Protein J (2008) 27:79–87
DOI 10.1007/s10930-007-9110-3

The Crystal Structure of a Trypsin-like Mutant Chymotrypsin: The Role of Position 226 in the Activity and Specificity of S189D Chymotrypsin

Balázs Jelinek · Gergely Katona · Krisztián Fodor · István Venekei · László Gráf

Published online: 6 September 2007

Abstract The crystal structure of the S189D+A226G rat chymotrypsin-B mutant has been determined at 2.2 Å resolution. This mutant is the most trypsin-like mutant so far in the line of chymotrypsin-to-trypsin conversions, aiming for a more complete understanding of the structural basis of substrate specificity in pancreatic serine proteases. A226G caused significant rearrangements relative to S189D chymotrypsin, allowing an internal conformation of Asp189 which is close to that in trypsin. Serious distortions remain, however, in the activation domain, including zymogen-like features. The pH-profile of activity suggests that the conformation of the S1–site of the mutant is influenced also by the P1 residue of the substrate.

Keywords Position 226 · Serine protease · Substrate specificity · Crystal structure

Abbrevations

AMC	Amino-4-Methyl Coumarin
BPTI	Bovine pancreatic trypsin inhibitor
CABS	4-(Cyclohexylamino)-1-butanesulfonic acid
CCP4	Collaborative Computational Project Number 4
CHES	2-(Cyclohexylamino)ethanesulfonic acid
DPI	Dispersion precision indicator
ESRF	European Synchrotron Radiation Facility
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
MES	2-(N-Morpholino)ethanesulfonic acid
MOPS	3-(N-Morpholino)propanesulfonic acid
MUGB	4-Guanidinobenzoic acid 4-methylumbelliferyl ester hydrochloride
PDB	Protein data bank
PEG	Polyethylene glycol
SBTI	Soybean trypsin inhibitor
SDS-PAGE	Sodium dodecyl sulfate - polyacrylamide gel electrophoresis
TLS	Translational, librational, screw
TRIS	Tris(hydroxymethyl)aminomethane

B. Jelinek (✉) · I. Venekei · L. Gráf
Department of Biochemistry, Eötvös Loránd University, Pázmány s. 1/C, Budapest 1117, Hungary
e-mail: jelinek@elte.hu

G. Katona
Institut de Biologie Structurale, UMR 5075, CEA/CNRS/UJF, 41 avenue Jules Horowitz, 38027 Grenoble Cedex 1, France

K. Fodor · L. Gráf
Biotechnology Research Group, Hungarian Academy of Sciences, Pázmány s. 1/C, Budapest 1117, Hungary

1 Introduction

Pancreatic serine proteases share the same catalytic mechanism and structural scaffold, but have dissimilar substrate binding pockets to accommodate different substrate residues. Trypsin, with the negative charge of Asp189 at the base of the S1 site, prefers Arg and Lys in the P1 substrate position (nomenclature of [1]), while the less polar S1 site of chymotrypsin with Ser189 has the highest affinity for Tyr, Phe and Trp substrate residues. Starting with D189S trypsin [2], a series of interconversion experiments of trypsin and chymotrypsin specificities were carried out by replacing different sets of residues at the substrate binding site with the corresponding amino acids in the other enzyme. To produce a trypsin with

chymotrypsin-like activity [2–6] the substitution of 15 residues in the S1 site and its surroundings was necessary, contrary to only three residues at positions 189, 216 and 226, as it had been proposed by a relatively simple text book model [7]. However, the reverse chymotrypsin→trypsin specificity conversion experiments based on the above complex model resulted in proteases with only very modest trypsin-like specificity and low catalytic activity [8]. These results suggest that a specific combination of extended structural elements determines specificity.

In the chymotrypsin→trypsin conversion experiments a B-type chymotrypsin was used, which contains an Ala at site 226 (A-type chymotrypsin contains Gly at this site). In trypsin, where Gly is conserved at this position, the methyl group of an Ala severely reduces the enzyme activity by shielding the negative charge of Asp189 from the positive charge of Arg and Lys substrate residues, as revealed by a G226A mutant [9, 10]. Therefore we tested the role of site 226 in the chymotrypsin→trypsin specificity conversion [11] by performing A226G substitution in a mutant of chymotrypsin-B, which contained the two S1 site loops L1 and L2 exchanged to the corresponding loops in trypsin and exhibited a weak trypsin-like specificity profile with a very low activity. Unexpectedly, the mutant became a better chymotrypsin. At the same time, the A226G substitution caused a 100-fold increase in trypsin-like activity (still 1000-fold less than wild type enzymes), and resulted in a trypsin-like specificity profile in the case of the S189D chymotrypsin mutant, which was a poorly active, slightly chymotrypsin-like enzyme. Moreover, the S189D + A226G double mutant gained an additional trypsin-like feature, the ability of autoactivation.

These results suggest that a small number of key residues might be sufficient to confer trypsin-like activity to chymotrypsin. Here we address this question with the investigation of the structural changes in S189D chymotrypsin upon the A226G substitution.

2 Materials and Methods

2.1 Expression and Purification of S189D + A226G Chymotrypsin

The mutant zymogen was expressed as described previously [11]. Autoactivation reaction was conducted overnight in 50 mM Tris-HCl pH 8.0, 10 mM $CaCl_2$, 0.1 M NaCl reaction buffer at room temperature. The active form was first purified on an SBTI-Sepharose column, and then further purified on a Pharmacia HiPrep Sephacryl-100 column in presence of 0.2 mg/ml benzamidine. The purity of the enzyme was analyzed by SDS-PAGE, enzyme concentration was determined by active site titration with MUGB [12].

2.2 Crystallization, Structure Solution and Refinement

S189D + A226G chymotrypsin was crystallized at 20 °C by the hanging drop, vapor diffusion method, mixed 1:1 with a 0.1 M HEPES, pH 7.0 precipitant solution containing 30% PEG6000 and equilibrated against 0.5 ml of precipitant solution. The final protein concentration in the drop was 7.5 mg/ml, and the solution contained benzamidine in twofold molar excess. Crystals were grown in 6 days. The crystals were flash cooled in liquid nitrogen after briefly soaking in cryosolution containing the crystallization solution and 20% glycerol.

Crystallographic data were collected from a single crystal at ESRF on beamline ID 14 EH2 (λ = 0.933 Å) at 100 K. The X-ray diffraction displayed high mosaicity (estimated to be 1.8°). Crystallographic intensities were integrated and scaled to a resolution of 2.2 Å using Mosflm and Scala of the CCP4 package (CCP [13]). The crystal structure belonged to space group of $P2_1$. Completeness of the data was 98.5% at 2.2 Å resolution. The structure was solved by molecular replacement using the program Amore [14] of the Collaborative Computing Project 4 (CCP [13]) and a search model derived from the X-ray structure of S189D chymotrypsin (PDB entry: 1 kdq) [15]. A single well-defined solution appeared with correlation coefficient of 0.55 and R-factor of 38.3%. After the initial rigid body refinement, it became obvious that regions of 186–190 and 218–221 were significantly different compared to the S189D single mutant structure. Therefore these loops were systematically rebuilt from scratch using iterative cycles of phase improvements in RESOLVE [16], manual rebuilding using the program Coot [17] and restrained TLS refinement with REFMAC [18]. For modeling anisotropic disorder in the crystal structure two TLS groups were defined: the alpha-chain domain (residue ranges 19–146, 151–157 and the covalently bound propeptide) and beta-chain domain (residue range 158–243). Introduction of more TLS groups did not significantly improve the Rfree value. The switch to TLS refinement was justified by a drop of 5.2% in Rfree value indicating that anisotropic disorder is an important characteristic of the diffraction data, final Rfree values, however, remained still unusually high. The final protein model contained 230 residues of rat chymotrypsin including the bound propeptide. The stereochemistry of the structure was assessed with Whatcheck [19] and PROCHECK [20]. Data collection and refinement statistics are summarized in Table 1. The crystallographic data and the refined model have been deposited at the Protein Data Bank (http://www.pdb.org) with accession code 2jet.

Table 1 Crystallographic data and refinement statistics

Resolution (Å)[a]	43.44–2.20 (2.32–2.20)		
Cell parameters	Space group $P2_1$		
	a = 34.7 Å, b = 64.4 Å, c = 44.2 Å, β = 102.1°, α,γ = 90°		
Number of observed reflections[a]	29867 (3843)		
Number of unique reflections[a]	9608 (1357)		
Completeness (%)[a]	98.5 (95.8)		
Mosaicity (°)	1.8		
Oscillation range per image (°)	1		
I/σ[a]	10.2 (2.3)		
R_{merge} (%)[c,a]	9.0 (44.5)		
R_{work} (%)[d]	27.6		
R_{free} (%)[d]	33.4		
R.m.s. bond length (Å)	0.005		
R.m.s. bond angles (°)	0.8		
No. of modeled atoms			
Protein	1679		
Solvent	67		
Ramachandran plot (by Procheck)			
Favoured (%)	84.5		
Additionally allowed (%)	14.4		
Generously allowed (%)	0.5		
Disallowed (%)	0.5		
Estimates of coordinate error (Å)			
Luzzati [21]	0.48		
DPI [22]	0.32		
Average B-factors (Å^2)	Protein[b]	Waters	Overall
	55.3	56.8	55.4

[a] Values in parentheses indicate statistics for the highest resolution shell

[b] Isotropic residual B-factors without TLS contribution

[c] $R_{merge} = \Sigma\ |I_o - \langle I\rangle|/\Sigma\ I_o$ x 100%, where I_o is the observed intensity of a reflection and $\langle I\rangle$ is the average intensity obtained from multiple observations of symmetry related reflections

[d] R factor = $\Sigma\ ||F_{obs}|-|F_{calc}||/\Sigma\ |F_{obs}|$ x 100%

2.3 Structure Analysis

Superposition of X-ray structures was improved with iterative sieve fitting [23] in LSQMAN [24] until about the 50% of the matched residues were under the final distance cutoff (typically 0.5–0.7 Å). Structural comparisons were performed against bovine trypsin (1tpo) [25], S189D mutant of rat chymotrypsin (1 kdq) [15] and bovine chymotrypsin (6 gch) [26]. The precision of the crystal structure was evaluated by the Luzzati plot [21] and the DPI method [22] using the program sfcheck [13]. Using the conservative Luzzati estimate of 0.48 Å for the coordinate error, the structural changes described for the residue ranges 185–191 and 219–223 are significant.

2.4 Enzyme Activity Measurements

pH dependency of amide hydrolysis was measured on 400 μM Succinyl-Ala-Ala-Pro-Phe-AMC and 400 μM Succinyl-Ala-Ala-Pro-Lys-AMC substrates, in 40 mM buffers of Na-Acetate (pH 4.0–5.0), MES (pH 5.5–6.5), MOPS (pH 7.0–7.5), Tris-HCl (pH 8.0–8.5) CHES (pH 9.0–9.5) and CABS(pH 10.0–10.5). The reaction buffers contained 10 mM $CaCl_2$ and 0.1 M NaCl, measurements were carried out at 37 °C using a Spex Fluoromax spectrofluorimeter. The excitation and emission wavelengths were 380 nm and 460 nm, respectively. The difference between the activity values of three measurements were within 10%. Unweighted curve fittings (Fig. 4) were performed with the Origin software (OriginLab), using the following equations:

(a) for chymotrypsin on Phe substrate and the mutant on Lys substrate:

$$y = \mathrm{Vm}/(1 + 10^{-x}/10^{-\mathrm{pKa}} + 10^{-\mathrm{pKb}}/10^{-x})$$

(b) for chymotrypsin on Lys substrate:

$$y = \mathrm{Vm}_1/(1 + 10^{-x}/10^{-\mathrm{pKa1}}) + \mathrm{Vm}_2/(1 + 10^{-x}/10^{-\mathrm{pKa2}})$$

(c) for the mutant on Phe substrate:

$$y = \mathrm{Vm}_1/(1 + 10^{-x}/10^{-\mathrm{pKa}} + 10^{-\mathrm{pKb}}/10^{-x}) + Vm_2/(1 + 10^{-\mathrm{pKb2}}/10^{-x})$$

Equation (a) describes a bell-shaped pH dependence [27], where x = pH and Vm is an amplitude parameter. We derived Eqs. (b) and (c) from Eq. (a) to fit a half bell curve

with two titration steps, and to fit a bell shaped curve with two titration steps on the right side, respectively.

3 Results

3.1 The Overall Structure

The structure of the S189D + A226G rat chymotrypsin-B double mutant was determined at 2.2 Å. Crystallization was conducted in presence of the small inhibitor benzamidine, which is not visible, however, in the structure. The activation domain (an autonomous folding unit composed of loops L1 and L2 and two sterically adjacent sections: residues from the N terminus to 19, and the autolysis loop (amino acids 142–152)) including the site of substitutions is well defined, whereas residues 11–18 at the activation cleavage site and amino acids 147–150 in the cleaved autolysis loop were not visible. While chymotrypsins have a high tendency of autolytic cleavage at Tyr146, which does not inactivate the enzyme, the double mutant was stable against autolysis in solution, probably because of its trypsin-like specificity profile. However, under the conditions of crystallization (at very high protein concentration and pH 7.0, where specificity becomes more chymotryptic (see below)), autolysis was observed in the mutant also.

The removal of the methyl group at site 226 generated substantial rearrangements at the S1 region, compared to the S189D single mutant [15], while there were minor differences in other regions. When compared to wild type trypsin and chymotrypsin, substantial differences were found in the loops of the activation domain, loops L1 and L2, and especially in the autolysis loop (see equivalent Cα distances after superposition in Fig. 1).

3.2 The S1 Specificity Site

Comparison of the S189D + A226G double mutant to the S189D single mutant shows that rearrangements induced by the removal of the methyl group from position 226 in the S1 site were, in some respects, towards a wild type trypsin conformation (Fig. 2). In S189D chymotrypsin, the Ala226 side chain forms a van der Waals contact with the 191–220 cystine. The interaction appears to stabilize the cystine in a conformation which is substantially different from that in active enzymes. As a result, the cystine and the 216–218 segment of loop L2 bulge into the S1 pocket occluding it seriously [15]. In the S189D + A226G mutant, where the Ala226 methyl—191–220 cystine interaction is missing, the cystine moved out from the pocket, close to the wild type position. The 191–220 cystine is an important determinant of the S1 site structure [28], connecting loops L1 and L2 (and thereby the two sides of the S1 pocket), and has an identical conformation in trypsin and chymotrypsin. Analysis of the 191–220 cystine chi angles in the wild type and mutant enzymes according to Katz and Kossiakoff [29] showed that chirality of the cystine is right handed, except for the S189D + A226G mutant, where it is left handed. Neither wild type nor mutant 191–220 cystines fit well into any of the defined subclasses, and the mutants showed high variance in their chi values, when compared to the wild types. Upon activation, the 191–220 cystine in chymotrypsinogen undergoes a chirality change from left handed to right handed, therefore the left handed 191–220 cystine in the S189D + A226G mutant is a chymotrypsinogen-like property.

The conformation of loops L1 and L2 was also reshaped, so that loop L1 was able to stabilize the negative charge of Asp189, which therefore turned towards the interior of the S1 site from its surface position in S189D

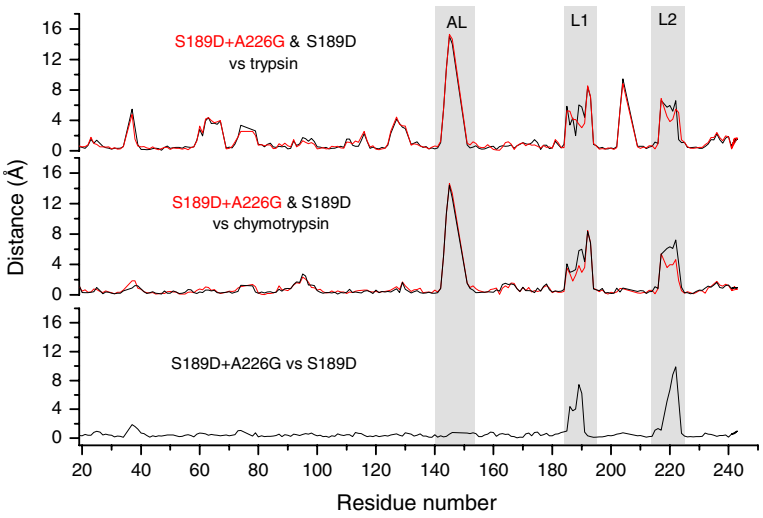


Fig. 1 Cα distance comparison plots. Cα distances of corresponding positions were compared using superimposed structures: the single and double mutant enzymes against wild type trypsin and chymotrypsin, and also S189D + A226G vs. S189D. Highlighted areas refer to the autolysis loop (AL), and loops L1 and L2

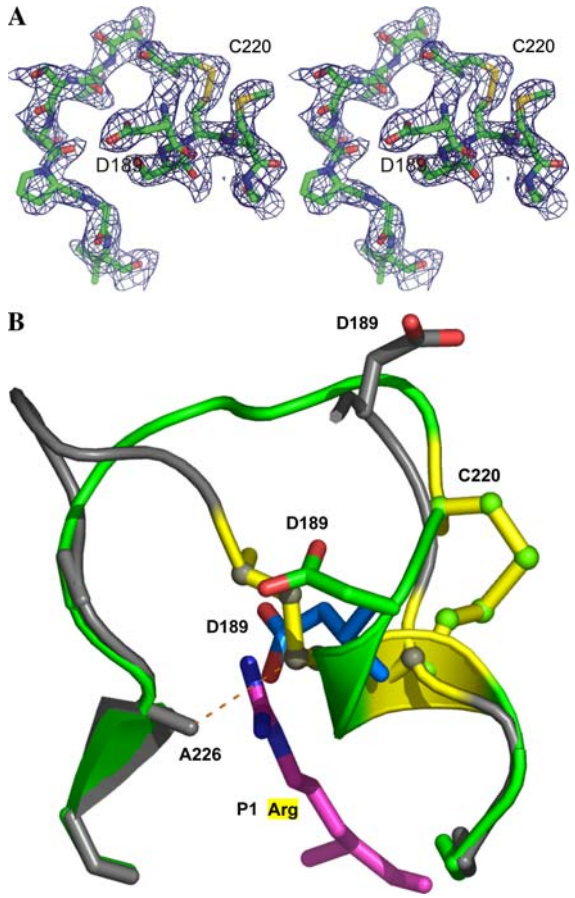


Fig. 2 S1 specificity site structures of mutants S189D (gray) and S189D + A226G (green) (**A**) Stereo view of the final 2Fo-Fc electron density maps of S189D + A226G, from 189 to 193, and from 220 to 227 regions contoured at 1σ level (blue). **(B)** The following parts of the S1 sites are shown and compared: the backbone from 189 to 193, and from 220 to 227; the 191–220 cystine (gold in S189D, yellow in S189D + A226G); residues 189 and 226. Asp189 (blue) and P1Lys (magenta) positions in wild type trypsin are also shown. The Ala226 side chain of S189D is in Van der Waals interaction (distance is 3.9 Å, blue dotted line) with the 191–220 cystine connecting the two loops. This figure was generated with the program Pymol (http://pymol.sourceforge.net/)

chymotrypsin. However, H-bonding partners that dissipate the half-buried charge of the Asp189 sidechain, the hydroxyl group of Ser186 and Ala185 main chain nitrogen atom, were still different from those in trypsin, which might contributed to the remaining distortions in loop L1 and, consequently, in loop L2. Thus, similarly to the S189D single mutant, part of loop L2 still bulged into the S1 pocket so that residue 217 was in the way of a P1 residue. The structure of S189D + A226G also retained the chymotrypsinogen-like features characteristic to S189D chymotrypsin at the S1 site, such as the oxyanion hole was not formed and the side chain of Asp194 formed a salt bridge with His40 instead of the N-terminus. Crystal contacts may also influence the X-ray structure of proteins. The S189D + A226G and S189D mutant chymotrypsin crystallized in different spacegroup, hence the crystal environment is different in the two crystal structure. Since the overall fold is almost identical in the two structures (except for the 185–191 and 219–223 regions) we can assume that the general structure is mostly unaffected by the crystal contacts. In the S189D + A226G mutant residue range 219–223 are stabilized by two hydrogen bond by a neighboring protein molecule ($Leu123_N$–$Ser221_O$ and $Ile47_O$—$Ser221_{O\gamma}$), while the 185–191 region is free of crystal contacts. In the S189D mutant structure, both regions are involved in crystal contacts through the $Phe114_N$—$Ser223_{O\gamma}$, $Glu49_{O\varepsilon1}$—$Ser223_N$, $Glu49_{O\varepsilon2}$—$Thr222_N$ and $Glu49_{O\varepsilon2}$—$Gly187_N$ hydrogen bonds. In these chymotrypsin mutants these flexible loops adopt various conformations, which might be explained on one hand by the structural differences genuinely attributed to the A226G mutation. On the other hand, there is little free energy difference between the alternative conformations and, unlike the rest of the structure; they can be easily influenced by the environment. In the crystalline form this means that weak crystal contacts are sufficient to alter the conformation whereas in solution changes in pH, ionic strength etc. may play a similar role.

3.3 The Activation Domain

The diagram of Cα distances between the mutant and wild type chymotrypsins showed large differences in the activation domain, especially in the autolysis loop (Fig. 1) which is cleaved in both crystal structures. In wild type enzymes, this loop makes contacts with loops L1 and L2 at the 191–220 cystine, and also with the N-terminal residues 16–17. Two H-bonds are formed between $Lys143_O$—$Ile16_N$ and $Lys143_N$—$Met192_O$. All these contacts seem to be crucial to the stabilization of the active conformation of the wild type enzymes. In both the S189D and the S189D + A226G mutants, however, the autolysis loop was moved towards the substrate binding site, occupying the P2–P2' substrate positions (Fig. 3). Contacts with the 191–220 cystine and the N-terminus were lost, and the new conformation was stabilized by 5 H-bonds to amino acids 194 and 195, as well as by van der Waals contacts with residues 40, 41, 57, 58, 217, 218. This position of the autolysis loop sterically blocked the formation of the oxyanion hole, and brought the amide nitrogen of Lys145 in H-bond distance from the active hydroxyl group of Ser195. These structural features of the activation domain showed that both the single and the double mutants crystallized in a conformation which is completely incompatible with canonical binding which is presumably a prerequisite for efficient activity [30–32 33].

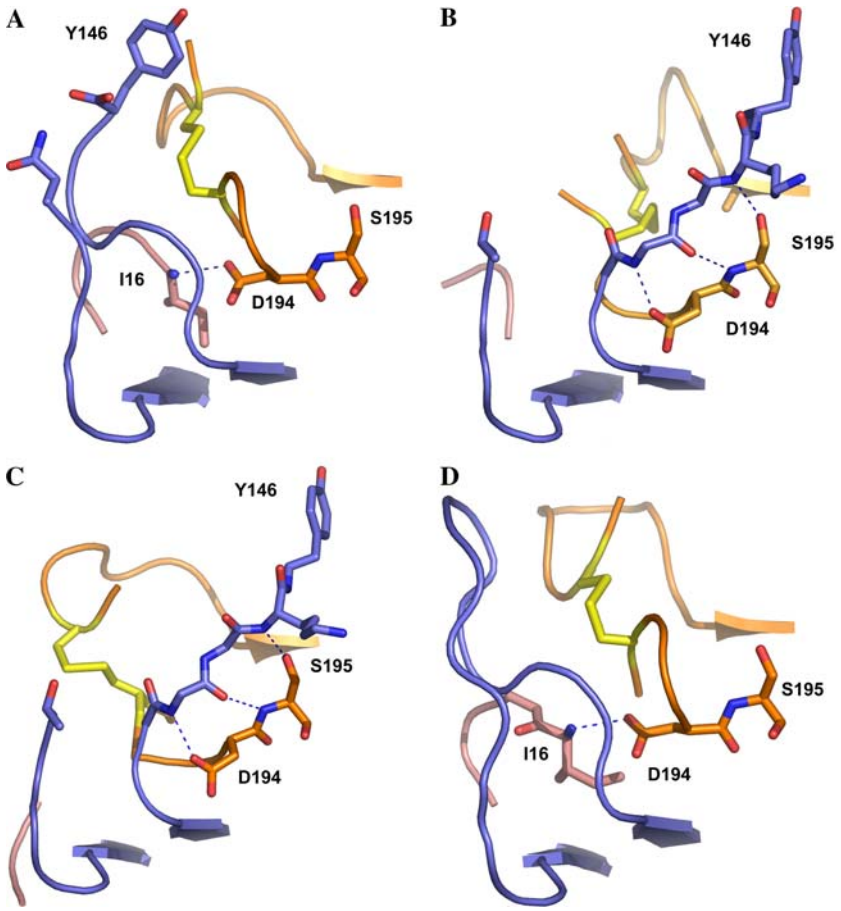


Fig. 3 Activation domain structures, **A**: chymotrypsin, **B**: S189D, **C**: S189D + A226G, **D**: trypsinThe following parts of the activation domains are shown and compared: loop L1 from 191 to 195 and sites 220 to 227 (orange); the 191 to 220 cystine (yellow); the autolysis loop (blue) and the N-terminus (salmon). Dotted blue lines represent H bonds

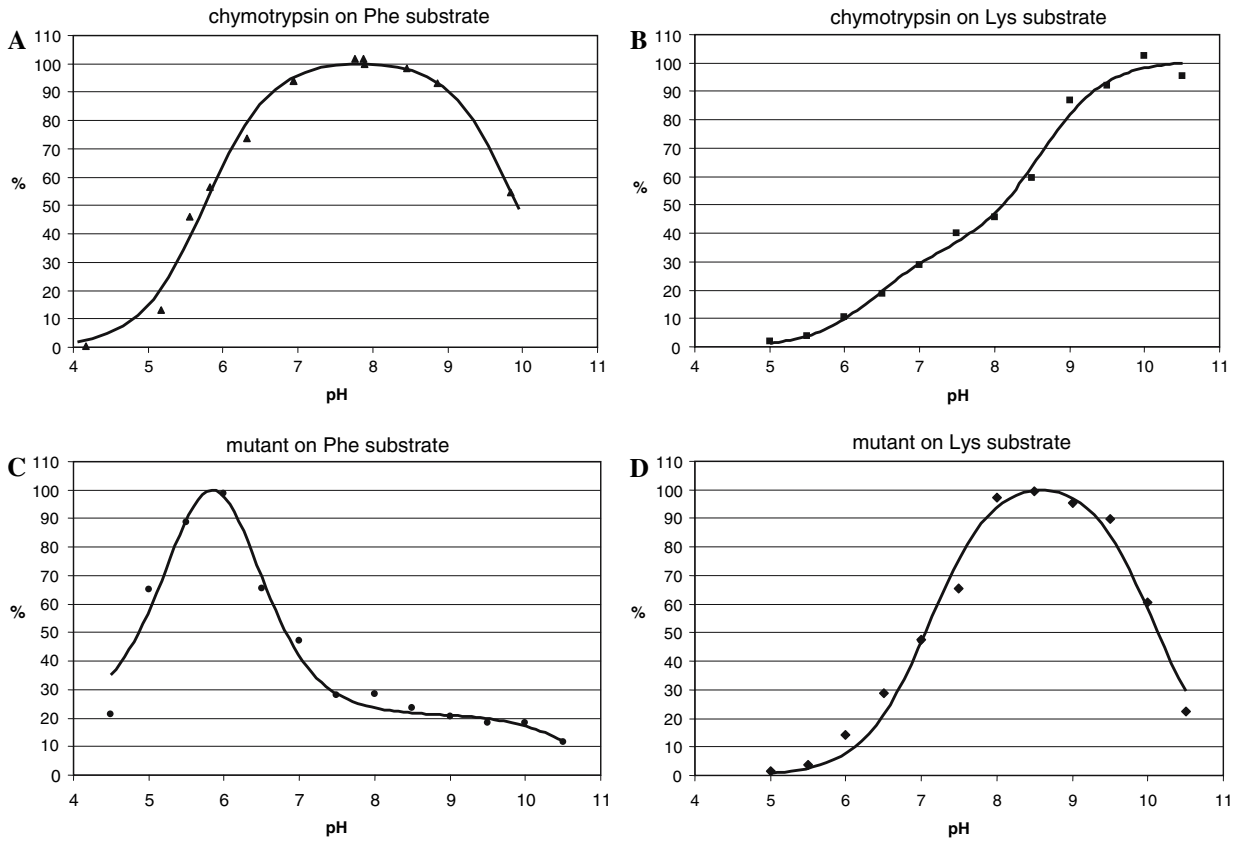
3.4 The pH Profile of Activities

While several changes in the structure of the double mutant might be favorable for the activity (e.g. the rearrangement of Asp189 and the 191–220 cystine), the remaining serious distortions in the S1 region are difficult to reconcile with the increase in enzyme activity relative to the S189D mutant [11]. It is known that the S1 region of wild type trypsin and chymotrypsin is sensitive to changes in the environment. For example, the effect of pH is due primarily to alterations in this region, to the loss of the salt bridge between the Asp194 and the N-terminal amino group (pK_b) which disrupts the oxyanion hole, and to the protonation of His57 (pK_a). Previous studies [34, 35] have found that there is a pH dependent equilibrium of the active and inactive states of wild type trypsin and chymotrypsin, depending on the structure of the S1 region. Thus altered parameters of the pH profile of the mutants can also be a useful indicator of perturbations in the solution structure of the S1 region due to amino acid substitutions and perhaps are informative about the congruence of the solution and the crystal structures. Therefore we compared the influence of pH on the activity of the S189D + A226G double mutant and wild type chymotrypsin using succAAPK-AMC and succAAPF-AMC (tryptic and chymotryptic) fluorogenic substrates.

On its preferred succAAPF-AMC substrate, wild type chymotrypsin showed wide activity range around a pH optimum at 7.9 (Fig. 4) with $pK_a = 5.8$ and $pK_b = 9.9$. At the same time, on the tryptic substrate succAAPK-AMC its pH optimum shifted upward to above 9.5, and in the acidic side two pH steps are evident ($pK_{a1} = 6.4$ and $pK_{a2} = 8.6$).

In the case of the S189D + A226G double mutant on the succAAPF-AMC substrate, the pH profile had a narrower range of activity, and showed a significant acidic shift in the parameters relative to the wild type enzyme with two pH steps on the basic side. (The apparent values of pH optimum, pK_a, pK_{b1} and pK_{b2} were 5.9, 4.9, 7.0 and 8.7, respectively.) The activity on tryptic substrate also showed narrower range of pH optimum relative to wild type chymotrypsin but, with a maximum at pH 8.5. The higher maximum was due solely to the basic shift in pK_a from 5.8 to 7.1. The pK_b (10.1) was the same as that of the wild type enzyme on the Phe substrate.

Taken together, the pH profiles on P1 Lys and Phe substrates show that, unlike wild type chymotrypsin, the mutant has either a trypsin- or chymotrypsin-like



E

Enzyme	Substrate	pK_a	pK_{a2}	pH_{opt}	pK_b	pK_{b2}
WT chymotr.	Phe ▲	5.8		7.9	9.9	
WT chymotr.	Lys ■	6.4	8.6	9.5<	-	
S189D+A226G	Phe ●	5.5		5.9	6.2	10.6
S189D+A226G	Lys ♦	7.1		8.5	10.1	

Fig. 4 pH dependency experiments A: pH dependence of the activity of the S189D + A226G mutant and wild type chymotrypsin was measured with 400 μM succAAPKAMC and succAAPFAMC fluorogenic substrates, in 40 mM buffers of Na-Acetate (pH 4.0–5.0), MES (pH 5.5–6.5), MOPS (pH 7.0–7.5), Tris-HCl (pH 8.0–8.5) CHES (pH 9.0–9.5) and CABS(pH 10.0–10.5). The reaction buffers contained 10 mM $CaCl_2$ and 0.1 M NaCl, measurements were carried out at 37 °C, enzyme concentration was 0.02 μM. Activity is represented as the percentage of the calculated maximum (= 100%) activity. (**A**): chymotrypsin on Phe substrate (from Venekei et al., unpublished) (**B**): chymotrypsin on the Lys substrate. (**C**): mutant on Phe substrate (**D**): mutant on Lys substrate. (**E**): Calculated pK and pH_{opt} values for the titration curves. Note that in the case of the mutant on Phe substrate, the calculated pK_a and pK_b values do not match the apparent pK values, see the discussion for an explanation

specificity in a pH dependent manner, such that the trypsin-like specificity at pH 8.0 becomes chymotrypsin-like at pH 6.0.

4 Discussion

The structural alterations in the S189D mutant upon the A226G substitution can partly be attributed to the loss of the interaction between the methyl group of Ala226 and the 191–220 cystine: the shift of the cystine and parts of loop L1 and L2 to positions which are closer to those in wild type trypsin, and the stabilization of the Asp189 residue in an internal conformation, closer to the S1 pocket (Fig. 2). With a direct effect on substrate binding, these rearrangements might explain the significant shift in specificity towards trypsin. However, according to the crystal structure a number of structural features remained which do not allow canonical substrate binding and catalysis, and therefore are inconsistent with the observed three orders of magnitude increase in catalytic efficiency. An explanation might be that the crystal structure does not represent fully the structure in solution, especially in the S1 region which, as part of the activation domain is known to be sensitive to

external effects which can easily be more pronounced in mutant enzymes [36].

Based on detailed analysis of the pH dependence of refolding and the knowledge of structure, the pK values can provide some information on the structure of the S1 region in solution, such as the protonation state of the catalytic His57, and the existence and strength of the salt bridges of Asp194 γ carboxyl with either His40 or the Ile16 N-terminal amine (the assigned pK values of wild type chymotrypsin are: $pK_{a1} = 5.8$, $pK_{a2} = 7.0$ and $pK_b = 9.9$, respectively [34, 35]). Of the latter two, the Asp194 γ carboxyl—Ile16 N-terminal amine interaction is characteristic to the active enzyme, while the Asp194 γ carboxyl—His40 interaction is a zymogen like state of Asp194.

In the case of wild type chymotrypsin, two of the titration steps, $pK_{a1} = 5.8$ and $pK_b = 9.9$, are visible (Fig. 4A), when the titration is followed by activity measurement on chymotryptic substrate. When the tryptic substrate, succ-AAPK-AMC is used $pK_{a1} = 6.4$ is probably due to the titration of the Asp194—His40 interaction [34], and the $pK_a = 5.8$ of the titration of His57 is not visible due to the very low activities below pH 6.0. The pK_{a2} at pH 8.6 in this titration curve results probably from the deprotonation of the P1 Lys residue, which makes it better accepted by the hydrophobic S1 site of chymotrypsin. This gain of activity seems to compensate the loss of it due to the titration of the Asp194—Ile16 N-terminal salt bridge, thus the titration step of wild type chymotrypsin at pH $\sim$ 10 is not seen.

In contrast, deprotonation of the P1 Lys residue is seemingly not favorable for the S189D + A226G double mutant, therefore a $pK_b = 10.1$ appears, which might correspond to the $pK_b = 9.9$ of the wild type enzyme measured on the Phe substrate above. Thus, at variance with the structural data, this titration indicates the existence of the Asp194 γ carboxyl—Ile16 N-terminal amine salt bridge, and also suggests that the interaction might be as strong as in the wild type enzyme.

The S189D + A226G mutant on the succAAPF-AMC chymotryptic substrate shows a narrow titration curve with unique pK values (Fig. 4B). It is tempting to interpret $pK_{b1} = 6.2$ as the titration of the Asp189 side chain, which—when loosing its charge—becomes more favorable for the activity on the substrate with hydrophobic P1 residue. We speculate that, when lowering the pH, this effect could give rise to a further gain of activity but it is limited by the loss of activity due to the parallel titration of His57 ($pK_a = 5.5$). It seems reasonable to suppose that due to their closeness, the two titrations influence each other thus the apparent pK_a and pK_{b1} of the curve are combination of two titration effects and, therefore, do not show a 50% titration of the residues. Data points between pH 7.5 and pH 10.0 might suggest an additional titration with a pK of $\sim$8.8, which is hard to interpret but if it is a downshifted pK of the titration of the Asp194 γ carboxyl—Ile16 N-terminal amine salt bridge, then it indicates a weaker interaction than in the wild type enzyme, or when the P1 substrate residue is Lys (see above).

This latter comparison would show a stabilization of the salt bridge and the S1 region in the double mutant by the positive charge of P1 residue. A similar effect was observed when trypsinogen was complexed with BPTI: the S1 region (partially) rearranged from the zymogen to an active conformation, which included the formation of the Asp194 γ carboxyl—Ile16 N-terminal amine salt bridge [37]. In the case of the S189D + A226G double mutant, an attraction between the P1 Lys and the slightly displaced Asp189 might induce similar rearrangements.

Taken together our data offer some structural explanation for the effect of the A226G substitution on the activity and specificity of the S189D chymotrypsin mutant. Whereas the crystal structure shows that, of the constituents of the S1 region, only the Asp189 and the 190–220 cystine are in better positions for activity but others are in very unfavorable ones, the pH profile of activity indicates alternative structures in solution, which seem to be dependent on the P1 residue of the substrate.

Thus the A226G substitution, on one hand, induces in itself rearrangements of the inherently flexible S1 region of the S189D chymotrypsin mutant towards trypsin-like conformation (seen in the crystal structure). On the other hand, in solution it appears to permit substrate induced conformational states, which enhance the activity, while such changes may be restricted in the S189D mutant. We conclude that Asp189 is indeed a key specificity determinant, however, the stabilization of this half-buried charge requires a unique interaction system involving even more extended structures.

Acknowledgements We are grateful for the beamline staff at the ESRF ID14–2 beamline for their expert assistance. This work was supported by the Hungarian Research Fund OTKA TS049812 and T047154 to L.G., G. K. acknowledges the support from EMBO and CEA.

References

1. Schechter I, Berger A (1967) Biochem Biophys Res Commun 27:157–162
2. Gráf L, Jancsó A, Szilágyi L, Hegyi Gy, Pintér K, Náray-Szabó G, Hepp J, Medzihradszky K, Rutter WJ (1988) Proc Natl Acad Sci USA 85:4961–4965
3. Hedstrom L, Szilágyi L, Rutter WJ (1992) Science 255: 1249–1253
4. Hedstrom L, Perona JJ, Rutter WJ (1994) Biochemistry 33: 8757–8763

5. Hedstrom L, Farr Jones S, Kettner CA, Rutter WJ (1994) Biochemistry 33:8764–8769
6. Gráf L (1995) In: Zwilling R (ed) Natural sciences and human thought. Springer-Verlag, Berlin, Heidelberg, pp 139–148
7. Steitz TA, Henderson R, Blow DM (1969) J Mol Biol 46: 337–348
8. Venekei I, Szilágyi L, Gráf L, Rutter WJ (1996) FEBS Lett 379:143–147
9. Craik CS, Largman C, Fletcher T, Roczniak S, Barr PJ, Fletterick RJ, Rutter WJ (1985) Science 228:291–297
10. Wilke ME, Higaki JN, Craik CS, Fletterick RJ (1991) J Mol Biol 219:525–532
11. Jelinek B, Antal J, Venekei I, Gráf L (2004) Protein Eng Des Sel 17:127–131
12. Jameson GW, Adams DV, Kyle WS, Elmore DT (1973) Biochem J 131:107–117
13. CCP 4 (1994) Acta Cryst D 50:760–763
14. Navaza J (2001) Acta Crystallogr D Biol Crystallogr 57: 1367–1372
15. Szabó E, Venekei I, Böcskei Zs, Náray-Szabó G, Gráf L (2003) J Mol Biol 331(5):1121–1130
16. Terwilliger TC (2000) Acta Cryst D56:965–972
17. Emsley P, Cowtan K (2004) Acta Cryst D60:2126–2132
18. Murshudov GN, Vagin AA, Dodson EJ (1997) Acta Cryst D53:240–255
19. Hooft RW, Vriend G, Sander C, Abola EE (1996) Nature 381:272
20. Laskowski RA, MAcArtur MW, Moss DS, Thornton JM (1993) J Appl Crystallog 26:283–291
21. Luzzati PV (1952) Acta Cryst 5:802–810
22. Cruickshank DWJ (1999) Acta Cryst D55:583–601
23. Lesk AM (1991) Protein architecture: a practical guide. IRL Press, Oxford
24. Madsen D, Kleywegt GJ (2002) J Appl Crystallog 35:137–139
25. Marquart M, Walter J, Deisenhofer J, Bode W, Huber R (1983) Acta Crystallogr Sect B v39:480
26. Brady K, Wei AZ, Ringe D, Abeles RH (1990) Biochemistry v29:7600–7607
27. Segel IH (1993) Enzyme Kinetics: Behavior and analysis of rapid equilibrium and steady-state enzyme systems, Wiley Classics Library Edition
28. Dementiev A, Dobó J, Gettins PGW (2006) J Biol Chem 281(6):3452–3457
29. Katz B, Kossiakoff A (1986) J Biol Chem 261(33):15480-15485
30. Fodor K, Harmat V, Neutze R, Szilágyi L, Gráf L, Katona G (2006) Biochemistry 21;45(7):2114–2121
31. Radisky ES, Lee JM, Lu CJ, Koshland DE Jr (2006) Proc Natl Acad Sci U S A 103(18):6835–6840
32. Fuhrmann CN, Daugherty MD, Agard DA (2006) J Am Chem Soc 128(28):9086–9102
33. Liu B, Schofield CJ, Wilmouth RC (2006) J Biol Chem 281(33):24024–24035
34. Fersht AR (1972) J Mol Biol 64(2):497–509
35. Verheyden G, Matrai J, Volckaert G, Engelborghs Y (2004) Prot Sci 13:2533–2540
36. Szabó E, Böcskei Zs, Náray-Szabó G, Gráf L (1999) Eur J Biochem 263(1):20–26
37. Bode W, Schwager P, Huber R (1978) J Mol Biol 118:99–112