complexes				Inhibitor complexes	
	ROOM	PH4	PH5	PH6	DFP	PMSF
Structure details						
PDB code	1xvm	1pq8	1pq5	1xvo	1ppz	1pqa
Crystallization pH	5.0	4.0	5.0	6.0	5.0	5.0
Temperature of data collection (K)	298	100	100	100	100	100
Data statistics						
Space group	*P*1	*P*1	*P*1	*P*1	$P2_1$	$P2_1$
Resolution range (Å)	10–1.1 (1.11–1.10)	20–1.0 (1.01–1.00)	35–0.85 (0.86–0.85)	22–0.84 (0.85–0.84)	35–1.22 (1.23–1.22)	25–1.23 (1.24–1.23)
No. of unique reflections	65556	86657	137323	142228	46357	44082
Multiplicity	2	3	2	2	4	4
R_{sym} (%)	9.4 (36.1)	7.0 (38.5)	3.3 (10.1)	4.6 (24.7)	4.5 (21.0)	4.6 (32.8)
$I/\sigma(I)$	6.9 (2.6)	11.0 (3.8)	34.0 (7.2)	31.9 (3.5)	28.0 (5.5)	11.0 (2.7)
Completeness (%)	92.5 (92.4)	94.9 (92.6)	92.1 (80.6)	92.1 (80.0)	99.6 (97.7)	93.8 (76.3)
Overall isotropic displacement parameter† (Å^2)	9.0	7.1	5.0	4.6	11.3	11.6
Refinement statistics						
R value (%)	14.4	12.8	9.8	10.1	14.1	14.1
No. of atoms in refinement	1713	2127	2229	2140	1978	1935
No. of solvent and heteroatoms	131	355	494	535	310	296
Mean isotropic *B* value for protein atoms	15.1	9.8	6.4	6.2	14.6	15.3
Mean anisotropy for protein atoms	1.45	1.46	1.46	1.39	1.61	1.70
Residues with alternate main-chain conformations	4 (2%)	40 (18%)	33 (15%)	19 (8%)	27 (12%)	16 (7%)
Substrate/inhibitor	Arginine	Arginine	Arginine	—	DFP	PMSF
Mode of substrate/inhibitor binding‡	Cov	Non	Non	—	Cov	Cov

† From the Wilson plot (Wilson, 1942). ‡ Non, non-covalent interaction with the catalytic serine; Cov, covalent interaction with the catalytic serine.

horseradish peroxidase, in which the catalytic reaction was induced by X-ray radiation and the reaction intermediates could be characterized (Berglund *et al.*, 2002).

Motion within a macromolecule in a crystal can to an extent be modelled by the so-called translation–libration–screw (TLS) refinement, describing either the rigid-body movement of regions in the molecule or the molecule as a whole (Winn *et al.*, 2001; Schomaker & Trueblood, 1968; Dunitz, Maverick *et al.*, 1988; Harata & Kanai, 2002). The calculated TLS parameters always contain contributions from both crystalline lattice disorder and lattice vibrations, which cannot as yet be reliably deconvoluted (Harata *et al.*, 1999). Atomic motion can be modelled using atomic displacement parameters that complement the positional three-dimensional coordinates and describe a probability density distribution for the location of the atom (Dunitz, Maverick *et al.*, 1988; Dunitz, Schomaker *et al.*, 1988; Cruickshank, 1956). Where the atomic displacement is not restricted to the simplified spherical model, 'thermal ellipsoids' are used to describe the anisotropic ADPs. This ADP tensor is defined by the six unique elements of a symmetric 3×3 matrix that determines the orientation and the lengths of the three principal axes of the thermal ellipsoid. The anisotropic ADPs can only be reliably derived if the X-ray data extend to sufficiently high resolution. Although in many cases they have so far only been used for a more precise fit of the structure to the data, considerable efforts have recently been invested in a fuller exploitation of directional motion in macromolecular crystallography (see, for example, Bernett *et al.*, 2004).

Here, we present a thorough exploitation and analysis of the information stored in the individual anisotropic ADPs for several X-ray crystal structures of *F. oxysporum* trypsin at atomic resolution. The main goal is to gain insight into the driving forces for substrate binding and the development of electrostatic charges during catalysis that still remain a mystery and a subject of active research (Schmidt *et al.*, 2003; Perona, Craik *et al.*, 1993; Wilmouth *et al.*, 2001; Rypniewski *et al.*, 2001; Perona, Tsu *et al.*, 1993; Tsilikounas *et al.*, 1992; Frey *et al.*, 1994) and to advance a technique for the extraction of molecular motion in crystal structure analysis.

2. Experimental procedures

2.1. Data collection and processing

Protein crystals were produced as described elsewhere (Rypniewski *et al.*, 1993). X-ray data for the ROOM and PH6 structures were collected at the EMBL Hamburg beamlines X11 and X13 equipped with a MAR CCD detector, single-axis goniostat and an Oxford 600 cryosystem. The crystals were briefly transferred into paraffin oil as a cryoprotectant prior to data collection. For the room-temperature experiment, the crystal was mounted in a 0.5 mm glass capillary directly from the drop. Data collection was carried out in two passes over 200° each: a high-resolution pass with a step width of 0.5° to the maximum resolution and a low-resolution pass with a step width of 1.5° to a cutoff of ∼2 Å. An exception was the ROOM structure, for which only one pass was used. None of

the data showed a detectable amount of radiation damage. Data were integrated using *DENZO* and scaled isotropically using *SCALEPACK* (Otwinowski & Minor, 1997). Intensities were truncated to structure-factor amplitudes for refinement in *REFMAC* (Murshudov *et al.*, 1997) using *TRUNCATE* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The structures PH4, PH5, DFP and PMSF have been determined previously (Schmidt *et al.*, 2003) and are available in the PDB (Bernstein *et al.*, 1977; Berman *et al.*, 2000). Crystallographic details are summarized in Table 1.

2.2. Refinement

The structure PH5 served as a template. After a few rounds of rigid-body refinement, the models were completed using *ARP*/*wARP* and *REFMAC* (Perrakis *et al.*, 1999; Murshudov *et al.*, 1997; mode 'solvent'). A randomly chosen 5% of the data set were used for cross-validation using $R_{\rm free}$ at this step. Double conformers and the substrate (where present) were included at this stage and inserted using the graphics program *O* (Jones *et al.*, 1991). Solvent molecules with displacement parameters higher than 35 Å^2 were removed before each building cycle. When the model was complete in terms of number of atoms (*i.e.* equal numbers of water atoms added/removed and converged *R* factors), the structures were finally refined with *SHELXL* (Sheldrick, 1997) against all data (no $R_{\rm free}$) of diffraction intensities including anisotropic atomic displacement parameters and riding H atoms. Default *SHELXL* restraints and weights were used for the refinement of the atomic thermal parameters. Water molecules with isotropic ADPs between 45 and 55 Å^2 were assigned occupancies of 0.5 and those with values higher than 55 Å^2 were removed. Occupancies of double conformers were not refined but set so that their respective isotropic ADPs were about equal. The electron-density maps were inspected using the graphics program and multiple conformations and solvent atoms were corrected manually where needed. The two new structures ROOM and PH6 have good overall agreement with the target stereochemistry. The r.m.s.d.s in bonded distances are 0.022 and 0.014 Å, respectively (target value 0.020). The r.m.s.d.s in angle bonded distances are 0.030 and 0.029 Å^2, respectively, with a target value of 0.040. The same refinement protocol was applied to all structures under study to ensure consistency in the comparison.

2.3. ADP analysis

The orientation and lengths of the principal axes of thermal motion were determined from the eigenvectors and eigenvalues of the U_{ij} tensor of the individual atoms in the ortho-

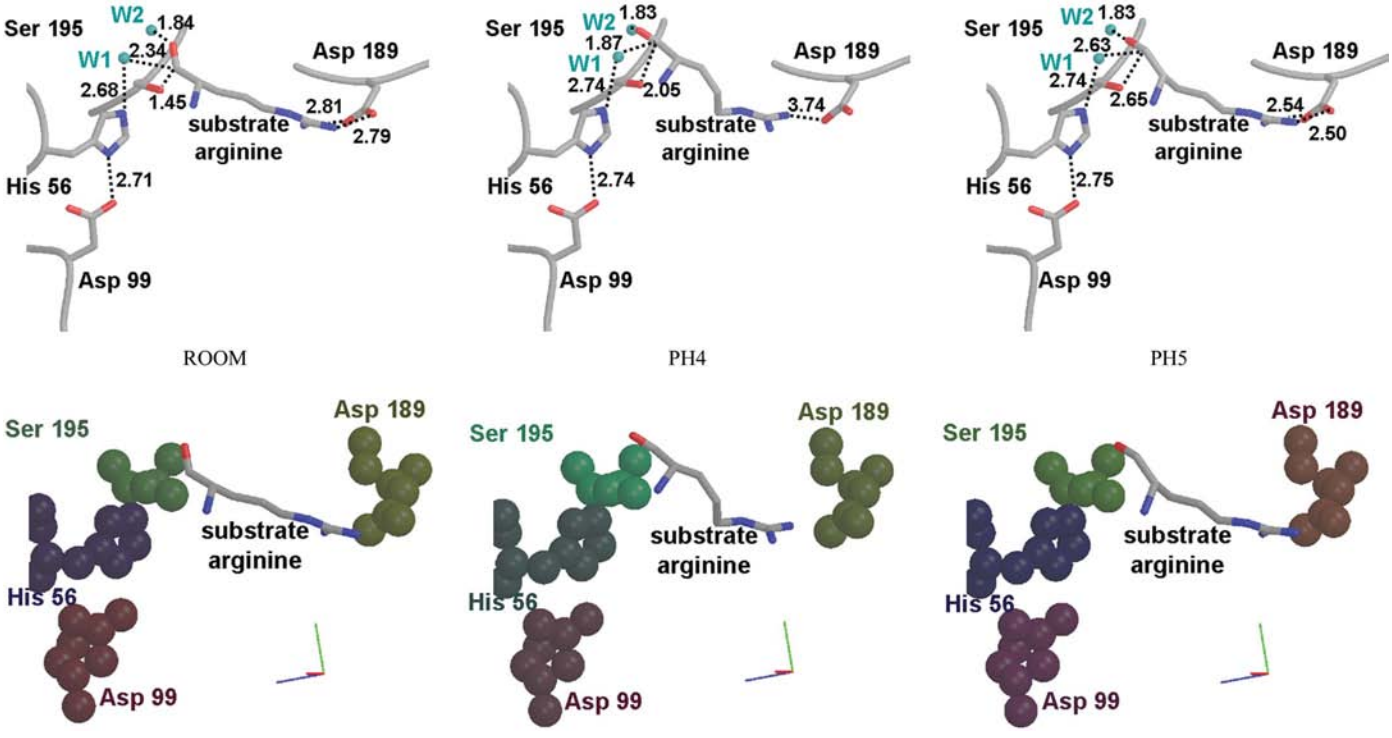

Figure 1
Active site in *F. oxysporum* trypsin. The top row shows the change in the interaction between the substrate and the catalytic serine and histidine, as well as Asp189 in the specificity pocket in the ROOM, PH4 and PH5 structures. The distances of interest around the substrate carbonyl group and in the specificity pocket are shown by dotted lines with their lengths given in angstroms. W1 and W2 are two water molecules located in the active site: W1 acts as the nucleophile and W2 as an activator by forming a strong hydrogen bond to the substrate carbonyl O atom. In the bottom row, the same arrangement of residues is shown. The protein atoms and coordinate axes are represented and coloured following the same scheme as in Fig. 3 to indicate their directional motion. The orientation is the same as in Fig. 3. The figure was created with *MOLSCRIPT*/*RASTER3D* (Kraulis, 1991; Merritt & Murphy, 1994).

gonal coordinate frame. The longest vector was chosen as that describing the principal motion of a given atom. The three ratios $R(i) = L_{max}/L(i)$ between the length of the longest vector and the other vector lengths were calculated. In this analysis, the ratio of the longest and shortest length is hereafter defined as the anisotropy of that atom. Atoms with differences in the axes lengths of 5% or less were flagged as 'spherical' and excluded from the analysis in order to avoid ambiguities in the assignment of direction of motion. For visualization, the vector components XYZ of the longest vector were translated into RGB colour codes: R, G and B values were assigned to the X, Y and Z values, respectively. In order to ensure maximum brightness of the colours, the largest vector component was set to 100% colour-component contribution (RGB value of 255 or 1.0). The other colour components were calculated relative to that value. In addition, the resulting RGB codes were scaled by a factor $F = 1 - 1/[R(1)R(2)R(3)]$, reflecting the anisotropy. This colour scheme yields the direction of motion from the hue and allows distinction between highly elongated (bright; two ratios >> 1.0), oblate spheroid (medium; 'pancake' shape: two ratios close to 1.0) and spherical (dark; all three ratios $\simeq$ 1.0) atoms.

3. Results

3.1. Structures under study

Several crystal structures of trypsin from *F. oxysporum* were analysed. They were determined to atomic resolution at three different pH values and two temperatures (Table 1). Four trypsin structures determined previously (PDB codes 1pq8, 1pq5, 1ppz and 1pqa; Schmidt *et al.*, 2003) have been complemented with two new structures, ROOM and PH6. The six *F. oxysporum* trypsin structures occur in different crystal forms. The change in the space group from P_1 to $P2_1$ is achieved by the rotation of every second layer in the P_1 lattice, thereby creating the screw dyad and almost doubling the length of the crystallographic b axis.

The structures can be grouped as follows (Table 1). The 'natural' structures (hereafter called substrate complexes) contain either an empty active site (PH6) or an arginine molecule as a remnant of the natural substrate (ROOM, PH4 and PH5). The second group is formed of complexes with inhibitors covalently attached to the catalytic serine (PMSF, DFP). The substrate complexes crystallized in space group $P1$ and showed binding states ranging from no substrate in PH6, non-covalently bound substrate in PH5 and PH4 to covalently

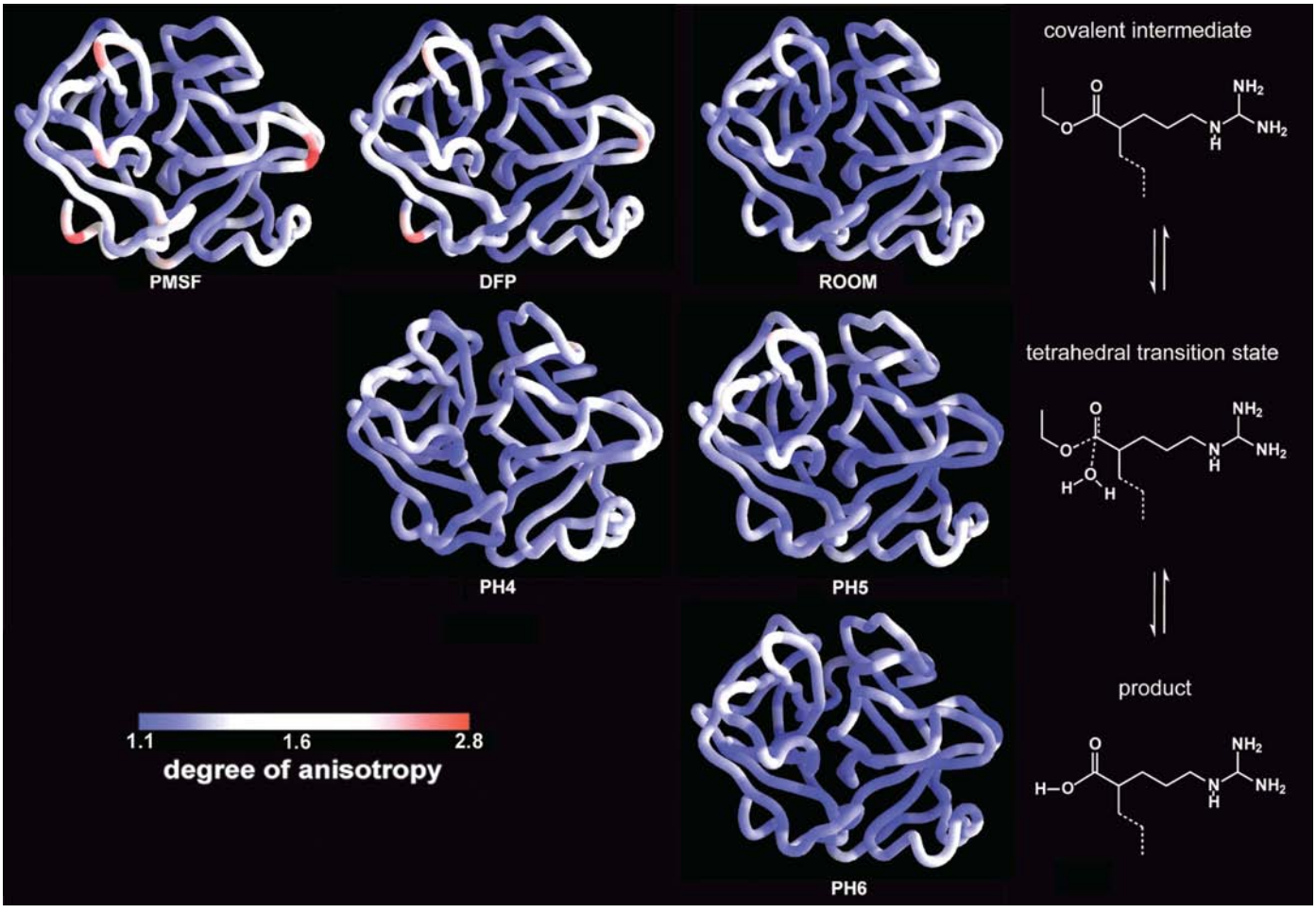

Figure 2
Trypsin structures with the degree of anisotropy (average per residue) mapped onto the C^{α} trace. The figure was produced with *GRASP* (Nicholls *et al.*, 1991). The deacylation pathway is depicted on the right-hand side of the figure.

attached substrate in ROOM (Fig. 1). The inhibitor complexes crystallized in a different space group, $P2_1$.

3.2. Mobility in the rigid crystal

The six structures of trypsin show varying conformational flexibility (Table 1) and also display pronounced differences in the mode (Fig. 2) and direction of motion (Fig. 3). Atoms within or close to the substrate-binding cleft tend to move less and in a more isotropic manner (although no atoms fulfilled the criteria for being 'spherical' as defined in §2.3). In the regions remote from the active site, especially in exposed loops, the shape of the ADPs becomes more elongated. The anisotropy, defined as the ratio between the longest and shortest lengths of the eigenvectors of the anisotropic ADPs, is expected to increase with the distance of atoms to the centre of mass (Schneider, 1996), but this varies from structure to structure and reflects the binding state of the ligand (Fig. 2). The structures PMSF and DFP with covalently attached inhibitors, ROOM with a covalently bound substrate and PH6 with no substrate display a moderate number of residues in alternate main-chain conformations. The anisotropy in the residues constituting the active site and the specificity pocket is generally increased when covalent attachment is observed. There is distinct concerted directional motion around the specificity pocket (Fig. 3), where many of the highly anisotropic atoms are located. The non-covalent substrate complexes (PH4, PH5) have a generally higher number of multiple conformers, lower anisotropy in the active site and different direction of atomic motion (Figs. 2 and 3). Overall, the number of alternate main-chain conformers, anisotropy in the active site and the directions of atomic motion gradually change starting from ROOM to PH4 and PH5 and finally ending at PH6, mapping the pathway from the covalent intermediate to the tetrahedral transition state and to product release (Fig. 3).

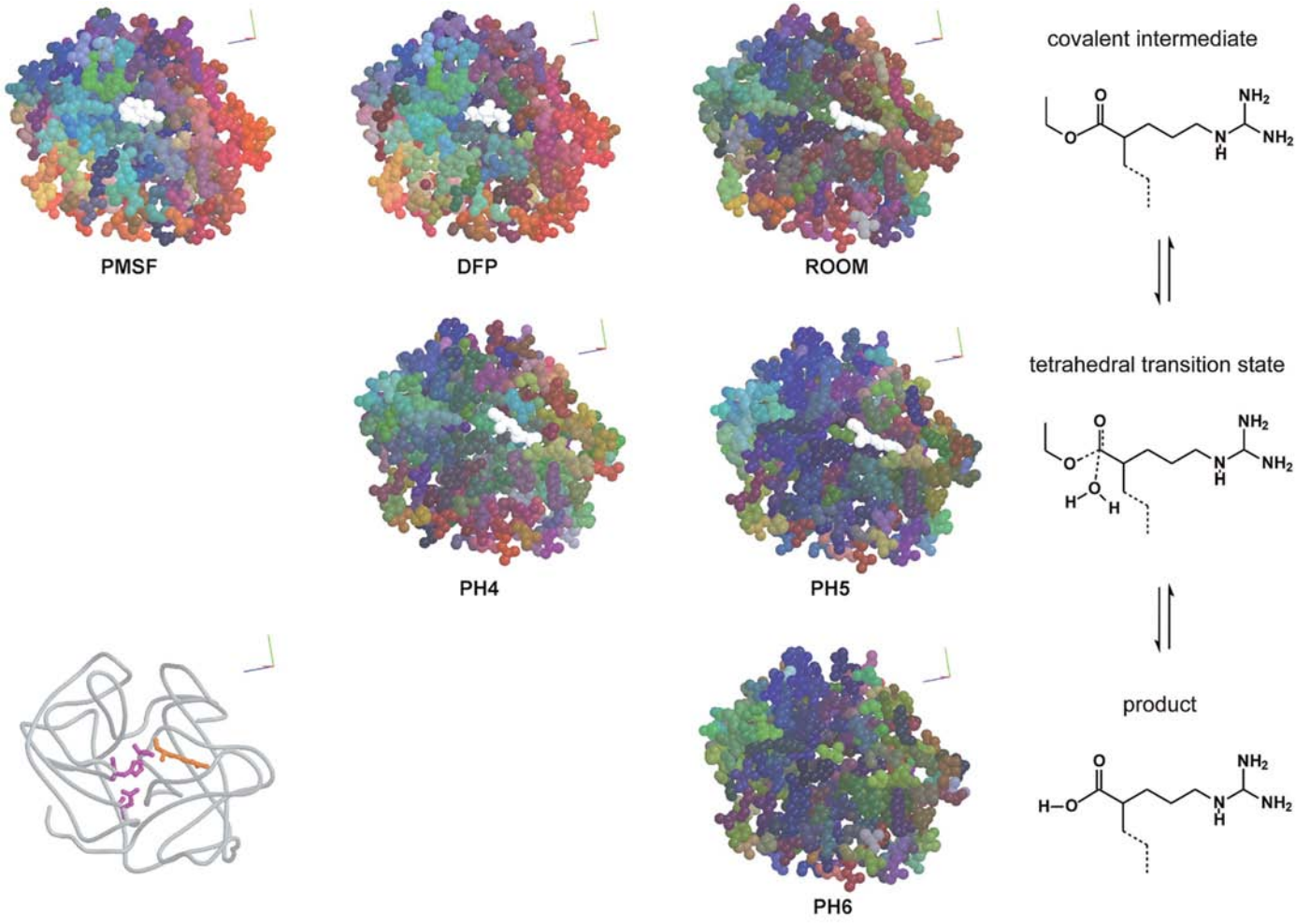

Figure 3
Trypsin structures representing the different binding states and the motion connected with them. The loop surrounding the specificity pocket (right-hand side of the molecule) shows the most pronounced changes in the direction of motion. The substrate-binding cleft (centre of the molecule) changes primarily its degree of anisotropy rather than the direction. Atoms are represented as 1 Å spheres with RGB colours generated from the direction of movement, where the atomic ADP tensors were averaged for each residue. Brighter colours indicate higher anisotropy. The coordinate axes are represented in the same colour code. See §2 for the description of the RGB colour scheme. The substrate or inhibitors are drawn in white. The grey backbone diagram shows the location of the active-site residues (magenta) and the substrate (orange). The figure was created with *MOLSCRIPT*/*RASTER3D* (Kraulis, 1991; Merritt & Murphy, 1994).

4. Discussion

4.1. Directed motion and induced fit model

Despite the fact that the *F. oxysporum* trypsin structures crystallize in two different space groups, there is an overall pattern of directional motion that is common to all six of them and that is apparently not affected by the crystal lattice contacts. The molecular motion, the direction of which is determined by its sign, can be interpreted as the opening of a hand in which the active site would be located at the centre of the palm. This leads to a breathing-like movement by changing the width of the active-site cleft, the specificity pocket and the oxyanion hole. This observation is in agreement with the presence of multiple conformers related to the occupancy of the substrate in the active site (Schmidt *et al.*, 2003) and supports spectroscopic studies (Tsilikounas *et al.*, 1992) and molecular-dynamics simulations (Topf, Varnai, Schofield *et al.*, 2002; Brandsdal *et al.*, 2001) on other serine proteases which showed that the active site is not rigidly pre-formed but rather displays a high degree of flexibility. The residues at the bottom of the substrate-binding cleft in the *F. oxysporum* trypsin structure remain relatively rigid and change their degree of atomic anisotropy rather than the direction of atomic motion (bright turquoise shades to dark blue in Fig. 3).

The differences in the direction of atomic motion between the structures reveal fine details of molecular flexibility. In the substrate complexes, which have the same crystalline arrangement, the atomic motion in the surroundings of the specificity pocket changes from an out-of-the-paper plane (red and brown shades) to an in-plane (green shades) direction (Fig. 3). The direction of motion apparently depends on the binding state and positioning of the substrate arginine (Figs. 1 and 3). These structures seem to represent binding states ranging from the covalent intermediate and the tetrahedral transition state to the empty state after product release. This is also supported by the characteristics of substrate binding depicted in Fig. 1.

In the inhibitor complexes, which crystallize in another space group, the motion in the active-site cleft is different compared with the substrate complexes. A pronounced movement emerges that folds the entire loop surrounding the specificity pocket 'upwards' and towards the substrate-binding cleft. The observed overall motion in DFP and PMSF is not unrelated to the directional motion in the PH6, PH5, PH4 and ROOM structures despite the fact that these two groups of structures belong to different crystal forms.

In addition to Fig. 3, which shows the RGB colour type representation of the residue-based molecular motion, Fig. 4 displays the averaged directions of overall molecular motion. The same ADP analysis was applied as for Fig. 3. A clear change occurs along the transition from ROOM (covalent intermediate) to PH6 (empty) structures. From these changes one could even hypothesize that the PH4 structure corresponds to a state before the transition state, while the PH5 structure represents a state closer to product release in the deacylation pathway. The overall direction of the motion in the inhibitor structures is almost orthogonal to that in the substrate complexes. Clarification of to what extent this could relate to a different lattice environment requires further investigation.

The above led us to look at the influence of the lattice contacts on the overall concerted movements in trypsin. We chose the atomic resolution structure of bovine trypsin (BOV; PDBcode 1h9j; Leiros *et al.*, 2001; 42% sequence identity to the *F. oxysporum* trypsin), which crystallized in space group $P2_12_12_1$. The specificity pocket in the BOV structure is occupied by an aniline molecule, but without direct interaction between the ligand and the catalytic serine. We applied the ADP analysis to the bovine trypsin structure as to the six *F. oxysporum* structures and detected a similar breathing-like motion (result not shown) in the structures from both organisms. We thus conclude that the presented ADP analysis is able to reveal general features of the molecular motion of trypsin and that, at least at the level of the whole macromolecule, the results can be compared across different lattice environments.

4.2. Implications for catalysis

Active sites in serine proteases reach their catalytically active conformation upon substrate binding (Tsilikounas *et al.*, 1992) and the shape and properties of the ligand binding into the specificity pocket has a major influence (up to 70%) on the conformation and binding energetics (Brandsdal *et al.*, 2001). The different directions of motion observed in the six *F. oxysporum* trypsin structures are related to the binding state of the substrate. This suggests spatial rearrangement upon substrate binding, which is also reflected in the occurrence of multiple main-chain conformers The relation between the substrate-binding state and the presence of alternate conformers has already been shown for *F. oxysporum* trypsin (Schmidt *et al.*, 2003).

A plausible reason for the changes in direction of movement along the reaction pathway could be steric strain and the generation of charges in the active site during the catalytic

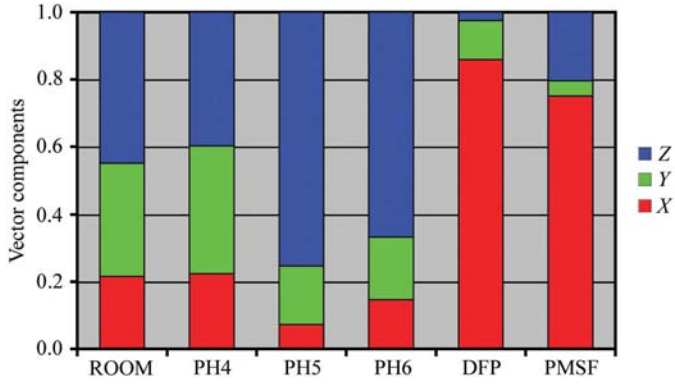

Figure 4
Overall direction of molecular motion. The ADP tensors were averaged over all protein atoms in each molecule. The *XYZ* components of the resulting principal direction of motion of the whole molecule are presented in the RGB colour scheme on a relative scale so that they sum up to 1.

reaction (Shan & Herschlag, 1996). While the electrostatic effects are expected to have most influence around the catalytic serine and at the bottom of the specificity pocket (Perona, Tsu *et al.*, 1993), steric effects are expected to occur mainly in and around the specificity pocket. The motion deduced from the ADP analysis indicates a change in the length of the specificity pocket and thereby implies a shift of the bound substrate with respect to the active-site serine. This is confirmed by the different contact distances of the substrate arginine in the active site in the ROOM, PH4 and PH5 structures (Fig. 1). The elongation of the covalent bond between the substrate carbonyl C atom and the catalytic serine O^{γ} is necessary for the deacylation process.

In the inhibitor structures (PMSF and DPF), the increased anisotropy in the active site hints at steric strain that could be induced by the bulky phosphate or sulfate group attached to the catalytic serine. While stereochemically these groups mimic the tetrahedral transition state, their binding properties place them closer to the covalent intermediate and they induce different electrostatic charges, which is reflected in the different properties of the anisotropic ADPs. For example, the loop around the specificity pocket has slightly changed the direction of its principal motion from PMSF to DFP (Fig. 3). It is also noticeable that the anisotropy in the inhibitor structures is not only increased in the active site, but throughout the whole molecule. This is visualized in colour in Figs. 2 and 3 and is also indicated by the values of overall anisotropy in Table 1.

In the substrate structures (ROOM, PH4, PH5) the specificity pocket is occupied by a bulky arginine side chain with a positively charged guanidinium group (Fig. 1). This positive charge is compensated for by a salt bridge with an aspartate side chain at the bottom of the specificity pocket. Relocation of the arginine side chain, caused by the attack of the nucleophilic water molecule and the break of the covalent bond to the serine, will consequently result in a rearrangement of the residues in the specificity pocket through the salt-bridge interaction. This may cause the changes in the direction of motion in the loop surrounding the specificity pocket. On the other hand, the energy gain from the formation of the salt bridge is used by the enzyme to fuel the conformational change necessary to bring the substrate carbonyl group closer to the catalytic serine. It is known that the empty specificity pocket represents a higher energy state. For further discussion on this topic, see the literature (Wilmouth *et al.*, 2001; Topf, Varnai, Schofield *et al.*, 2002; Brandsdal *et al.*, 2001; Topf, Varnai & Richards, 2002; Futumura *et al.*, 1998; Otlewski *et al.*, 1999; Otlewski *et al.*, 2001).

The charges around the active-site serine have been computed with quantum-chemical calculations (Schmidt *et al.*, 2003). The formation of the covalent intermediate may induce steric (or electrostatic) stress to favour the attack of the nucleophilic water molecule and the proceeding of the catalytic reaction. Quantum-chemical calculations have indicated that these charges, dependent on the location of the observed structure along the reaction coordinate, can become quite large (up to the equivalent of one excess electron in the oxyanion hole in the substrate complex PH4, which represents a state close to the tetrahedral transition state) and reach a maximum for the tetrahedral transition state (Schmidt *et al.*, 2003). As indicated by the distorted geometry and the change in contact distances in the active sites of the ROOM, PH4 and PH5 structures (Fig. 1), conformational adaptations occur for the need of stabilization of the structure, in particular the tetrahedral transition state, and can be related the change of the degree and the direction of the anisotropic atomic motion (Fig. 3).

4.3. General applicability and impact on structure analysis

The analysis of the anisotropic ADPs reveals subtleties of protein function which may not be attainable if a structure is analysed only on the basis of the atomic coordinates. Not only local tendencies of motion but also concerted movement of atoms within protein domains can be detected. Such domain motions could be expected for the formation of any type of complex with proteins, cofactors or substrates. A thorough analysis in terms of accessible conformational states, which can be deduced from the directional motion, provides insight into the energetics of complex formation and the driving forces for allosteric mechanisms. The ADP analysis is therefore a useful tool to reveal steric strain in the molecular structure. Furthermore, the knowledge of the atomic motion may hint at functionally relevant plasticity and the behaviour of the protein in solution (Wilson & Brunger, 2000).

A related topic, concerned with the interrelation between lattice disorder, lattice vibration, motion of the molecule as a whole and individual atomic motion, is becoming of increasing interest (Hinsen *et al.*, 1999; Kidera & Go, 1992). Since it is the environment that determines the stereochemistry and contact networks around a residue as well as its possible modes of movement, one can envision mapping of such three-dimensional properties in order to infer motion and identify or predict mobile regions in structures of biological molecules at lower resolution.

We thank Garib Murshudov and Wojciech Rypniewski for helpful discussions.

References

Berglund, G. I., Carlsson, G. H., Smith, A. T., Szoke, H., Henriksen, A. & Hajdu, J. (2002). *Nature (London)*, **417**, 463–468.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Bernett, M. J., Somasundaram, T. & Blaber, M. (2004). *Proteins*, **57**, 626–634.
Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *Eur. J. Biochem.* **80**, 319–324.
Brandsdal, B. O., Aqvist, J. & Smalås, A. O. (2001). *Protein Sci.* **10**, 1584–1595.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cruickshank, D. W. J. (1956). *Acta Cryst.* **9**, 754–756.
Dunitz, J. D., Maverick, E. F. & Trueblood, K. N. (1988). *Angew. Chem. Int. Ed. Engl.* **27**, 880–895.

Dunitz, J. D., Shomaker, V. & Trueblood, K. N. (1988). *J. Phys. Chem.* **92**, 856–867.
Frey, P. A., Whitt, S. A. & Tobin, J. B. (1994). *Science*, **264**, 1927–1930.
Futumura, A., Stratikos, E., Olson, S. T. & Gettins, P. G. W. (1998). *Biochemistry*, **37**, 13110–13119.
Harata, K., Abe, Y. & Muraki, M. (1999). *J. Mol. Biol.* **287**, 347–358.
Harata, K. & Kanai, R. (2002). *Proteins*, **48**, 53–62.
Hinsen, K., Thomas, A. & Field, M. J. (1999). *Proteins*, **34**, 369–382.
Jones, T. A., Zou, J. Y., Cowan, S. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kidera, A. & Go, N. (1992). *J. Mol. Biol.* **225**, 457–475.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Leiros, H.-K. S., McSweeney, S. M. & Smalås, A. O. (2001). *Acta Cryst.* D**57**, 488–497.
Merritt, E. A. & Murphy, M. E. (1994). *Acta Cryst.* D**50**, 869–873.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins*, **11**, 281–296.
Otlewski, J., Jaskolski, M., Buczek, O., Cierpicki, T., Czapinska, H., Krowarsch, D., Smalas, A. O., Stachowiak, D., Szpineta, A. & Dadlez, M. (2001). *Acta Biochim. Pol.* **48**, 419–428.
Otlewski, J., Krowarsch, D. & Apostoluk, W. (1999). *Acta Biochim. Pol.* **46**, 531–565.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perona, J. J., Craik, C. S., Fletterick, R. J., Singer, P. T., Smalås, A., Carty, R. P., Mangel, W. F. & Sweet, R. M. (1993). *Science*, **261**, 620–622.
Perona, J. J., Tsu, C. A., McGrath, M. E., Craik, C. S. & Fletterick, R. J. (1993). *J. Mol. Biol.* **230**, 934–949.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Rypniewski, W. R., Hastrup, S., Betzel, C., Dauter, M., Dauter, Z., Papendorf, G., Branner, S. & Wilson, K. S. (1993). *Protein Eng.* **6**, 341–348.
Rypniewski, W. R., Ostergaard, P. R., Norregaard-Madsen, M., Dauter, M. & Wilson, K. S. (2001). *Acta Cryst.* D**57**, 8–19.
Schlichting, I., Almo, S. C., Rapp, G., Wilson, K., Petratos, K., Lentfer, A., Wittinghofer, A., Kabsch, W., Pai, E. F., Petsko, G. A. & Goody, R. S. (1990). *Nature (London)*, **345**, 309–315.
Schlichting, I. & Chu, K. (2000). *Curr. Opin. Struct. Biol.* **10**, 744–752.
Schmidt, A., Jelsch, C., Ostergaard, P., Rypniewski, W. & Lamzin, V. S. (2003). *J. Biol. Chem.* **278**, 43357–43362.
Schneider, T. R. (1996). *Proceedings of the CCP4 Study Weekend. Macromolecular Refinement*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey, pp. 133–144.Warrington: Daresbury Laboratory.
Schomaker, V. & Trueblood, K. N. (1968). *Acta Cryst.* B**24**, 63–76.
Shan, S. O. & Herschlag, D. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 14474–14479.
Sheldrick, G. M. (1997). *SHELXL97. Program for Crystal Structure Solution.* University of Gottingen, Gottingen, Germany.
Topf, M., Varnai, P. & Richards, W. G. (2002). *J. Am. Chem. Soc.* **124**, 14780–14788.
Topf, M., Varnai, P., Schofield, C. J. & Richards, W. G. (2002). *Proteins*, **47**, 357–369.
Tsilikounas, E., Kettner, C. A. & Bachovchin, W. W. (1992). *Biochemistry*, **31**, 12839–12846.
Wilmouth, R. C., Edman, K., Neutze, R., Wright, P. A., Clifton, I. J., Schneider, T. R., Schofield, C. J. & Hajdu, J. (2001). *Nature Struct. Biol.* **8**, 689–694.
Wilson, A. J. C. (1942). *Nature (London)*, **150**, 151–152.
Wilson, M. A. & Brunger, A. T. (2000). *J. Mol. Biol.* **301**, 1237–1256.
Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.