PROTEINS: Structure, Function, and Genetics 40:207–217 (2000)

# Electrostatics of Mesophilic and Psychrophilic Trypsin Isoenzymes: Qualitative Evaluation of Electrostatic Differences at the Substrate Binding Site

**Alemayehu A. Gorfe, Bjørn O. Brandsdal, Hanna-Kirsti S. Leiros, Ronny Helland, and Arne O. Smalås***
*Department of Chemistry, University of Tromsø, Tromsø, Norway*

***ABSTRACT*** **A qualitative evaluation of electrostatic features of the substrate binding region of seven isoenzymes of trypsin has been performed by using the continuum electrostatic model for the solution of the Poisson-Boltzmann equation. The sources of the electrostatic differences among the trypsins have been sought by comparative calculations on selective charges: all charges, conserved charges, partial charges, unique cold trypsin charges, and a number of charge mutations. As expected, most of the negative potential at the $S_1$ region of all trypsins is generated from $Asp^{189}$, but the potential varies significantly among the seven trypsin isoenzymes. The three cold active enzymes included in this study possess a notably lower potential at and around the $S_1$-pocket compared with the warm active counterparts; this finding may be the main contribution to the increased binding affinity. The source of the differences are nonconserved charged residues outside the specificity pocket, producing electric fields at the $S_1$-pocket that are different in both sign and magnitude. The surface charges of the mesophilic trypsins generally induce the $S_1$ pocket positively, whereas surface charges of the cold trypsins produce a negative electric field of this region. Calculations on mutants, where charged amino acids were substituted between the trypsins, showed that mutations in Loop2 (residues 221B and 224) and residue 175, in particular, were responsible for the low potential of the cold enzymes. Proteins 2000;40:207–217.** © 2000 Wiley-Liss, Inc.

**Key words: Poisson-Boltzmann; cold-adapted trypsins; potential; specificity pocket; substrate binding**

## INTRODUCTION

Psychrophilic (cold-adapted) enzymes have adapted to cold environments by possessing a high catalytic efficiency ($k_{cat}/K_m$-ratio) at all temperatures, compared with their mesophilic counterparts. The source of the increased efficiency does, however, seem to vary between enzyme systems (for a review, see Ref. 1). For example, the α-amylase from an Antarctic bacteria (*A. haloplanctis*) possesses an approximately 4- to 7-fold increase in the turnover number compared with the porcine counterpart, whereas the Michaelis-Menten constant is virtually unchanged.[2] In contrast, the increased efficiency of some trypsins and other pancreatic serine proteinases from cold-adapted fish species seems to arise from a combination of reduced $K_m$ and increased $k_{cat}$.[3] The 20- to 40-fold increase in catalytic efficiency of the anionic form of trypsin (AST) from Atlantic salmon, compared with mammalian trypsins, is due to a 2–5 times higher $k_{cat}$ in the temperature range 4–37°C combined with a 10-fold reduction of $K_m$ at all temperatures.[4] The anionic form of salmon trypsin has also been found to be less thermostable and nonresistant to changes in pH relative to the mesophilic trypsins. Similarly, Asgeirsson et al.[5] and Genicot et al.[6] have determined reactivities and stabilities of cod trypsin (CT1), *Gadus morhua,* and Antarctic fish trypsin (AFT), *Paranotothenia magellanica*, respectively, and found them catalytically more efficient and less thermostable than their mammalian counterparts. A cationic salmon trypsin (CST), however, resembles the mammalian trypsins in activity and stability and thus cannot be characterized as psychrophilic.

The increased binding affinity of anionic salmon trypsin compared with bovine trypsin (BT) has, in addition to measurements of $K_m$-values, been demonstrated from studies on the association to $P_1$ mutants (primary binding, nomenclature of Schechter and Berger[7]) of the bovine pancreatic trypsin inhibitor (BPTI).[8] In this study the lysine side chain of the $P_1$ position of wild-type BPTI was mutated to 17 coded amino acids. The succeeding association constant measurements showed that there were only small differences between salmon and bovine trypsin for the noncognate $P_1$ mutants, whereas the anionic salmon trypsin showed an approximately 100-fold stronger association for the variants of BPTI with the cognate lysine and arginine side chains. Furthermore, because the association constants for $P_1$ Gly, which has no side chain penetrating the pocket, were almost identical for BT and AST, the differences for cognate $P_1$ residues may be predominantly due to stronger electrostatic interactions of AST.[8] Further support for the stronger electrostatic interactions was obtained from comparative crystal structure analysis of AST and BT in complex with BPTI, were the $P_1$ Lys was found with a direct interaction to $Asp^{189}$ in the AST-complex, whereas via a water molecule for BT.[9]

*Correspondence to: Arne O. Smalås, Department of Chemistry, University of Tromsø, N-9037, Tromsø, Norway. E-mail: Arne.Smalas@chem.uit.no

Received 14 October 1999; Accepted 9 March 2000

**TABLE I. Trypsins Selected for Comparative Electrostatic Potential Calculations†**

| Abbreviation | Species | pI | Temperature adaptation | PDB entry | Reference |
|---|---|---|---|---|---|
| AST | Atlantic salmon | 5.6 | Psycrophilic | 2TBS | 13 |
| AFT | Antarctic fish | 6.0 | Psycrophilic | Model | 12 |
| CT1 | Atlantic cod | 6.6 | Psycrophilic | Model | 12 |
| CST | Atlantic salmon | 8.1 | Mesophilic | 1A0J | 26 |
| RT | Rat | 4.7 | Mesophilic | 1DPO | 27 |
| PT | Pig | 7.9 | Mesophilic | 1MCT | 28 |
| BT | Cattle | 8.3 | Mesophilic | 3PTB | 29 |

†Origin, temperature adaptations and calculated iso-electric nature of the trypsins are shown.

Trypsins from different species have generally been thoroughly studied, comparing amino acid sequences[10] and 3D structures, and sources of the cold-adaptation behavior have been sought.[11–13] The reduced temperature stability of the cold-adapted trypsins is likely to originate from a combination of reduced core packing and a lower number of hydrogen bonds in certain regions of the structures, as found by Leiros et al.[12] The source of the increased activity is, however, more difficult to predict. The lack of several interdomain hydrogen bonds close to the catalytic site in the contact region between the two domains is likely to impose higher flexibility and, thereby, possible increased catalytic efficiency. Furthermore, all the cold-adapted trypsins studied so far are anionic, but several mesophilic forms have an even more negative overall charge. However, the distribution of negatively charged residues seems to be unique for the cold trypsins. The charge of the N-terminal halves is relatively similar among all trypsins, whereas the C-terminal domains of the cold active forms are more heavily populated with acidic residues than any of the warm trypsins. Thus, the cold trypsins are much more polar, and because the C-terminal domain comprises the specificity pocket, additional negative charge in this region is likely to affect substrate affinity and binding. This result, in combination with the results from association constant measurements, led us to suspect that the electrostatics play a major role, at least for the binding component of the increased activity of the cold-adapted trypsins.

The present study has, by using the continuum electrostatic model for the solution of Poisson-Boltzmann equation, aimed at exploring the roles of electrostatic potentials in governing the binding affinities of seven trypsin isoenzymes. The anionic form of salmon trypsin (AST), along with Antarctic fish trypsin (AFT) and Atlantic cod trypsin (CT1) represent the cold-adapted trypsins. All three are anionic and are classified as cold active on the basis of their relative higher catalytic activity and lower heat and acid pH stability. Rat trypsin (RT), bovine trypsin (BT), porcine trypsin (PT), and the cationic salmon trypsin (CST) represent the mesophilic forms. The latter three are cationic, whereas the rat trypsin (RT) included in the study represents the anionic form of the mesophilic trypsins.

## MATERIALS AND METHODS

### Calculation and Parameters

The finite difference Poisson-Boltzmann calculations were performed by the DelPhi software package (DelPhi version 3.0[14,15]). The boundary of the protein-solvent interface was generated by a probe radius of 1.4 Å, as defined by Richards[16], and an electrolyte-exclusion radius of 2 Å has been used. To visualize surface potentials and to compare the trypsins largely in qualitative terms, the linearized form of the Poisson-Boltzmann equation was used for speed. However, the effect of nonlinearity was checked with sample calculations and was insignificant for the purpose of this work. The proteins electronic polarizability was accounted for by assigning a dielectric constant ($\epsilon_p$) of value 2 in the interior, assuming that the protein molecules are static and no major conformational reorganization occur.[17] The bulk solvent dielectric constant ($\epsilon_s$) of value 80 and physiological ionic strength (i.e., 0.145 M) was used for all the calculations. Rotational averaging and two-step focusing was used to improve the accuracy.[18] An initial run with a coarse grid of 20% fill was focused on to a final finer grid of 90% fill on a cube of $65^3$ grid size with a final resolution of 1.1 Å/grid. Atomic radii were assigned according to Connolly.[19] Note that radii of hydrogen atoms are 0. Charges were taken from two different force fields: surface (permanent) charges and full partial charge sets from the AMBER force field.[20] For all calculations involving surface charges of the trypsins, the assignment of charges on acidic and basic residues and terminals has been made assuming neutral pH; Asp and Glu, −0.5 on each side chain oxygen atom; Arg, +0.5 on the two guanido nitrogens; Lys, +1.0 on the Nε nitrogen atom; +0.25 on the two side chain nitrogen atoms of His; terminal nitrogen, +1.0; terminal oxygen atoms, −0.5 on each; and +2.0 on the calcium ($Ca^{2+}$) ion.

### Structure Selection and Modeling

Five structures from the Protein Data Bank (PDB) at Brookhaven national laboratory[21] and two homology models have been chosen (Table I). The trypsins were selected on the basis of criteria that there is a high primary and tertiary structure homology while they have different catalytic efficiencies, isoelectric nature, and temperature adaptations. The availability of kinetic and stability data from previous studies is also taken into consideration.

Atlantic salmon trypsin (AST) is the first and the only of the trypsins characterized as cold active (psychrophilic) for which atomic level X-ray structure determination has been performed. Therefore, homology models were built for Antarctic fish trypsin (AFT) and Atlantic cod trypsin (CT1), so that the catalytically more active psychrophilic trypsins are represented by AST, AFT, and CT1. Both models were built from the AST crystal structure because of the high-sequence identity. X-ray structures of all the mesophilic trypsins were obtained from the protein data bank (Table I).

To achieve a common basis for the comparisons and to avoid bias from, for example, crystal packing, the coordinates (X-ray and models) were subjected to a short molecular dynamic simulation (MD) simulation. The software package CHARMm[22] was used for the simulations, and the SQUID program[23] was used for analysis of the molecular dynamic trajectories. All the starting models were superimposed onto AST, and 22 internal waters from AST and the calcium ion were included. This was to ensure that effects of the internal waters to the electrostatic potentials of protein charges would be the same (or nearly so) for all seven trypsins. The individual systems were energy minimized by 100 cycles "steepest decent" (SD). The molecular systems were solvated by overlapping each residue with a sphere of 10 Å preequilibrated waters. Subsequently, water molecules that overlapped with the protein or previously placed waters were removed. The overlap criteria was $<2.8$ Å from the water oxygen atoms to another nonhydrogen atom, resulting in 1,533–1,631 water molecules for each of the seven models.

The solvated systems were further energy minimized (200 steps of SD) followed by the dynamic simulations: 6 ps thermalization from 0 to 300 K in 5-K steps with the assignment of velocities from a Gaussian distribution, 30 ps equilibration with rescaling of the velocities if the temperature deviated by more than 10 K from 300 K, and 20 ps productive simulation with no rescaling or reassignment of velocities. Integration steps of 0.001 ps were used, and the bond lengths were fixed by the constraint algorithm SHAKE. The models used for the electrostatic calculations were made from the average of the last 10 ps of the production, first minimizing only the water probe 500 cycles SD, and then only the protein 500 cycles SD and 500 cycles conjugate gradient (CG), and finally the complete system 500 cycles SD and 500 cycles adopted basis Newton-Rahpson (ANBR). The root-mean-squares (RMS) deviations for main-chain atoms from the starting structures varied around 1 Å through the equilibration and productive simulations for all seven models. Visualization and analysis of the calculated electrostatic potentials was made with GRASP.[24] Mutations of amino acids have been made with Graphics software known as "O".[25]

## RESULTS

The asymmetric distribution of charged groups on the protein surface (see Refs. 30–32) and specific charged groups at functionally important sites[33–36] affect the function of proteins. Therefore, we were interested in looking into possible differences caused by the cumulated effects of charged surface residues and, in particular, the charge distribution at the substrate binding sites around $Asp^{189}$. With this in view, potentials were calculated for seven trypsins, first with all charges included and then with all charges except the charge of $Asp^{189}$. The differences in molecular surface potential contours produced from all atomic charges (including $Asp^{189}$) are small, but still it appears that the mesophilic trypsins have less negative potentials around the specificity pocket than the cold-adapted forms have (results not shown). However, to be able to interpret possible differences between the trypsins more easily, all further calculations were performed with the charge on $Asp^{189}$ turned off. Because of the additivity of electrostatic potentials in the linear Poisson-Boltzmann equation, the difference between the two potentials at atoms and surfaces can be taken as the sole effect of $Asp^{189}$.

### Potentials With $Asp^{189}$ Uncharged

The 3D constant surface potential contours at +/− 1kT/e (blue and red, respectively) in Figure 1 illustrate the potentials when the charge of $Asp^{189}$ is excluded. The specificity pocket is negative (red) in the cold trypsins and positive (blue) in the other trypsins. This is especially pronounced at the base of the pocket comprising backbone of residues 217, 219, and 189 where the most important enzyme-substrate interactions take place. The mesophilic trypsin contours occupy the upper regions of the crevice and are relatively far from the base of the pocket. The cold enzyme contours cover most of the cleft area and correspond well to the walls of the specificity pocket, whereas the mesophilic enzyme contours cover only a fraction of the crevice. This is particularly surprising for RT, given its net negative charge and higher number of Asp and Glu residues. In fact, rat trypsin produces the smallest 2D negative contour area (not shown) when Asp189 is uncharged. In contrast, CST with a net positive charge of +8 units, produced a contour area larger than the other mesophilic enzymes. If these two trypsins are selected as markers, the potentials produced in the specificity pocket do not seem to be functions of net charges of the trypsins considered.

The average potentials of the residues forming the specificity pocket (residues 189–193, and 216–221; Table II, Fig. 2) also indicate that all trypsins have potentials rather similar for one wall of the pocket (residues 189–193), whereas the cold-adapted trypsins have a significantly lower potential than the mesophilic counterparts for the other wall of the binding pocket (residues 216–221). $Asp^{189}$ is a conserved residue, and there is a general sequence and structural homology among the trypsins. The structural differences among the specificity pockets are small, i.e., the orientations of $Gln^{192}$ are slightly different, and position 217 being Ser in BT while Tyr in the other trypsins. The potentials of the specificity pocket were expected to be relatively similar among the trypsins, and the observed differences must, therefore, originate from charges distant from it.

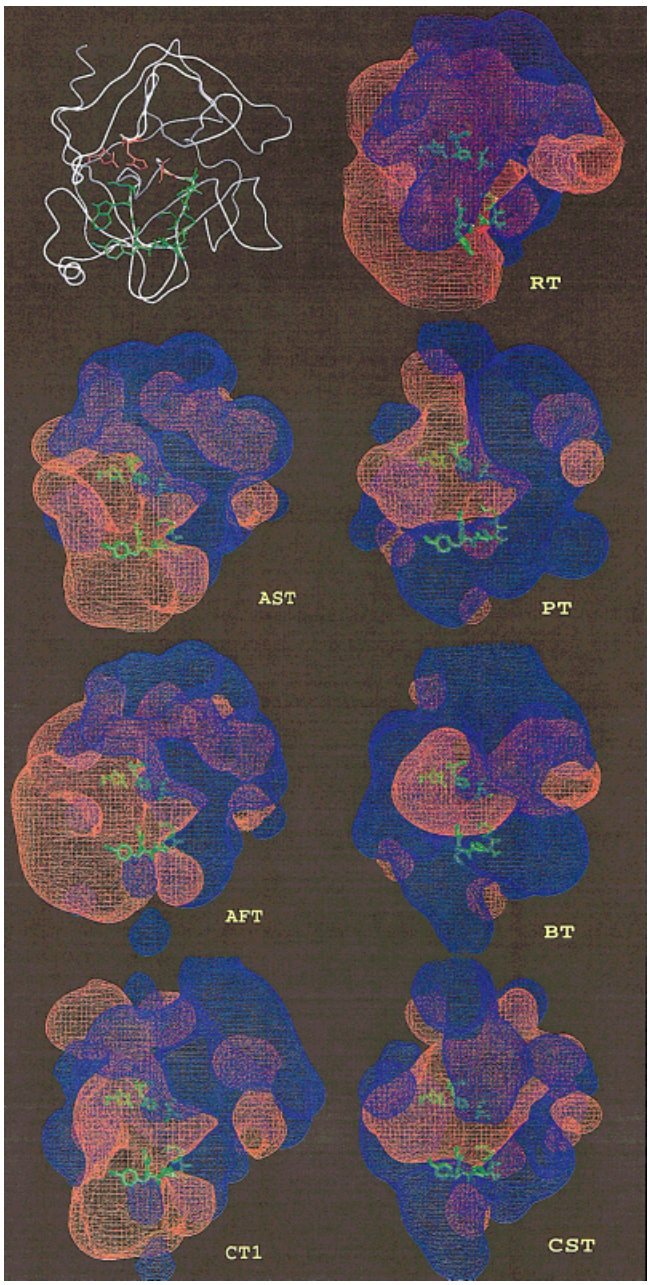


Fig. 1. Electrostatic surface potentials of trypsins presented as 3D constant surface potential contours at +/− 1 kT/e in blue and red, respectively. The orientation of the molecules is shown in the top left figure. Four mesophilic trypsins [rat trypsin (RT), porcine trypsin (PT), bovine trypsin (BT), and cationic salmon trypsin (CST)] and three cold-adapted forms [anionic salmon trypsin (AST), Antarctic fish trypsin (AFT), and cod trypsin (CT1)] are included. The potentials are calculated by using all the atomic charges except on Asp[189]. The specificity pocket residues 189–193 and 216–219 together with the catalytic triad are indicated in greed rods.

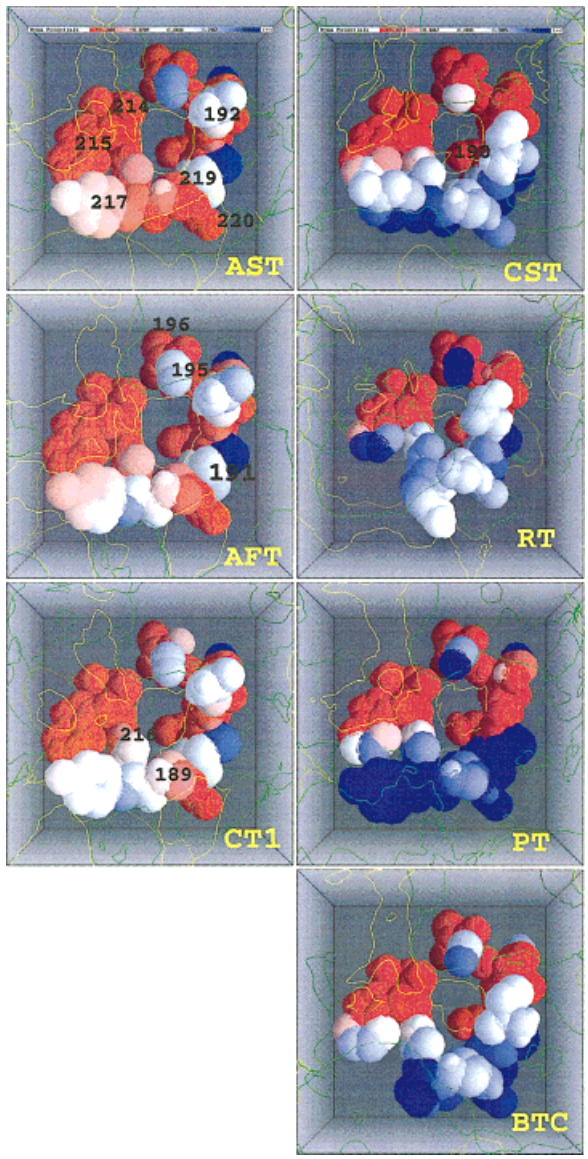


Fig. 2. Atomic potentials of the specificity pocket residues (189–197 and 214–220). Atoms are represented by CPK spheres and colored by the calculated potentials (charges on Asp[189] turned off) in red (negative) and blue (positive). The difference in potential at the base and residues 216–219 of the pocket is illustrated. Molecular surface equipotential contours at +/− 1 kT/e, in green and yellow, respectively, are also shown for comparison.

## Origins of the Differences in Potentials

Comparison of the anionic salmon trypsin and the cationic bovine trypsin,[9,13] along with an extensive sequence analysis of 27 trypsins,[12] suggest that the distribution of the charged residues on the enzyme surfaces is one of the reasons for the reactivity differences observed for the cold-adapted and mesophilic trypsins. However, the difference in potentials cannot be related solely to the total number of charged residues, because if this was the case, RT with the highest number of Asp and Glu residues should be similar to the cold trypsins, whereas CST with the lowest content of Asp and Glu residues should be similar to BT and PT. However, small structural differences involving the whole molecule, and particularly at the substrate-binding site, can result in different screening effects of the protein and of the solvent dielectric. The homologous

**TABLE II. Averaged Electrostatic Potentials of Residues of the Specificity Pocket (189–193, 216–221) Produced From Calculations With Different Charge Assignments†**

| Res. No. | | Potentials (kT/e) AST | AFT | CT1 | CST | RT | PT | BT | Residue | Except |
|---|---|---|---|---|---|---|---|---|---|---|
| 189 | 1 | *2.455* | *3.42* | *2.92* | *3.95* | *2.80* | *7.13* | *4.98* | Asp | |
| | 2 | **−99.00** | **−92.54** | **−92.24** | **−94.02** | **−97.78** | **−91.00** | **−98.49** | | |
| | 3 | **−86.52** | **−89.77** | **−90.98** | **−98.17** | **−108.7** | **−93.75** | **−96.69** | | |
| 190 | 1 | *−4.591* | *−4.05* | *−4.12* | *−5.00* | *−5.13* | *−2.47* | *−2.82* | Ser | |
| | 2 | **−32.17** | **−32.00** | **−31.82** | **−32.39** | **−35.10** | **−29.09** | **−30.24** | | |
| | 3 | **−31.23** | **−31.13** | **−31.23** | **−32.66** | **−32.30** | **−28.21** | **−32.15** | | |
| 191 | 1 | *−9.737* | *−9.83* | *−10.04* | *−9.18* | *−8.46* | *−7.23* | *−7.53* | Cys | |
| | 2 | **−17.60** | **−17.97** | **−18.12** | **−17.19** | **−17.37** | **−14.51** | **−16.24** | | |
| | 3 | **−16.91** | **−17.81** | **−17.92** | **−18.36** | **−16.57** | **−14.76** | **−17.19** | | |
| 192 | 1 | *−0.845* | *−0.79* | *−0.98* | *−0.70* | *−0.39* | *−0.08* | *−0.40* | Gln | |
| | 2 | **−2.83** | **−2.50** | **−3.00** | **−3.28** | **−3.28** | **−2.52** | **−2.89** | | |
| | 3 | **−2.77** | **−2.86** | **−2.92** | **−3.63** | **−3.25** | **−2.84** | **−3.29** | | |
| 193 | 1 | *−0.474* | *−1.09* | *−0.56* | *−6.66* | *−1.71* | *−0.43* | *−2.70* | Gly | |
| | 2 | **−2.91** | **−2.84** | **−3.15** | **−8.38** | **−5.12** | **−2.29** | **−4.85** | | |
| | 3 | **−3.93** | **−3.59** | **−4.02** | **−8.83** | **−4.40** | **−3.07** | **−4.67** | | |
| 216 | 1 | *−1.82* | *−1.70* | *−1.10* | *0.19* | *0.89* | *0.83* | *0.50* | Gly | |
| | 2 | **−6.82** | **−7.78** | **−6.41** | **−5.21** | **−5.04** | **−4.70** | **−5.14** | | |
| | 3 | **−6.13** | **−6.1•** | **−6.18** | **−6.06** | **−6.47** | **−5.96** | **−7.07** | | |
| 217 | 1 | *−1.14* | *−0.14* | *−0.01* | *2.06* | *0.85* | *4.53* | *2.85* | Tyr | Ser (BT) |
| | 2 | **−3.76** | **−3.13** | **−2.70** | **−0.96** | **−2.11** | **0.67** | **−1.67** | | |
| | 3 | **−2.96** | **−3.03** | **−3.04** | **−3.13** | **−3.07** | **−3.14** | **−4.41** | | |
| 219 | 1 | *−1.14* | *−0.74* | *−0.77* | *0.84* | *0.69* | *1.10* | *0.97* | Gly | |
| | 2 | **−3.76** | **−5.85** | **−5.96** | **−4.53** | **−5.06** | **−4.26** | **−4.79** | | |
| | 3 | **−5.18** | **−5.33** | **−5.20** | **−5.58** | **−5.61** | **−5.07** | **−5.61** | | |
| 220 | 1 | *−1.94* | *−1.25* | *−1.40* | *1.66* | *1.03* | *2.36* | *2.05* | Cys | |
| | 2 | **−13.52** | **−13.39** | **−13.49** | **−10.84** | **−12.21** | **−10.32** | **−11.93** | | |
| | 3 | **−11.49** | **−11.78** | **−11.18** | **−12.33** | **−11.80** | **−10.86** | **−12.98** | | |
| 221 | 1 | *−3.10* | *−1.35* | *−2.00* | *1.76* | *0.07* | *3.82* | *2.35* | Ala | |
| | 2 | **−28.78** | **−27.18** | **−27.71** | **−23.39** | **−27.16** | **−22.87** | **−26.11** | | |
| | 3 | **−25.66** | **−23.99** | **−25.24** | **−24.83** | **−25.02** | **−23.50** | **−28.84** | | |

†1, potentials due to all atomic charges except $Asp^{189}$; 2, potentials due to all atomic charges; 3, potentials due to all conserved charges. Mean potentials shown in bold are for calculations when charges on the oxygen atoms of $Asp^{189}$ side chain are included, whereas those shown in italic are calculations when $Asp^{189}$ is uncharged.

bovine and rat trypsins have a net charge difference of about 12 e, but the catalytic site potentials due to surface charges were effectively shielded, producing a nearly zero potential.[35] The same effect is, however, not observed for the specificity pockets of the seven trypsins. The potentials of the mesophilic trypsins are generally screened to a greater extent than the cold trypsins are, particularly for one part of the crevice (Fig. 1). Difference in screening effects from the solvent may arise from the different shapes of the dielectric boundary. The protein solvent interface is related to the 3D structure of the molecules, and hence, the amino acid sequence, and the nature of the interaction of the surface residues. Larger water-accessible surface area might mean that the solvent molecules are relatively distant from the internal residues, and screening would accordingly be less. Surface areas, as measured by the method of Lee and Richards,[37] show that the fish enzymes generally have larger water-accessible (molecular) surface areas than the mammalian trypsins.

### *Effects from conserved charges*

To investigate the screening effects of the dielectric boundary resulting from differences in the amino acid sequences, potentials were calculated by using charges on the conserved charged residues only. Fourteen charged residues are conserved in all the trypsins (see Fig. 3): five basic, six acidic, and three histidine residues (Lys-60, 107, 188, and 230; Arg-66; Glu-70, 77, and 80; Asp-102, 189, and 194; and His-40, 57, and 91). Ten of the conserved charged residues are in the N-terminal domain. All the histidines and the glutamates are found here. Four residues are conserved in the C-terminal domain.

Using only the conserved charges resulted in potentials in the cleft of the specificity pocket that are more similar among the trypsins, but again, the potential contours are extended to the surfaces of residues 189–193 for the three cold trypsins, whereas the mesophilic trypsins contours fail to do so (Fig. 4 and Table II). In contrast, the potentials of residues 216–221 of the binding pocket are more similar among the molecules than when also nonconserved charges

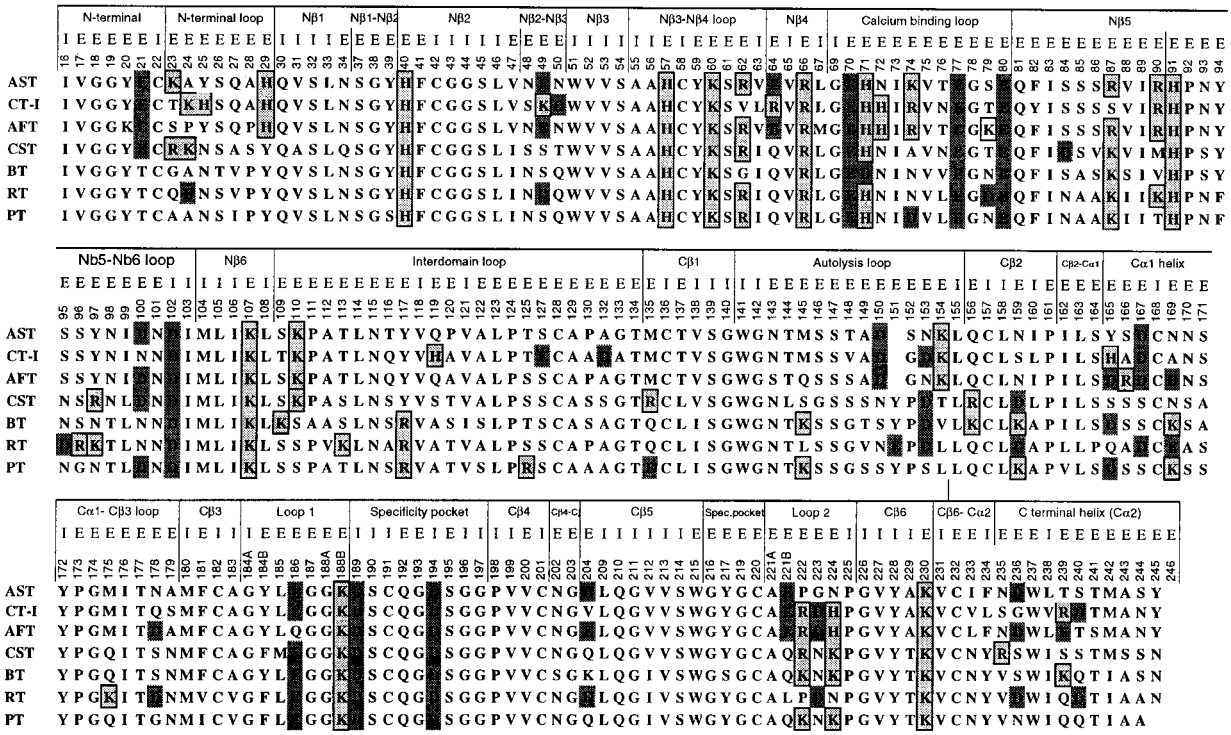


Fig. 3. Sequence alignment of the trypsins describing the primary structure charge distribution. The negatively charged amino acids (Asp and Glu) are shaded, whereas the positively charged residues (Lys and Arg) along with His are boxed. The structural regions are named according to their structural or functional relevance. E and I denote external and internal residues, respectively.

are included. Still, comparing the potentials of the conserved charges with those provided by all charges revealed features that are unique for each enzyme or group of enzyme. The potentials of the cold trypsins increased when using only the conserved charges, whereas they decreased for the mesophilic trypsins. This finding suggests that the effect of the nonconserved charges is opposite for the cold and the mesophilic enzymes. The fact that the potentials produced by the conserved charges are uniform and lower for the cold trypsins and higher for the mesophilic trypsins must imply that particular charged residues are contributing a positive potential in this region of the mesophilic enzymes and a negative potential of the cold enzymes.

The reason for the observed differences in potentials at the residues 189–193 region of the specificity pocket is, therefore, most likely caused by the different shapes of the molecular surfaces, because the same differences are observed when the same number and position of charged residues are used (i.e., only the conserved charges). The previously observed differences in potentials at residues 216–221 are not reproduced by the conserved charges. Nonconserved charged residues are thus responsible for these differences.

### *Effects from nonconserved charges*

In an attempt to find the origins of the observed differences in potentials, particular charged residues were searched for, which could have long-range effects on the potentials of the pocket in their position and orientation relative to the $S_1$ cleft. A few residues were identified and subjected to subsequent modeled mutations. There are four charged amino acids uniquely conserved among the cold trypsins (residues 29, 150, 154, and 221B), whereas there are no unique conservation of charged residues among the mesophilic trypsins (Fig. 3). However, a lysine is conserved at position 224 for the cationic trypsins. A lysine occupies position 175 in RT, whereas this position is Met or Gln in the other trypsins. Figure 4b shows that the charges of residues 221B and 224 largely affect the potentials of residues 216–221 and the base of the $S_1$ pocket of the cationic mesophilic trypsins. The same region of the anionic RT is influenced by the charges of residues 175 and 221B. The positive charge of $Lys^{224}$ in BT, PT, and CST contribute to a positive potential of the active site of the cationic trypsins. Similarly, the +0.5 charge of $His^{224}$ of AFT and CT1 causes a slightly more positive potential than in AST, where an Asn occupies position 224. A positive potential arise from $Lys^{175}$ in RT and from $Gln^{175}$ in the cationic mesophilic trypsins, whereas $Met^{175}$ generates a negative potential in the cold-adapted trypsins. Residue 221B is a conserved Glu in the cold trypsins and Gln or Leu in the warm trypsins, thereby affecting the potentials of the pocket differently.

### *Partial charges*

Also, the distribution of the polar amino acids is different among the trypsins in this study. To study the effect this may have on the specificity pocket and the enzyme-substrate interface, potentials were calculated by using only the partial charges from the AMBER charge set.[20] The surface potentials contours of the specificity pockets (data not shown) were strikingly different for the seven trypsins. A negative potential was found for AST, CST, and RT, a nearly zero potential for CT1, AFT, and PT, but a very strong positive potential for BT. However, the differences observed are relatively small and do not seem to be consistent with type of trypsin.

## DISCUSSION

The main results of the calculations are as follows: (a) The electrostatic potentials at the substrate binding residues are generally significantly lower for the cold-adapted trypsins than those of the mesophilic counterparts. (b) At most of the residues, the potentials produced from $Asp^{189}$ are uniform among the molecules. (c) Most of the potential in the $S_1$ pocket is produced from the conserved charges. The effect of the nonconserved charges seems to be opposite in the cold and the warm active trypsins. (d) Residues such as, among others, 221B, 224, and 175 (RT), are the main sources of the differences in potentials of the $S_1$ pockets of the mesophilic and the cold active trypsins.

### Potentials After Mutations of Selected Residues

In addition to the residues at the $S_1$ site, residues at the two specificity pocket extension loops (184A–188B and 221A–225) and residue 172 are important for the specificity of trypsins as general structural specificity determinants.[38–41] As has been outlined above, the electrostatic potentials associated with these sites are also different among the trypsins, correspondingly affecting the potentials at the $S_1$ substrate binding site differently. Hence, it was tempting to mutate some of the variable residues at the extension loops for purely electrostatic reasons. Residue 224 in CST, CT1, AFT, BT, and PT was mutated to the corresponding residue in AST (Asn), and residue $Lys^{175}$ of RT was mutated to $Met^{175}$. Then, another Loop2 residue (221B) of the mesophilic trypsins was mutated to a Glu that is conserved in the cold active enzymes. The stronger negative potential observed in the $S_1$ pocket of AST was expected to be (at least partly) reproduced by these mutations.

The potential of AST is taken as a model due to the 20- to 40- fold higher catalytic efficiency of this enzyme compared with the mammalian trypsins,[4] and more importantly, the 100-fold stronger association to BPTI of AST relative to BT.[7] Figure 5 displays the equipotential contours and atoms of selected residues colored by the potentials for the different single and multiple mutants of the trypsins. The mutation H224N in AFT and CT1 resulted in a negative potential as strong as in AST at the base of the pocket. The shapes of the negative potential contours in the pocket also became identical in the three enzymes. However, the H224N mutation had only a small effect on the potential of the $S_1$ site compared with the K224N mutations of the cationic trypsins. Mutation of residue 224 in the cationic mesophilic trypsins (BT, PT, and CST) produced a stronger negative potential within the substrate binding crevice and a nearly zero potential on atoms and surfaces at the base of the pocket. The potential in and around the pocket became identical in all three of the cationic trypsins. However, the shape of the contours and the atomic potentials around the base of the pocket and the extension loops are still not similar to that of AST. The K175M mutation in RT produced a significant change in the potential of the residues 216–221 wall of the pocket. The shape of the contours became similar to the K224N mutants of the cationic mesophilic trypsins. Therefore, in terms of electrostatic potentials of the substrate binding site, $Lys^{175}$ in RT and $Lys^{224}$ in BT, PT, and CST have the same purpose of providing a positive polarization. The two residues are located at different sites, but both are pointing toward the same target: residue 217. It can be seen that mutation of these residues to the corresponding residues in AST is significant but not sufficient to generate the strong negative potential observed at $S_1$ site of AST.

Mutations of residue 221B (Q/L221E) in the mesophilic trypsins resulted in significant changes of the potentials (data not shown), but the influence was not strong enough to reproduce the potentials of the cold trypsins. It was, therefore, interesting to see the effects of the double mutations involving residues 224 and 221B for the cationic mesophilic trypsins (BT, PT, and CST) and 175 and 221B for RT. Figure 5 shows that the electrostatic potentials of AST at the pocket and the extension loops are successfully reproduced by the double mutations. The atomic potentials and the shape of the contours of the mesophilic trypsins became similar to that of AST or to the cold-adapted trypsins in general. In fact, the negative potential of the mutated trypsins seems to be even stronger than the model trypsin (AST). In addition, multiple mutations involving residues 159 and 160 to the corresponding residues in AST have been attempted for porcine trypsin. The results (data not shown) showed that these residues do not have significant effects on the potential of the substrate-binding site. Mutations of residues 221B and 188B to hydrophobic residues (Q221A and K188G) and residue 224 to a polar amino acid (K224N) have been attempted to see the effect of introducing a hydrophobic residue at 221B and perturbing the positive charge at 188B. The potential in the pocket is affected, and an example is shown for CST and PT at the bottom right of Figure 5. The shape of the contour is similar to the single mutation K224N, but the area slightly increases. The increase in area of the contour can be the result of the perturbation of the conserved positive charge at 188. However, the specific significance of the conserved $Lys^{188}$ is not definitively illustrated by the mutations, and it seems that it is not as important as for example residue 221B. This is in line with the experimental observation that K188B-Trp, Phe, and Tyr mutants of trypsin conserve capacity of the native trypsin to degrade peptide bonds containing lysyl and arginyl residues.[42] In electrostatic

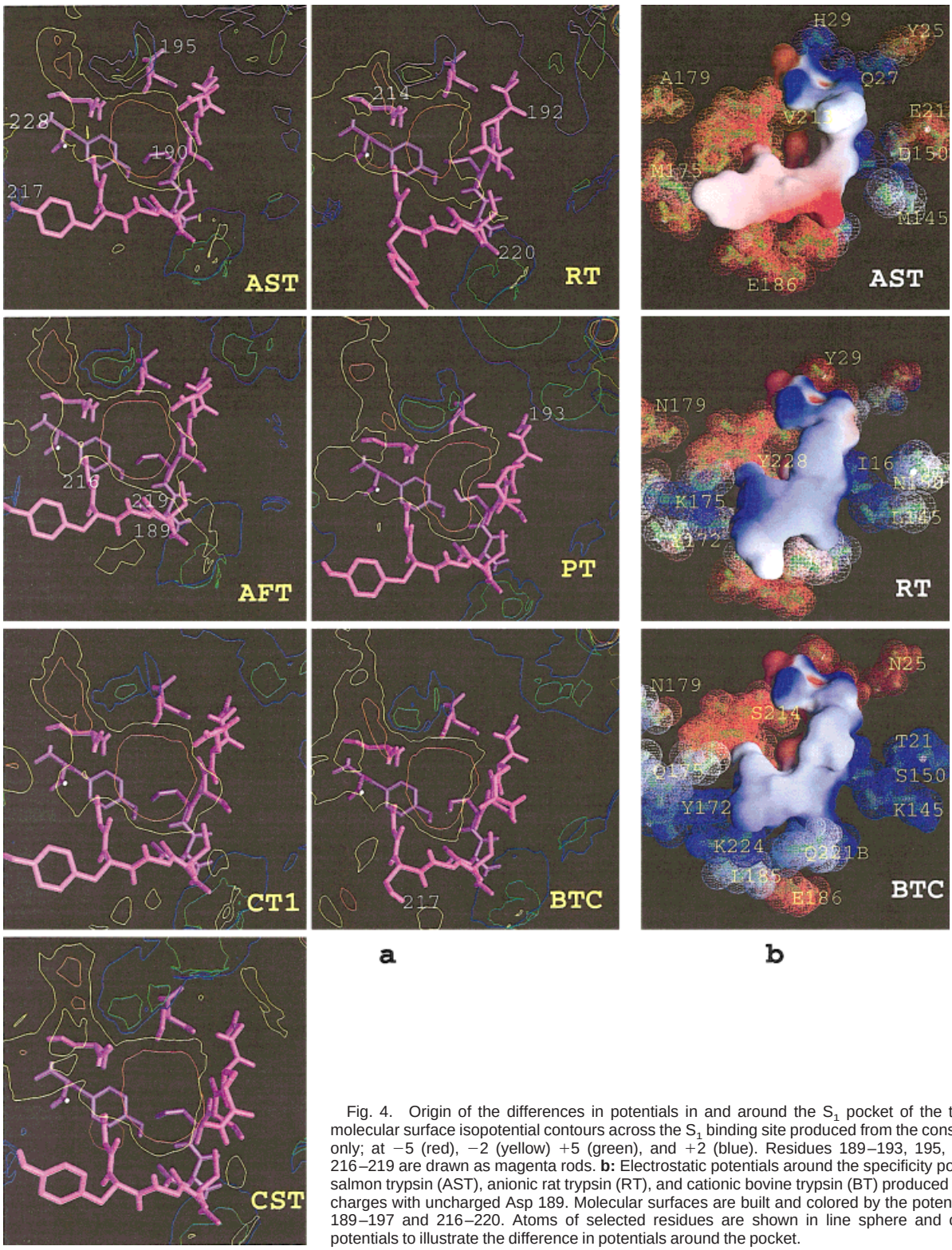


Fig. 4. Origin of the differences in potentials in and around the $S_1$ pocket of the trypsins. **a:** 2D molecular surface isopotential contours across the $S_1$ binding site produced from the conserved charges only; at −5 (red), −2 (yellow) +5 (green), and +2 (blue). Residues 189–193, 195, 214, 228, and 216–219 are drawn as magenta rods. **b:** Electrostatic potentials around the specificity pocket of anionic salmon trypsin (AST), anionic rat trypsin (RT), and cationic bovine trypsin (BT) produced from all atomic charges with uncharged Asp 189. Molecular surfaces are built and colored by the potential of residues 189–197 and 216–220. Atoms of selected residues are shown in line sphere and colored by the potentials to illustrate the difference in potentials around the pocket.

potential, residues, such as at positions 221B, 224, and 175 (RT), seem to be crucial for the enhanced reactivity of AST (or all the cold trypsins) relative to the warm trypsins. The mutations showed that the observed strong electrostatic potential at the $S_1$ site of the cold enzymes is a result of charged residues not directly in contact with substrate atoms. It should be noted that the potential of $Tyr^{172}$, which was shown to have structural influence on substrate

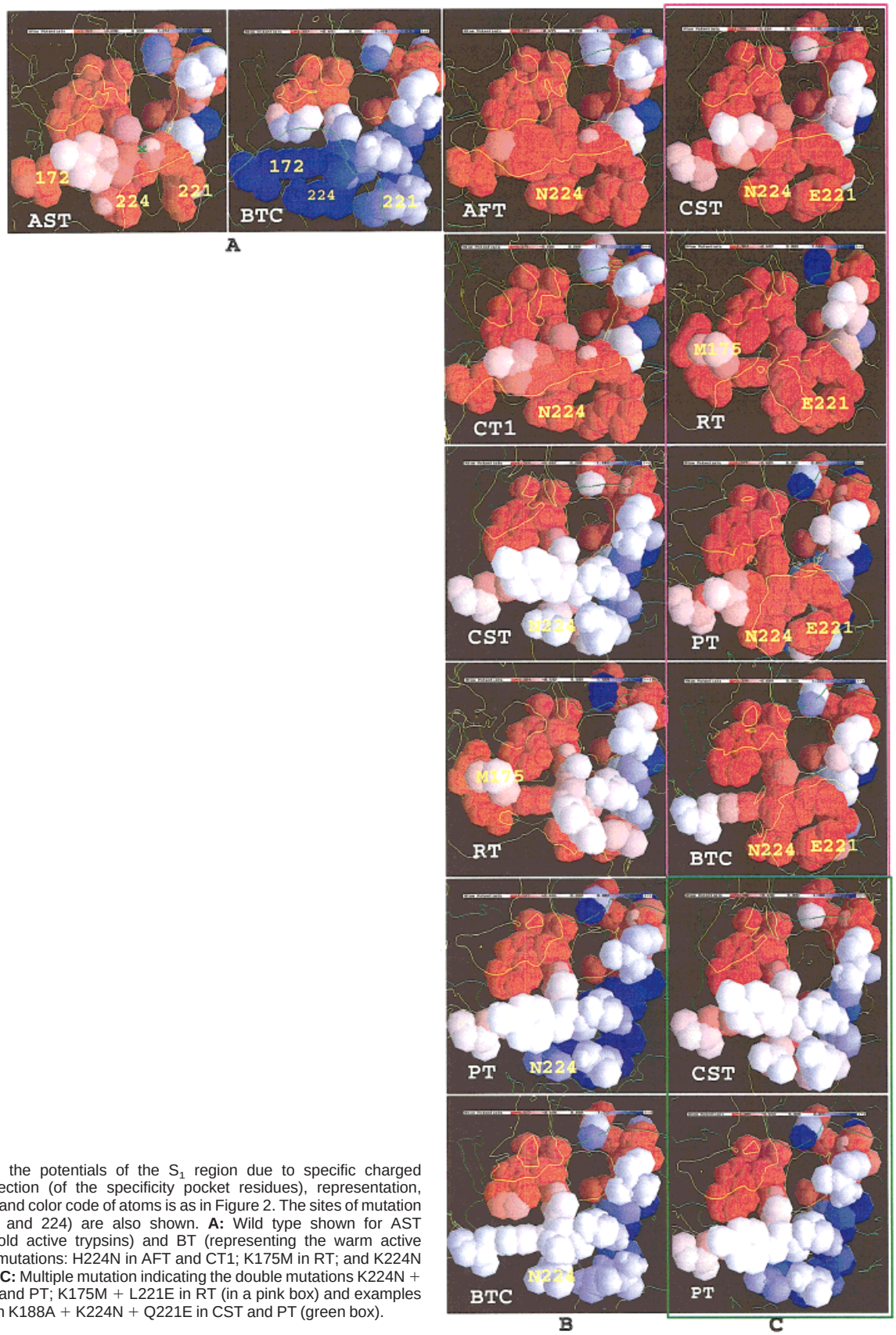


Fig. 5. Effect in the potentials of the $S_1$ region due to specific charged mutations. The selection (of the specificity pocket residues), representation, charge assignment, and color code of atoms is as in Figure 2. The sites of mutation (residues 175, 221 and 224) are also shown. **A:** Wild type shown for AST (representing the cold active trypsins) and BT (representing the warm active trypsins). **B:** Single mutations: H224N in AFT and CT1; K175M in RT; and K224N in BT, CST, and PT. **C:** Multiple mutation indicating the double mutations K224N + Q221E in BT, CST, and PT; K175M + L221E in RT (in a pink box) and examples for the triple mutation K188A + K224N + Q221E in CST and PT (green box).

specificity, is strongly affected by the charges of the three residues mentioned.

## CONCLUSIONS

As expected, most of the negative potential at the $S_1$ region of the trypsins is generated from $Asp^{189}$. However, the potentials at the $S_1$ site of the trypsins do not seem to be equivalent for all, and the main source of the differences are charged residues located outside the specificity pocket that do not have any direct contact with substrate atoms. Surface charges produce an electric field at the $S_1$ pocket of the trypsins, which is different in sign and magnitude among the seven trypsins. The surface charges of the mesophilic trypsins generally induce the $S_1$ pocket positively, whereas surface charges of the cold active trypsins produce a negative electric field at this site. Some particular nonconserved charged residues were responsible for a major part of the differences in electrostatic potential of the $S_1$ pocket of the cold-adapted and mesophilic trypsins. In the cationic mesophilic trypsins, the presence of a lysine residue at position 224 and the substitution of glutamic acid by glutamine at 221B seem to be the major sources of the higher potential at the $S_1$ site of these trypsins compared with the cold active counterparts. Similar effects arise from a lysine residue at 175, combined with a leucine residue at 221B in the case of anionic rat trypsin. The results of the electrostatic potential calculations leave residues 175, 221, and 224 as interesting targets for future mutagenesis experiments. The result of the calculations using only the conserved charges were uniform among the molecules, and it represented most of the potential compared with all the real charges. This finding may suggest that the conservation of these residues is important for the generation of most of the negative electrostatic potential of the $S_1$ pocket, whereas the nonconserved charges induce the pocket differently.

The nonconserved charges seem to be sources of the stronger negative potential of the cold active trypsins, whereas their effect is to reduce the negative potential of the mesophilic trypsins. The effect of the partial charges on the potential of the substrate-binding site do not seem to follow the reactivity pattern of the molecules. The present study has also shown that the potential of the $S_1$ pocket does not depend on the overall charge of the trypsin molecule. All the cold active fish trypsins are anionic and posses a similar potential of the $S_1$ site, but rat trypsin (RT), which has an even lower overall charge, possesses electrostatic potentials of the binding region that is similar to those of the other mesophilic trypsins with overall cationic nature. In a similar manner, the cationic salmon trypsin, which has many structural features in common with the cold active equivalents, also resembles the other mesophilic trypsins in electrostatic potentials. The lower potential at this region, produced by the surface charges of the cold trypsins, can be related to their observed increased substrate binding affinity. The positive charge of the substrates can be stabilized by the stronger negative potential more efficiently in the cold trypsins than in the mesophilic trypsins. The qualitative trends of the calculated potentials, thus, also reflects the association strength of trypsin-inhibitor complexes.

## REFERENCES


1. Gerday C, Aittaleb M, Arpigny JL, et al. Psychrophilic enzymes: a thermodynamic challenge. Biochem Biophys Acta 1997;1342:119–131.
2. Feller G, Narinx E, Arpigny JL, et al. Enzymes from psychrophilic organisms. FEMS Microbiol Rev 1996;18:189–202.
3. Kristjansson MM, Ásgeirsson B, Bjarnason JB. Serine proteinases from cold-adapted organisms. Adv Exp Med Biol 1997;415:27–46.
4. Outzen H, Berglund GI, Smalås AO, Willassen NP. Temperature and pH sensitivity of trypsins from atlantic salmon *(Salmo salar)* in comparison with bovine and porcine trypsin. Comp Biochem Biophys 1996;115B:33–45.
5. Asgeirsson B, Fox JW, Bjarnason JB. Purification and characterization of trypsin from the poikilotherm *Gadus morhua*. Eur J Biochem 1989;180:85–94.
6. Genicot S, Feller G, Gerday C. Trypsin from the antarctic fish (*Paranotothenia magellanica* forster) as compared with trout (*Salmo gairdneri*) trypsin. Comp Biochem Physiol 1988;90B:601–609.
7. Schechter I, Berger A. On the size of the active site in proteases. I. Papain. Biochem Biophys Res Commun 1967;27:157–162.
8. Krowarsch D, Dadlez M, Buczek O, Krokoszynska I, Smalås AO, Otlewski J. Interscaffolding additivity: binding of P1 variants of bovine pancreatic trypsin inhibitor to four serine proteinases. J Mol Biol 1999;289:175–186.
9. Helland R, Leiros I, Willassen NP, Berglund GI, Smalås AO. The crystal structure of anionic salmon trypsin in complex with bovine pancreatic trypsin inhibitor. Eur J Biochem 1998;256:317–324.
10. Rypniewski WR, Perrakis A, Vorgias CE, Wilson KS. Evolutionary divergence and conservation of trypsin. Protein Eng 1994;7:57–64.
11. Leiros H-KS, Willassen NP, Smalås AO. Residue determinants and sequence analysis of cold-adapted trypsins. Extremophilies 1999;3:205–219.
12. Leiros H-KS, Willassen NP, Smalås AO. Structural comparison of psychrophilic and mesophilic trypsins: elucidating the molecular basis of cold-adaptation. Eur J Biochem 2000;267:1039–1049.
13. Smalås AO, Heimstad ES, Hordvik A, Willassen NP, Male R. Cold adaption of enzymes: structural comparison between salmon and bovine trypsins. Proteins 1994;20:149–166.
14. Sharp KA, Nichols A. Delphi Version 3.0.® 1988 Colombia University.
15. Sharp KA, Honig B. Electrostatic interactions in macromolecules: theory and applications. Annu Rev Biophys Biophys Chem 1990; 19:301–332.
16. Richards FM. Areas, volumes, packing and protein structure. Annu Rev Biophys Bioeng 1977;6:151–176.
17. Gilson MK, Rashin A, Fine R, Honig B. On the calculation of electrostatic interactions in proteins. J Mol Biol 1985;184:503–516.
18. Gilson MK, Honig B. Calculation of the total electrostatic energy of a macromolecular system: solvation energy, binding energies, and conformational analysis. Proteins 1988;4:7–18.
19. Connolly ML. The molecular surface package. J Mol Graphics 1993;11:139–141.
20. Weiner SJ, Kollman PA, Case DA, et al. A new force field for molecular mechanical simulation of nucleic acids and proteins. J Am Chem Soc 1984;106:765–783.
21. Bernstein FC, Koetzle TF, Williams GJB, et al. The Protein Data Bank: a computer-based archival file for macromolecular structures. J Mol Biol 1977;112:535–542.
22. Brooks CL, Karplus M. Deformable stochastic boundaries in molecular dynamics. J Chem Phys 1983;79:6312–6325.
23. Oldfield TJ. SQUID: a program for the analysis and display of data from crystallography and molecular dynamics. J Mol Graphics 1992;10:247–252.
24. Nicholls A, Sharp K, Honig B. Protein folding and association: insight from the interfacial and thermodynamic properties of hydrocarbons. Proteins Struct Funct Genet 1991;11:281–296.
25. Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr 1991;A47:110–119.
26. Schrøder H-K, Willassen NP, Smalås AO. Structure of a non-

psychrophilic trypsin from a cold-adapted fish species. Acta Crystallogr 1998;D54:780–798.
27. Earnest T, Fauman E, Craik CS, Stroud R. 1.59 Å structure of trypsin at 120 K: comparison of low temperature and room temperature structures. Proteins 1991;10:171–181.
28. Huang Q, Liu S, Tang Y. Refined 1.6 Å resolution crystal structure of the complex formed between porcine β-trypsin and MCTI-A, a trypsin inhibitor of the squash family: detailed comparison with bovine β-trypsin and its complex. J Mol Biol 1993;229:1022–1036.
29. Marquart M, Walter J, Deisenhofer J, Bode W, Huber R. The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. Acta Crystallogr 1983;B39:480–490.
30. Klapper I, Hagstrom R, Fine R, Sharp K, Honig B. Focusing of electric fields in the active site of Cu-Zn superoxid dismutase: effects of ionic strenght and amino-acid modification. Proteins 1986;1:47–59.
31. Sharp K, Fine R, Honig B. Computer simulations of diffusion of a substrate to an active site of an enzyme. Science 1987;236:1460–1463.
32. Warwicker J, Ollis D, Richard FM, Steitz TA. Electrostatic field of the large fragment of *Escherichia coli* DNA polymerase I. J Mol Biol 1985;186:645–649.
33. Gilson MK, Honig B. Calculation of electrostatic potentials in an enzyme active site. Nature 1987;330:84–86.
34. Sitkoff D, Lockhart DJ, Sharp KA, Honig B. Calculation of electrostatic effects at the amino terminus of an alpha helix. Biophys J 1994;67:2251–2260.
35. Soman K, Yang AS, Honig B, Fletterick R. Electrostatic potentials in trypsin isoenzymes. Biochemistry 1989;28:9918–9926.
36. Warshel A, Naray-Szabo G, Sussman F, Hwang JK. How do serine proteases really work? Biochemistry 1989;28:3629–3637.
37. Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. J Mol Biol 1971;55:379–400.
38. Hedstrom L. Trypsin: a case study in the structural determinants of enzyme specificity. Biol Chem 1996;377:465–470.
39. Hedstrom L, Farr-Jones S, Kettner CA, Rutter WJ. Converting trypsin to chymotrypsin: ground state binding does not determine substrate specificity. Biochemistry 1994;33:8764–8769.
40. Hedstrom L, Szilagyi L, Rutter WJ. Converting trypsin to chymotrypsin: the role of surface loops. Science 1992;255:1249–1253.
41. Hung SH, Hedstrom L. Converting trypsin to elastase: substitution of the S1 site and adjacent loops reconstitutes esterase specificity but not amidase activity. Protein Eng 1998;11:669–673.
42. Chobert JM, Briand L, Tran V, Haertle T. How the substitution of K188 of trypsin binding site by aromatic amino acids can influence the processing of beta-casein. Biochem Biophys Res Commun 1998;246:847–858.