Structure and Mechanism of an Aspartimide-Dependent Peptide Ligase in Human Legumain**

Elfriede Dall, Julia C. Fegg, Peter Briza, and Hans Brandstetter*

Abstract: Peptide ligases expand the repertoire of genetically encoded protein architectures by synthesizing new peptide bonds, energetically driven by ATP or NTPs. Here, we report the discovery of a genuine ligase activity in human legumain (AEP) which has important roles in immunity and tumor progression that were believed to be due to its established cysteine protease activity. Defying dogma, the ligase reaction is independent of the catalytic cysteine but exploits an endogenous energy reservoir that results from the conversion of a conserved aspartate to a metastable aspartimide. Legumain’s dual protease–ligase activities are pH- and thus localization controlled, dominating at acidic and neutral pH, respectively. Their relevance includes reversible on–off switching of cystatin inhibitors and enzyme (in)activation, and may affect the progression that were believed to be due to its established immunological processes such as antigen processing and TLR maturation[1] as well as in tumor progression.[2] The cysteine protease legumain can develop asparaginyl endopeptidase (AEP) and carboxypeptidase (ACP) activities in a pH- and context-dependent manner.[3] These diverse roles call for a delicate regulation of its protease activity which is in part conferred by type 2 cystatins C and F, and the most potent cystatin E/M.[4] When studying the mechanistic basis of the AEP inhibition by cystatin E/M, we found a peptide ligase activity in human AEP, apparently independent of ATP. Whereas bond-conserving modifications, such as those seen in inteins and sortases, are energetically balanced,[6] genuine peptide ligations require a source of energy.[6] Ligase and cyclase activities have been reported for plant legumains,[5,b,7] but the mechanism of action remained unclear. These puzzling observations prompted us to crystallize human cystatin E (hCE) alone and in complex with legumain.

The crystal structure of hCE revealed the typical five-stranded antiparallel β-sheet wrapped around a central α-helix. Two disulfide bridges additionally stabilize the structure by clamping strands β4 and β5 and the hCE-specific appending structure that is inserted between strands β3 and β4 (Figure 1a; Figure S1 and Table S1 in the Supporting Information). Parts of the appendix structure (Trh76–His82) were flexible in the electron density map.

The interaction with papain-like proteases has been described by the so-called elephant trunk model,[3] involving cystatin’s N-terminus (“the trunk”) and two characteristic loops L1 and L2. hCE loop L2 deviated considerably from the conformations seen in cystatins C and F (hCC and hCF) (PDB entries 3gax and 2h9, respectively; Figure 1b). This deviation was centered around glyco-Asn108I and resulted in a frameshift in the segment Pro105I–Met110I in hCE relative to the conformation in hCC and hCF (cystatin C numbering with subscript I is used for cystatin inhibitors). We compared the affinity of glycosylated hCE and non-glycosylated (E. coli produced) hCE towards cathepsin B and found that the affinity of cathepsin B towards glyco-hCE (IC50 = 5.7 ± 1.3 nm) is two times higher than towards E. coli hCE (IC50 = 9.3 ± 2.4 nm).

The legumain reactive center loop (RCL) of hCE exposed Asn39, in a conformation similar to that seen in hCC and hCF (Figure 1a and Figure S1c), consistent with its suggested role for active site (P1–S1) interaction.[5,b] The conserved conformation of the RCL suggests a canonical, substrate-like binding mode that is shared within the type 2 cystatin family; differences in binding affinity within hCC/E/F should be related to exosite interactions. We considered Cys73–Cys83 as one candidate for a legumain exosite loop (LEL) because it was stabilized relative to the reactive center loop (RCL) via charged interactions mediated by Lys75I to the P2 (Ser38I) and P1’ (Ser40I) carbonyls of the RCL (Figure 1a and Figure S1c). The legumain reactive center loop (RCL) of hCE exposed Asn39, in a conformation similar to that seen in hCC and hCF (Figure 1a and Figure S1c), consistent with its suggested role for active site (P1–S1) interaction.[5,b] The conserved conformation of the RCL suggests a canonical, substrate-like binding mode that is shared within the type 2 cystatin family; differences in binding affinity within hCC/E/F should be related to exosite interactions. We considered Cys73–Cys83 as one candidate for a legumain exosite loop (LEL) because it was stabilized relative to the reactive center loop (RCL) via charged interactions mediated by Lys75I to the P2 (Ser38I) and P1’ (Ser40I) carbonyls of the RCL (Figure 1a and Figure S1c).

In the cocystal structure of the legumain–hCE complex we indeed found both the RCL and the LEL to contribute to AEP binding, utilizing substantial AEP contact areas that were previously found in promodulin binding within prolegumain (Figure 1b and Figure S2a,c).[5] This observation pro-
voked the question whether the structural mimicry between the AEP–cystatin complex and prolegumain would reflect functional analogies. One striking property of prolegumain is its stability at neutral pH, whereas isolated AEP becomes irreversibly denatured at pH > 6.\[9\]

Indeed, when complexed with either hCC or hCE, AEP remained stable at neutral pH, as shown by differential scanning fluorometry measurements (Figure S2b). The stabilization of AEP upon complexation with cystatin could also be monitored by its enzymatic activity towards a chromogenic substrate: Whereas isolated AEP became rapidly and irreversibly inactivated at pH 6.5, the preceding complex formation of AEP with hCC resulted in a basal AEP activity (Figure S2c). This cystatin-induced activity is easily explained by the continuous dissociation of AEP from the stabilizing hCC–AEP complex, thus accounting for the observed agonistic activity. A further shift from pH 6.5 to pH ≤ 4 led to quantitative dissociation of the AEP–hCC complex, accompanied by the recovery of approximately 80% of the initial activity (Figure S2c).

This pH profile was specific to hCC (Figure 2a) and can be explained by the different chemical nature of the exosite loop in hCE and hCC: While the hCE legumain exosite loop (LEL) is overall hydrophobic, the hCC LEL carries characteristic charged residues, e.g., Arg70I and Arg93 I which form salt bridges with Glu190. Consistently, charge reversal by an E190K mutation abolished the AEP–hCC, but not the AEP–hCE complex (Figure S2d).

We found the hCE RCL to bind to the active site in a substrate-like (“canonical”) manner, with the Asn391 fully inserted into AEP’s S1 recognition site and the scissile peptide bond intact in the electron density map (Figure 1c and Figure S2e). The geometry of the AEP active site as well as that of the hCE RCL were virtually identical in the AEP–hCE complex and in the structures of the isolated protein components, with the only exception being the catalytic Cys189 and Glu190: The thiol of Cys189 was rotated by...
and formed a zwitterionic pair with the carboxylate of Glu190, representing a resting protease state (Figure 1c).

Canonical inhibitors often act as slowly converting substrates.[10] Therefore, we incubated AEP with a twofold molar excess of hCE and hCC at pH 4.0/5.0 for 2 h and analyzed the sample by gel filtration chromatography (Figure 2b and Figure S3a). The corresponding elution profile was bimodal with the heterodimeric AEP–hCE complex followed by the monomeric (excess) hCE. Interestingly, partly cleaved hCE was eluted in both peaks (Figure 2b). Mass spectrometry showed that the cystatin cleavage occurred after Asn39I, consistent with the proposed canonical binding mode.

Given the substrate-like binding mode of cystatin, one should expect a continuous accumulation of cleaved cystatin. Intriguingly, this was not observed. Instead, the ratio of cleaved versus intact cystatin remained largely constant over time. To better understand this puzzling behavior, we analyzed the pH dependency of hCE processing by AEP (Figure 2c). The cleavage was favored at more acidic pH, with complete cleavage at pH 4. This pH dependence is in conflict with the known activity profile of AEP which shows highest proteolytic activity at pH 5.5.[15] Furthermore, the processing of hCC and hCE was in marked contrast to that of family 1 cystatins (stefins A and B) where processing was progressive and occurred preferentially at Asn107I and Asn61I (stefin B only), and resulted in rapid degradation (Figure S3b).

In an attempt to reconcile the puzzling observations, we hypothesized that AEP could catalyze a peptide bond ligation in addition to the peptide bond hydrolysis. Thus, the incomplete cleavage of hCC/E (Figure 2b,c and Figure S3a) would result as the pH-dependent equilibrium of two opposing reactions, with the protease activity and the ligase activity prevailing at pH 4 and pH 6, respectively.

We tested this hypothesis by preincubation of hCE/C with AEP at pH 4.0/5.0, resulting in fully cleaved hCE[1] (or hCC[1]). When the hCE[1] (hCC[1])–AEP complex was incubated at different pH values, the corresponding equilibria of cleaved hCE[1] (hCC[1]) and religated hCE (hCC) were established. Peptide bond resynthesis by AEP was most efficient at pH ≥ 6, resulting in equimolar amounts of religated hCE (Figure 3a and Figure S4a,b).

The AEP–hCE crystal structure was determined at pH 6.5 and should thus represent the ligase state of AEP. Hereby, the thiol (SH) of the catalytic Cys189 is rotated away from the hCE peptide bond (Figure 1c), suggesting that it is not directly involved in the ligation reaction. To investigate its role in ligation, we oxidized the Cys189 by adding S-methyl methanethiosulphonate (MMTS), resulting in the mixed disulfide Cys189-S-CH₃, herein referred to as oxidized AEP, AEPox. The modification was confirmed by crystallography (Figure S5a) and expectedly suppressed the AEP protease activity. Upon incubation of AEPox with cleaved hCE[1], we observed the resynthesis to intact hCE at pH 4.0 to pH 6.5 (Figure 3a and Figure S4a,b). This experiment confirmed that 1) the ligase and protease activities are superimposed, 2) the ligase activity of AEP is independent of Cys189, and as a consequence 3) the peptide ligation reaction is mechanistically not the reverse of protease catalysis via a thioester intermediate,[11] but must instead follow a distinct reaction mechanism. As a consequence, AEPox cannot employ an intein- or sortase-like mechanism for peptide bond ligation which would be bond and energy conserving.

Genuine peptide bond synthesis requires coupling to an energy-rich reagent that typically activates the carboxylic acid of the P1 residue.[12] In search for a suitable coupling reagent, we noted an aspartimide (succinimide) at position 147...
Proposed reaction scheme of the legumain ligase activity. The $\text{O}^-$ of the new C-terminus of processed hCE (Asn39I) attacks the energy-rich Suc147, resulting in an activated carboxylic acid at Asn39I (a). Next, the carbamoyl nitrogen of the P1-Asn39I side chain will attack the electrophilic carbon of the carboxylic anhydride (b), thereby generating Suc39I (c). The succinimide ring can now be opened via a nucleophilic substitution by the P1$'$ nitrogen, resulting in the intact P1–P1$'$ peptide bond (d). Additionally, the Suc147 may be regenerated via a condensation reaction.
Keywords: endergonic reactions · enzyme catalysis · hydrolysis · ligases · protein modification

[1] a) C. B. Chan, M. Abe, N. Hashimoto, C. Hao, I. R. Williams, X. Liu, S. Nakao, A. Yamamoto, C. Zheng, J. J. Henter, M. Meeths, M. Nordenskjöld, S. Y. Li, I. Hara-Nishimura, M. Asano, K. Ye, Proc. Natl. Acad. Sci. USA 2009, 106, 468–473; b) S. E. Ewald, A. Engel, J. Lee, M. Q. Wang, M. Bogoy, G. M. Barton, J. Exp. Med. 2011, 208, 643–651; c) S. Maschaidi, S. Hassler, E. Fantin, F. E. Sepulveda, M. Tohme, M. Chignard, P. van Endert, M. Si-Tahar, D. Descamps, B. Manoury, PLoS Pathog. 2012, 8, e1002841; d) K. Wolk, G. Grutz, K. Witte, H. D. Volk, R. Sabat, Genes Immun. 2005, 6, 452–456.

[2] a) M. H. Haugen, H. T. Johansen, S. J. Pettersen, R. Solberg, K. Brix, K. Flatmark, G. M. Maelandsmo, PLoS One 2013, 8, e52980; b) Y. Ohno, J. Nakashima, M. Izumi, M. Ohori, T. Hashimoto, M. Tachibana, World J. Urol. 2013, 31, 359–364; c) L. Wang, S. Chen, M. Zhang, N. Li, Y. Chen, W. Su, Y. Liu, D. Lu, S. Li, Y. Yang, Z. Li, D. Stupack, P. Qu, H. Hu, R. Xiang, J. Cell. Biochem. 2012, 113, 2679–2686.

[3] E. Dall, H. Brandstetter, Proc. Natl. Acad. Sci. USA 2013, 110, 10940–10945.

[4] a) M. Alvarez-Fernandez, A. J. Barrett, B. Gerhartz, P. M. Dando, J. Ni, M. Abrahamson, J. Biol. Chem. 1999, 274, 19195–19203; b) J. J. Briggs, M. H. Haugen, H. T. Johansen, A. I. Riker, M. Abrahamson, O. Fodstad, G. M. Maelandsmo, R. Solberg, BMC Cancer 2010, 10, 17; c) J. A. Hardy, J. Lam, J. T. Nguyen, T. O’Brien, J. A. Wells, Proc. Natl. Acad. Sci. USA 2004, 101, 12461–12466; d) M. Donepudi, A. Mac Sweeney, C. Briand, M. G. Grutter, Mol. Cell 2003, 11, 543–549.

[5] a) F. B. Perler, N. M. Allewell, J. Biol. Chem. 2014, 289, 14488–14489; b) S. K. Mazmanian, G. Liu, T. T. Hung, O. Schneewind, Science 1999, 285, 760–763.

[6] K. Tanaka, T. Suzuki, T. Chiba, Mol. Cells 1998, 8, 503–512.

[7] a) I. Hara-Nishimura, K. Inoue, M. Nishimura, FEBS Lett. 1991, 294, 89–93; b) G. Basurto-Islas, I. Grundke-Iqbal, Y. C. Tung, F. Liu, K. Iqbal, J. Biol. Chem. 2011, 288, 17495–17507.

[8] W. Machleidt, U. Thiele, B. Laber, I. Assfalg-Machleidt, A. Escherl, G. Wiegand, J. Kos, V. Turk, W. Bode, FEBS Lett. 1989, 247, 234–238.

[9] E. Dall, H. Brandstetter, Acta Crystallogr. Sect. F 2012, 68, 24–31.

[10] a) W. Bode, R. Huber, Biochim. Biophys. Acta Protein Struct. Mol. Enzymol. 2000, 1477, 241–252; b) O. D. Ekici, M. G. Gotz, K. E. James, Z. Z. Li, B. J. Rukamp, J. L. Asgian, C. R. Caffrey, E. Hansell, J. Dvorak, J. H. McKerrow, J. Potemja, T. Travis, J. Mikolajczyk, G. S. Salvesen, J. C. Powers, J. Med. Chem. 2004, 47, 1889–1892.

[11] a) D. Y. Jackson, J. Burnier, C. Quan, M. Stanley, J. Tom, J. A. Wells, Science 1994, 266, 243–247; b) L. Abrahamsen, J. Tom, J. Burnier, K. A. Butcher, A. Kossiakoff, J. A. Wells, Biochemistry 1991, 30, 4151–4159.

[12] V. R. Pattabiraman, J. W. Bode, Nature 2011, 480, 471–479.

[13] P. Fuentes-Prior, G. S. Salvesen, Biochem. J. 2004, 384, 201–232.

[14] a) L. Stryer, Biochemistry, Vol. 4, Spektrum, Heidelberg, 2003; b) K. Tabata, H. Ikeda, S. Hashimoto, J. Bacteriol. 2005, 187, 5195–5202.

[15] N. D. Rawlings, A. J. Barrett, A. Bateman, J. Biol. Chem. 2011, 286, 38321–38328.

[16] L. Zhao, T. Hua, C. Crowley, H. Ru, X. Ni, N. Shaw, L. Jiao, W. Ding, L. Ou, L. Wu, H. Huang, L. Liu, K. Ye, S. Ouyang, G. Cheng, Z. J. Liu, Cell Res. 2014, 24, 334–358.

[17] a) D. J. Craik, U. Malik, Curr. Opin. Chem. Biol. 2013, 17, 546–554; b) T. Kondo, Y. Higashiyama, S. Goto, T. Iida, S. Cho, M. Iwanaga, K. Morii, M. Tani, Y. Urata, Free Radical Res. 1999, 31, 325–334.

[18] a) B. Manoury, E. W. Hewitt, N. Morrice, P. M. Dando, A. J. Barrett, C. Watts, Nature 1998, 396, 695–699; b) B. Manoury, D. Mazzuco, D. N. Li, J. Billson, K. Loak, P. Benaroch, C. Watts, Immunity 2003, 18, 489–498.