An enigmatic peptide ligation reaction: Protease-catalyzed oligomerization of a native protein segment in neat aqueous solution

SANGARALINGAM KUMARAN, DEBJANI DATTA, AND RAJENDRA P. ROY

National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110067, India

(RECEIVED October 26, 1999; FINAL REVISION January 20, 2000; ACCEPTED January 21, 2000)

Abstract

We report an enigmatic peptide ligation reaction catalyzed by Glu-specific *Staphylococcus aureus* V8 protease that occurs in neat aqueous solution around neutral pH utilizing a totally unprotected peptide substrate containing free α -carboxyl and α -amino groups. V8 protease catalyzed a chain of ligation steps between pH 6 and 8 at 4 °C, producing a gamut of covalent oligomers (dimer through octamer or higher) of a native protein segment TAAAKFE (S39) derived from ribonuclease A (RNase A). Size-exclusion chromatography suggested the absence of strong interaction between the reacting peptides. The circular dichroism spectra of monomer through pentamer showed length-dependent enhancement of secondary structure in the oligomers, suggesting that protease-catalyzed ligation of a monomer to an oligomer resulted in a product that was more structured than its precursor. The relative conformational stability of the oligomers was reflected in their ability to resist proteolysis, indicating that the oligomerization reaction was facilitated as a consequence of the “conformational trapping” of the product. The ligation reaction proceeded in two phases—slow formation and accumulation of the dimer followed by a fast phase of oligomerization, implying that the conformational trap encountered in the oligomerization reaction was a two-step process. The Gly substitution at any position of the TAAAKFE sequence was deleterious, suggesting that the first step of the conformational trap, namely the dimerization reaction, that proceeded very slowly even with the parent peptide, was quite sensitive to amino acid sequence. In contrast, the oligomerization reaction of an Ala analog, AAAAKFE, occurred in much the same way as S39, albeit with faster rate, suggesting that Ala substitution stabilized the overall conformational trapping process. The results suggest the viability of the product-directed “conformational trap” as a mechanism to achieve peptide ligation of totally unprotected peptide fragments in neat aqueous solution. Further, the study projects the presence of considerable innate synthetic potential in V8 protease, baring rich possibilities of protein engineering of this enzyme to generate a “V8 peptide ligase.”

Keywords: peptide ligation; peptide oligomerization; proteosynthesis; reverse proteolysis; V8 protease

The hydrolysis equilibria of a protease-catalyzed reaction may be reversed to peptide bond synthesis in accord with the principle of microscopic reversibility of reactions. Although the reverse proteolytic condensation of unprotected native protein segments under physiological conditions is thermodynamically unfavorable, the protease-catalyzed peptide bond equilibria is shifted toward synthesis in the presence of organic cosolvents (Homandberg et al., 1978) and facilitated in situations where the reacting termini are brought into stereochemical proximity through noncovalent inter-

actions (Homandberg & Laskowski, 1979; Homandberg & Chaiken, 1980; Komoriya et al., 1980; Chaiken, 1981; Graf & Li, 1981; Wallace, 1995). The enthalpic barrier to peptide bond synthesis is lowered by cosolvent-mediated enhancement of carboxyl pK_a , while the proximity effect helps overcome the entropic barrier. Furthermore, a protease-catalyzed reaction is also driven to synthesis with the use of a “molecular trap.” In such cases, “trapping” of the product shifts the equilibrium to synthesis due to continuous removal of the spliced material from the chemical equilibrium (Homandberg et al., 1982; Nyberg, 1988; Roy et al., 1992; Kumaran et al., 1997; Sahni et al., 1998).

The catalysis of amide bond formation by proteases proceeds as well in neat aqueous solution, but only with peptide esters as substrate. The systematic protein engineering of subtilisin by chemical modification, site-directed mutagenesis, and phage-display technology has yielded a variant enzyme, subtiligases, that can efficiently

Reprint requests to: Rajendra P. Roy, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110067, India; e-mail: rproy@nii.res.in.

Abbreviations: CD, circular dichroism; ESMS, electrospray mass spectrometry; RNase A, ribonuclease A; RP-HPLC, reversed-phase high-performance liquid chromatography; S39, residues 3–9 (TAAAKFE) of RNase A; S-peptide, sequence representing the first 20 residues of RNase A.

catalyze peptide ligation in aqueous solution (Nakatsuka et al., 1987; Wu & Hilvert, 1989; Abrahmsen et al., 1991; Chang et al., 1994; Braisted et al., 1997; Atwell & Wells, 1999). The total synthesis of ribonuclease A (RNase A) by stepwise condensation of appropriate protein segments closely matching the substrate specificity of subtiligase has been accomplished (Jackson et al., 1994). It turns out that besides the requirement of protease specificity (primary as well as subsite specificity), activation of the carboxyl group by protonation or esterification, and the proximity of the peptide reacting ends seem to be essential elements for catalysis of peptide bond formation between native peptide segments by proteases.

We report herein an enigmatic and a novel protease-catalyzed ligation reaction that proceeds around neutral pH in neat aqueous solutions but without the benefit of carboxyl group activation and strong interaction between the reacting peptides. V8 protease catalyzed a chain of peptide ligations between pH 6 and 8 when provided with a free α -carboxyl group and free α -amino group containing a heptapeptide (TAAAKFE) segment derived from RNase A and produced a gamut of oligomers (dimer through octamer) of the heptapeptide.

Results

V8 protease-catalyzed peptide ligation in aqueous buffer

We were interested in generating peptides that would be suitable for calibrating columns for size-exclusion chromatography of small peptides. Toward this, we wished to fragment S-peptide (N-terminal fragment spanning residues 1–20) of RNase A (Richards & Vithayathil, 1959) by *Staphylococcus aureus* V8 protease. S-peptide contains two Glu residues at positions 2 and 9, respectively (Fig. 1). The Glu-specific V8 protease is expected to generate three peptides. Consistent with this, RP-HPLC analysis of a V8 protease digest (pH 8, 37 °C) of S-peptide showed only three peaks. These peaks were identified by ESMS. The mass data (275.19, 1,189.29, and 736.4 Da) were in agreement with the peptide compositions corresponding to residues 1–2, 10–20, and 3–9, respectively. However, an aliquot of the same V8 protease digest of S-peptide, which was stored at 4 °C for about 144 h or more, when analyzed by HPLC displayed a curious chromatographic profile; a major abnormal peak and traces of minor components were seen in addition to the above three peptides. The major peak when subjected to ESMS analysis yielded a mass of 1,454.7 Da that could be fitted to the composition of a dimer of residues 3–9 of S-peptide. Ten cycles of Edman degradation provided the sequence TAAAKFETAA, confirming the presence of Glu-Thr connectivity establishing the covalent nature of the dimer. Furthermore, a synthetic 14-mer peptide TAAAKFETAAAKFE showed identical RP-HPLC retention time as that of the material isolated from the V8 protease digest.

Fig. 1. Sequence of the S-peptide. The N-terminal fragment of RNase A spanning residues 1–20 and referred to as S-peptide is illustrated. S-peptide contains two Glu residues (shown by arrows) at positions 2 and 9, respectively, which are V8 protease cleavage sites.

The oligomerization reaction

To further probe that the dimer and other components present in the HPLC profile were formed as a result of V8 protease-catalyzed ligation reactions, we generated substantial quantities of S39 by chemical synthesis. A solution of the purified peptide prepared in ammonium bicarbonate buffer (pH 8) was incubated with and without V8 protease at 4 °C. The RP-HPLC analysis of an aliquot of the sample that did not contain the protease revealed the presence of only the starting peptide S39 after 240 h of incubation. In contrast, the sample incubated with V8 protease yielded a complex product profile; nearly 85–90% of S39 was converted into dimer and other chromatographically distinct components, as shown in Figure 2A. These peaks were collected separately and subjected to ESMS. From the mass data summarized in Table 1, peaks A–H of

Fig. 2. Oligomerization reaction. **A:** The ligation reaction of peptide S39 in the presence of V8 protease was carried out in aqueous ammonium bicarbonate (pH 8) or ammonium acetate (pH 6) buffer as described in Materials and methods. An aliquot of the reaction mixture after 240 h was analyzed by RP-HPLC on an Aquapore RP300 column (4.6 × 250 mm, 7 μ M) using an acetonitrile–TFA–water solvent system (4 to 56% acetonitrile containing 0.1% TFA in 130 min at a flow rate of 0.7 mL/min). **B:** Time course of the ligation reaction (pH 8). Aliquots of the reaction mixture were analyzed at various time points by RP-HPLC on a Zorbax C8 column (9.4 × 250 mm, 5 μ M) employing the same gradient as above but at a flow rate of 1.5 mL/min.

Table 1. Calculated and experimental mass of peptides^a obtained from protease-mediated ligation reaction

HPLC peak	Experimental mass	Calculated mass	Oligomeric status
A	736.40 ± 0.00	736.82	Monomer
B	1,454.70 ± 0.00	1,455.63	Dimer
C	2,175.13 ± 0.57	2,174.63	Trimer
D	2,893.10 ± 0.26	2,893.44	Tetramer
E	3,612.56 ± 0.34	3,612.05	Pentamer
F	4,330.76 ± 0.95	4,330.86	Hexamer
G	5,050.10 ± 0.57	5,049.67	Heptamer
H	5,768.70 ± 0.51	5,768.47	Octamer

^aThe mass of peptides (expressed in Da) were obtained by ESMS and are computed based on natural isotopic abundance.

Figure 2A could be readily assigned as monomer (A) through the octamer (H) of S39.

The oligomerization reaction described above occurred at pH 8. However, considering that the lowering the pH of the reaction would facilitate the transfer of proton from the protonated amino group to the carboxylate ion to generate the unionized substrates required for peptide ligation, the ligation reaction was carried out in ammonium acetate buffer (pH 6) at 4 °C. Interestingly, HPLC analysis of the reaction mixture exhibited a similar profile as that obtained when the reaction was carried out at pH 8.

Kinetics of oligomerization reaction

The time kinetics of ligation reaction (Fig. 2B) revealed that dimer was formed slowly and accumulated for about 144 h or more until the appearance of the trimer. The formation of higher oligomers occurred at a considerably faster rate than the dimerization reaction, suggesting that the ligation reaction occurred in two phases—a slow phase relating to the dimer formation, followed by a relatively fast phase of oligomerization. The propagation of the ligation reaction to the second phase was noticeable only after about 50% of the monomer had been converted into dimer.

Status of noncovalent association

To examine the role of stereochemical proximity of amino and carboxyl reacting ends that might emanate from noncovalent association of the peptides and facilitate the catalysis of an oligomerization reaction, the possible interactions between starting reactant S39 and oligomers formed early in the reaction pathway were investigated by gel-permeation chromatography (data not shown). The peptide S39 eluted as a single component at a position distinct from that of the covalent dimer when chromatographed on a size-exclusion column in the ligation reaction buffer. Also, co-mixtures of the monomer and trimer of peptide S39 eluted as two components at their respective retention times. The trimer eluted later than the tetramer. Further, the retention time of the trimer (21-mer peptide) exactly matched with the retention of S-peptide (20-mer peptide) that was chromatographed under identical conditions, corroborating the absence of a 1:1 complex of monomer and trimer or higher molecular components in the chromatographic profile. Taken together, the results suggested the absence of a

strong interaction between the reacting peptides, and that enzymatic coupling occurs even if the peptides do not interact to form a rather stable complex. Nonetheless, given the resolution between the monomer and trimer in the gel filtration experiment, the possible presence of 10 to 25% sample as a dimer in the equilibrium with the monomer cannot be ruled out.

Conformational features of the oligomers

The formation of larger oligomers of S39 was intriguing because the increase in chain length is also accompanied by the introduction of additional sites for proteolytic action that would oppose chain elongation. Importantly, addition of a S39 monomer unit to the pre-existing chain will generate symmetric protease-specificity sites on the oligomer that are likely to be hydrolyzed with similar rates. However, differences in hydrolysis rates of the E–T peptide bond or the degree of resistance to proteolysis may arise due to length-dependent conformational properties of the oligomers. It is quite conceivable that oligomers begin to adopt a more ordered conformation that enhances their ability to resist proteolysis. The use of the product-generated “conformational trap,” wherein a new equilibrium is established between the “folded” proteolytically resistant product and the starting peptides, has been found to drive protease-catalyzed ligation reactions in the aqueous-organic milieu. Although the V8 protease-catalyzed oligomerization of S39 proceeded in the absence of any cosolvent, the possibility that conformation of the oligomers may drive the proteosynthetic reaction in aqueous solution must be ascertained.

To see if the oligomerization reaction was a consequence of the contiguity-mediated generation of the conformation of peptides, we studied the circular dichroic behavior of the peptides ranging from monomer to pentamer (Fig. 3). The far-ultraviolet (UV) CD spectrum of the monomer was characterized by a positive band at 217 nm. The dimer exhibited two minima in the region of 235 and 206 nm in addition to a positive band around 217 nm. The spectrum of trimer was qualitatively similar to that of the dimer, albeit with a distinct minimum at 208 nm. In contrast, tetramer and pentamer spectra showed the presence of double minima at 208 and 228 nm.

The CD data of isolated oligomers described above suggested that chain elongation from the monomer to the pentamer by successive addition of a monomer unit induced conformational changes in the oligomers. Although the dimer and trimer did not display signature CD spectrum of a well-recognized secondary structure, the presence of a negative band in the region of 206–208 nm does suggest rudiments of helical conformation in these peptides. This is consistent with the fact that the TAAAKFE sequence assumes helical conformation in RNase A and, hence, should be endowed with helical propensity. The tetramer and pentamer were clearly structured and contained a significant amount of helical conformation. However, it is worth noting that the CD double minima in these spectra were found to be at 208 and 228 nm instead of 208 and 222 nm usually found for an α -helical conformation. The origin of one of the minima at 228 nm in the case of a tetramer and pentamer may be due to the presence of four or five phenylalanine residues in these oligomers. The contribution from aromatic side chains to near- and far-UV CD spectra of peptides and proteins have been recognized in theoretical as well as experimental studies (Woody, 1978; Manning & Woody, 1989).

The analyses of the above CD data revealed peptide length-dependent enhancement of secondary structure in the oligomers

Fig. 3. CD spectra of peptides. Far-UV CD spectra of peptides were recorded (pH 6 and 4°C, 10 mm cell) on a Jasco-710 instrument equipped with a Peltier-type constant temperature cell holder. The peptide concentration was $\sim 40 \mu\text{M}$.

suggesting that protease-catalyzed ligation of a monomer to an oligomer resulted in a product that was more structured than its precursor. This can be interpreted to mean that the oligomerization reaction was facilitated as a consequence of the “conformational trapping” of the product.

Proteolysis of isolated oligomers

To see if the length-dependent induction of helical conformation in oligomers was reflected in their abilities to resist hydrolysis by V8 protease, proteolysis experiments were performed on the isolated

Fig. 4. Time course of proteolysis of isolated oligomers. The V8 protease-catalyzed hydrolysis of respective oligomers was carried out in 50 mM ammonium bicarbonate buffer (pH 8). The respective peptides were incubated with the protease as described in Materials and methods. Aliquots of the reaction mixture were analyzed by HPLC as described in Figure 2A. The extent of proteolysis was quantitated from the peak areas calculated using a computer program.

trimer and pentamer. The proteolysis experiments with these oligomers were carried out under identical conditions, except that the concentration of the enzyme was normalized (see Materials and methods) to account for the presence of four scissile E-T bonds in the pentamer compared with only two such bonds in the trimer.

The kinetics of proteolysis was monitored (pH 6, 4°C) by RP-HPLC. The time course of proteolysis of the trimer and pentamer with reference to the disappearance of the parent oligomer is depicted in Figure 4. The rate of pentamer proteolysis ($t_{1/2} = 45 \text{ min}$) was considerably slower than the proteolysis of the trimer ($t_{1/2} = 26 \text{ min}$). The HPLC profiles revealed that both the trimer and pentamer were proteolyzed to lower molecular weight oligomers during the early course of the reaction. The trimer yielded equal amounts of monomer and dimer until a 30 min time point, subsequent to which monomer predominated. The trimer was quantitatively converted to monomer in $\sim 4 \text{ h}$. The pentamer proteolysis produced oligomers ranging from the tetramer to the monomer in the initial period of 1 h, with the tetramer as the major product. However, HPLC analysis of the samples that were digested for 2 h or more showed only the presence of the monomer and pentamer. The samples analyzed between 3 and 24 h of digestion suggested that nearly 35 to 40% of the pentamer was resistant to proteolysis. These results imply that an increase in chain length enhanced the propensity of oligomers to resist proteolysis and thus facilitated the oligomerization reaction.

Protease-catalyzed ligation reactions of Gly analogs of S39

We studied the oligomerization reaction of Gly analogs of S39 with the idea that the presence of Gly might interfere with the generation of a product-directed conformational trap and, hence, affect the V8 protease-catalyzed ligation reaction. For this purpose, six synthetic analogs of S39 containing Gly substitutions at each

Table 2. Influence of amino acid sequence on the V8 protease-catalyzed peptide oligomerization reaction^a

Peptide	Sequence	Product profile	Yield (%)
S39	TAAAKFE	Dimer-octamer	85
S3G	GAAAKFE	Nil	0
S4G	TGAAKFE	Dimer	5
S5G	TAGAKFE	Dimer-pentamer	30
S6G	TAAGKFE	Dimer	15
S7G	TAAAGFE	Nil	0
S8G	TAAAKGE	Nil	0

^aThe V8 protease-catalyzed ligation reaction of each heptapeptide substrate (Gly analogs of S39) was carried out under similar conditions as that employed for S39 (see Materials and methods). Aliquots of the respective reaction mixture were analyzed after 240 h by RP-HPLC, and the extent of synthesis was quantified from the peak areas obtained through a computer program.

position except for the primary specificity determining site (Glu) were generated by solid-phase synthesis (Table 2). Interestingly, the ligation reaction did not occur when S3G, S7G, and S8G were used as substrates, while only the dimer was obtained with analogs S4G and S6G with yields of 5 and 15%, respectively. In contrast, analog S5G produced oligomers. However, the oligomer chain length did not extend beyond the pentamer (Fig. 5). Also, only

Fig. 5. V8 protease-mediated oligomerization of analog S5G. The protease-catalyzed ligation reaction of S5G (TAGAKFE) was carried out in ammonium acetate buffer (50 mM, pH 6) in a similar way as that described for S39. An aliquot of the reaction mixture was analyzed by RP-HPLC as described in Figure 2A. The numbers at the peak tops refer to the oligomeric status. The peptide peaks 1, 2, and 3 were characterized by ESMS as monomer (722.65 Da), dimer (1,427.45 Da), and trimer (2,132.30 Da), respectively. Peaks 4 and 5 were not characterized by ESMS. They are assumed to represent tetramer and pentamer by analogy with the oligomerization reaction of S39 described in Figure 2.

about 30% of the starting monomer S5G was converted into products within the reaction time of 240 h compared with 80% for the parent peptide S39. These results suggest that the conformational trap of the ligation reaction was sensitive to the presence of Gly residue.

Oligomerization of Ala analog, AAAAKFE

V8 protease-catalyzed oligomerization reaction of an Ala analog of S39 (S3A, AAAAKFE) was carried out to probe if the inhibitory effect of the Gly residue on peptide ligation was due to its deleterious effect on the generation of a product-directed conformational trap. Further, this would help clarify the chemical role of α -amino group or Thr side chain, if any, in the ligation reaction. The hydroxyl group in the Thr side chain might be involved in the splicing reaction, as it may generate an ester linkage that could participate in a transesterification reaction.

Under identical reaction conditions as that employed for S39, the oligomerization of S3A occurred in a similar fashion, albeit with slightly faster kinetics. S3A produced oligomers ranging from the dimer through the octamer or higher. Thus, Gly and Ala substitutions, respectively, for Thr in TAAAKFE have opposite effects on the generation of the conformational trap of the splicing reaction. These results are consistent with the helix forming propensities of Gly and Ala residues. The results also preclude any role of the α -amino group and hydroxyl side chain of Thr in the ligation reaction.

Discussion

The ability of proteases to catalyze the amide bond formation between totally unprotected native protein/peptide segments in the presence of organic cosolvents is well documented. However, catalysis of peptide bond formation between native protein/peptide segments by proteases in neat aqueous solution usually do not occur appreciably even in favorable situations entailing stereochemical proximity of the reacting ends, as in proteolytically nicked and thermodynamically stable noncovalent complexes of proteins, primarily because the protease-catalyzed peptide bond hydrolytic equilibria is not adequately shifted in the direction of ligation in 55.5 M water. Nonetheless, proteases catalyze the condensation of peptides in aqueous solution with the use of C-terminal active esters (Schellenberger & Jakubke, 1991; Wong & Wang, 1991; Braisted et al., 1997). Here we report that generation of product-directed conformational traps can act as a driving force for the protease-catalyzed splicing of native peptide segments in neat aqueous solutions, thereby obviating the necessity of active esters or organic cosolvents or strong interaction between the reacting peptides.

Together, the data of kinetics of oligomerization and the CD of isolated oligomers suggest that the "conformational trap" of the ligation reaction encountered in the oligomerization of the peptides is a two-step process; the first step leads to the slow formation of the dimer, while second step promotes the rapid oligomerization process. While the formation of helical conformation is clearly seen in the CD spectra of larger oligomers (tetramer and pentamer), the presence of significant amount of readily recognizable elements of secondary structure could not be ascertained from the spectra of the dimer and trimer. Nonetheless, the influence of increasing chain length on conformation is clearly noticeable. The slow formation of the dimer is presumably related to the availabil-

ity of fewer conformer population, which can potentiate the generation of a product-directed conformational trap. The kinetic data demonstrate that a buildup of the concentration of the dimeric species is required to facilitate the second phase of the ligation reaction. The energetic barrier for the generation of the trimer by ligation of the monomer to the dimer is possibly overcome by concentration of the dimer.

The Gly substitutions in the S39 peptide had a profound inhibitory effect on the ligation reaction. Previously it has been shown that peptide amide nucleophiles containing Gly residues at P1' or P2' positions can be efficiently transferred to an acyl donor Glu ester by V8 protease (Schuster et al., 1990), suggesting that the inability of Gly analogs of the TAAAKFE peptide to undergo ligation reaction were not related to the subsite specificity preferences of the enzyme. The abrogation or inhibition of the ligation reaction in the Gly analogs may be interpreted as that of the first step of the "conformational trap," namely, the formation of the dimer that proceeds very slowly even with the parent peptide is quite sensitive to the amino acid sequence. It is conceivable that the presence of Gly residue in analogs depopulate a potential conformational state that is present in TAAAKFETAAAKFE. Although conformational differences between dimers of Gly analogs and the parent peptide were not apparent at the level of CD spectra (data not shown), the stabilization of the conformational trap by an unusual (*i, i + 5*) Glu-Lys salt bridge remains a plausible possibility. The Gly residues may exert their negative influence on the generation of the first phase of the conformational trap by interfering with the stereochemistry of this salt bridge formation.

The rapidity of the oligomerization process observed in the second phase of the ligation reaction could be a consequence of helical or higher order structure of oligomers. Although helical conformation is described for tetramer and pentamer, it is likely that larger oligomers begin to adopt coiled-coil or helix bundle type of higher order structures. The coiled-coil structures are characterized by their primary sequence of heptad repeats (*abcdefg*), where positions *a* and *d* tend to be hydrophobic, and positions *e* and *g* are predominantly charged residues. An inspection of the heptapeptide sequence suggests that TAAAKFE fulfills this criterion at all positions except that Thr and Phe residues have low frequency of occurrence in heptad repeats (Cohen & Parry, 1990). The coiled-coil predicting program, MultiCoil (Wolf et al., 1997), suggested significant total coiled-coil potential (about 60%) to heptad repeats of TAAAKFE but not to GAAAKFE or TAGAKFE. Multicoil predictions of coiled-coil probability for heptad repeats of AAAAKFE (80%) was higher than that of the TAAAKFE repeats. The above predictions are consistent with the results of protease-catalyzed oligomerization of the heptapeptides.

The synthetic potential of V8 protease in organic cosolvents has been utilized in many studies concerning peptide ligation and semi-synthesis of proteins that include synthesis of fragments of thermolysin and α -globin, a large domain of thermolysin containing noncoded Aib residues and hemoglobins containing chimeric α -globins (Sahni et al., 1989; De Filippis & Fontana, 1990; Roy et al., 1993; Roy & Acharya, 1994; Kumaran et al., 1997; Natcharaju et al., 1997; De Filippis et al., 1998). There also exists an example of a V8 protease-mediated transpeptidation reaction in aqueous buffer, albeit the fragments were held in close proximity by disulfide linkage (Canova-Davis et al., 1991). Recently, replacement of the primary specificity site determining Glu residue by carboxymethyl thioester derivatives of other amino acids at the P1' site has been shown to act as a "substrate mimetic" in V8 protease-

catalyzed peptide ligation through acyl transfer reactions (Wehofskey & Bordusa, 1999). However, the peptide ligation reaction discussed here is distinct from the above-reported reactions in that it constitutes, to the best of our knowledge, the first example of a protease-catalyzed reaction synthesizing two to seven or more peptide bonds from a native monomeric heptapeptide substrate in aqueous solution. Protease-mediated ligations are generally limited to a single peptide bond formation between two peptide fragments in organic cosolvents except for triosephosphate isomerase where the enzyme was regenerated by formation of three specific peptide bonds by subtilisin-catalyzed reaction in aqueous-acetonitrile solutions (Vogel & Chmielewski, 1994; Ray et al., 1999). The formation of multiple peptide bonds is not favored, as the protease-mediated synthesized bonds are also sites of proteolytic attack. The successful synthesis of triosephosphate isomerase was attributed to the tendency of the fragments to remain self-assembled through noncovalent complexation.

The last decade has seen *in vivo* generation of many functional proteins through the phenomenon of protein splicing wherein an internal segment, intein, of the precursor protein excises itself and autocatalytically ligates the amino and carboxy terminus precursor fragments (Perler et al., 1997). However, *in vivo* splicing of concanavalin A does not follow the usual "self-catalyzed" protein splicing mechanism (Wang & Hugh Jones, 1994). The splicing of the precursor fragments of concanavalin A are catalyzed by an asparagine specific protease. Interestingly, the protease-catalyzed reaction turns around the sequence of the precursor protein in the matured concanavalin A by forming a peptide bond between the carboxyl group of the C-terminus and the amino group of the N-terminus of the precursor. Although the mechanism of this curious asparaginyl endoproteinase-mediated splicing reaction is not very clear, it has been hypothesized that the newly synthesized peptide bond may increase the structural stability of the mature product relative to its unligated precursor. The present studies, demonstrating that the stabilized structure of the product could act as a driving force in protease-catalyzed splicing reactions in aqueous solution, corroborates the above hypothesis. Additionally, our studies implicate a greater role for proteases in *in vivo* generation of proteins; natural protein splicing through protease-mediated ligations may be a more common occurrence than is currently realized.

The development of facile ligation approaches that allow specific joining of unprotected peptide segments remain an area of intense investigation (Gaertner et al., 1992; Dawson et al., 1994; Canne et al., 1997; Muir et al., 1997; Camarero et al., 1998; Evans et al., 1998; Wilken & Kent, 1998; Paulus, 1998; Muir et al., 1999). The need for a facility to join synthetic peptide modules has emanated from the size limitations of solid phase synthesis of proteins as also the current unavailability of a readily accessible procedure for site-specific incorporation of nonstandard amino acids or spectroscopic labels in proteins (Balaram, 1992; Xu et al., 1999). Although several synthetic and biosynthetic procedures are now available for this purpose (Cotton & Muir, 1999, and references therein), protease-mediated ligations still remain an attractive alternative because they are regio- and stereo-selective, require minimal or no side-chain protection, and proceed under mild reaction conditions. The peptide oligomerization reaction described here suggests that V8 protease is endowed with considerable synthetic potential in aqueous solution. The enhancement of innate peptide ligation activity of V8 protease by protein engineering is conceivable in light of the successful development of engineered subtilisins as peptide ligases (Abrahmsen et al., 1991; Chang et al.,

1994; Braisted et al., 1997; Atwell & Wells, 1999). The broad pH-activity profile of V8 protease (Birktoft & Breddam, 1994) and its stability in the presence of moderate concentrations of denaturants (Drapeau, 1977) as well as its high synthetic propensity both in aqueous and aqueous-organic milieu qualify V8 protease as an ideal candidate to be engineered into a peptide ligase useful for stepwise synthesis or semisynthesis of proteins. Notwithstanding this, the results presented here suggest that ligation of native peptide fragments in neat aqueous solution may be achieved with the wild-type enzyme through judicious engineering of product-directed conformational traps.

Materials and methods

Materials

Staphylococcus aureus V8 protease was purchased from Pierce Chemical Company (Rockford, Illinois). The chemicals used in peptide synthesis were obtained from Novabiochem (Läufelfingen, Switzerland). All other chemicals were of reagent grade and procured from standard commercial sources.

Peptide synthesis and purification

Peptides were synthesized by a solid-phase method using 9-fluorenylmethoxycarbonyl (Fmoc) strategy on a peptide synthesizer (Model 90, Advanced Chemtech, Louisville, Kentucky). The peptides were cleaved from the resin and deprotected with 95% trifluoroacetic acid (TFA) for 2 h at room temperature. The resin was removed by filtration, and the crude peptides were precipitated in cold anhydrous ether. The peptides were purified by RP-HPLC on an RP300 (250 × 7 mm, 7 μ m; Applied Biosystems, Foster City, California) column employing a linear gradient of 5–70% solvent B (solvent B; 80% acetonitrile containing 0.1% TFA in water, solvent A; 0.1% TFA in water) in 130 min at a constant flow rate of 2 mL/min. The identity of the peptides was established by ESMS.

ESMS analysis of peptides

The RP-HPLC purified peptides were taken in 50% of acetonitrile containing 1% formic acid and analyzed by positive ion ESMS. The instrument was calibrated with standard solutions of gramicidin S or myoglobin.

Protease-catalyzed peptide ligation reactions

The lyophilized peptide samples were dissolved in 0.05 M ammonium bicarbonate (pH 8) or ammonium acetate (pH 6) buffers to obtain a stock solution. Appropriate amount of peptides was taken and incubated with V8 protease [3 mM peptide, enzyme to substrate ratio, 1: 50 (w/w)] at 4 °C. The progress of reaction was monitored by RP-HPLC. The extent of synthesis were quantified from the peak areas obtained using a computer program.

Peptide sequence analysis

The sequencing of peptides were done by sequential degradation (Edman degradation) of amino acids from the amino terminus as phenylthiohydantion (PTH) derivatives on an automated protein

sequencer (ABI, Model 470A) equipped with an on-line PTH analyzer.

Size-exclusion chromatography

The gel filtration of peptides was carried out using a Superdex peptide column PC 3.2/30 (Pharmacia, Uppsala, Sweden) by HPLC. The column was equilibrated and run with 0.05 M ammonium acetate buffer (pH 6). Twenty microliters of the 3 mM peptide sample prepared in the above buffer was loaded onto the column and chromatographed at a flow rate of 0.05 mL/min.

CD spectroscopy

CD spectra were recorded on a Jasco-710 instrument equipped with a Peltier-type constant temperature cell holder (PTC-348W). The instrument was calibrated with (+)-10-camphorsulfonic acid. All measurements were performed in ammonium acetate buffer (5 mM, pH 6) at 4 °C. The spectra were an average of 10 scans and data presented as mean residue ellipticity (MRE) expressed in deg cm² dmol⁻¹.

Proteolysis of oligomers

V8 protease-catalyzed hydrolysis of respective oligomers was carried out in 50 mM ammonium bicarbonate buffer (pH 8, 4 °C). The respective peptides (2 mM) were incubated with the protease, and aliquots of the reaction mixture were analyzed at various time points by RP-HPLC. The extent of proteolysis was quantitated from the peak areas calculated using a computer program. The enzyme to substrate ratio for proteolysis was 1:200 (w/w) and 1:100 (w/w), respectively, for the trimer and the pentamer.

Acknowledgments

We thank Prof. Sandip K. Basu for his interest in this work. This work was supported by grants to the National Institute of Immunology from the Department of Biotechnology, Government of India. D.D. received post-doctoral fellowships from the Council of Scientific and Industrial Research, India. We acknowledge many helpful suggestions of the reviewers.

References

- Abrahmsen L, Tom J, Burnier J, Kossiakoff A, Butcher KA, Kossiakoff A, Wells JA. 1991. Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution. *Biochemistry* 30:4151–4159.
- Atwell S, Wells JA. 1999. Selection for improved subtiligases by phage display. *Proc Natl Acad Sci USA* 96:9497–9502.
- Balaran P. 1992. Nonstandard amino acids in peptide design and protein engineering. *Curr Opin Struct Biol* 2:845–851.
- Birktoft JJ, Breddam K. 1994. Glutamyl endopeptidases. *Methods Enzymol* 244:114–126.
- Braisted AC, Judice JK, Wells JA. 1997. Synthesis of proteins by subtiligase. *Methods Enzymol* 289:298–313.
- Camarero JA, Cotton GJ, Adeva A, Muir TW. 1998. Chemical ligation of un-protected peptides directly from a solid support. *J Pept Res* 51:303–316.
- Canne LE, Thompson DA, Simon RJ, Kent SBH. 1997. Solid phase protein synthesis by chemical ligation of unprotected peptide segments in aqueous solution. In: Tam JP, Kauyama PTP, eds. *Peptides, frontiers in peptide science; proceedings of the 15th American peptide symposium*. Nashville, Tennessee. pp 301–302.
- Canova-Davis B, Kessler JT, Ling VT. 1991. Transpeptidation during proteolysis of proteins. *Anal Biochem* 196:39–45.
- Chaiken IM. 1981. Semisynthetic peptides and proteins. *Crit Rev Biochem* 11:255–301.

- Chang TK, Jackson DY, Burnier JP, Wells JA. 1994. Subtiligase: A tool for semisynthesis of proteins. *Proc Natl Acad Sci USA* 91:12544–12548.
- Cohen C, Parry DA. 1990. Alpha-helical coiled coils and bundles: How to design an alpha-helical protein. *Proteins Struct Funct Genet* 7:1–15.
- Cotton GJ, Muir TW. 1999. Peptide ligation and its application to protein engineering. *Chem Biol* 6:247–256.
- Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. 1994. Synthesis of proteins by native chemical ligation. *Science* 266:776–769.
- De Filippis V, De Antoni F, Frigo M, deLaureto P, Fontana A. 1998. Enhanced protein thermostability by Ala → Aib replacement. *Biochemistry* 37:1686–1696.
- De Filippis V, Fontana A. 1990. Semisynthesis of carboxy-terminal fragments of thermolysin. *Int J Pept Protein Res* 35:219–227.
- Drapeau GR. 1977. Cleavage at glutamic acid with staphylococcal protease. *Methods Enzymol* 47:189–191.
- Evans TC Jr, Benner J, Xu MQ. 1998. Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci* 7:2256–2264.
- Gaertner HF, Rose K, Cotton R, Timms D, Camble R, Offord RE. 1992. Construction of protein analogs by site-specific condensation of unprotected peptides. *Bioconjug Chem* 3:262–266.
- Graf L, Li CH. 1981. Human somatotropin: Covalent reconstitution of two contiguous fragments with thrombin. *Proc Natl Acad Sci USA* 78:6135–6138.
- Homandberg GA, Chaiken IM. 1980. Trypsin-catalyzed conversion of staphylococcal nuclease-T fragment complexes to covalent forms. *J Biol Chem* 255:4903–4909.
- Homandberg GA, Komoriya A, Chaiken IM. 1982. Enzymatic condensation of non-associated peptide fragments using a molecular trap. *Biochemistry* 21:3385–3389.
- Homandberg GA, Laskowski M Jr. 1979. Enzymatic resynthesis of hydrolyzed peptide bond(s) in ribonuclease S. *Biochemistry* 18:586–592.
- Homandberg GA, Mattis JA, Laskowski M Jr. 1978. Synthesis of peptide bonds by proteinases: Addition of organic cosolvents shifts peptide bond equilibria toward synthesis. *Biochemistry* 17:5220–5227.
- Jackson DY, Burnier J, Quan C, Stanley M, Tom J, Wells JA. 1994. A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science* 266:243–247.
- Komoriya A, Homandberg GA, Chaiken IM. 1980. Enzyme-catalyzed formation of semisynthetic staphylococcal nuclease using a new synthetic fragment, [48-glycine] synthetic-(6-49). *Int J Pept Protein Res* 16:433–439.
- Kumaran S, Datta D, Roy RP. 1997. Conformationally driven protease-catalyzed splicing of peptide segments: V8 protease-mediated synthesis of fragments derived from thermolysin and ribonuclease A. *Protein Sci* 6:2233–2241.
- Manning MC, Woody RW. 1989. Theoretical study of the contribution of aromatic side chains to the circular dichroism of basic bovine trypsin inhibitor. *Biochemistry* 28:8609–8613.
- Muir TW, Dawson PE, Kent SBH. 1997. Protein synthesis by chemical ligation of unprotected peptides in aqueous solution. *Methods Enzymol* 289:266–298.
- Muir TW, Sondhi D, Cole PE. 1999. Expressed protein ligation: A general method of protein engineering. *Proc Natl Acad Sci USA* 95:6706–6710.
- Nacharaju P, Roy RP, White SP, Nagel RL, Acharya AS. 1997. Inhibition of sickle β -chain dependent polymerization by nonhuman α -chains: A super-inhibitory mouse-horse chimeric α -chain. *J Biol Chem* 272:27869–27876.
- Nakatsuka T, Sasaki T, Kaiser ET. 1987. Peptide segment coupling catalyzed by the semisynthetic enzyme thiosubtilisin. *J Am Chem Soc* 109:3808–3810.
- Nyberg F. 1988. Substance P synthesis by enzymatic fragment condensation using product-directed antibodies as molecular trap. *J Mol Recogn* 1:59–62.
- Paulus H. 1998. The chemical basis of protein splicing. *Chem Soc Rev* 27:375–386.
- Perler FB, Xu MQ, Paulus H. 1997. Protein splicing and autoproteolysis mechanisms. *Curr Opin Chem Biol* 3:292–299.
- Ray SS, Balaram H, Balaram P. 1999. Unusual stability of a multiply nicked form of plasmodium falciparum triosephosphate isomerase. *Chem Biol* 9:625–637.
- Richards FM, Vithayathil PJ. 1959. The preparation of subtilisin-modified ribonuclease and the separation of the peptide and the protein components. *J Biol Chem* 234:1459–1465.
- Roy RP, Acharya AS. 1994. Semisynthesis of hemoglobin. *Methods Enzymol* 231:194–215.
- Roy RP, Khandke KM, Manjula BN, Acharya AS. 1992. Helix formation in enzymatically ligated peptides as a driving force for the synthetic reaction: Example of α -globin semisynthetic reaction. *Biochemistry* 31:7249–7255.
- Roy RP, Nagel RL, Acharya AS. 1993. Molecular basis of sickle β^s -chain-dependent polymerization by mouse α -chains: Semisynthesis of chimeras of mouse mouse α -chains. *J Biol Chem* 268:16406–16412.
- Sahni G, Cho YJ, Iyer KS, Khan SA, Seetharam R, Acharya AS. 1989. Semisynthetic hemoglobin A: Reconstitution of functional tetramer from semisynthetic α -globin. *Biochemistry* 28:5456–5461.
- Sahni G, Khan SA, Acharya AS. 1998. Chemistry of the “molecular trap” of protease-catalyzed splicing reaction of complementary segments of alpha subunit of hemoglobin A. *J Protein Chem* 17:669–678.
- Schellenberger V, Jakubke H-D. 1991. Protease-catalyzed kinetically controlled peptide synthesis. *Angew Chem Intl Ed Engl* 30:1437–1449.
- Schuster M, Aaviksaar A, Schellenberger V, Jakubke H-D. 1990. Characterization of the S'-subsite specificity of V8 proteinase via acyl transfer to added nucleophiles. *Biochim Biophys Acta* 1036:245–247.
- Vogel K, Chmielewski J. 1994. Rapid and efficient resynthesis of proteolyzed triosephosphate isomerase. *J Am Chem Soc* 116:11163–11164.
- Wallace CJA. 1995. Peptide ligation and semisynthesis. *Curr Opin Biotechnol* 6:403–410.
- Wang M, Hugh Jones D. 1994. In vitro splicing of concanavalin A is catalyzed by asparaginyl endopeptidase. *Nat Struct Biol* 1:502–504.
- Wehofsky N, Bordusa F. 1999. Programming of enzyme specificity by substrate mimetics: Investigations on the Glu-specific V8 protease reveals a novel general principle of biocatalysis. *FEBS Lett* 443:220–224.
- Wilken J, Kent SBH. 1998. Chemical protein synthesis. *Curr Opin Biotechnol* 9:412–426.
- Wolf E, Kim PS, Berger B. 1997. MultiCoil: A program for predicting two and three stranded coiled coils. *Protein Sci* 6:1179–1189.
- Wong C-H, Wang K-T. 1991. New developments in enzymatic peptide synthesis. *Experientia* 47:1123–1129.
- Woody RW. 1978. Aromatic side-chain contributions to the far ultraviolet circular dichroism of peptides and proteins. *Biopolymers* 17:1451–1467.
- Wu Z, Hilvert D. 1989. Conversion of a protease into an acyl transferase: Selenosubtilisin. *J Am Chem Soc* 111:4513–4514.
- Xu R, Ayers B, Cowburn D, Muir TW. 1999. Chemical ligation of folded recombinant proteins: Segmental isotopic labeling of domains for NMR studies. *Proc Natl Acad Sci USA* 96:388–393.