Engineering specificity and activity of thermolysin-like proteases from Bacillus

de Kreij, Arno

The effects of modifying the surface charge on the catalytic activity of a thermolysin-like protease.

Abstract.

The impact of long-range electrostatic interactions on catalysis in the thermolysin-like protease from *Bacillus stearothermophilus* was studied by analysing the effects of inserting or removing charges on the protein surface. Various mutations were introduced at six different positions and double-mutant cycle analysis was used to study the extent to which mutational effects were interdependent. The effects of single point mutations on the k_{cat}/K_m were non-additive, even in cases where the point mutations were located 10 Å or more from the active site Zn^{2+} and separated from each other by up to 25 Å. This shows that catalysis is affected by large electrostatic networks that involve major parts of the enzyme. The interdependence of mutations at positions as much as 25 Å apart in space also indicates that other effects, such as active site dynamics, play an important role in determining active site electrostatics. Several mutations yielded a significant increase in the activity, the most active (quadruple) mutant being almost four times as active as the wild-type. In some cases the shape of the pH-activity profile was changed significantly. Remarkably, large changes in the pH-optimum were not observed.

Introduction.

The acceleration of reaction rates by enzymes is one of the essential prerequisites for life as we know it, and the multitude and diversity of enzymes shows that it should be possible to design an enzyme that will catalyse almost any reaction under almost any set of conditions. To achieve a high rate-acceleration, enzymes rely on charged groups in their active site that stabilise the transition state and function as acid and base catalysts in the reaction. The kinetic parameters of enzymes therefore display a significant pH-dependence, which is determined by the pKa values of the active site groups.

Since catalysis depends on intricate electrostatic interactions, which may be noticeable over relatively large distances, as compared to other short ranged interactions such as H-bonds and hydrophobic interactions, larger parts of an enzyme may be involved in optimizing its catalytic centre than previously thought. The long-range character of electrostatic effects is illustrated by a, very limited, number of examples in the literature, showing that changes in surface charge at locations as far as 15Å from a catalytic centre may affect enzyme activity (74, 191). Unfortunately, electrostatic interactions are hard to handle theoretically, not only because of their long-range character, but also because of intrinsic theoretical difficulties. For example, most electrostatic models still use a single rigid protein structure, and at most two dielectric constants to account for all the dynamics of the protein. This clearly is an oversimplification of reality (192).

We have studied the contribution of long range electrostatic interactions to catalysis by analyzing the effects of a series of charge-mutations scattered over a larger part of the surface of a thermolysin-like protease from *Bacillus stearothermophilus* (TLP-ste). Thermolysin-like proteases (TLPs) are members of the peptidase family M4 (51) of which thermolysin (TLN, EC 3.4.24.27) is the prototype. One of their characteristics is a zinc ion bound in the catalytic centre. The amino acid sequences of several TLPs have been

determined [see (51), or the Merops data base at http://www.merops.co.uk/merops/famcards/m4.htm], and the three-dimensional structure of TLPs isolated from several bacteria have been solved *i.e.* those from TLPs from *Bacillus thermoproteolyticus* (52), *Bacillus cereus* (54), *Pseudomonas aeruginosa* (55) and *Staphylococcus aureus* (127). TLPs consist of an α-helical C-terminal domain and a β-rich N-terminal domain. These two domains are connected by a central α-helix, which is located at the bottom of the active site cleft and which contains several of the catalytically important residues. From X-ray structures of TLN-inhibitor complexes (52, 57-61), the active site residues have been identified and a mechanism has been proposed (61, 62, 155). Recently, an alternative mechanism has been proposed which has gained some support (63, 64). In both proposed mechanisms residues Glu143, His231, Tyr157 and a Zn^{2+} bound water play important roles during catalysis.

Mutations were introduced at six surface positions in TLP-ste, located at 10 - 15 Å from the catalytic centre. The single and multiple mutants that were obtained displayed varying effects on catalytic efficiency, including considerable increases in activity. Double-mutant cycle analysis (193) was used to study the additivity of mutational effects, which showed a remarkable interdependence of the mutated residues. The results provide insight in the complexity of predicting and interpreting electrostatic effects in catalysis.

Materials and Methods

Modelling and mutant design.

A three-dimensional model of TLP-ste was built with WHAT-IF (140) using the crystal structure of thermolysin as template, as described elsewhere (78). The high sequence identity between thermolysin and TLP-ste (86%) indicates that the TLP-ste model is sufficiently reliable for prediction and analysis of the effects of most amino acid substitutions (78, 141). Indeed, the TLP-ste model has been used successfully for the design of various stabilising mutations (43, 82, 177). Throughout this paper, residues are numbered according to the corresponding residues in thermolysin.

All Glu, Gln, Asp and Asn residues that are close to but not in the active site cleft and are positioned at the surface of the protein were selected for mutagenesis. Position-specific rotamer searches (180) for the residues to be introduced showed that all new side chains could adopt a favourable rotamer without concomitant introduction of steric overlap.

Molecular Biology.

The *nprT* gene encoding the TLP of *B. stearothermophilus* CU21 (136) (TLP-ste) was cloned, subcloned, and expressed as described previously (178). Site-directed mutagenesis was performed on a subcloned fragment of the TLP-ste gene by the QuikChange site-directed mutagenesis kit from Stratagene, La Jolla, USA. The nucleotide sequences of mutated fragments of the *nprT* gene were verified by DNA sequencing and the mutated fragments were subsequently cloned into the *Bacillus* expression vector pGE501 (178) containing the TLP-ste gene with a deletion of the previously subcloned fragment.

Production and Characterisation of Mutant Enzymes.

Production and purification of the enzymes were performed as described elsewhere (144). Prior to determining the kinetic parameters, protease preparations were desalted using pre-packed PD-10 gel filtration columns supplied by Amersham Pharmacia, Uppsala, Sweden.

Determination of kinetic constants.

The k_{cat}/K_m values of the enzymes for the furylacryloylated tripeptide 3-(2-furylacryloyl)-L-glycyl-L-leucine-L-alanine (FaGLA) obtained from Bachem AG, Bubendorf, Switzerland were determined at 37°C, in a thermostated Perkin-Elmer Lambda 11 spectrophotometer. The reaction mixture (1 ml) contained 50 mM 2-amino-2-(hydroxymethyl)-1,3-propanediol (Tris·HCl), 50 mM 4-morpholineethanesulfonic acid (MES), pH 4.4 to 8.4 with 0.4 units interval, 5 mM $CaCl_2$, 1% Me_2SO, 1% 2-propanol, 0.01% Triton X-100 and 100 μM of substrate. The reaction was followed by measuring the decrease in absorption at 345 nm ($\Delta\varepsilon_{345}$ = -317 $M^{-1}\cdot cm^{-1}$)(114). The stock solution of the furylacryloylated tripeptide was prepared by dissolving the peptide in Me_2SO. Apparent second order rate constants (k_{cat}/K_m) were determined by varying the enzyme concentrations over a 50-fold range under pseudo-first-order conditions and measuring the initial activity, essentially according to the method described by Feder (114). All k_{cat}/K_m values are the result of a linear regression analysis of at least 14 independent measurements. The error margins, defined by the highest and lowest value of the 95% confidence interval of the linear regression analysis, were at most 15% of the values given.

Double-mutant cycle analysis.

Determining the contribution of a single amino acid to the activity or stability of a protein by mutating that residue alone is often misleading. This is because the neighbouring residues often change their position slightly to compensate for the mutation. A very illustrative recent example of this effect is provided by Albeck & Schreiber (194, 195) for the TEM-BLIP complex. To overcome this problem the method of double-mutant cycle analysis has been recommended (193, 196, 197).

Double-mutant cycle analysis uses the energy difference between free enzyme plus substrate and the enzyme-substrate complex in the transition state (193, 196, 198). This energy difference, referred to as $\Delta G^{\ddagger}$, reflects binding energy in the transition state and can be derived from measured k_{cat}/K_m values using

$$\Delta G^{\ddagger} = -RT \ln (k_{cat}/K_m \times h/kT).$$

Now, consider two residues A and B in the enzyme which do not interact. In that case the effects of mutating these residues will be independent and therefore additive. Expressed in the thermodynamic terms of Fig. 5.1:

$$\Delta\Delta G^{\ddagger}_1 = \Delta\Delta G^{\ddagger}_1{}' \qquad \text{and} \qquad \Delta\Delta G^{\ddagger}_2 = \Delta\Delta G^{\ddagger}_2{}'$$

$\Delta\Delta G^{\ddagger}_1$ stands for the change in $\Delta G^{\ddagger}$ upon mutation A→A'. $\Delta\Delta G^{\ddagger}_1$ is calculated as follows (using the annotation of Fig. 5.1):

$$\begin{aligned}\Delta\Delta G^{\ddagger}_1 &= \Delta G^{\ddagger}_{A'B} - \Delta G^{\ddagger}_{AB} \\ &= -RT \ln [(k_{cat}/K_m)_{A'B} \times h/kT] \\ &\quad + RT \ln [(k_{cat}/K_m)_{AB} \times h/kT]\end{aligned}$$

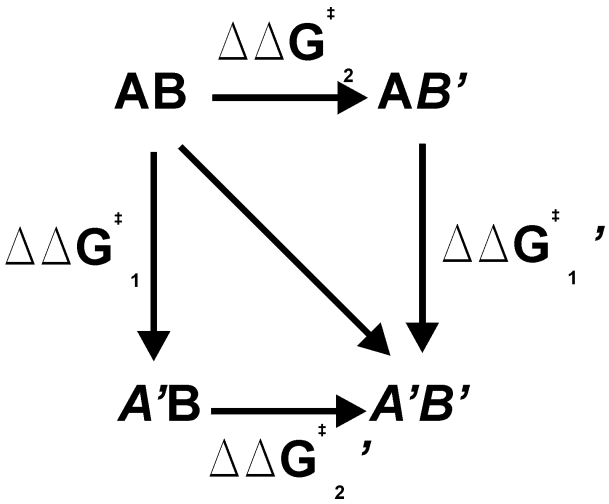

Figure 5.1. Double-mutant cycles. Residue A is mutated to A′ and residue B to B′. $\Delta G^{\ddagger}$ is the energy difference between free enzyme plus substrate and the enzyme-substrate complex in the transition state. $\Delta\Delta G_1^{\ddagger}$ is the change in $\Delta G^{\ddagger}$ upon mutating residue A, $\Delta\Delta G_2^{\ddagger}$ is the change in $\Delta G^{\ddagger}$ upon mutating residue B. See text for details.

$\Delta\Delta G^{\ddagger}_1 = -RT \ln [(k_{cat}/K_m)_{A'B}/(k_{cat}/K_m)_{AB}]$

If the residues symbolized by A and B in Fig. 1 somehow influence each other, the effect of mutating one may become dependent on whether or not the other is mutated too. Expressed in the terms of Fig. 1:

$|\Delta\Delta G^{\ddagger}_1 - \Delta\Delta G^{\ddagger}_1'| = |\Delta\Delta G^{\ddagger}_2 - \Delta\Delta G^{\ddagger}_2'| \neq 0$

In which the term $|\Delta\Delta G^{\ddagger}_1 - \Delta\Delta G^{\ddagger}_1'|$ is defined as the coupling energy. This term is zero if the effects of the mutations are independent, but not so if they are dependent.

The experimental error in the calculated $\Delta\Delta G^{\ddagger}$ values is derived by inserting the highest and lowest values of the 95% confidence interval of the k_{cat}/K_m values in the equation for $\Delta\Delta G^{\ddagger}$. The experimental error of the coupling factor is the sum of the errors in the $\Delta\Delta G^{\ddagger}$ values. The coupling factor is significantly non-zero when the error is smaller than the value of the coupling factor.

Electrostatic calculations.

The change in the electrostatic potential at the Nδ1 of His231, the Oε1 of Gly143, and at the oxygen of the water molecule bound to the catalytic zinc ion was calculated using WHAT IF (140) interfaced to DelPhi II (199). A dielectric constant of 4 was applied in the interior of the protein and a dielectric constant of 80 was assigned to the solvent phase (200). The ionic strength in the calculations was set to match the experimental conditions at pH 7.0 (210 mM). The ΔpKa can be calculated from the electrostatic potential difference Φ using the following formulas:

$$\frac{G}{Q} = \Phi \qquad \text{and} \qquad \Delta G = -RT \ln K_a$$

In which Q is the charge, Φ the electrostatic potential energy in V, and G the free-energy. Rearangement of these formula's leads to an equation for the ΔpKa:

$$\Delta pKa = \frac{-\Delta\Phi}{RT \ln(10)}$$

Results.

Mutant design and production of mutant proteins.

All mutants were constructed as described in the Materials and Methods section. Fermentation and purification yields were normal for all single and multiple mutants. Table V.I summarizes the characteristics of the single mutants, compared to the wild-type. Fig. 5.2 gives an overview of the stereochemical relationship of the mutated surface residues to each other and to the catalytically active residues and the Zn^{2+} . The closest contact distance between the mutated residues and the active site Zn^{2+} is generally between 10 and 15 Å. The closest contact distance between the most remote mutated residues is approximately 25 Å between residue 225 and residues 116 and 119.

Characterisation of mutant proteases.

To examine the pH-activity profiles of the mutant enzymes, the kinetic parameters for the reaction of wild-type and mutant TLP-stes with the tripeptide substrate FaGLA were determined at different pH-values. Fig. 5.3 shows the pH-activity profile of the wild-type TLP-ste and of all the single mutants that were constructed. The pH-activity profiles generally show minimal changes in the pH optimum, the only exception being the relatively inactive mutant N227D, for which the profile shows a small acidic shift. The profiles do show some conspicuous changes in shape, in particular around the “second” pH optimum near pH 6,0.

Table V.I. Characteristics of the TLP-ste variants.

Mutation	k_{cat}/K_m [a]	$\Delta G^{\ddagger}$	ΔCharge	ΔpKa		
				Glu143	Zn – H_2O	His231
	$M^{-1}\cdot s^{-1} \times 10^{-4}$	$kJ\cdot mol^{-1}$				
Wild-type	6.1	47.6				
N116D	10.4	46.3	-1	0.4	0.2	0.1
Q119R	12.8	45.7	+1	-0.1	-0.1	-0.1
D150N	2.4	50.0	+1	-0.5	-0.6	-0.2
D150E	11.5	46.0	0	0.1	0.2	0.1
D150Q	7.6	47.0	+1	-0.5	-0.5	-0.2
D213E	0.8	53.0	0	-0.1	-0.1	0.1
Q225E	6.3	47.5	-1	0.1	0.1	0.1
Q225R	8.2	46.9	+1	-0.1	-0.2	-0.3
N227D	1.4	51.5	-1	0.2	0.2	0.3

[a]Measurements were performed at pH 6.8, experimental errors are less than 15% of the values given.

Charge effects.

To analyze the possible electrostatic effects on the active site, we calculated mutational effects on the (calculated) pKa of groups in the catalytic centre (Table V.I). Generally, the calculated effects on pKa's were small and there is no clear overall correlation between the effects on activity and the effects on pKa's in the active site. However, the

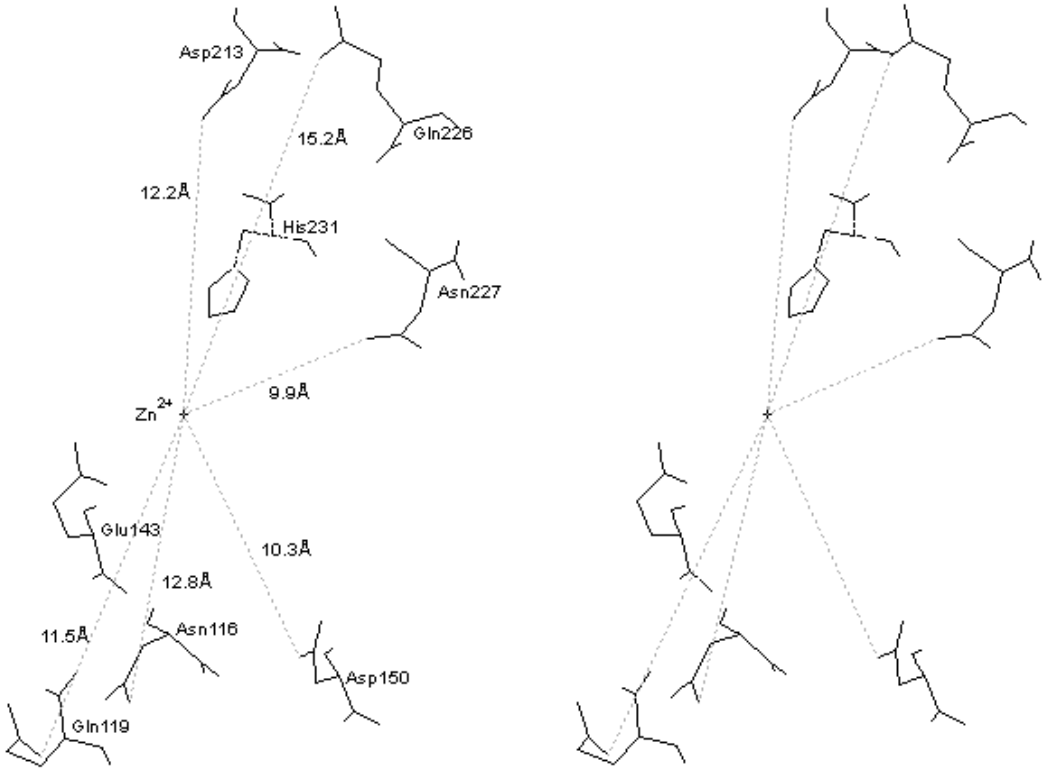

Figure 5.2. Stereo plot of the stereochemical relationship between the mutated residues and the catalytically active residues. The active site Zn^{2+} (small cross at centre), and active site residues Glu143 and His231 are indicated.

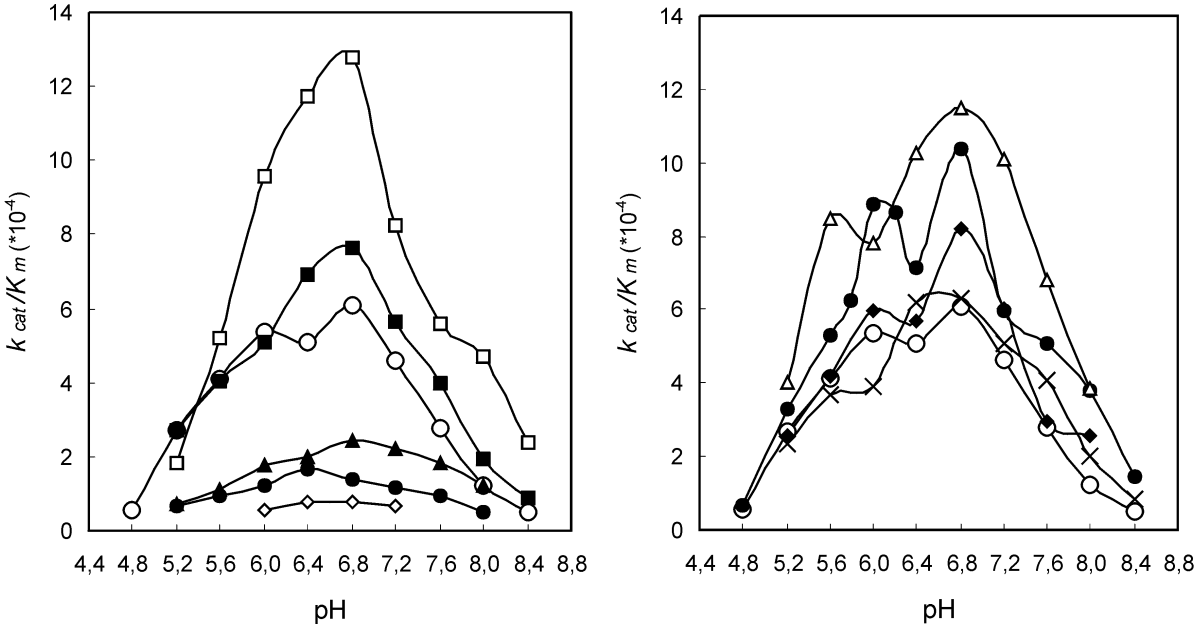

Figure 5.3. pH-dependent activity profiles of TLP-ste and the single surface charge mutants. Left panel; ○ TLP-ste, ●N227D, □ Q119R, ■ D150Q, ▲D150N, and ◊ D213E. Right panel; ○ TLP-ste, ● N116D, Δ D150E, ♦ Q225R, and × Q225E. Experimental errors are less than 15% of the values given.

various mutations at position 150 showed a relatively clear correlation; comparison of D150D (wild-type) with D150N and comparison of D150E with D150Q (Table V.I) shows that removal of charge at position 150 leads to a relatively large decrease in active site pKa's which is accompanied by a decrease in activity. Multiple mutations were also made at position 225, but here no clear correlation is observed: although replacement of the uncharged Q225 by either a negative (Glu) or a positive (Arg) amino acid had opposite effects on the calculated pKa's, the effects on the pH-activity profile and on the activity at the pH optimum were marginal and almost identical.

Hydrogen-bonds and loop flexibility; possible effects on dynamics.

Residues 116 and 119 are located in the same surface loop and share a hydrogen bond in the wild-type enzyme. Interestingly, changing the interaction between residues 116 and 119 by replacement of either one of those residues leads to an increase in activity, regardless of the net change in charge. This suggests that the increase in activity is due to effects other than electrostatics. An appealing hypothesis is that the increased activity is due to increased mobility of the surface loop.

The mutations at position 150 also indicate that other effects may play a role, in addition to expected effects on the pKa's of active site groups. This is clearly seen in a comparison of the activities of D150N and D150Q at pH 6.8. Both mutations remove the charge at this position and are expected to yield nearly equal ΔpKa's. However, D150N displays a 60% drop in activity, whereas D150Q shows a 25% increase in activity. Repositioning the charge at position 150 as in D150E resulted in a 90% increase in activity at pH 6.8.

Taken together, these observations show that other factors, such as changes in loop flexibility due to disrupted or changed hydrogen-bonds and relatively small changes in the spatial localization of relevant charges, contribute to the observed mutational effects.

Table V.II. Characteristics of multiple mutants.

Mutation	k_{cat}/K_m [a]	$\Delta G^{\ddagger}$	ΔCharge	ΔpKa		
				Glu143	Zn–H_2O	His231
	$M^{-1}\cdot s^{-1} \times 10^{-4}$	$kJ\cdot mol^{-1}$				
Wild-type	6.1	47.6				
N116D+D150Q	7.2	47.2	0	-0.1	-0.3	-0.1
N116D+Q225R	6.0	47.7	0	0.3	0.0	-0.2
D150E+Q225R	21.9	44.3	+1	0.0	0.0	-0.2
D150Q+Q225R	6.3	47.5	+1	-0.6	-0.7	-0.5
N116D+Q119R+Q225R	21.9	44.3	+1	0.2	-0.1	-0.3
N116D+Q119R+D150E+Q225R	12.8	45.7	+1	0.3	0.1	-0.2
N116D+Q119R+D150Q+Q225R	22.2	44.3	+2	-0.3	-0.6	-0.5

[a]Measurements were performed at pH 6.8, experimental errors are less than 15% of the values given.

Combining mutations.

To examine whether more active enzymes could be obtained and to study the additivity of the mutational effects, double, triple, and quadruple mutants were constructed. Table V.II summarizes the characteristics of the purified multiple mutants and Fig. 5.4 shows pH-activity profiles. The most active mutants displayed a 3.7-fold increase in activity at pH 6.8. Like for the single mutants, changes in pH optimum were marginal, but some multiple mutants displayed distinct changes in the shape

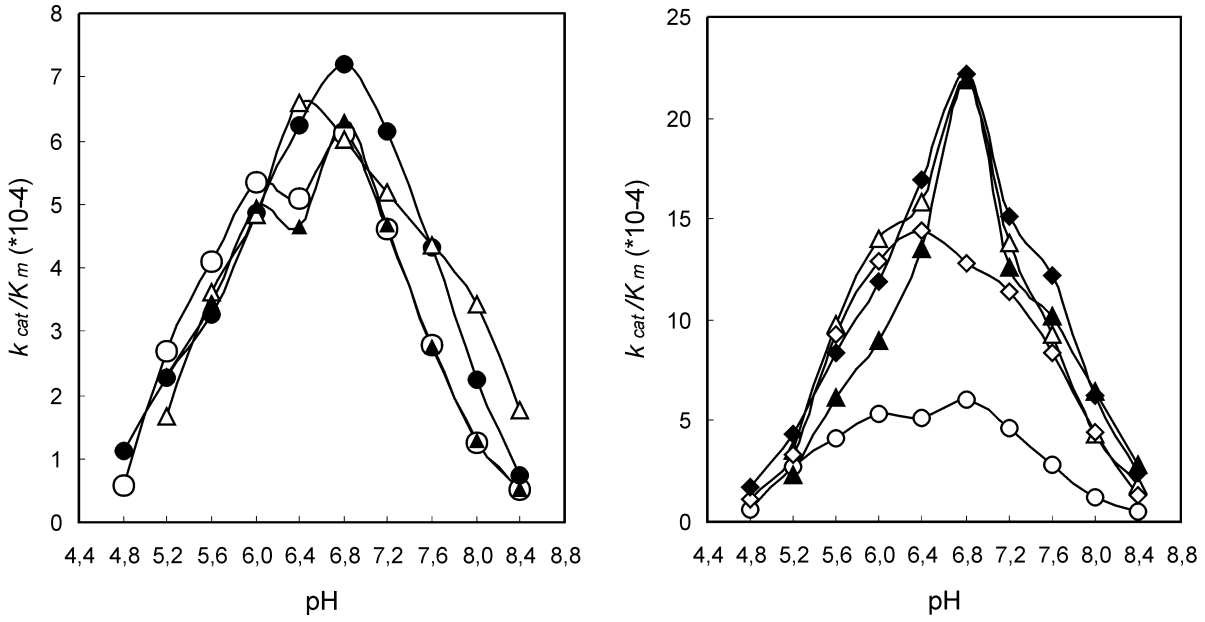

Figure 5.4. pH-dependent activity profiles of TLP-ste and the multiple surface charge mutants. Left panel; ○ TLP-ste, ● N116D+D150Q, Δ N116D+Q225R, and ▲ D150Q+Q225R. Right panel; Δ D150E+Q225R, ▲N116D+Q119R+Q225R, ◊ N116D+Q119R+D150E+Q225R, and ♦ N116D+Q119R+D150Q+Q225R. Experimental errors are less than 15% of the values given.

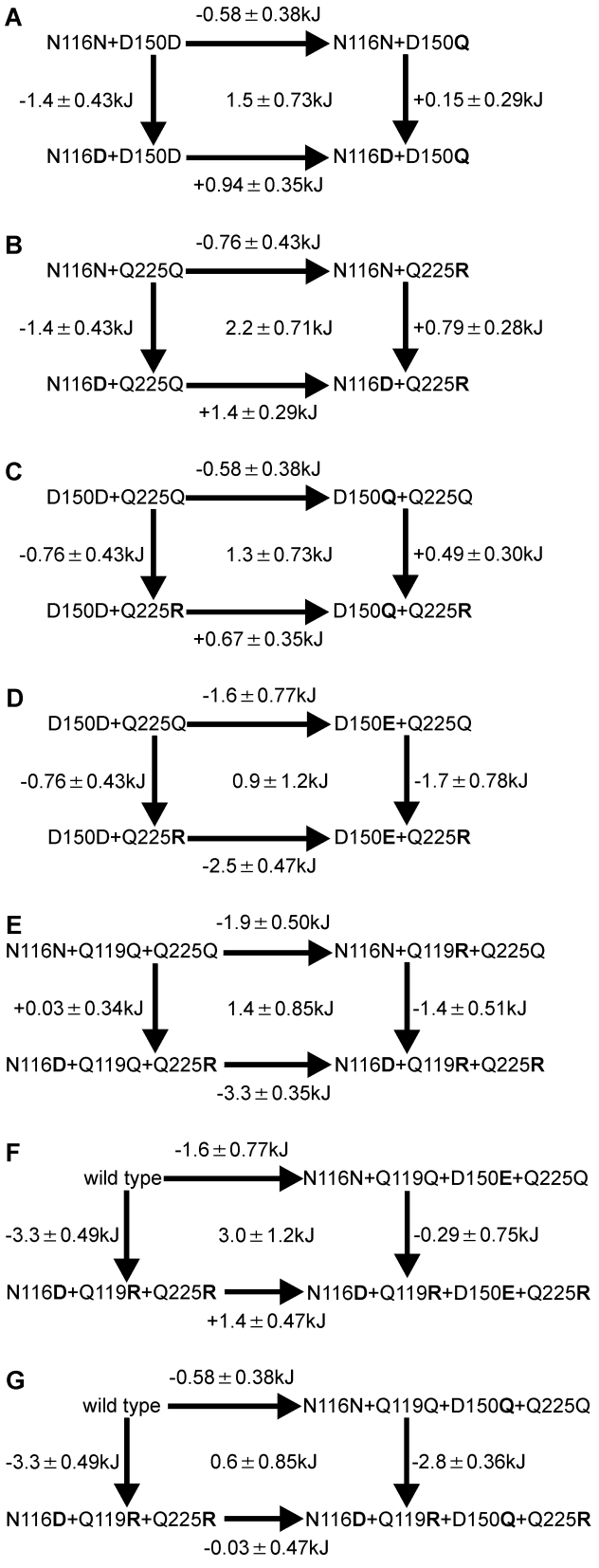

Figure 5.5. Double-mutant cycle analysis. The $\Delta\Delta G^{\ddagger}$ values indicated, were calculated from the k_{cat}/K_m values shown in Tables I and II. The coupling factor $|\Delta\Delta G^{\ddagger}_1 - \Delta\Delta G^{\ddagger}_1'|$ is indicated in the centre of each cycle. A non zero coupling factor indicates that the effects of mutations are dependent on each other. See Materials and Methods for further details.

of the pH-activity profile. The "second" pH optimum at pH 6,0 that is observed in the wild-type enzyme, has completely disappeared in some of the most active mutants, in particular in the triple mutant N116D+Q119R+Q225R and the quadruple mutant N116D+Q119R+D150Q +Q225R. Since the double optimum must be a result of the ionisation constants of the catalytic residues, the disappearance of this double optimum indicates a change in active site electrostatics. However, as noted above with respect to the changes in activity is no clear correlation existed with either ΔpKa or Δcharge. Interestingly, the results indicate considerable non-additivity of mutational effects. For example, addition of the Q225R mutation to D150E increases activity considerably, whereas the Q225R mutation had only marginal effects when introduced into the wild-type enzyme.

Calculation of coupling energies.

To determine the interdependence of the residues mutated in this study, double-mutant cycle analyses were performed as described in the Materials and Methods section. Fig. 5.5 shows the double-mutant cycles which can be constructed from the available data. The $\Delta\Delta G^{\ddagger}$ values and the coupling factors, were calculated from the k_{cat}/K_m values presented in Tables V.I and V.II.

The coupling factor in Fig. 5.5A indicates that the effects of mutations at position 116 and 150 are dependent on each other. The effects of mutating positions 116 and 225 are dependent on each other, as indicated by their coupling factor in Fig. 5.5B. Fig. 5.5C shows the dependence of mutations at position 150 and 225. From these three cycles it can be concluded that the effects of mutations at the positions 116, 150 and 225 are dependent on each other.

Since the contributions of residues 116, 150 and 225 to the pH-activity profile are dependent on each other, demonstrating that the effects of residue 119 are dependent on any

combination of these residues, provides strong evidence that the effects of all four positions are dependent on each other. Fig. 5.5E indicates that the effect of mutating residue 119 is indeed dependent on the residues present at positions 116 and 225. Therefore, the effects of mutating 116, 119, 150 and 225 are all dependent on each other. This is a remarkable observation considering the closest contact distance of almost 25 Å between some of the mutated amino acids and the fact that the closest contact distance of the individual amino acids to the active site residues are all 10 to 15 Å..

Fig. 5.5F and 5.5G both show a double-mutant cycle to determine whether the effect of mutating residue 150 depends on the presence of the combined mutations N116D+Q119R +Q225R. Surprisingly, whereas the effect of D150Q depends on the individual mutations N116D (Fig. 5.5A) and on Q225R (Fig. 5.5C), Fig. 5.5G shows no significant dependence of the D150Q mutation on the presence of the 116+119+225 combination. Even more surprising, Fig. 5.5F, showing the dependence of the effect of D150E on the 116+119+225 combination, shows the largest coupling factor. These results indicate that even though the effects of mutating certain positions are dependent on each other, some mutations might appear to be independent of each other, depending on which combination of amino acids is present.

Discussion

Modification of the surface charge should, according to electrostatic theory, lead to a change in the active site electrostatics (74). Therefore, the catalytic performance of an enzyme may be modified without interfering with the structure of the active site. Here, we have engineered a considerable increase in activity by modifying surface charges in TLP-ste. The most active mutants were approximately four times more active than the wild-type. It is noteworthy that this increase was achieved by mutations which are all far from the active site. The distances between the catalytically important Zn^{2+} and the mutated residues varied from 10 to 15 Å.

This paper has provided several indications that the effects of surface charge mutations on catalysis result from factors which are independent of the charge changes per se. In general, with few exceptions, the observed changes in pH-activity profile and activity do not correlate with the expected change in pKa values. This suggests that other than Δcharge-induced ΔpKa effects, that is, effects that are not accounted for in the software used for calculating pKa values, are important.

A second indication for the occurrence of other than charge effects comes from the studies on additivity of mutational effects. As has been observed before (191), the effects of surface charge mutations are expected to be additive (201). Accordingly, we observed increased effects upon combination of several mutations. However, the mutational effects were far from additive, as shown by the double-mutant cycle analysis. This observation is not surprising for residues 116 and 119, which share a hydrogen-bond in the wild-type enzyme. However, it is highly surprising to find that all residues mutated in this study affect each other, that is, they seem to interact over distances as long as 25 Å. The fact that the interdependent residues are so far apart excludes the possibility of direct contacts and also makes it unlikely that charge effects alone account for the mutational effects.

The third indication of other than charge effects is the discrepancy between the observed effects on the activity, and the observed effects on the pH optimum. If the change in activity is the result of a change in active site electrostatics then a change in the pH-optimum would also be expected. Although changes in the shape of the pH-profile were observed, the pH optimum

itself was not changed in the most active enzyme variants. These three examples probably indicate that other than electrostatic effects play a role, as previously suggested by Nielsen *et al.* (72).

A change in the active site dynamics could be responsible for a large part of the observed effects. This suggestion is supported by earlier work in which replacing uncharged residues by other uncharged residues resulted in changes in activity and in the pH-activity profile that were similar to, or larger than, the changes observed after introducing or removing charges (72). Alternatively, a complex electrostatic network may exist on the surface of the enzyme which makes the effect of a surface charge mutation unpredictable.

The fact that residues 25 Å apart are coupled has serious consequences for modelling and calculation of the effects of charge mutations, since it implies that the effect of any mutation on the surface of an enzyme is dependent on the rest of the surface residues. The coupling also implies that larger parts of an enzyme are involved in optimizing its catalytic centre. If this is true, then nature probably optimised considerably more than just the active site of enzymes during evolution. The large size of enzymes could be explained by the need to balance all the interactions on the surface of an enzyme, and their influence on the catalytic centre.

Considerable increases in catalytic activity can be obtained by modification of the surface charge. The most active mutant obtained was almost four times as active than the wild-type TLP-ste. However, the non-additivity of the mutations and their small effect on the pH-optimum show that important contributions such as active site dynamics and unknown electrostatic network effects exist that are not included in current electrostatic models. Reliable models and predictions as to how to modify the pH profile of an enzyme requires better understanding of these contributions.