The α1/2 helical backbone of the prodomains defines the intrinsic inhibitory specificity in the cathepsin L-like cysteine protease subfamily

Ying Lan Guo[a], Ursula Kurz[a], J.E. Schultz[a], Chin Chia Lim[b], B. Wiederanders[b], K. Schilling[b,*]

[a] *Fakultät für Chemie und Pharmazie, Universität Tübingen, Morgenstelle 8, D-72076 Tübingen, Germany*
[b] *Institut für Biochemie I, Klinikum der Friedrich-Schiller-Universität, Nonnenplan 2, D-07740 Jena, Germany*

Received 29 November 1999; received in revised form 10 February 2000

Edited by Felix Wieland

Abstract Proregions of papain-like cysteine proteases are potent and often highly selective inhibitors of their parental enzymes. The molecular basis of their selectivity is poorly understood. For two closely related members of the cathepsin L-like subfamily we established strong selectivity differences. The propeptide of cathepsin S was observed to inhibit cathepsin L with a K_i of 0.08 nM, yet cathepsin L propeptide inhibited cathepsin S only poorly. To identify the respective structural correlates we engineered chimeric propeptides and compared their inhibitory specificity with the wild-types. Specificity resided in the N-terminal part, strongly suggesting that the backbone of the prodomain was the underlying structure.

Key words: Papain family; Propeptide; Specificity; Inhibition kinetics; *Paramecium*

1. Introduction

As papain-like cysteine proteases, cathepsin L and S are synthesised as proenzymes. The proregions, comprising almost one third of the zymogen, are strictly required for expression of native protease, i.e. for correct folding, stability and intracellular targeting [1–4]. Their most evident function, however, is inhibition of the parental enzymes in a pH-dependent manner [5–7]. Fox et al. [8] were the first to show that a covalent linkage between the mature part of the enzyme and the proregion is not essential for this function. A synthetic peptide consisting of 56 N-terminal residues of procathepsin B specifically inhibited cathepsin B with a K_i of 0.4 nM, whereas the K_i for papain was 5.6 μM. Similar results have been published for members of the cathepsin L-like subfamily, e.g. papain and cathepsins L and S [9–12]. In each case, the propeptide binds to the cognate protease with high affinity, albeit not with absolute specificity. Maubach et al. [12] found that the propeptide of cathepsin S inhibits human cathepsin S (CS), as well as cathepsins L from human and *Paramecium* (CL) with similar potency. On the other hand, the propeptide of human cathepsin L was shown to be more selective, discriminating cathepsin S by a factor of 500, compared to its cognate enzyme [10]. The structural basis for this discrimination has not yet been investigated, albeit it would be important for the modulation of the activity of specific cysteine proteases in biological systems by proregion-derived inhibitors, as suggested by Visal et al. [13].

X-Ray data for the mature cathepsins S and L revealed a very similar fold [14]. Thus, the selectivity differences in propeptide enzyme interaction of the two cathepsins were attributed to structural differences between their proregions. Consequently, construction of propeptide chimeras was considered to be a useful attempt to identify the specificity determining parts of the cathepsin L-like zymogens.

2. Material and methods

2.1. Reagents

Radioisotopes and Sequenase Version 2.0 were from Amersham Life Sciences, the substrates *Z*-Phe-Arg-NH-Mec and *Z*-Val-Val-Arg-NH-Mec were from Bachem Biochemica, E-64 from Serva. IgA protease, all restriction enzymes and nucleases were from either Roche Diagnostics or New England Biolabs. Nylon membranes were obtained from Spectrum Medical Industries. pEV41C was used as the expression vector. Oligonucleotides were synthesised by MWG-Biotech. Ni-NTA resin was from Qiagen. Preparation of CL, CS and ppCS was carried out according to Völkel et al. [11] and Maubach et al. [12], respectively.

2.2. Expression and purification of recombinant propeptides

The cDNA fragments coding for ppCL and ppCS were amplified by PCR. Plasmids containing the proregions of CL [11] and CS [16] served as templates, the respective sense/antisense primers were 5′AGGTCGTCATATGAATCTTTATGCAAATTGG/5′ATCCTCGAGTCACTTGTATTGGAAGTTAG and 5′AGGTCGTCATATGGTGGCACAGTTGCATAAAG/5′ATCGTCGACTCATATCCGATTAGGGTTTGAC. PCR cycles (20) were run at 94°C (50 s), 45–55°C (depending on the primer pair used; 50 s) and 72°C (60 s). Products were digested with *Nde*I and *Xho*I or *Sal*I, blunt-ended by Klenow treatment and inserted into the *Sma*I site of pEV41C, a heat-inducible vector fusing a $(His)_6$-tag and an IgA-protease (Igase) digestion site with the N-terminus of the recombinant protein. Fidelity of all steps was ascertained by sequencing. Using these vectors as templates, the cDNA segments for the chimeras, specified in Fig. 1, were PCR-generated using primer pairs covering the 5′-*Nco*I or 3′-*Eco*RI site in the pEV41C multiple cloning site and the linking boundaries, respectively [S/L-chimera: 5′CATCACCACCATGGCCAGC (vector sequence)/5′CATTCCCATTGAATGCTCCAGG and 5′GCCACATTCACTTTGGAATTG/antisense primer for ppCL, above specified; L/S-chimeras: ppCL sense primer/5′TTCTTCTGGACTTTCATAGAAAGC and 5′CACTCATACGATCTGGGCATG (L_{1-43}/S_{55-98}) or 5′GGAATGCACTCATACGATC (L_{1-43}/S_{53-98})/5′GCAGAATTCTCATATCCGATTAGGGTTTGAC]. Triple ligation led to vectors with chimeric inserts. Expression and purification of the propeptides was carried out according to Völkel et al. [11], except that protein was eluted with 300 μl 50 mM phosphoric acid from the Ni-NTA-column. The $(His)_6$-tag was cleaved off at an Igase/protein ratio

*Corresponding author. Fax: (49)-3641-938612.
E-mail: ksch@mti-n.mti.uni-jena.de

Abbreviations: NH-Mec, 7-amido-4-methylcoumarin; CS, human cathepsin S; ppCS, propeptide of human cathepsin S; CL, *Paramecium* cathepsin L; ppCL, propeptide of *Paramecium* cathepsin L

PII: S0014-5793(00)01281-3

of 1:250 at 22°C in 13 mM potassium phosphate pH 8.0, 80 mM NaCl, 20 mM KCl, 1 mM EDTA, 1 M urea at 22°C for 2–3 days. Due to the Igase specificity, all propeptides bear the N-terminal extension Tyr-Pro. Non-cleaved material was bound to Ni-NTA resin. Dialysis (1 M NaCl) led to a precipitate which was dissolved in 50 mM phosphoric acid. Yields varied between 0.3–5.5 mg of purified propeptides/200 ml culture. Protein concentration was calculated from A_{280}, using amino acid composition based absorbance coefficients. The purity and identity of all peptides was assessed by SDS/PAGE and MALDI-TOF.

2.3. Kinetic measurements

Enzyme activities (kinetic measurements with a LS50B Perkin Elmer fluorimeter) and dissociation constants (K_d) for the enzyme propeptide complexes (competitive Michaelis–Menten-like model) were determined according to Kirschke and Wiederanders [15], as modified by Maubach et al. [12], at 37°C. The Michaelis constants (K_m), essential for K_d calculations, are 13.8 and 3.6 μM for the substrates *Z*-Val-Val-Arg-NH-Mec (CS) and *Z*-Phe-Arg-NH-Mec (CL), respectively. The software 'PRISM' (GraphPad) was used for all calculations.

3. Results and discussion

3.1. Inhibition of cathepsins S and L by the recombinant wild-type propeptides

Tight binding of ppCL and ppCS to their parental enzymes as well as their lower affinity to papain and to the cathepsins B and H have been reported [11,12]. To establish reference values for the purpose of our studies, we first determined the kinetic parameters of CL and CS inhibition by the wild-type propeptides under the experimental conditions. Product formation (P) as a function of time (t) at inhibitor concentrations [I] indicated at the curves (Fig. 2A,B) showed the typical slow-binding behaviour of high-affinity ligands employed at very low concentrations. Fitting the data to $P = v_s t + (v_z - v_s)(1 - e^{-\lambda t})/\lambda$ led to values for the three parameters: initial velocity (v_z), steady-state velocity (v_s) and apparent first-order rate constant for the pre-steady state exponential phase (λ). Plots of the parameter values versus [I] (Fig. 2C,D) revealed the reaction mechanism and the kinetic constants for inhibition [17,18]. The primary data for CL showed that v_z decreased with increasing [I] (Fig. 2A) which indicated instant inhibition. The initial complex between CL and ppCL had only moderate stability [initial K_d (K_i^*) = 21 nM], which time-dependently increased by nearly one order of magnitude [overall K_d (K_i) = 2.6 nM]. Together with the saturation of λ with increasing [I], ensuing from the non-linearity of the respective curve in Fig. 2C (right *Y*-axis), this indicated a two-step reaction mechanism, most probably caused by an induced conformational fit of ppCL to a more productive conformation after its initial binding to the enzyme, as supported by the fit of $\lambda = f(I)$ with the equation for this mechanism in Fig. 2C. The reasons for the initial misfit between the inhibitor and its target were beyond the scope of this investigation; even certain assay conditions can not be ruled out definitely, e.g. the temperature may be unphysiological for protozoan proteins. However, 37°C was the set measuring the steady state propeptide inhibition of the canonical papain and other plant-members of the family [19]. Except for papain (K_i = 1.9 nM), all these constants were higher than our K_i for *Paramecium* cathepsin L (2.6 nM, Table 1).

In contrast, ppCS instantly inhibited CS with high affinity (K_i = 0.05 nM; Fig. 2B,D), λ was a linear function of [I] and v_z remained nearly constant (Fig. 2D). These results are readily fitted by a one-step reaction and agree with those from

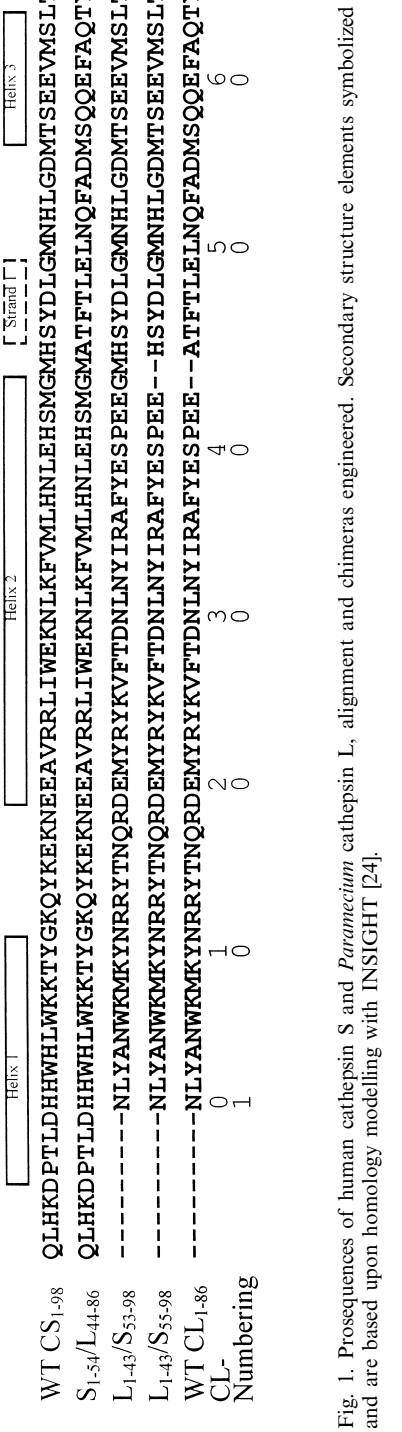

Fig. 1. Prosequences of human cathepsin S and *Paramecium* cathepsin L, alignment and chimeras engineered. Secondary structure elements symbolized in the first line refer to the CS proregion and are based upon homology modelling with INSIGHT [24].

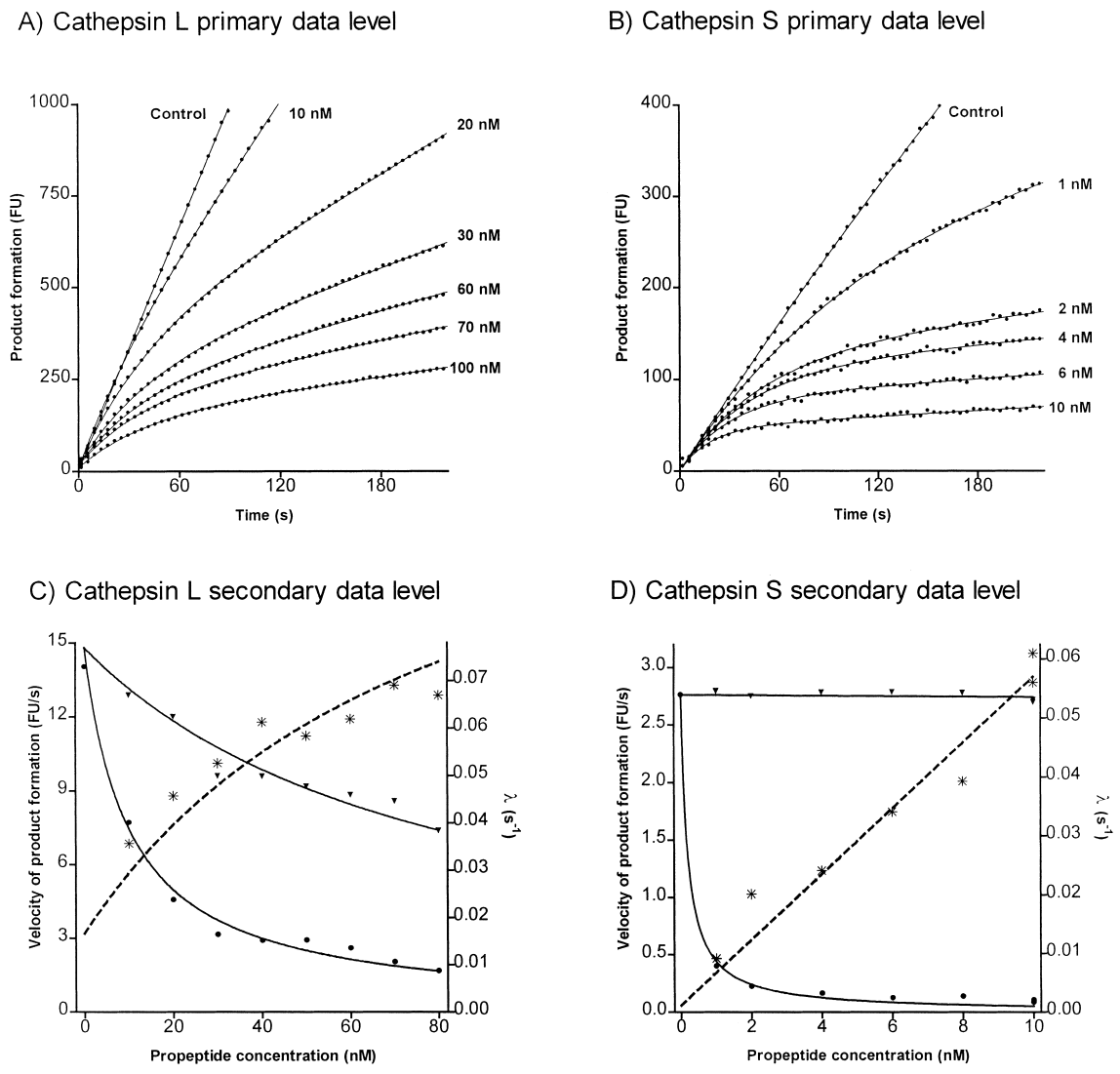

Fig. 2. Inhibition of the cathepsins L and S by their cognate propeptides. Measurements were performed at 37°C and pH 6.5 as specified in Section 2. Concentrations of the reactants were: (A) 0.25 nM CL, 10 μM *Z*-Phe-Arg-NH-Mec; (B) 0.05 nM CS, 40 μM *Z*-Val-Val-Arg-NH-Mec; and propeptides as indicated. The curves at the primary data level (upper part) correspond to $P = v_s t + (v_z - v_s)(1\text{*minus};e^{-t\lambda})/\lambda$ as fitted to the respective series of experimental data points. At the secondary data level (lower part), initial (K_i^*) and overall (K_i) dissociation constants were estimated by fitting the calculated initial (▼) and steady-state (●) velocities, respectively, to $v = V_{max}[S]/[K_m(1+[S]/K_m+[I]/K_d)]$ using the corresponding K_m and [S]. Results are given in Table 1. The left ordinate scales λ, the apparent first-order rate constant for the pre-steady-state exponential phase (*); the dotted lines correspond to $\lambda = f([I])$ using the right set of [S], K_m and K_d values for the one- and two-step mechanism, respectively: $\lambda = k_{off}(1+[I]K_m/(K_i(K_m+[S])))$ and $\lambda = k_{off}((1+([I]/(K_i(1+[S]/K_m))))/(1+([I]/(K_i^*(1+[S]/K_m)))))$ [18].

analogous studies using human cathepsin L ($K_i = 0.088$ nM) [10].

Similar inhibition experiments were carried out using ppCS in conjunction with CL and vice versa (primary data not shown, K_d values in Table 1). ppCS inhibited CL in a one-step reaction with about the same kinetics and affinity as its cognate enzyme ($K_i = 0.08$ nM). However, ppCL bound CS at least three orders of magnitude less effective than its cognate enzyme ($K_i > 150$ nM). Thus, the striking difference in inhibitory specificity makes these two propeptides ideally suited to examine the molecular basis of propeptide specificity in the CL-like protease subfamily.

Table 1
Inhibition of the cathepsins S and L by wild-type and chimeric propeptides, initial (*) and overall dissociation constants

Propeptide	Target enzyme	
	Cathepsin S (nM)	Cathepsin L (nM)
WT CS_{1-98}	0.049 ± 0.005	0.079 ± 0.009
WT CL_{1-86}	> 150	20.9 ± 1.5*, 2.6 ± 0.2
S_{1-54}/L_{44-86}	0.1 ± 0.006	1.4 ± 0.08
L_{1-43}/S_{55-98}	> 150	13.6 ± 0.7*
L_{1-43}/S_{53-98}	> 150	not determined

S.D. values as given by the non-linear regression program PRISM.

3.2. Inhibition of cathepsin S and L by L/S-propeptide chimeras

Elucidation of the tertiary structures of several cathepsin L-like proenzymes has demonstrated that all proregions fold to a compact mini-domain, whose core is defined by the intersection point of two longer helices, the N-terminal α1- and the central α2-helix. The C-terminus of the α2-helix establishes a hairpin-like structure with the adjacent short β1-strand, the

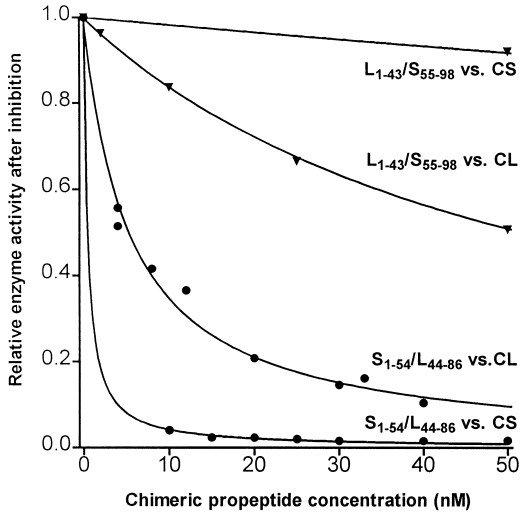

Fig. 3. Inhibition of *Paramecium* cathepsin L and human cathepsin S by the chimeric propeptides. Measuring conditions are the same as in Fig. 2. Initial (▼) or steady-state (●) velocities, determined from the time course of product formation, were used to fit K_d as explained in the legend to Fig. 2. Results are listed in Table 1.

super-secondary structure responsible for anchoring and positioning the proregion at the propeptide binding loop of the enzyme. A short stretch connects the β1-strand with the very short α3-helix preceding the extended C-terminus, sterically preventing substrate access to the active site cleft [20–23]. Although the α1/2 helices of the proregion do not interact with the enzyme directly they are essential for high-affinity enzyme binding, as shown for human cathepsin L in inhibition experiments with a series of truncated propeptides [10]. To investigate their importance for binding specificity, we constructed chimeras in which the α1/2 helices were swapped between ppCL and ppCS. Considering amino acid alignment and available structural information, the short band from α2 to β1 was chosen as the linker region to the C-terminal fragment of the proregion, which is in a direct contact to the enzyme. As the alignment of ppCL and ppCS, given in Fig. 1, revealed in this region two additional amino acids in the ppCS sequence, we constructed the two L/S chimeras, L_{1-43}/S_{55-98} and L_{1-43}/S_{53-98}.

Compared to wild-type ppCS, chimera S_{1-54}/L_{44-86} inhibited CS and CL with identical kinetics (primary data not shown) and only somewhat reduced efficiency, quantified by the increase of the K_i values from 0.05 to 0.1 nM and 0.08 to 1.4 nM, respectively (Fig. 3 and Table 1). Thus, the N-terminal fragment dominated the binding behaviour of S_{1-54}/L_{44-86}. Also chimera L_{1-43}/S_{55-98} behaved more like the donor of the N-terminus, since it inhibited CS only weakly ($K_i > 150$ nM) and CL with an initial K_i^* of 13.6 nM, i.e. in between the initial and steady-state values of CL inhibition by ppCL (Table 1). We were unable to establish a steady-state K_i value for this chimera because of a slow time-dependent loss of inhibitory potency. This could be explained by proteolytic degradation of the L/S chimera, analogous to the known pH-dependent ppCS cleavage by CL [12]. The L_{1-43}/S_{53-98} chimera with the prolonged linker region (additional -GM-), built to check the influence of the above mentioned gap in the alignment of the wild-type prosequences, did not inhibit CS with higher affinity than L_{1-43}/S_{55-98} (Table 1). Consequently, the lack of effective CS inhibition by both L/S chimeras could be attributed to the fused fragments itself and not to their impaired mutual orientation due to an inappropriate length of the linker region.

4. Conclusions

Our chimeric approach clearly indicates that in the cathepsin L-like subfamily the crossed α1/2-helices, establishing the backbone of the small prodomain, contribute decisively to the selectivity of propeptide–enzyme interactions. This result extends the finding by Carmona et al. [10], that N-terminal truncation of the primary structure of human cathepsin L propeptide abolishes high-affinity binding to the parental enzyme. Discovering the three-dimensional structure of the respective zymogen, Coulombe et al. [20] discussed the surprising fact that a part of the molecule without any direct contact to the target strongly determines the binding behaviour. They discussed that at least the α2-helix may act via conformational stabilisation of the adjacent structures, β1-stretch and α3-helix, considered to be the principal anchors towards the enzyme [24]. Recently, evidences for ordered, pH-dependent secondary and tertiary structures of the propeptides of the cathepsins L and S in the absence of the parental enzymes have been presented [7,25]. This suggests that, at least at about neutral pH, the α2-helix may not only stabilise but even determine the position of the two anchor sites. According to Groves et al. [24], comparing the X-ray data of procathepsin L and procaricain, the region flanked by the anchors establishes numerous specific contacts to the cognate enzymes, probably involved in specificity determination. Because these interactions clearly are distal to the substrate-binding cleft, the data may open new perspectives for inhibitor design, too.

References

[1] Smith, S.M. and Gottesman, M.M. (1989) J. Biol. Chem. 264, 20487–20495.
[2] Tao, K., Stearns, N.A., Dong, J., Wu, Q.L. and Sahagian, G.G. (1994) Arch. Biochem. Biophys. 311, 19–27.
[3] Nomura, T. and Fujisawa, Y. (1997) Biochem. Biophys. Res. Commun. 230, 143–146.
[4] McIntyre, G.F., Godbold, G.D. and Erickson, A.H. (1994) J. Biol. Chem. 269, 567–572.
[5] Mach, L., Mort, J.S. and Glossl, J. (1994) J. Biol. Chem. 269, 13030–13035.
[6] Menard, R., Carmona, E., Takebe, S., Dufour, E., Plouffe, C., Mason, P. and Mort, J.S. (1998) J. Biol. Chem. 273, 4478–4484.
[7] Jerala, R., Zerovnik, E., Kidric, J. and Turk, V. (1998) J. Biol. Chem. 273, 11498–11504.
[8] Fox, T., de Miguel, E., Mort, J.S. and Storer, A.C. (1992) Biochemistry 31, 12571–12576.
[9] Taylor, M.A., Briggs, G.S., Baker, K.C., Cummings, N.J., Pratt, K.A., Freedman, R.B. and Goodenough, P.W. (1995) Biochem. Soc. Trans. 23, 80S.
[10] Carmona, E., Dufour, E., Plouffe, C., Takebe, S., Mason, P., Mort, J.S. and Menard, R. (1996) Biochemistry 35, 8149–8157.
[11] Völkel, H., Kurz, U., Linder, J., Klumpp, S., Gnau, V., Jung, G. and Schultz, J.E. (1996) Eur. J. Biochem. 238, 198–206.
[12] Maubach, G., Schilling, K., Rommerskirch, W., Wenz, I., Schultz, J.E., Weber, E. and Wiederanders, B. (1997) Eur. J. Biochem. 250, 745–750.
[13] Visal, S., Taylor, M.A. and Michaud, D. (1998) FEBS Lett. 434, 401–405.
[14] McGrath, M.E., Palmer, J.T., Brömme, D. and Somoza, J.R. (1998) Protein Sci. 7, 1294–1302.

[15] Kirschke, H. and Wiederanders, B. (1994) Methods Enzymol. 244, 500–511.
[16] Wiederanders, B., Brömme, D., Kirschke, H., von Figura, K., Schmidt, B. and Peters, C. (1992) J. Biol. Chem. 267, 13708–13713.
[17] Cha, S. (1975) Biochem. Pharmacol. 24, 2177–2185.
[18] Baici, A. (1998) Biol. Chem. 379, 1007–1018.
[19] Taylor, M.A., Baker, K.C., Briggs, G.S., Connerton, I.F., Cummings, N.J., Pratt, K.A., Revell, D.F., Freedman, R.B. and Goodenough, P.W. (1995) Protein Eng. 8, 59–62.
[20] Coulombe, R., Grochulski, P., Sivaraman, J., Menard, R., Mort, J.S. and Cygler, M. (1996) EMBO J. 15, 5492–5503.
[21] Groves, M.R., Taylor, M.A., Scott, M., Cummings, N.J., Pickersgill, R.W. and Jenkins, J.A. (1996) Structure 4, 1193–1203.
[22] LaLonde, J.M., Zhao, B., Janson, C.A., D'Alessio, K.J., McQueney, M.S., Orsini, M.J., Debouck, C.M. and Smith, W.W. (1999) Biochemistry 38, 862–869.
[23] Sivaraman, J., Lalumiere, M., Menard, R. and Cygler, M. (1999) Protein Sci. 8, 283–290.
[24] Groves, M.R., Coulombe, R., Jenkins, J. and Cygler, M. (1998) Proteins 32, 504–514.
[25] Kreusch, S., Fehn, M., Maubach, G., Nissler, K., Rommerskirch, W., Schilling, K., Weber, E., Wenz, I. and Wiederanders, B. (1999) submitted for publication.