Impact of the lysine-188 and aspartic acid-189 inversion on activity of trypsin

Loïc Briand[a], Jean-Marc Chobert[a,*], René Gantier[a], Nathalie Declerck[b], Vinh Tran[c], Joëlle Léonil[d], Daniel Mollé[d], Thomas Haertlé[a]

[a]*Institut National de la Recherche Agronomique, LEIMA, P.O. Box 71627, 44316 Nantes Cedex 3, France*
[b]*INA-PG, Thiverval-Grignon, France*
[c]*LPCM, Nantes, France*
[d]*LRTL, Rennes, France*

Received 17 November 1998

Abstract The impact of the charge rearrangement on the specificity of trypsin was tested by an inversion of sequence K188D/D189K maintaining the integrity of the charges of the substrate binding pocket when switching their polarity. In native trypsin, aspartate 189 situated at the bottom of the primary substrate binding pocket interacts with arginine and lysine side chains of the substrate. The kinetic parameters of the wild-type trypsin and K188D/D189K mutant were determined with synthetic tetrapeptide substrates. Compared with trypsin, the mutant K188D/D189K exhibits a 1.5- to 6-fold increase in the K_m for the substrates containing arginine and lysine, respectively. This mutant shows a ~30-fold decrease of its k_{cat} and its second-order rate constant k_{cat}/K_m decreases ~40- and 150-fold for substrates containing arginine and lysine, respectively. Hence, trypsin K188D/D189K displays a large increase in preference for arginine over lysine.

Key words: Trypsin; Site-directed mutagenesis; Synthetic substrate; Protease

1. Introduction

Site-directed mutagenesis can provide a powerful approach to design enzymes with novel specificities, which is one of the challenges of protease engineering [1].

Trypsin (EC 3.4.21.4) is a well studied endopeptidase. This enzyme belongs to the large and diverse family of serine proteases sharing a catalytic mechanism involving three residues: seryl 195, histidyl 57 and aspartyl 102 (chymotrypsinogen numbering system). Members of this protease family differ in hydrolytic specificities. This is determined by the character of the amino acid side chains lining the surface of the substrate binding pocket. Trypsin, the most specific of the serine proteases, catalyzes the hydrolysis of proteins and peptides at the carboxy-terminus of the α-carbon of arginyl and lysyl residues. Chymotrypsin preferentially attacks peptide bonds following an aromatic amino acid. Elastase does not display very distinct specificity, but generally attacks an uncharged non-aromatic amino acid.

The analysis of the available crystal structures of trypsin indicates that amino acids Asp-189 and Ser-190 are located at the bottom of the primary substrate binding pocket (S1). Substrate arginine and lysine side chains (P1) interact in a different way with the primary binding site determinant Asp-189 [2,3]. The guanidinium group of P1-arginine substrate engages in an ion-pair interaction with Asp-189 (S1), whereas the interaction of P1-lysine is stabilized solely by a water-mediated interaction. Both substrate arginine and lysine side chain form putative hydrogen bonds with Ser-190 OH.

Modulation of the specificity for substrates containing arginine and lysine was shown previously to depend on the amino acid at position 190 and it was proposed that the precise location of the negative charge is essential for proper orientation of the substrate [4]. The removal of the negative charge from position 189 reduced the overall catalytic rate by a factor of at least 10^5 [5–7]. However, the activity was restored partially by the substitution of a Glu at position 189, as well as by the introduction of either Asp or Glu at position 190 [7]. This suggests that the presence of a negative electrostatic field at the base of what is usually named the S1 binding subsite is sufficient to enhance the hydrolysis rate, even in cases where the geometry of the interaction is far from optimal. Alignment of the known trypsin sequences reveals that the basic character of the amino acid residue at position 188 is highly conserved. The substitution of this crucial position with histidine generates a very efficient copper binding site and transition metal switch [8]. In another series of experiments, Lys-188 was replaced with aromatic amino acid residues in order to amplify local hydrophobic interactions at the substrate binding site [9]. The kinetic analysis revealed that all the mutants (K188F, K188Y, and K188W) still conserved their capacity to split the peptide bonds involving arginyl and lysyl residues. As shown only by proteolysis of a natural substrate (β-casein), these mutants displayed a considerable amplification of cleavage of peptide bonds involving asparagine and glutamine amino acids.

To test to what extent the charges can be displaced, an inversion of sequence K188D/D189K was realized by site-directed mutagenesis. This allows maintaining the integrity of global charges of the substrate binding pocket of the trypsin when their polarity is switched.

The kinetic parameters of the wild-type trypsin and its K188D/D189K mutant were determined by hydrolysis of synthetic substrates.

2. Materials and methods

Trypsin mutant was prepared by oligonucleotide-directed mutagenesis carried out by the method of Kunkel [10], using single-stranded, uracil-containing DNA templates as previously described [8].

The 30-base oligonucleotide primer 5′-GCCCTGGCAGGACTTATCGCCTCCCTCTAG-3′ directing the mutation was synthesized

*Corresponding author. Fax: (33) 2.40.67.50.84.
E-mail: chobert@nantes.inra.fr

PII: S0014-5793(98)01611-1

with a Biosearch Cyclone DNA synthesizer (mismatched bases are underlined).

Wild-type and K188D/D189K trypsins were secreted in the periplasm space of *Escherichia coli* and purified by chromatography on CM-Sepharose fast-flow column and benzamidine-Sepharose column as previously described [11].

2.1. Kinetic measurements

Succinyl-Ala-Ala-Pro-Arg-*pNA* and succinyl-Ala-Ala-Pro-Lys-*pNA* were used for the kinetic studies. Their hydrolysis was monitored spectrophotometrically at 405 nm ($\varepsilon = 13\,500$) in a Varian Cary 13E spectrophotometer. Twelve substrate concentrations from the appropriate range were used for determining initial rates. The initial slopes of the reaction curves were measured to yield ΔA/min. Michaelis-Menten parameters were calculated from Eadie-Hofstee plots. Hydrolysis was performed at 37°C using 50 mM Tris-HCl, 20 mM $CaCl_2$, pH 8.

2.2. Computer modeling method

Molecular modeling studies were carried out on Silicon Graphics computers with Biosym/MSI packages. Molecular displays and energy minimization were performed with Insight II, Biopolymer and Discover modules. For all calculations, the CFF91 forcefield was selected which represents an efficient evolution of the initial CVFF with essentially a quartic polynomial for bond stretching and angle bending terms. The van der Waals term was reconsidered using an inverse 9th-power term for the repulsive part and more interestingly for these docking studies, the balance between Coulombic and van der Waals terms was considerably improved.

The starting structure of trypsin was extracted from X-ray data of its complex with ecotin [12] accessible from the Protein Data Bank (1SLU). This structure was chosen because of its high crystallographic resolution (1.8 Å) but since two segments were missing (113–117 and 146–148), they were replaced by those found in the trypsin belonging to the complex with the benzamidine (1TRM) resolved by Sprang et al. [13]. For each segment, this artificial replacement was done by superimposing the adjacent residues (112, 118 and 145, 149, respectively) before the minimization procedures. In fact, this restoration of the trypsin has no real effect on the conformation of the substrate binding pocket delimited by the two segments 182–195 and 213–228.

The calculations were carried out in two steps. The first one consisted in building the mutant K188D/D189K and performing energy comparisons with the wild-type trypsin in the absence of substrate. Then, the effect of inversion was estimated on the binding of the substrate. In the first stage, the mutant model was obtained by replacing Lys-188a with Asp and Asp-189 with Lys using Insight II facilities. For each replacement, several side chain orientations were tested and the best solution was kept for further calculations (Biopolymer facility). Finally, the addition of hydrogen atoms was done with respect to pH 8 (Insight II facility) and both wild-type and mutant molecules were optimized with the same protocol as follows: a minimization of 3000 iterations fixing all heavy atoms except those belonging to the side chains of residues involved in the substrate binding site. Since in β-casein the first peptide bond cleaved by trypsin is Arg-25–Ile-26 [14], during the second step of calculations, the wild-type was initially minimized in the presence of a substrate fragment modeled with only three residues (Thr-Arg-Ile, simulating segment Thr-24–Arg-25–Ile-26 of β-casein). The initial orientation of the central residue inside the binding pocket was derived from that extracted from the complex trypsin/BBI (Bowman-Birk inhibitor) resolved by Huang et al. [15] and accessible in the PDB (1SMF). In this latter case, the binding occurs with the side chain of a lysyl residue (necessitating its replacement by Arg) but interestingly, two water molecules were identified in the binding pocket. Therefore, these water molecules were added in all our calculations. Then, for the wild-type and the mutant trypsins, the docking calculations were performed using the same minimization conditions: 1500 iterations only fixing the backbone heavy atoms of the trypsin. Therefore, the substrate as well as the two water molecules were allowed to vary. Finally, the interaction energies (trypsin+water/substrate) were used to estimate the enzyme/substrate affinity.

3. Results and discussion

In this study, rat anionic trypsin was modified by site-directed mutagenesis leading to an inversion of Asp-189 and Lys-188 residues. The purpose of this strategy was to modulate the specificity of trypsin by a simple exchange of charged residues when maintaining the global charge in the substrate binding pocket.

3.1. Kinetic analysis of wild-type and K188D/D189K trypsins

The trypsin mutant was characterized by analysis of the kinetic parameters k_{cat}, K_m, and k_{cat}/K_m for comparable tetrapeptide substrates containing arginine and lysine. Compared with trypsin, the mutant K188D/D189K exhibits a 1.5- to 6-fold increase in K_m values for substrates containing arginine and lysine, respectively (Table 1). This mutant has reduced ~30-fold its k_{cat} values and has decreased the second-order rate constant k_{cat}/K_m ~40- and 150-fold for substrates containing arginine and lysine, respectively. Wild-type trypsin has a 6-fold arginyl/lysyl preference whereas trypsin K188D/D189K prefers the substrate containing arginine over the lysine by a surprising factor of 22.

Trypsin has a negative electrostatic field in the substrate binding pocket. This negative field attracts the solvated arginine and lysine substrate side chains and helps to stabilize their positive charges in the enzyme-substrate complex in which solvent is excluded.

Besides their identical charges in most simple cases at pH below ~8.5, arginine and lysine side chains are structurally different. Thus, an enzyme that has evolved to cleave both arginyl and lysyl substrates may take advantage of the electrostatic similarity and yet have additional features allowing it to be structurally specific for each one of the side chains. Trypsin binding pocket has a cylindrical shape which can accommodate easily both arginine and lysine side chains. Therefore, the precise location of a negatively charged group within or with respect to the trypsin S1 site is important for precise positioning of the scissile bonds in catalytic distance from nucleophilic components of the triad, Ser-195 and His-57.

Although the requirements for productive arginine binding are more restrictive, due to its greater size and reduced flexibility, the variant K188D/D189K shows an enhanced prefer-

Table 1
Kinetic parameters of trypsin variants

Enzyme	Substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})	Arg/Lys[b]
Wild-type	**Arg**[a]	32.8 ± 7.1	163 ± 31	4.2 ± 1.2	5.9
K188D/D189K		50.8 ± 4.2	4.7 ± 0.1	0.11 ± 0.01	22.5
Wild-type	**Lys**[a]	105 ± 15	74 ± 15	0.71 ± 0.21	
K188D/D189K		651 ± 36	3.2 ± 0.7	0.0049 ± 0.0006	

[a]Arg and Lys denote the tetrapeptide amide substrates succinyl-Ala-Ala-Pro-Arg-*pNA* and succinyl-Ala-Ala-Pro-Lys-*pNA*.
[b]Arg/Lys preference is calculated as the ratio of k_{cat}/K_m values.

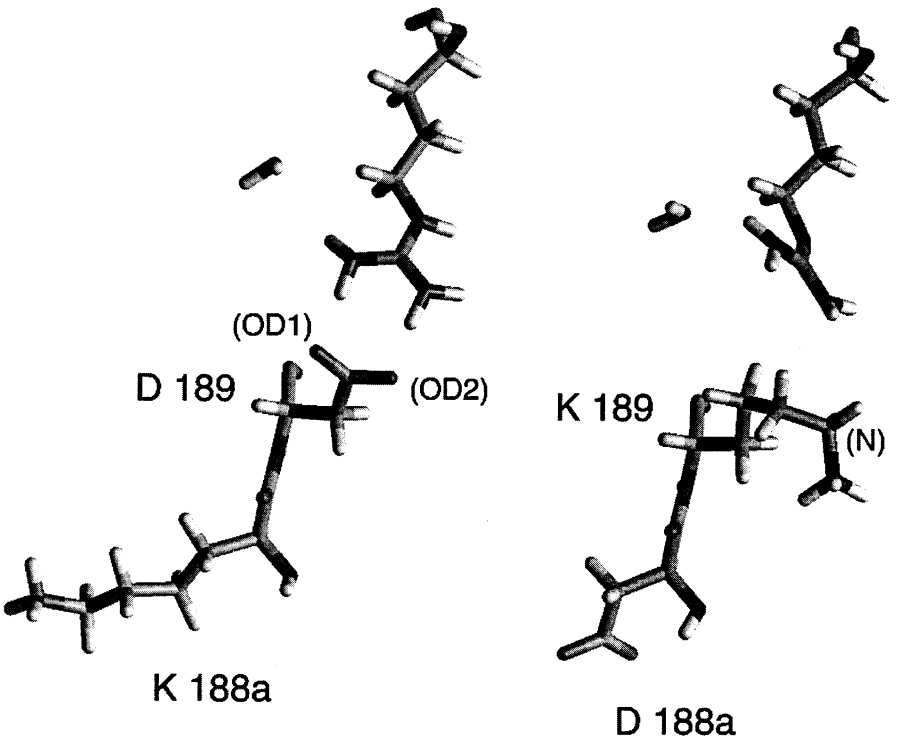

Fig. 1. Left, Lys-188 side chain pointing outward and Asp-189 side chain, with (upper part) the arginyl side chain of the substrate in wild-type trypsin. Right, molecular modeling of K188D/D189K mutation.

ence for succinyl-Ala-Ala-Pro-Arg-*pNA* substrate. Since a similar decrease in k_{cat} is observed for both substrates, this preference is principally caused by the increase in K_m for succinyl-Ala-Ala-Pro-Lys-*pNA*.

3.2. Comparison of the kinetic parameters of K188D/D189K with other trypsin mutants

While the importance of a negative charge has been proved by the incapacity of the D189N and D189N/S190T trypsins to cleave artificial substrates containing arginine and lysine [7] – activity toward peptidyl arginine or lysine containing amide substrates being partially restored by the presence of an Asp or Glu residue at positions 189 or 190 –, a recent study [16] showed that a triple trypsin mutant G187W/K188F/D189Y surprisingly conserved its activity towards tetrapeptide substrates containing arginine and lysine. This trypsin mutant exhibited a 1.3-fold increase in K_m values for substrates containing arginine and lysine. Its k_{cat} values decreased 30–40-fold compared with wild-type trypsin (Table 2). The second-order rate constant k_{cat}/K_m of this triple trypsin mutant, measuring catalytic efficiency, decreased ~40- and 55-fold for substrates containing arginine and lysine, respectively. The arginyl/lysyl preference was similar to that observed with wild-type trypsin.

In a previous work, Evnin et al. [7] showed that trypsin mutant D189E presented an increase in K_m and a decrease in k_{cat} values, and exhibited a 7000- and 200-fold decrease in k_{cat}/K_m for substrates containing arginine and lysine, respectively. Mutant D189E preferred the lysyl over the arginyl substrate ~40-fold when compared with wild-type trypsin. Shifting the negative charge in the mutant D189S/S190D, these authors reported a smaller increase in K_m and more

Table 2
Relative influence (%) of mutation on kinetic parameters of trypsin variants

Enzyme	Substrate	K_m	k_{cat}	k_{cat}/K_m	Arg/Lys[c]
Wild-type	Arg[a]	100	100	100	6
K188D/D189K	Arg[a]	156	2.9	2.60	22
D189S/S190D[d]	Arg[b]	2900	1.3	0.04	2
G187W/K188F/D189Y[e]	Arg[a]	140	3.0	2.60	8
Wild-type	Lys[a]	100	100	100	
K188D/D189K	Lys[a]	620	4.3	0.70	
D189S/S190D[d]	Lys[b]	1800	1.4	0.09	
G187W/K188F/D189Y[e]	Lys[a]	120	2.3	1.80	

[a]Arg and Lys denote the tetrapeptide amide substrates succinyl-Ala-Ala-Pro-Arg-*pNA* and succinyl-Ala-Ala-Pro-Lys-*pNA*.
[b]Arg and Lys denote the tripeptide amide substrates N^{α}-*p*-tosyl-Gly-Pro-Arg-AMC and N^{α}-*p*-tosyl-Gly-Pro-Lys-AMC.
[c]Arg/Lys preference is calculated as the ratio of k_{cat}/K_m values.
[d]According to Evnin et al. [7].
[e]According to Chobert et al. [16].

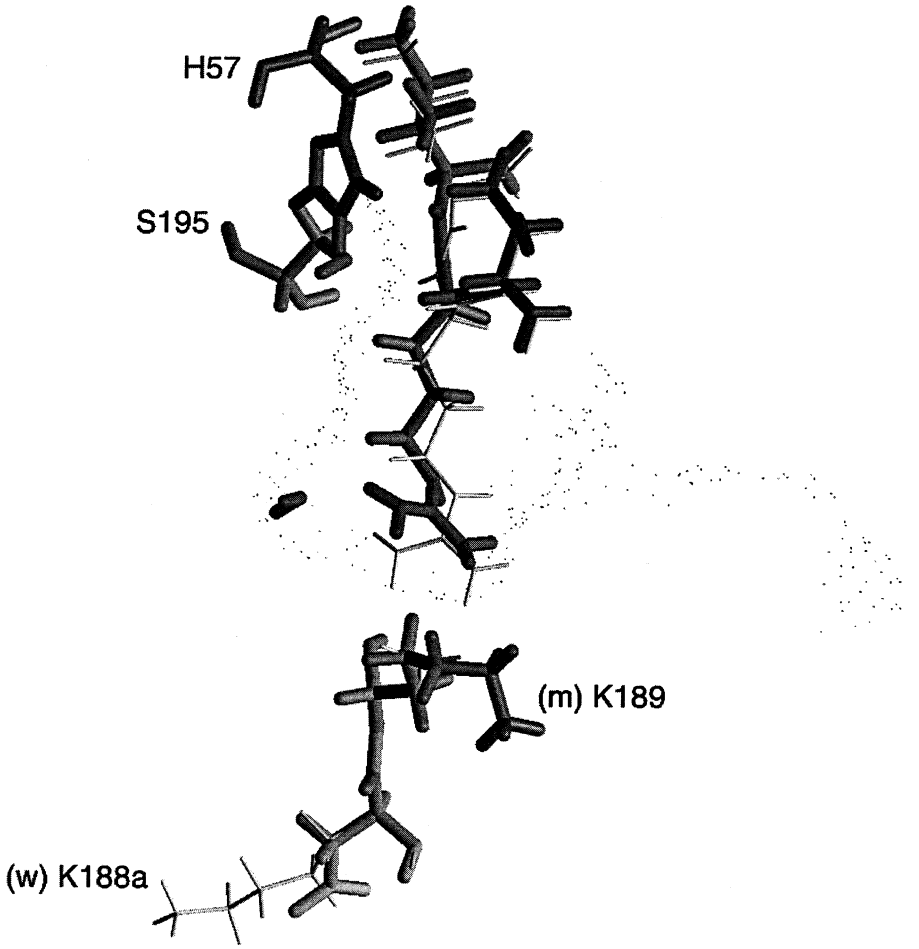

Fig. 2. Superposition of wild-type (thin line) and K188D/D189K mutant (thick line) trypsins. The upper part shows positioning of the Thr-Arg-Ile substrate in the substrate binding pocket of wild-type (thin line) and mutant (thick line) trypsins.

affected k_{cat}, when compared to the D189E mutant. Thus, the kinetic analysis revealed that mutant D189S/S190D had reduced k_{cat}/K_m values $\sim$3 orders of magnitude (Table 2). The arginyl/lysyl preference was attenuated only two-fold.

When the catalytic properties of the obtained mutant K188D/D189K were compared with those previously described by Evnin et al. [7] and Chobert et al. [16] it appeared that the K_m values for cleavage of the synthetic peptides increased in the following order: wild-type $<$ G187W/K188F/D189Y $\leq$ K188D/D189K $\ll$ D189S/S190D (Table 2). When the K_m values increased, k_{cat} and k_{cat}/K_m values for the mutant trypsins decreased in the same order. Taken together, the results imply that the catalytic function was more affected by inversion than substrate recognition by the trypsin.

The present study confirms that a negatively charged amino acid situated at the bottom of the substrate binding pocket is important for both specificity and activity. In contrast to the mutant D189S/S190D [7] showing decreased preference for arginyl/lysyl, trypsin K188D/D189K increases this preference. Thus, mutant D189S/S190D, wild-type, mutant G187W/K188F/D189Y and mutant K188D/D189K trypsins exhibit a 2-, 6-, 8- and 22-fold higher arginyl preference over lysyl, respectively. This suggests that the shift of the negative charge substantially modulates the selectivity of arginyl/lysyl. The increase in the relative arginine specificity for mutant K188D/D189K trypsin may reflect the rise of selective stabilization of the transition state involving the arginine-containing substrates.

3.3. Molecular modeling

By molecular modeling, in the absence of substrate in the binding pocket, the total energies of the wild-type and the mutant are quite comparable. The difference of 82 kcal/mol (in favor of the wild-type) can be explained by the rather strict minimization conditions perturbing the full relaxation of the mutant after its construction. However, no significant steric constraints are revealed and the concerned backbones (188a–189; amino acid labeling according to the X-ray crystal structure of Brinen et al. [12]) of the two molecules are well superimposed. The side chain of residue 189 (Asp and Lys for wild-type and mutant, respectively) is pointing toward the binding site contrary to that of residue 188a (Lys and Asp for wild-type and mutant, respectively). In both cases, there is room enough but the binding pocket of the mutant is significantly smaller.

In the presence of the substrate, the docking calculations yielded quite different results even if the substrate backbone had not significantly moved in both cases. For the wild-type, the steric complementarity is excellent in the binding site – especially at the bottom of this pocket – as shown in Fig. 1 (left) which explains the good enzyme/substrate affinity. For the K188D/D189K trypsin, the side chain of Lys-189, initially

pointing toward the binding site, has to accommodate the presence of the Arg side chain of the substrate and therefore is distorted as shown in Fig. 1 (right), as well as the substrate. The superimposition of wild-type and mutant (Fig. 2) illustrates the deformation of the substrate side chain. In Fig. 2, the dot contour delimits the new binding pocket of the mutant. The difference of interaction energies (84.2 kcal/mol) between the wild-type and the mutant confirms the preferential affinity of the substrate for the wild-type as found experimentally (see above). However, other conformational changes elsewhere in the molecule and long range effects may influence the activity.

Additional structural and functional characterization of this enzyme would be required for a complete understanding of the functional consequences of the mutations in the Asp-189 site and in its neighborhood. Together with long distance interactions, the electrostatic forces are the least understood factor influencing enzyme structure and functions. Until recently, the effect of trypsin mutations on catalytic activity was determined mostly using synthetic substrates or small peptides. Synthetic substrates allow rapid calculation of the catalytic constants of an enzyme. However, the major part of conformational states induced by the medium, characteristic of longer peptides and proteins, such as changes in electrostatic charges of the targeted peptide bond, and created steric hindrance, are neglected during such tests. Since these factors are clearly critical, the hydrolysis of a long and well defined natural substrate, such as β-casein, is under investigation. The use of a full-size native protein as substrate should allow us to understand better the validity and scope of the results obtained with synthetic substrates.

Acknowledgements: The funding by a joint fellowship from the Institut National de la Recherche Agronomique and the Région Pays de Loire to Loïc Briand is acknowledged. The funding of this work by INRA and the Région Pays de Loire within the framework of the VANAM program 'Conception de nouveaux outils protéolytiques' is also greatly acknowledged. We also thank Dr. C.S. Craik (Department of Pharmaceutical Chemistry, University of California, San Francisco, CA, USA) for providing the DNA encoding the anionic rat trypsin. The authors are very grateful to Marie-Georgette Nicolas and Andrée Lépingle for their experimental assistance.

References

[1] Perona, J.J. and Craik, C.S. (1995) Protein Sci. 4, 337–360.
[2] Ruhlmann, A., Kukla, D., Schwager, P., Bartels, K. and Huber, R. (1973) J. Mol. Biol. 77, 417–436.
[3] Bode, W., Walter, J., Huber, R., Wenzel, H.R. and Tschesche, H. (1984) Eur. J. Biochem. 144, 185–190.
[4] Perona, J.J., Tsu, C.A., McGrath, M.E., Craik, C.S. and Fletterick, R.J. (1993) J. Mol. Biol. 230, 934–949.
[5] Graf, L., Craik, C.S., Pathy, A., Roczniak, S., Fletterick, R.J. and Rutter, W.J. (1987) Biochemistry 26, 2616–2623.
[6] Graf, L., Jansco, A., Szilagyi, L., Hegyi, G., Pinter, K., Naray-Szabo, G., Hepp, J., Medzihradsky, K. and Rutter, W.J. (1988) Proc. Natl. Acad. Sci. USA 85, 4961–4965.
[7] Evnin, L.B., Vasquez, J.R. and Craik, C.S. (1990) Proc. Natl. Acad. Sci. USA 87, 6659–6663.
[8] Briand, L., Chobert, J.M., Tauzin, J., Declerck, N., Léonil, J., Mollé, D., Tran, V. and Haertlé, T. (1997) Protein Eng. 10, 551–560.
[9] Chobert, J.M., Briand, L., Tran, V. and Haertlé, T. (1998) Biochim. Biophys. Res. Commun. 246, 847–858.
[10] Kunkel, T.A. (1985) Proc. Natl. Acad. Sci. USA 82, 488–492.
[11] Pouvreau, L., Chobert, J.M., Briand, L., Quillien, L., Tran, V., Guéguen, J. and Haertlé, T. (1998) FEBS Lett. 423, 167–172.
[12] Brinen, L.S., Willett, W.S., Craik, C.S. and Fletterick, R.J. (1996) Biochemistry 35, 5999–6009.
[13] Sprang, S., Standing, T., Fletterick, R.J., Stroud, R.M., Finer-Moore, J., Xuong, N.-H., Hamlin, R., Rutter, W.J. and Craik, C.S. (1987) Science 237, 905–909.
[14] Léonil, J., Mollé, D. and Maubois, J.L. (1988) Lait 68, 281–294.
[15] Huang, R., Li, Y., Zhang, S., Tang, Y. and Qi, C. (1992) in: Protein Data Bank, 24 October.
[16] Chobert, J.M., Briand, L. and Haertlé, T. (1998) J. Food Biochem. 22, 529–545.