THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 44, Issue of October 31, pp. 43357–43362, 2003

Printed in U.S.A.

Trypsin Revisited

CRYSTALLOGRAPHY AT (SUB) ATOMIC RESOLUTION AND QUANTUM CHEMISTRY REVEALING DETAILS OF CATALYSIS*

Received for publication, June 30, 2003, and in revised form, August 21, 2003
Published, JBC Papers in Press, August 22, 2003, DOI 10.1074/jbc.M306944200

Andrea Schmidt‡§, Christian Jelsch¶, Peter Østergaard‖, Wojciech Rypniewski, and Victor S. Lamzin‡**

*From the ‡European Molecular Biology Laboratory (EMBL) Hamburg c/o DESY, D-22607 Hamburg, Germany, the **Institute of Bioorganic Chemistry, Polish Academy of Sciences, 61–704 Poznan, Poland and Institute of Biochemistry and Molecular Biology, University Hospital, Hamburg-Eppendorf, c/o DESY, 22603 Hamburg, Germany, the ¶Laboratoire de Cristallographie et Modelisation des Matériaux Mineraux et Biologiques LCM3B CNRS UHP Faculté des Sciences, 54506 Vandoeuvre-les-Nancy Cedex France, and the ‖Novozymes Protein Design, 2cs.01 DK-2880 Bagsvaerd, Denmark*

A series of crystal structures of trypsin, containing either an autoproteolytic cleaved peptide fragment or a covalently bound inhibitor, were determined at atomic and ultra-high resolution and subjected to *ab initio* quantum chemical calculations and multipole refinement. Quantum chemical calculations reproduced the observed active site crystal structure with severe deviations from standard stereochemistry and indicated the protonation state of the catalytic residues. Multipole refinement directly revealed the charge distribution in the active site and proved the validity of the *ab initio* calculations. The combined results confirmed the catalytic function of the active site residues and the two water molecules acting as the nucleophile and the proton donor. The crystal structures represent snapshots from the reaction pathway, close to a tetrahedral intermediate. The de-acylation of trypsin then occurs in true S_N2 fashion.

Serine proteases are among the most abundant proteins. A large subgroup comprising trypsin and chymotrypsin is called trypsin-like serine proteases. Trypsin cleaves peptides at the C-terminal side of arginine or lysine.

Catalysis takes place in a cleft from which a binding pocket (specificity pocket) protrudes into the interior of the enzyme. A catalytic triad (Ser^{195}-His^{56}-Asp^{99}) constitutes the core of the reactive center. For the stabilization of the charges developing on the reaction intermediates during catalysis an oxyanion hole is formed by the main-chain N-H of residues 192–195 in an arrangement known as a nest (1). The specificity pocket accommodates long, basic amino acid side chains. At its bottom, an aspartate residue forms a tight salt bridge with the substrate. The difference in specificity between various serine proteases is determined by the shape and the width of the specificity pocket (2, 3).

The catalytic mechanism of trypsin isolated from various sources has been investigated in great depth (4–10). Yet several questions remain unanswered. One is the identification of the catalytically active water molecules. Also, the catalytic triad appears to function in different ways in different proteins and display different features, particularly the presence of the so-called short hydrogen bonds (11).

Trypsin from *Fusarium oxysporum* shows auto-proteolysis activity even at crystallization conditions. A peptide fragment can always be found in the active site of the crystalline enzyme (12). This together with the fact that its crystals diffract extremely well has made it especially interesting for mechanistic studies and offered an excellent tool for the investigation of the catalytic activity at the electronic level (13, 14).

EXPERIMENTAL PROCEDURES

Crystallization—Protein samples of *Fusarium oxysporum* trypsin were supplied by Novozymes at a concentration of 10 mg/ml. Crystals were grown as described elsewhere (15), with the pH lowered to 5.0 instead of 6.0 using 100 mM sodium citrate buffer. A combination of streak seeding and macroseeding was applied to initiate nucleation. Crystals grew to a size of 0.3–0.5 mm within a few days.

Co-crystallization with the Inhibitor DFP[1]—DFP (diisopropyl fluorophosphate, Mw = 184 g/mol) was purchased from Aldrich Chemicals. 5 μl of the inhibitor were dissolved in 250 μl of isopropyl alcohol. 1 μl of this solution was added to a drop containing 10 μl of protein solution and 20 μl of precipitant solution. The approximate molar ratio of inhibitor to protein was 20:1.

Co-crystallization with the Inhibitor PMSF—An aqueous stock solution of a concentration of 8 mg/ml PMSF (Mw = 174 g/mol) from Sigma was prepared, and 1 μl of this solution was added to the drops containing 10 μl of protein and 20 μl of precipitant solution. The molar ratio of inhibitor to protein was again 20:1. Nucleation was initiated as for the native crystals in both cases.

Soaking—Native trypsin crystals were soaked in a drop containing the precipitant solution at pH 4 for several minutes prior to cryo-freezing.

Cryo-freezing Conditions—In previous publications, the cryo-freezing conditions involved a gradual increase of glycerol content in the mother liquor up to 20% (v/v) glycerol (12). Now, we transferred the crystals into light paraffin oil directly from their drops. In order to avoid ice formation, all traces of the mother liquor were carefully removed from the crystal surface by moving it around in the oil drop. Crystals were then flash frozen directly in the cryo-stream at 100 K.

Data Collection and Processing—Data were collected on beamlines X11 and X13 at EMBL Hamburg, equipped with MAR CCD detectors, at wavelengths of 0.8110 and 0.8020 Å, respectively. A summary of the data statistics is given in Table I. Data were processed with *denzo/scalepack* or *HKL2000* (16), intensities were truncated to structure factors using the *CCP4* suite of programs (17).

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (codes 1PPZ, 1PQ5, 1PQ7, 1PQ8, and 1PQA) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ To whom correspondence should be addressed. E-mail: andrea@embl-hamburg.de.

[1] The abbreviations used are: DFP, diisopropyl fluorophosphate; PMSF, phenylmethylsulfonyl fluoride; rmsd, root-mean-square-deviation.

TABLE I
Data collection statistics

Summary of the five datasets collected for native and inhibited trypsin crystals.

Data set	Beamline	Resolution	R_{sym}	$\langle I/\sigma(I)\rangle$ (outer shell)	Wilson B factor	Space group
		Å	%		Å^2	
pH 4	X13	22–1.0	7.3	3.8	7.1	P1
pH 5	X11	30–0.83	3.3	7.2	5.0	P1
DFP	X11	31–1.22	4.5	5.5	11.3	$P2_1$
PMSF	X11	25–1.23	4.6	2.7	11.6	$P2_1$
pH 5/borax	X11	22–0.80	3.3	6.2	4.8	P1

Refinement—Initial refinement and addition of solvent were carried out with *ARP/wARP* (18) and *refmac* (19). Subsequent steps were performed with *SHELXL* (20), including the addition of hydrogen atoms and anisotropic *B* factors, until convergence was reached. Refinement was carried out against diffracted intensities. The same refinement protocol was used for all structures described here. Heterocompounds were introduced in the refinement as early as possible. Occupancies for alternate conformers, and the substrate were set according to their anisotropic atomic displacement parameters (ADP) and individually refined in a last round. Solvent molecules with ADP higher than 60 Å^2 were removed; those with ADP higher than 45 Å^2 were assigned occupancy of 0.5.

Complementary Methods—Ab initio quantum chemical calculations on the pH 4 structure were performed using GAMESS (21) with a 6–31G** base set and a Hückel model for the construction of the initial molecular orbitals. The active site residues, the substrate and side chains involved in direct contact with the substrate were chosen as the geometrical template. Only relevant parts of the interacting compounds (*i.e.* atoms at an H-bonding distance) were included in the calculations. The substrate arginine was treated as alanine since only its main chain atoms interacted with the active site. Aspartate 99 was modeled as acetic acid. The system included a total of 64 atoms (265 electrons, with a net charge of −1). Atoms of chemically inert groups not involved in direct contact with the substrate were fixed at their positions. Solvent effects were neglected as the occupied active site is a closed cavity without considerable contact to the solvent space. Hydrogen atoms (at the two water molecules in the active site) were inserted into the structure according to standard stereochemistry and existing H-bonding partners. The following cases were subjected to *ab initio* quantum chemical calculations: QC1, active site in the native state; QC2, Ser^{195} de-protonated; QC3, W1 as hydroxyl ion; PMSF, covalently bound.

Multipolar Atoms Refinement—The 0.80 Å structure was refined using the charge density refinement program MoPro (20). The program describes the electron density of the atoms with the Hansen and Coppens model (22), which includes atomic charges, expansion/contraction coefficients, and multipoles.

The structure refined with SHELXL was used as the starting model. The charge density parameters were transferred to the structure from our data bank describing the electron density of atoms in proteins (23, 24). Protein atoms with ADPs higher than 15 Å^2 and multiple conformers were considered spherical and were only assigned a charge and a contraction/expansion coefficient. Water molecules were described as a neutral oxygen atom and were refined in conventional manner.

The contribution of the bulk solvent to the crystal diffraction was added to the structure factors (25, 26) in Equation 1.

$$F_c = F_c(\text{prot}) + K_{sol} \exp(-B_{sol}\, s^2)\, F_c(\text{sol}) \qquad \text{(Eq. 1)}$$

These parameters were also refined in MoPro. The following weighting scheme was applied to the reflections shown in Equation 2.

$$W_{hkl} = [(4\sigma_F)^2 + (0.02F)^2]^{-1/2} \qquad \text{(Eq. 2)}$$

The three different types of structural parameters were refined in an iterative way until convergence: scale factor with solvent parameters K_{sol} and B_{sol}, atomic coordinates, ADPs. Hydrogen atoms were not refined and their coordinates were constrained according to standard stereochemistry while their isotropic *B*-factors were set to the B_{eq} of the parent atom. Usual distance restraints were applied on the structure according to the Engh & Huber stereochemical dictionary (target rmsd deviations of 0.02 and 0.05 Å for bond and angle distances, respectively).

RESULTS

Despite a considerable degree of disorder in the structures, all of the models refined to crystallographic *R* factors well below 15%. A summary of the structure refinements is given in Table II. The active site was never empty. In the native cases, an arginine molecule was bound into the specificity pocket with its main-chain atoms located in a good position for catalysis near Ser^{195}. Arginine is the C-terminal residue of a tripeptide, which is the remaining fragment of auto-proteolysis. The substrate tripeptide shows only partial occupancy, which seems to correlate with the occupancy distribution of the multiple conformers in the structure. The arginine lacks the second oxygen atom of its carboxylic group and the C=O distance of 1.48 Å is unusually long. The arginine side chain was observed in at least two conformations, in the pH 4 structure even overlapping with a lysine. Ser^{195} was always in a single conformation in well defined electron density. Water molecule W2 was located close to the substrate carbonyl oxygen atom in the oxyanion hole. The substrate carbonyl carbon had contacts to W1 (1.87 Å) and Ser^{195} Oγ (2.05 Å) (Fig. 1), which implies covalent interaction. The geometry around the substrate carbonyl carbon was roughly tetrahedral. No electron density could be found for the N-terminal fragment of the substrate. The water molecules W1 and W2 were fully occupied as opposed to the substrate, which showed partial occupancy with the total ranging from 0.4 (pH 5) to 0.8 (pH 5/borax and pH 4).

DFP and PMSF were bound covalently to Ser^{195}. The phosphate groups were always well visible. The respective isopropyl and phenyl moieties displayed weaker electron density and did not point into the specificity pocket but in the direction of His^{56}. As an effect, the histidine side chain is statically disordered showing two conformations, one of which is rotated away from the inhibitors thus disrupting the catalytic triad (Fig. 1). In the DFP and PMSF structures, one oxygen atom in the sulfonyl or phosphate groups takes the position of W1, another one the position of W2.

Like the other native structures, the pH 5/borax structure also had a peptide bound into the active site. This structure had the highest quality data set and was used for multipole refinement.

Co-crystallization with DFP and PMSF changed the space group from P1 to $P2_1$ but left the internal protein structure unchanged and the crystal contacts mostly preserved.

Disorder and Mobility—One-fifth of the total number of residues of *F. oxysporum* trypsin shows at least two conformers. Their occupancies relate to the occupancy of the substrate or of the inhibitor side chain. In all the structures except native pH 5, the substrate- or inhibitor-bound state seems to correspond to the protein conformer with higher occupancy, with an average value of 0.65 in all structures. The conformers with corresponding lower occupancy (0.35) correspond to the empty active site. The rmsd for the two visible main chain conformers of disordered residues lies between 1 and 2 Å There was no evidence for the presence of other, intermediate conformational states.

Most of the disordered residues are spatially linked to each other. The effect starts in the loops of the active site cleft and is then transferred to the adjacent regions. The patterns of disorder are different for the native and the inhibited structures.

TABLE II
Crystallographic refinement

Summary for the different trypsin structures. The observed structure factors, phases and derived coordinates have been deposited to the PDB with access codes 1PQ5 (pH5), 1PQ7 (pH5/borax), 1PQ8 (pH4), 1PPZ (DFP), and 1PQA (PMSF).

	pH 4	pH 5	PMSF	DFP	p H5/borax
Resolution	1.0	0.83	1.23	1.22	0.80
R (%; all data)	12.8	9.8	14.1	14.0	10.7
Number of unique reflections	86,663	146,839	44,081	47,457	165,882
Number of atoms total	2,139	2,263	1,935	1,979	2,162
Number of water molecules	330	489	269	265	495
Overall atomic displacement (all atoms)	11.8	9.2	16.8	16.0	9.4
Overall atomic displacement (protein)	9.8	6.6	15.3	14.6	6.7
Number of disordered residues	40	33	16	27	14

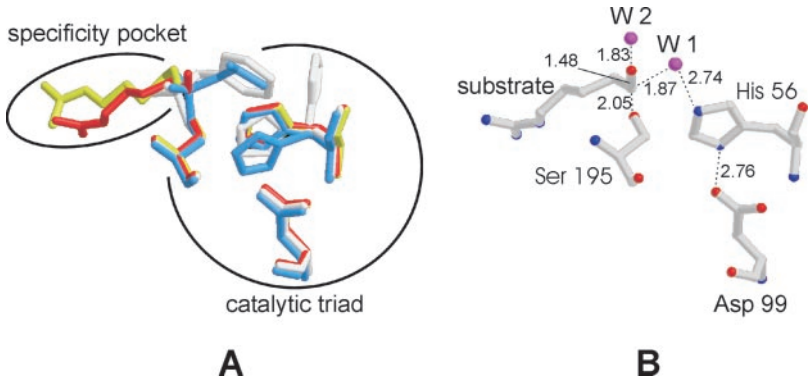

FIG. 1. *A*, overlay of the active sites in the structures at pH 4 (*red*), pH 5 (*yellow*), with PMSF (*gray*) and DFP (*blue*), showing the effects of the covalent inhibitors on the catalytic triad. It also shows that different orientations are possible for the substrate. *B*, the active site in the pH 4 structure, as used for the *ab initio* calculations. The oxyanion hole is omitted for clarity. The geometry around the substrate arginine carbonyl is roughly tetrahedral and shows unusual interatomic distances. Figures were produced with Molscript/Raster3D (36).

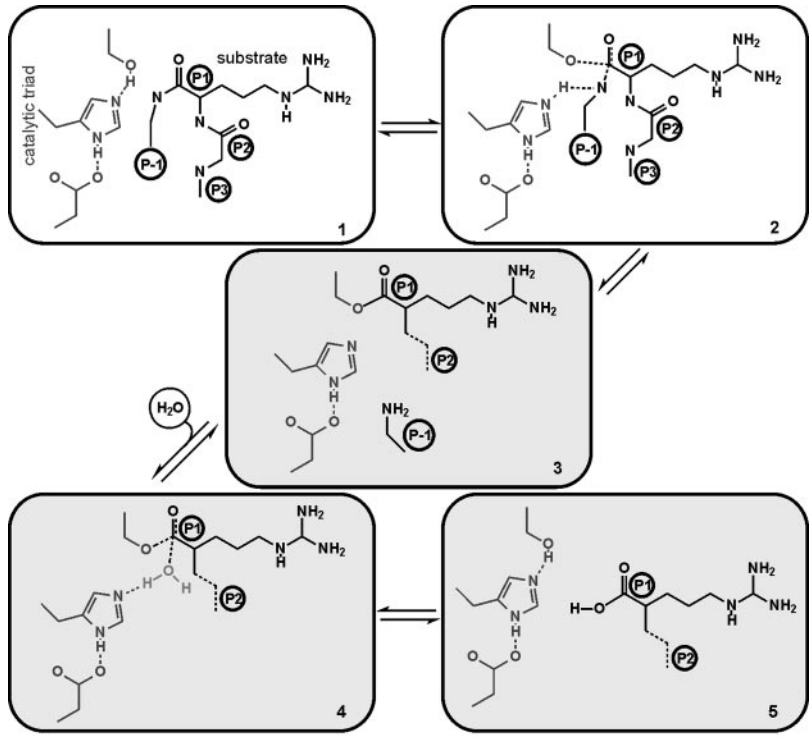

FIG. 2. **Reaction scheme of the enzymatic reaction divided into 5 steps.** Those in the *gray panels* (3–5) belong to the de-acylation and are mapped by the structures described in this study.

The inhibited structures, mimicking a reaction intermediate, show higher anisotropy of the active site residues. In the native structures at pH 4 and pH 5, disorder arises in the specificity pocket, which is not the case in the DFP and PMSF structures where this site is empty. Despite the apparent movement in and around the active site, Ser^{195} shows only one conformation as mentioned above.

An analysis of corresponding conformers showed that there is good correlation between rmsd and ADP in the native structures indicating disorder of mostly dynamic origin. Disorder in the inhibited structure appears to be mostly static.

Quantum Chemical Calculations on the pH 4 Structure—A stable match between the crystal structure and the theoretical model could only be achieved under the assumption that the catalytic serine is de-protonated and the attacking water molecule W1 is actually a water molecule and not a hydroxyl ion. In

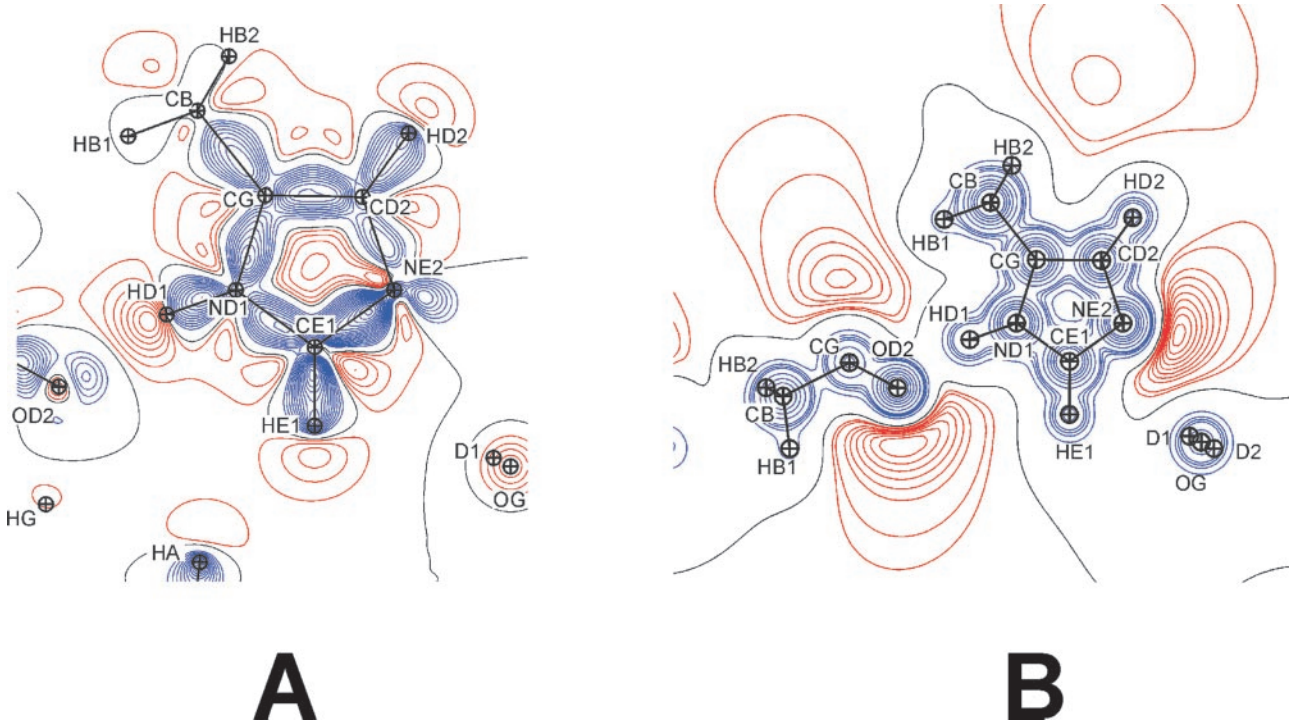

FIG. 3. **Deformation density for the His[56] ring showing its electronic state (*A*) and electrostatic potential (*B*).** Atom OD1 belongs to Asp[99], OG to Ser[195].

all other cases, the geometry was directed toward either the products of peptide cleavage (QC3) or the covalent intermediate (see Fig. 2 for an overview of the catalytic pathway).

This implies that the crystal structure represents a state close to a tetrahedral transition state. Bonds of respective order 0.99 and 0.47 are formed between the substrate C and the Ser[195] Oγ and from the substrate C to W1, which is held in place from the other side by His[56]. His[56] is deprotonated and acts as a Lewis base, activating the catalytic water molecule, which is indicated by a strong H-bonding contact of bond order 0.2–0.3.

Water molecule W2 in the oxyanion hole forms a very strong H bond to the carbonyl oxygen of the arginine substrate in the model matching the crystal structure. This is accompanied by a high negative charge (−0.99) on the W2 oxygen atom.

Multipole Refinement—For the ultra-high resolution structures (pH 5 and pH 5/borax), the protonation state in the active site could be visualized directly in residual electron density maps. His[56] was indeed not protonated (Nϵ1 is non-protonated while the hydrogen atom Hδ1 is bound to Nδ2). The multipole charge density parameters were transferred from the data base and Fig. 3*A* shows the corresponding deformation electron density on the catalytic histidine. The small peak near Nϵ2 indicates the electron lone-pair on the nitrogen atom or a partial protonation (which at this pH would be quite probable), or a rotation of the histidine ring by 180° around the Cβ–Cγ bond. The electrostatic potential derived from the multipoles transfer is shown for the His[56]-Asp[99] interaction (Fig. 3*B*).

DISCUSSION

Comparison to the Previously Refined Structures—Remnants of the substrate have been observed in the *F. oxysporum* trypsin structures before (12). Lowering the pH to 5.0 resulted in lower occupancy and higher mobility of the substrate arginine side chain and the disappearance of two of the three water molecules arranged around the arginine carbonyl in the previous structures (12). Only water W1 remained. The structures showed no other significant differences.

Mobility in the Crystal—The walls of the active site cleft at the subsites S2, S3 (residues 144–150) are always in double conformation, subsites S2′ and S3′ (residues 93–95) only in the DFP and PMSF structures. This can be understood as an effect of the release of the N-terminal fragment of the substrate (sites S2, S3) and, for the C-terminal part of the active site cleft, adjustments in the positioning of the substrate just before the nucleophilic attack of water W1.

In the pH 4 and pH 5 structures, Cys[191], Tyr[225], and the stretch between amino acids 215 and 220 show main chain double conformations. Cys[191] and Tyr[225] are part of the wall of the specificity pocket. Cys[216] forms a disulfide bridge with Cys[191] and hence moves together with it. Apparently, this is related to the binding and release of the substrate arginine side chain, which also shows multiple conformers. Movement upon binding or release is supported by the fact that in the case of the covalently bound inhibitors, which did not reach into the pocket, the specificity pocket was completely rigid. An induced fit model was also supported by interpretation of the structures with inhibitors. In both cases the protein showed two different conformers in the subsites S2, S3, S2′, and S3′, whose occupancies relate to the occupancies of DFP (0.7) and PMSF (0.6).

Two important residues involved in this movement are Asp[189] and Ser[190], both located at the very bottom of the pocket, highly conserved and responsible for the preference of trypsin for basic residues (2, 3). Asp[189] forms a strong ion pair with the substrate arginine side chain and its movement as well as the disorder of the other residues in the specificity pocket appears to be induced directly by the substrate. Changes in the substrate position would be transmitted to the catalytic center or *vice versa*. Re-positioning must occur during the formation and breaking of the covalent bond to Ser[195]. NMR experiments on complexes showed that neither active site nor specificity pocket is fully pre-formed in the native structures. Only binding to specificity subsites creates an environment suitable for a tetrahedral transition state (27). This seems to be the case also in *F. oxysporum* trypsin,

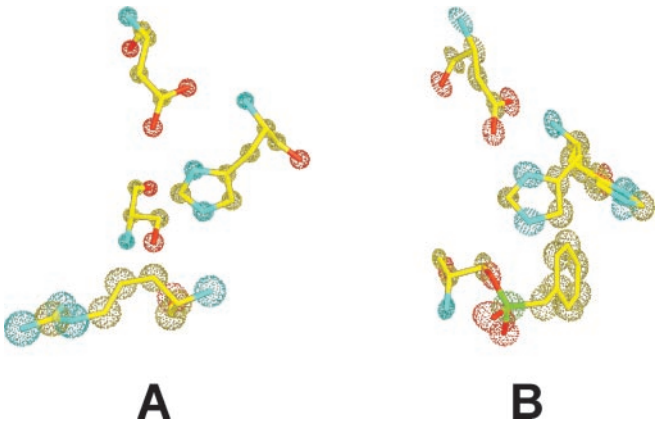

FIG. 4. **Thermal ellipsoids with 35% probability for the catalytic triad of the pH 5 and PMSF structures showing the difference in anisotropy.** The figure was produced in XtalView (37).

in contrast to the widely accepted lock-and-key model for serine protease substrate binding (28).

Increased anisotropy was observed for the residues of the catalytic triad in the inhibited structures. The shape of the thermal ellipsoids of the atoms in His^{56} and Asp^{99} indicated movement toward Ser^{195} and the atoms in Ser^{195} tend to move perpendicular to the plane formed by Cβ, Oγ, and P or S (Fig. 4). In addition, some residues at sites S2′ and S3′ show considerable shift of one conformer toward the active site, shortening the cleft. This is not observed in the structures with the substrate. We therefore assume that a re-arrangement in the substrate binding region is required before de-acylation. In the covalent intermediate, the Ser^{195} ester carbonyl can have any orientation (10). It has to be adjusted, fixed, and a water molecule has to approach the covalent intermediate and position itself in a location suitable for catalysis. The water molecule W1 in the pH 4 and pH 5 structures is already at its nucleophilic attack and steric effects might explain why the reaction did not proceed further. Embedding of protein molecules in crystals makes large movement impossible and impedes product release or binding: the active site of trypsin was in principle accessible in the crystals, but it was impossible to bind DFP to the protein by soaking; DFP could not be seen in the active site of trypsin in soaked crystals. This again strongly indicates an induced fit mechanism for substrate binding and the presence of an "open" and "closed" state or an "empty" and "loaded" conformation of the protein molecule.

Reaction Partners and Water Molecules—W1, attached to His^{56}, and W2 in the oxyanion hole were the two candidates for participation in catalysis. The role of the water molecule W1 as the nucleophile is easy to see. The location of water molecule W2 in the oxyanion hole seemed at first glance a bit exotic for steric reasons and in view of the commonly accepted direct stabilization of the carbonyl oxygen by the oxyanion hole. However, in the bovine trypsin, a water molecule has been found in the same position (29). Its proposed role was to protonate the carbonyl oxygen atom of the substrate, to take up the excess electrons and facilitate the shift of electrons toward the oxyanion hole (5). This role was controversial, as this water molecule does not exist in all known serine protease structures.

In *F. oxysporum* trypsin the distance of the substrate carbonyl to W2 is much shorter than expected for even the strongest of H bonds. This is exactly the property that can help facilitate catalysis. In short H bonds, the energy barrier for the proton transfer from one binding partner to the other becomes very small (30, 31). There has been a very lively debate over the influence of "low barrier hydrogen bonds" (30, 32, 33) on catalysis through an energy contribution of up to 20 kcal/mol. In these studies, a strong hydrogen bond was defined as of an O–O distance of 2.4 Å, while the hydrogen bond between W2 and the substrate carbonyl oxygen atoms is 1.83 Å long. The proton between these two oxygen atoms should be able to fluctuate easily. *Ab initio* calculations show that the bond order to the carbonyl oxygen is 0.82 (see Table III). This means that the proton has almost been transferred from the water molecule to the substrate. The hydrogen atoms of the water molecules W1 and W2 were not seen in the electron density maps.

TABLE III
Quantum chemical calculations

Results from the *ab initio* calculations on the pH 4 structure show good agreement (low rmsd) to the crystal structure for the model in which Ser^{195} is deprotonated (QC2).

	QC 1	QC 2	QC 3	Crystal
Distances (Å)				
Ser^{195} Oγ-Arg C	2.29	2.05	n/a	2.05
Arg C-Arg O	1.14	1.47	1.20	1.48
W1 O-Arg C	1.80	1.87	1.31	1.87
W1 O-His^{56} Ne2	2.65	2.92	2.06	2.76
W2 O-Arg O	2.51	1.83	1.86	1.83
R.m.s.d. to the crystal structure	0.24	0.07	0.47	
Bond orders				
Ser^{195} Oγ-Arg C	0.00	0.99	0.00	
Arg C-Arg O	1.79	1.07	1.63	
W1 O-Arg C	0.70	0.47	1.05	
W1 H2-His^{56} Nε2	0.37	0.21	0.05	
W2 H1-Arg O	0.05	0.82	0.06	
Charges				
Arg C	0.824	0.939	0.809	
Arg O	−0.550	−0.541	−0.656	
Ser^{195} Oγ	0.856	−0.668	−0.791	
W2 O	−0.744	−0.990	−0.757	
W1 O	−0.700	−0.720	−0.637	

Quantum Chemistry—The QC2 theoretical model matching the crystal structure shows the most extreme charge distribution as compared with the other models. The structure is close to a tetrahedral transition state. Its stability can be attributed mostly to the capability of the water molecule in the oxyanion hole to take up a large excess of electrons. It achieves that by attachment to the carbonyl oxygen via a very strong H bond with bond order 0.82. As the bond order between the shared hydrogen atom and the W2 oxygen atoms is only 0.11, the water molecule should be pictured rather as a HO^{-}–H^{+}-O^{-} arrangement at an extremely close distance with the internal structure of the water molecule maintained. In the other models (QC1 and QC3), W2 remains a water molecule with strong internal bonds and external H bridges of normal strength. This indicates that, after the nucleophilic attack of W1, W2 regains its original state and functions as the Henderson water is supposed to function (29). The QC3 model corresponds to the end-state of peptide cleavage, step 5 in Fig. 2.

W1 has been identified before in crystal structures (4, 29) but its chemical state was never clear. Quantum chemical calculations show that W1 performs its nucleophilic attack as a water molecule. When a hydroxyl ion was input into the calculations, the system was immediately converted into the products and the arginine pushed out of the active site. This is understandable as a hydroxyl ion is among the hardest nucleophiles. Its proximity to a carbon carrying a positive partial charge would naturally result in immediate attachment, excluding the chance to observe such a state in a crystal structure. Our results show that not only one, but both water molecules are catalytically active at the same time.

Electronic Details Revealed at Ultra-high Resolution—The beauty of ultra-high resolution protein crystallography is that one can attain electronic detail by crystallography alone (34). At ultra-high resolution the individual atoms in a molecule can

no longer be considered spherical. The presence of lone-pair and bonding electrons has then to be directly included into the refinement.

The electron density in the active site of the ultra-high resolution trypsin structures shows that His^{56} is definitely de-protonated. Because of a relatively high thermal motion on the Ser^{195} Oγ, the electron density cannot give evidence for an attached proton on this side chain. However, the serine 195 Cβ–Oγ distance is 1.22 Å (pH 5-borax structure), which is significantly smaller than the standard protonated serine Cβ–Oγ distance of 1.417 ± 0.020 in the Engh and Huber (35) stereochemistry dictionary. This implies the following. First, the pK_a of His^{56} is lower than 6.5. This is already hinted at by the presence of enzymatic activity under crystallization conditions despite a basic pH optimum and proves that His^{56} acts as a general base. A rotation of the imidazole ring by 180° could play a role in the substrate binding and product release steps. Second, Ser^{195} is de-protonated as well and is capable of covalent attachment to the substrate. And third, the assumptions made on the protonation state in the active site were correct and therefore the *ab initio* calculations have produced a realistic model.

Implications for Catalysis and Conclusions—Using the atomic resolution structures and the results from the *ab initio* calculations, it could be shown directly that trypsin de-acylation does occur via a tetrahedral intermediate with subsequent hydrolysis in true S_N2 fashion. The water molecule W1 is the obvious candidate to perform hydrolysis, while the preparation for this step is performed by W2 by protonating the carbonyl oxygen. Despite the requirement for W1 to carry a negative partial charge in order to perform a nucleophilic attack, complete de-protonation does not occur. Its attachment to the substrate and proton abstraction is concerted.

The lowered pK_a of His^{56} explains why the enzyme was still active in the crystallization solution. The fact that the His^{56} side chain swings out upon contact with the inhibitors implies that the catalytic triad is not rigid and tight. The inhibitors are larger than the substrate intermediate and remove the water molecules W1 and W2 from their normal positions. Therefore, no activator (W2) and no nucleophile (W1) for hydrolysis are present. Hence, with their modifications (fluorine instead of oxygen and therefore strong withdrawal of electrons), the inhibitors have triple effect. *Ab initio* calculations on the PMSF structure (result not shown) yielded a charge distribution around the Ser^{195} Oγ similar to the native, which indicates that inhibitors are good models for the tetrahedral intermediates. *F. oxysporum* trypsin is a dynamic enzyme, in which substrate binding and product release involves an induced fit mechanism and where the "loaded" and "empty" states can be structurally distinguished.

Atomic and ultra-high resolution crystal structure determination and *ab initio* quantum chemical calculations have now become indispensable tools for comprehensive structure interpretation, acquiring snapshots along the reaction pathway and the assignment of function to the residues involved in catalysis. The motion of the macromolecule in the crystal structure, revealed from the anisotropic atomic displacement parameters, complements the time-average structural picture with dynamics thus opening a new prospective in the structure-functional analysis of biological processes.

Acknowledgments—We thank Richard J. Morris, Rob Meijers, and Benoit Guillot for helpful discussions.

REFERENCES

1. Pal, D., Suhnel, J., and Weiss, M. S. (2002) *Angew. Chem. Int. Ed. Engl.* **41,** 4663–4665
2. Lu, W., Apostol, I., Qasim, M. A., Warne, N., Wynn, R., Zhang, W. L., Anderson, S., Chiang, Y. W., Ogin, E., Rothberg, I., Ryan, K., and Laskowski, M. J. (1997) *J. Mol. Biol.* **266,** 441–461
3. Brandsdal, B. O., Aqvist, J., and Smalas, A. O. (2001) *Protein Sci.* **10,** 1584–1595
4. Wilmouth, R. C., Edman, K., Neutze, R., Wright, P. A., Clifton, I. J., Schneider, T. R., Schofield, C. J., and Hajdu, J. (2001) *Nat. Struct. Biol.* **8,** 689–694
5. Perona, J. J., Craik, C. S., Fletterick, R. J., Singer, P. T., Smalas, A., Carty, R. P., Mangel, W. F., and Sweet, R. M. (1993) *Science* **261,** 620–622
6. Katona, G., Wilmouth, R. C., Wright, P. A., Berglund, G. I., Hajdu, J., Neutze, R., and Schofield, C. J. (2002) *J. Biol. Chem.* **277,** 21962–21970
7. Lin, J., Cassidy, C. S., and Frey, P. A. (1998) *Biochemistry* **37,** 11940–11948
8. Marquart, M., Walter, J., Deisenhofer, J., Bode, W., and Huber, R. (1983) *Acta Crystallogr.* **39,** 480–490
9. Otlewski, J., Jaskolski, M., Buczek, O., Cierpicki, T., Czapinska, H., Krowarsch, D., Smalas, A. O., Stachowiak, D., Szpineta, A., and Dadlez, M. (2001) *Acta Biochim. Pol.* **48,** 419–428
10. Topf, M., Varnai, P., Schofield, C. J., and Richards, W. G. (2002) *Proteins* **47,** 357–369
11. Frey, P. A., Whitt, S. A., and Tobin, J. B. (1994) *Science* **264,** 1927–1930
12. Rypniewski, W. R., Ostergaard, P. R., Norregaard-Madsen, M., Dauter, M., and Wilson, K. S. (2001) *Acta Crystallogr. Sect. D* **57,** 8–19
13. Meijers, R., Morris, R. J., Adolph, H. W., Merli, A., Lamzin, V. S., and Cedergren-Zeppezauer, E. S. (2000) *J. Biol. Chem.* **276,** 9316–9321
14. Schmidt, A., and Lamzin, V. S. (2002) *Curr. Opin. Struct. Biol.* **12,** 698–703
15. Rypniewski, W. R., Hastrup, S., Betzel, C., Dauter, M., Dauter, Z., Papendorf, G., Branner, S., and Wilson, K. S. (1993) *Protein Eng.* **6,** 341–348
16. Otwinowsky, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
17. CCP4 (1994) *Acta Crystallogr. Sect. D* **50,** 760–763
18. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463
19. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D* **53,** 240–255
20. Sheldrick, G. M., and Schneider, T. R. (1997) *Methods Enzymol.* **277,** 319–343
21. Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S., Windus, T. L., Dupuis, M., and Montgomery, J. A., Jr. (1993) *J. Comput. Chem.* **14,** 1347–1363
22. Hansen, N. K., and Coppens, P. (1978) *Acta Crystallogr. Sect. A* **34,** 909–921
23. Pichon-Pesme, V., Lecomte, C., and Lachekar, H. (1995) *J. Phys. Chem.* **99,** 6242–6250
24. Jelsch, C., Pichon-Pesme, V., Lecomte, C., and Aubry, A. (1998) *Acta Crystallogr. Sect. D* **54,** 1306–1318
25. Tronrud, D. E. (1997) *Methods Enzymol.* **277,** 306–319
26. Fokine, A., and Urzhumtsev, A. (2002) *Acta Crystallogr. Sect. D* **58,** 1387–1392
27. Tsilikounas, E., Kettner, C. A., and Bachovchin, W. W. (1992) *Biochemistry* **31,** 12839–12846
28. Talhout, R., and Engberts, J. B. (2001) *Eur. J. Biochem.* **268,** 1554–1560
29. Singer, P. T., Smalas, A., Carty, R. P., Mangel, W. F., and Sweet, R. M. (1993) *Science* **259,** 669–673
30. Schiott, B., Iversen, B. B., Madsen, G. K., Larsen, F. K., and Bruice, T. C. (1998) *Proc. Nat. Acad. Sci. U. S. A.* **95,** 12799–12802
31. Cleland, W. W., and Kreevoy, M. M. (1994) *Science* **264,** 1887–1890
32. Warshel, A., Papazyan, A., and Kollman, P. A. (1995) *Science* **269,** 102–106
33. Cleland, W. W., Frey, P. A., and Gerlt, J. A. (1998) *J. Biol. Chem.* **273,** 25529–25532
34. Lamzin, V. S., Morris, R. J., Dauter, Z., Wilson, K. S., and Teeter, M. M. (1999) *J. Biol. Chem.* **274,** 20753–20755
35. Engh, R. A., and Huber, R. (1991) *Acta Crystallogr. Sect. A* **47,** 392–400
36. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
37. McRee, D. E. (1992) *J. Mol. Graphics* **10,** 44–47