ACS OMEGA

Article

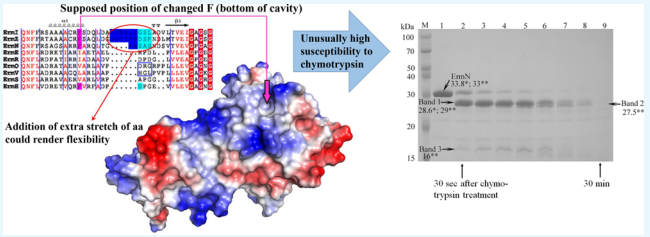
http://pubs.acs.org/journal/acsodf

Recognition Site Generated by Natural Changes in Erm Proteins Leads to Unexpectedly High Susceptibility to Chymotrypsin

Tien Le,[†] Hak Jin Lee,[‡] and Hyung Jong Jin*[,†]

[†]Department of Bioscience and Biotechnology, The University of Suwon, Hwaseong City, Gyeonggi-Do 18323, Republic of Korea
[‡]Department of Life Science, Korea University Graduate School, Seoul 02841, Republic of Korea

ABSTRACT: Erms are proteins that methylate the adenine (A2058) in *Escherichia coli* 23S rRNA, which results in resistance to macrolide, lincosamide, and streptogramin B antibiotics. In a previous report, ErmN appeared to be more susceptible to contaminating proteases in DNase I. To determine the underlying mechanism, cleavage with chymotrypsin over time was investigated. ErmN possesses unusually high-susceptibility recognition site (F45) as evidenced by a band (band 1) that represented greater than 80% of the total band intensity at 30 s. The exposure rate of the hydrophobic core was more than 67-fold and 104-fold faster in ErmN than those in ErmS and ErmE, respectively. After cleavage at F45, some of the hydrophobic interactions were disrupted. Further digestion of band 1 occurred through the exposed F163 with a half-life of 3.18 min. After 30 min, less than 1% of ErmN remained. On the basis of the structure of ErmC′, the location of F45 was presumed to be in an α helix at the bottom of a cavity. Both substitution of most common amino acids such as isoleucine, valine, or leucine with phenylalanine (ErmH, ErmI, ErmN, and ErmZ out of the 37 known Erms) and the apparent added flexibility, which could result from the additional loop region attached to phenylalanine that is four to nine amino acids longer (ErmI, ErmN, and ErmZ, which form one cluster in the phylogenetic tree), could cause unusually high susceptibility. The unexpectedly high susceptibility among the homologous proteins could indicate that caution should be taken not to misinterpret the observations when conducting any procedure in which protease or protease contamination is involved.



INTRODUCTION

Proteolysis serves a variety of functional roles in organisms including the digestion of proteins in food, the conversion of precursor proteins to active proteins, involvement in the regulation of some physiological and cellular processes, prevention of the accumulation of unwanted or abnormal proteins in cells, and association with diseases as a result of an abnormal proteolytic activity. However, some proteins exhibit increased susceptibility to proteolytic degradation that requires the implementation of special measures to inhibit potential proteolysis during isolation or purification, which could otherwise result in severe degradation.[1,2] A likely explanation for this observation is the intrinsic flexibility[3,4] of certain site(s) in the proteins that might otherwise be resistant to proteolytic cleavage when the proteins are in their native, fully folded state. Because of this physical property, the protein could undergo local unfolding, which could provide certain length of peptide thread to the active site of the protease and result in cleavage. Once this initial cleavage occurs and the overall folding of the protein is affected, additional susceptible sites could be exposed, leading to complete degradation of the protein.

Erythromycin ribosome methylation (Erm) proteins methylate the single adenine residue within the 23S rRNA to reduce the affinity of antibiotics to a region around the peptidyl transferase circle, thereby conferring resistance to macrolide, lincosamide, and streptogramin B (MLS_B) antibiotics in various microorganisms.[5,6] Approximately 40 different *erm* methyltransferase genes have been isolated and sequenced from diverse sources, ranging from antibiotic producers to pathogens. A comparison of the Erm methyltransferases that have been characterized revealed that sufficient homology exists in the regions of alignable amino acid sequences,[7] suggesting that they might share a very similar structure. On the basis of the previously determined structures, Erm proteins are believed to be composed of two domains, a catalytic domain and a substrate-binding domain.[8−10] The larger N-terminal catalytic domain exhibits a typical $\alpha/\beta/\alpha$ sandwich architecture that contains the *S*-adenosyl-L-methionine-binding site. The smaller C-terminal domain consists of three α-helices that reportedly function as an rRNA-binding domain.

ErmS and ErmN methyltransferases are two of the four gene products synthesized by *Streptomyces fradiae* (*S. fradiae*) to be resistant to its autogenous antibiotic, tylosin.[11,12] ErmS specifically dimethylates the adenine residue (A2058) in 23S rRNA, conferring a high level of resistance to MLS_B antibiotics. Unlike the dimethylation by ErmS, ErmN causes the monomethylation of the same A2058 position, which confers

Received: April 12, 2017
Accepted: July 10, 2017
Published: November 20, 2017

DOI: 10.1021/acsomega.7b00446
ACS Omega 2017, 2, 8129−8140

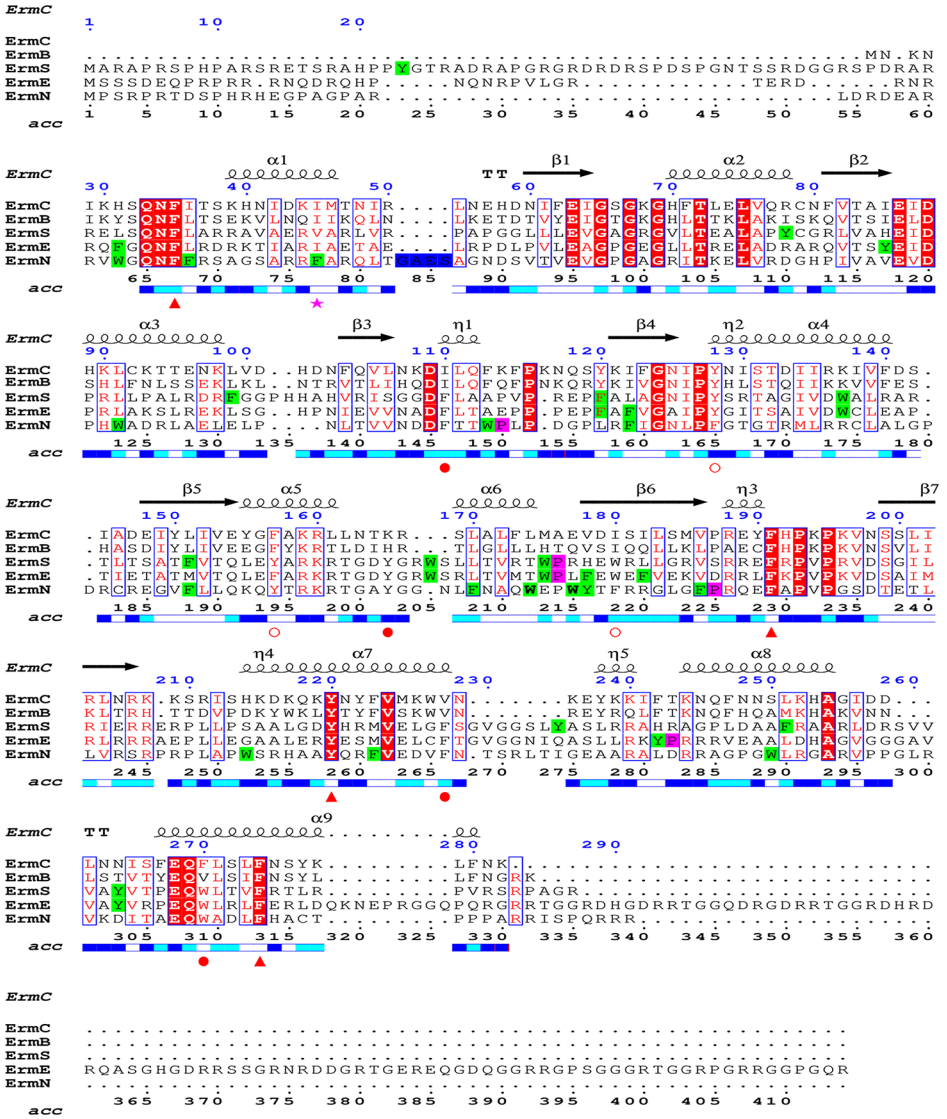


Figure 1. Sequence alignment of the representative Erm proteins: ErmB, ErmC′, ErmE, ErmN, and ErmS. The conserved residues are shown in white on a red background, and the conservative substitutions are shown in red. The structure of ErmC′ (PDB entry 1QAM) was used to assign the secondary structure on top of the alignment. The blue numbers are in accord with the ErmN sequence, and the black numbers are based on the column numbering. The relative accessibility is indicated below the sequence alignment with blue, cyan, and white representing accessible, intermediate, and buried, respectively. The preferential chymotrypsin cleavage sites [tryptophan (W), tyrosine (Y), and phenylalanine (F)] are marked as follows: a red triangle is shown below the sequence when the cleavage sites are conserved in all five aligned sequences, a solid red circle is shown below the sequence to indicate conservation among the Erm sequences tested (ErmE, ErmN, and ErmS), and an open red circle is shown below the alignment to indicate that W, Y, or F is located at the same position in the alignment among the Erm sequences. The other preferential cleavage sites are indicated in green boxes. F in ErmN, which exhibited unexpectedly high susceptibility to chymotrypsin, is marked with a pink star below the sequence. Four extra amino acids in the loop region of ErmN are indicated in blue boxes, which might help to exhibit the unusually high susceptibility to chymotrypsin. The figure was generated using Easy Sequencing in PostScript (ESPript).[17]

a high level of resistance to lincosamide antibiotics but a lower level resistance to macrolide and streptogramin B antibiotics.

A previous report[13] revealed that ErmN was seemingly more susceptible than ErmS to contaminating proteases in a commercial DNase I preparation. Accordingly, the current study aimed to elucidate the difference in proteolytic susceptibility between ErmS and ErmN and to discuss the consequences of the apparent increased susceptibility of ErmN and its effects on our daily experiments. To confirm our hypothesis that ErmN is more susceptible to proteolysis than other Erm proteins, we also employed ErmE,[14] a dimethylase from *Saccharopolyspora erythraea*, as a reference.

RESULTS

Number of Putative Cleavage Sites That Satisfy the Chymotrypsin Specificity Requirement among the Erm Proteins Tested. In the previous report, ErmN was more readily cleaved than ErmS by contaminating proteases in a commercially available DNase I, mainly chymotrypsin and trypsin. All of the cleavable bonds in ErmN, ErmS, and ErmE (using ErmC and ErmB as reference sequences, the structures of which have been elucidated) were searched for the potential alignment of these sequences. With trypsin, large variations in the number of susceptible sites among ErmE, ErmN, and ErmS were apparent (64, 42, and 54, respectively). The extent of the protease contamination in the DNase I, which was prepared

DOI: 10.1021/acsomega.7b00446
ACS Omega 2017, 2, 8129−8140

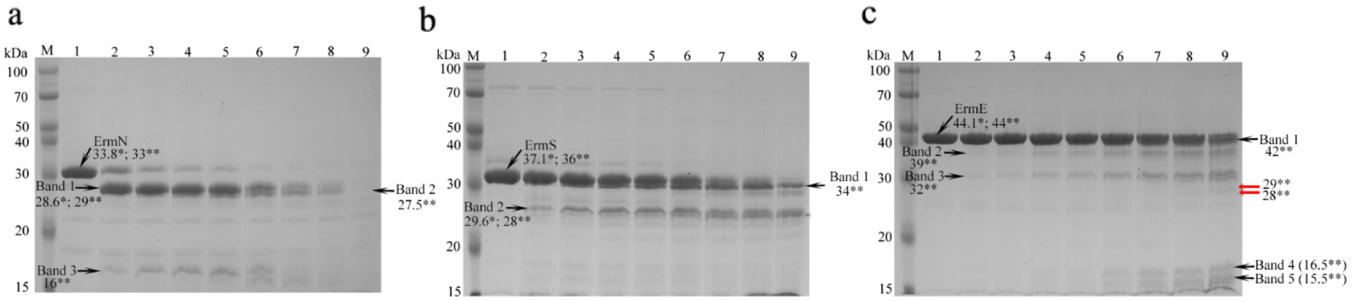


Figure 2. SDS-PAGE profiles of the proteolytic degradation of ErmN (a), ErmS (b), and ErmE (c) by chymotrypsin. Purified Erm protein (10 μg) was treated with sequencing grade chymotrypsin at an enzyme/substrate ratio of 1:50 in 10 μL of 50 mM Tris-HCl buffer (pH 8.0) and 50 mM $CaCl_2$ at 37 °C. The digestion was stopped by adding phenylmethylsulfonyl fluoride (PMSF) at a final concentration of 5 mM, and the mixture was boiled for 5 min. The samples were then analyzed on a 12% SDS-PAGE gel. The major products that resulted from the treatment with chymotrypsin were named and marked with arrows. The molecular weights of the bands were then deduced from the peptide sequence using the ExPASy ProtParam tool (*) provided by expasy.org[19] or were estimated from a calibration curve constructed by plotting the log molecular weight versus the migration distance of the protein markers and the Erm proteins/their identified peptides (**).[20] The red arrows indicate the theoretical location of the products that were generated but degraded rapidly after the removal of bands 4 and 5 from the native protein by digestion at Y135 and W144, which might have disproportionately affected the two resultant products and rapidly degraded the other products (see Results). To facilitate the identification of the cleavage site that generated the major degradation products after chymotrypsin treatment, the sample was subjected to N^{2+} affinity chromatography purification to obtain peptides containing the His-tag at the C-terminal end, which were subsequently exhibited on an SDS-PAGE gel (data not shown). Following this procedure, bands 1 and 2 from all of the Erm proteins investigated were confirmed as products that resulted from N-terminal cleavage, whereas bands 3, 4, and 5 from ErmN and ErmE were derived from cleavage at the C-terminal end (M: molecular weight marker; 1: untreated protein; and lanes 2–9 contain the products following incubation for 0.5, 1, 2, 3, 5, 10, 15, and 30 min).

from the bovine pancreas, was on the order of chymotrypsinogen > chymotrypsin > trypsin.[15,16] Therefore, chymotrypsin could have a greater effect on the increased susceptibility of ErmN. Furthermore, the number of preferential cleavage sites (W, Y, and F) among ErmE, ErmN, and ErmS was quite similar at 20, 24, and 21, respectively, even though the positions of the sites did not match exactly among the Erm proteins (Figure 1). However, 11 of the sites matched in position (marked with ▲ or ●) and with respect to preferential cleavability (○), which is indicated by a red triangle, a solid circle, or an open circle depending on the degree of consensus (Figure 1). The actual number of preferential cleavage sites was 22, 25, and 22. The positioning of proline (P; colored pink in Figure 1) on the carboxy-terminal side of the scissile bond blocked the cleavage almost completely, independent of the amino acids [even with the most preferable amino acids (W, F, and Y)] found on the amino-terminal side.[18] Consequently, the putative preferential cleavage sites for ErmE, ErmN, and ErmS were reduced to 20, 24, and 21, respectively. When ErmN was treated with chymotrypsin, greater than 80% of the protein was converted to a 28.6 kDa fragment that was produced by the action of chymotrypsin at the phenylalanine 45 (F45) residue of ErmN (Figure 2). Therefore, three other preferential cleavage sites before F45 (F77, column numbering) were, likewise, not considered. Thus, despite the ability to locate each site in different environments, such as accessibility and the secondary structure element shown in Figure 1, which could affect the cleavability by chymotrypsin, the number of sites susceptible to chymotrypsin was nearly the same among the Erm proteins tested at 20, 20, and 21.

Preferred Cleavage Sites in Each Erm Protein. The digestion of the purified Erm proteins with chymotrypsin at an enzyme/substrate ratio of 1:50 was monitored over time by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 2). Undigested Erm proteins, including the C-terminal His-tag, yielded a band on SDS-PAGE that corresponded to a molecular mass of 44.1 kDa (ErmE), 37.05 kDa (ErmS), or 33.8 kDa (ErmN; Figure 2). As previously stated, the three Erm proteins exhibited almost the same number of susceptible sites with regard to preferential cleavage sites with high specificity for chymotrypsin (W, F, and Y). As well, sequencing grade chymotrypsin, which is known to exhibit higher specificity compared with the other chymotrypsin grades, was used. However, at a lower rate, the peptide bonds after leucine (L), methionine (M), alanine (A), aspartic acid (D), and glutamic acid (E) could be degraded with chymotrypsin. With ErmS (Figure 2a), three bands could be observed in all lanes (0.5–30 min), including intact ErmS. Among these, band 2 was purified using the His-tag located at the C-terminal end of the protein, separated by SDS-PAGE, extracted, and subjected to N-terminal sequencing. On the basis of the results, it was determined that the band originated from a cleavage at F67 (similar to the column numbering).[13] Band 1 was detected after a 1 min incubation and after Ni^{2+} affinity chromatography, implying that the cleavage occurred near the N-terminus. Therefore, this band was thought to originate from a cleavage after Y23 (similar to the column numbering), showing a mobility similar to that of the deduced molecular weight of the cleaved fragment (34.4 kDa). Many other small bands (more than those in ErmN and ErmE) appeared only at certain time points or throughout the reaction. According to the structure of ErmC′, two cleavage sites that form bands 1 and 2 are located outside of the main body of ErmS, do not form any secondary structures that could be easily recognized by chymotrypsin, and become the main cleavage site. The possibility exists for the region around and preceding F67 to form an α helix (unpublished result), but it would not form an extensive core interaction with the protein body and likely would not influence the stability of the protein. A fragment from the truncation of ErmS at F67 survived intact even when it was treated with chymotrypsin for 30 min, which supports the aforementioned theory. However, the fragments appeared to be gradually degraded more by chymotrypsin in sites located within the protein body, reducing the total amount of staining of ErmS and its degraded fragments.

DOI: 10.1021/acsomega.7b00446
ACS Omega 2017, 2, 8129–8140

There were six noticeable bands corresponding to ErmE that could be detected in all of the SDS-PAGE lanes (Figure 2b). One of them represented intact ErmE and another (band 2) was believed to result from digestion after F39 (F63, column numbering) or F43 (F67), which could exhibit a mobility similar to that of the molecular weight calculated from the truncated fragment (39.2 and 38.8 kDa, respectively). Band 3, which was estimated to be 32 kDa, could be observed in all SDS-PAGE lanes in this study but was not recovered by Ni^{2+} affinity chromatography, suggesting that this fragment might have originated from the native protein and resulted from a cleavage near the C-terminal end. Thus, band 3 might have been derived from a digestion after Y271 (Y303, generating a band of 30.7 kDa), after W277 (W309, generating a band of 31.5 kDa), or after F281 (F313, generating a band of 32 kDa). Bands 4 and 5 could be detected from the onset of the reaction with chymotrypsin until the point that most of the native protein disappeared (after 30 min incubation with chymotrypsin), suggesting that these bands might result from the native protein, although the likelihood exists that the digestion product of band 3 could contribute to the intensity of the two bands to some extent. Because bands 4 and 5 could not be recovered by the affinity chromatography, these bands might have resulted from the digestion after W144 (generating a 16.3 kDa band) and after Y135 (generating a 15.3 kDa band), respectively. The other bands generated from the cleavage that occurred after W144 and Y135 in the native band were barely detected, presumably because of the rapid degradation (red arrows in Figure 2). This rapid degradation could be true for the C-terminal end band that occurred after the cleavage of band 3 at Y135 or W144, presumably because this rapid degradation could be the result of the ensuing structural perturbation after the cleavage at Y135 and W144 that disproportionately affected the two resultant bands. The estimated molecular weight of band 1 of ErmE was approximately 42 kDa. The ErmE sequence did not contain a preferential recognition site near the N-terminal end that would enable chymotrypsin to generate a band similar to the estimated molecular weight of band 1, despite the fact that band 1 could be recovered by affinity chromatography. However, considering the location of band 1, the 40 kDa marker, and band 2 (39 kDa), the cleavage site might be between D16 and R17 (generating a band of 42.1 kDa) or between L27 and G28 (generating a band of 40.9 kDa), which would indicate the recognition of a less preferable site.

Unlike the other Erm proteins, an unusually drastic cleavage was observed as band 1 in ErmN (Figure 2c). A fragment corresponding to this band was previously purified using a C-terminal end His-tag, transferred to a membrane, and subjected to N-terminal sequencing to verify that the N-terminus contained A46 (ErmN numbering, A78 in the column numbering) generated by the cleavage at F45.[13] Even after only a 30 s incubation, band 1 (28.6 kDa) represented more than 80% of the total density of the SDS-PAGE gel with the concomitant appearance of bands as small as 27.5 kDa (band 2) and 16 kDa (band 3). Bands 2 and 3 both presumably originated from band 1. This inference could be explained as follows: Although increasing the incubation time by 2 min (from 3 to 5 min) or by 1 min (from 1 min through 2−3 min) increased the densities of bands 2 and 3, respectively, the apparent decrease in the intact band density of ErmN could not cope with this increase. These observations strongly suggested that the structural disruption caused by the truncation after F45 resulted in the exposure of the inner susceptible sites and facilitated the cleavage of these sites by chymotrypsin. Although band 2 could be recovered by affinity chromatography, band 3 was not recovered by affinity chromatography, implying that band 2 resulted from digestion on the N-terminal side but band 3 resulted from digestion on the C-terminal side. Band 2 presumably resulted from the further degradation of band 1 at one of the following sites: A52−E53 (84 and 85, generating a 27.9 kDa band), E53−S54 (85 and 86, generating a 27.8 kDa band), A55−G56 (87 and 88, generating a 27.7 kDa band), or D58−S59 (90 and 91, generating a 27.4 kDa band). In addition, band 3 might have been derived from the further degradation of band 1 at F191−A192 (229 and 230, 163 and 164 in ErmC′ numbering, generating a 16.3 kDa band). On the basis of the results of the SDS-PAGE analysis of degradation patterns of the three proteins, the extent of digestion of the native protein, both outside and inside the core region, was greater in ErmS than that in ErmE. This degradation pattern was also observed in the kinetic analysis of the degradation of the three proteins (Figure 5).

Structural Consideration of F45 and the Truncation Fragment after F45. After the marked degradation of the native ErmN protein between F45 and A46, the resulting fragment (band 1 in Figure 2a) appeared to immediately degrade further into 27.5 and 16 kDa fragments (bands 2 and 3 in Figure 2a, respectively) because the two bands appeared on the gel with band 1 after only a 0.5 min incubation with chymotrypsin. Among the intermediate degradation bands detected from the three Erm proteins on the SDS-PAGE gels (Figure 2), the degradation rate of these bands was thought to be second only to fragments generated from cleavage after W144 and Y135 in intact ErmE and its band 3, which could not be observed in any lane of the gel. The effect of cleavage after F45 [ErmN numbering and is substituted with isoleucine (Ile) 22 in ErmC′] was analyzed based on the structure of ErmC′ (presented below in the Discussion and Conclusions section) because unfortunately the structure of ErmN is not yet solved. As expected, the truncation after F45 could remove most of the first α helix and expose a portion of the ErmN hydrophobic core for rapid degradation by chymotrypsin (Figures 2 and 3). On the basis of the ErmC′ structure, the following residues (all are numbered according to ErmC′ numbering) could be exposed by the loss of hydrophobic interaction residues in the first α helix and the preceding loop [the criteria for recognizing the hydrophobic interaction were based on the Protein Interactions Calculator (PIC),[21] which indicated that two hydrophobic residues fell within 5 Å]: Lys41, Phe44, Ile126, and Val158 (Figure 3b). Lys41 (located in the loop between β1 and α2) interacted with Phe12 in the loop preceding α1, Val158 (at the end of β6) interacted with Ile13 in the preceding loop, Phe44 (in α2) interacted with Ile19 in α1, Phe44 and Ile126 (in β5) interacted with Ile22 in α1 (replaced by F45 in ErmN), and Phe99 (in β4) interacted with Ile22, but the distance between these two residues was 5.4 Å. Phe44 interacted with Phe99 and Ile126 as well; thus, Phe44 interacted with four different residues (Ile19, Ile22, Ile126, and Phe99) in total. If Ile19 and Ile22 are lost following the truncation of α1, the interactions of Phe44 with Phe99 and Ile126 and the interaction between Phe99 and Ile126 might be affected. Among these, interactions that cause the disruption of the protein core structure might be Ile13 with Val158, Ile19 with Phe44, Ile22 with Ile126, or Ile22 with Phe44. Additionally, interactions of Phe44 with Phe99 and Ile126

DOI: 10.1021/acsomega.7b00446
ACS Omega 2017, 2, 8129−8140

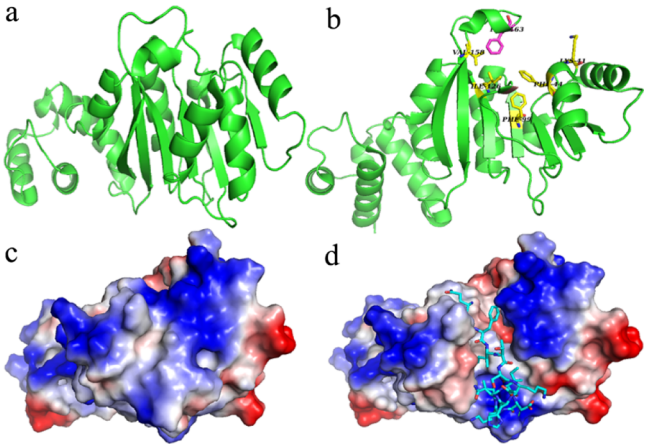


Figure 3. Exposed residues (Lys41, Phe44, Ile126, and Val158; ErmC′ numbering, colored yellow) generated by the loss of the preceding loop to α1 and most of α1 after the truncation at F45 (Ile22 in ErmC′) of ErmN and residue F163 (ErmC′ numbering, colored magenta), which generated band 3 of ErmN following digestion, as illustrated using the ErmC′ structure (PDB 1QAM). As described in the Results section, Lys41 (in α2) and Val158 (at the C-terminal end of β6) interact with Phe12 and Ile13 in the preceding α1 loop, respectively. In addition to the interaction with the truncated Ile19 and Ile22, Phe44 interacts with Phe99 and Ile126. When the residues and the associated interactions are lost, the interactions of Phe44 with Phe99 and Ile126 might be affected (a further reason for this is that Ile22 interacts with Ile126 as well). Furthermore, the interaction between Phe99 and Ile126 could be affected. Therefore, Phe99 (colored yellow) might be exposed by indirect effects. Thus, after the exposure of the hydrophobic core and the loss of the interaction, the structurally and sequentially nearby F163 (ErmC′ numbering; F191, ErmN numbering; and F229, column numbering) was recognized and digested to generate band 3 of ErmN. The remaining M23 in α1 after truncation at I22 is colored brown. (a) Ribbon structure of ErmC′; (b) ribbon structure of ErmC′ truncated at I22 (replaced by F45 in ErmN) with the exposed residues and F163 highlighted; (c) surface structure of ErmC′; and (d) surface structure of ErmC′ truncated at I22 but with the truncated amino acids represented with sticks.

and between Phe99 and Ile126 might likewise be disturbed. Interestingly, following the exposure and loss of the interactions, the nearby (both structurally and sequentially) preferential recognition site (F191 in ErmN numbering and F163 in ErmC′ numbering) was recognized and digested to generate band 3 from band 1 in ErmN (Figure 3b). Notably, this exposure of a new preferential site and its cleavage appeared to occur immediately after the cleavage at F45 and the subsequent disruption of the complicated interaction network described above, because band 3 was observed with band 1 within 30 s of incubation with chymotrypsin. According to the structure of ErmC′, the F45 residue in ErmN is located at the bottom of the cavity that developed opposite the active site (Figure 4). Hence, it is quite interesting that as a member of the α helix, the F45 in this location could be recognized and cleaved rapidly by chymotrypsin (see below). Furthermore, the immediate digestion of band 1 to generate band 2 was also noticeable because it involves less preferable recognition sites (see the Discussion and Conclusions section).

Kinetics of Proteolysis. To quantify the degradation kinetics of each Erm protein investigated in the presence of chymotrypsin, the residual band intensity, which was calculated by densitometry, was plotted against the incubation time (Figure 5). The residual band intensity could correspond to

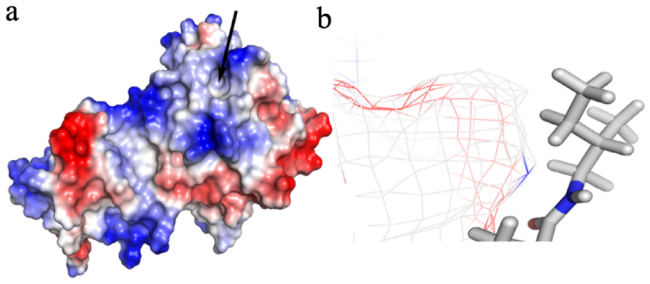
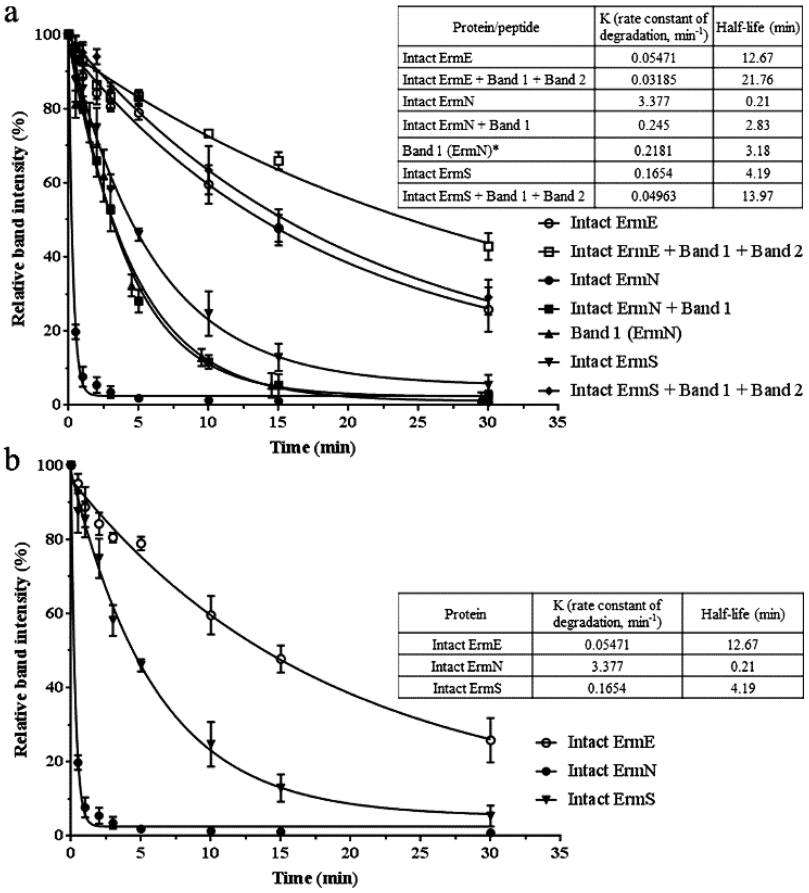


Figure 4. Location of the isoleucine residue (I22) in ErmC′ that is replaced by the phenylalanine residue (F45), which is highly susceptible to chymotrypsin in ErmN. In the ErmC′ structure, I22 is located at the bottom of the cavity that develops opposite the active site of ErmC′, the location of which is presumed to be difficult for chymotrypsin to penetrate. (a) Surface structure of ErmC′ illustrating the cavity containing I22 at the bottom, which is indicated with an arrow, and (b) location of I22 in the cavity.

that of the native protein, which remained after a certain time of incubation with chymotrypsin. However, the comparison could be made based on the core structure of each protein, with the exception of the N- or C-terminal unstructured region or the region that is structured but does not interact extensively with the core structure, because the N- or C-terminal region could be digested much more easily than the protein core because of the increased flexibility of these regions. Therefore, the residual band intensity would rather represent the combined intensity of the native band plus the bands that resulted from a truncation at the N- and/or C-terminal region outside the core structure, such as bands 1 and 2 in ErmS and bands 1 and 2 in ErmE, which could serve as a means to compare the degradability of the core structure of each protein. However, when the degradation rate of ErmN band 1 was considered, the residual band intensity was deemed equivalent to the intensity of ErmN band 1 because the degradation of ErmN band 1 was inherently involved. When only the native protein was considered to be the residual band, intact ErmN (●) disappeared almost immediately within a half-life of 0.21 min. Additionally, ErmE was the most difficult protein to degrade using chymotrypsin, as it exhibited a half-life of 12.67 min, which was approximately 60-fold higher compared with the half-life of ErmN (inset table in Figure 5a). The degradation rate observed in this step primarily reflected the digestion that occurred outside the core structure of each protein. Although both ErmE and ErmS harbor two preferred truncation sites outside the core structure, ErmS (▼) appeared to exhibit a higher digestion rate, as reflected in the curves in Figure 5a, which indicates a greater than three-fold shorter half-life than that of ErmE (○). This observation could be confirmed in Figure 5b because the digestion rate of the ErmS core structure was similar to that of the digestion that occurred outside the core structure of ErmE. The slower digestion rate of ErmE compared with that of ErmS, in what was assumed to be the unstructured region, might be caused by the fact that ErmE contained one less preferable recognition site out of the possible two sites in that region compared with ErmS, which contained the two preferential recognition sites that generated bands 1 and 2 of each protein.

Band 1 of ErmN (▲) was degraded faster than the region outside the core structure of ErmS (slightly more than three-fourths in the half-life), the rate of which would be marginally accelerated by digestion at the core of ErmS because the

DOI: 10.1021/acsomega.7b00446
ACS Omega 2017, 2, 8129−8140



Protein/peptide	K (rate constant of degradation, min^{-1})	Half-life (min)
Intact ErmE	0.05471	12.67
Intact ErmE + Band 1 + Band 2	0.03185	21.76
Intact ErmN	3.377	0.21
Intact ErmN + Band 1	0.245	2.83
Band 1 (ErmN)*	0.2181	3.18
Intact ErmS	0.1654	4.19
Intact ErmS + Band 1 + Band 2	0.04963	13.97

Protein	K (rate constant of degradation, min^{-1})	Half-life (min)
Intact ErmE	0.05471	12.67
Intact ErmN	3.377	0.21
Intact ErmS	0.1654	4.19

Figure 5. Kinetics of the proteolytic degradation of Erm proteins with chymotrypsin. The proteolytic degradation was quantified by densitometry after scanning the SDS-PAGE gels shown in Figure 2 using the GelQuant.NET software. The intensity of the bands corresponding to the untreated proteins served as a control for 100% protection against chymotrypsin action, unless otherwise indicated. The percentage of the intact protein (and truncated peptides as indicated) was plotted as a function of treatment time to monitor the degradation kinetics. The rate constant (K, min^{-1}) and half-life (min) of proteolytic degradation (inset table) were calculated from at least three different experiments by the exponential fitting of one-phase decay curve to the above data using the GraphPad Prism software. (a) Kinetics of the major pathways corresponding to the degradation of both the inside and outside core structures of the Erm proteins, whereas (b) kinetics of the pathways corresponding to the degradation of the outside core structure only. * To investigate the degradation of ErmN band 1, the intensity of band 1 in lane 2 in Figure 2a was used as a control. The contribution of intact ErmN to band 1 was calculated and subtracted after estimating the intensity of band 1 (from lanes 3−9, Figure 2a) because the contribution of ErmN to band 1 could be greater than 15%.

structures both outside and inside the core are simultaneously attacked by chymotrypsin. As reasoned above, if the bands generated from the digestion outside the core structure were included in the residual band, the degradation rate decreased, which reflected the relative insusceptibility of the recognition sites in the core structure of each protein to the protease. As observed in the digestion rate at the region outside the core structure, the digestion rate of the ErmE (□) core structure was slower than the digestion rate of the ErmS (◆) core structure, as indicated by the half-life ratio of 1:0.64. As shown in Figure 2, ErmS appeared to have an increased number of recognition sites and higher subsequent digestion rates than ErmE in the absence of a distinctive degradation pathway, evidenced by the increased number of bands on the SDS-PAGE gel. Furthermore, the degradation of ErmN band 1, when compared with the cleavage of ErmS plus bands 1 and 2 and to ErmE plus bands 1 and 2, showed approximately a 4.4-fold and 6.8-fold shorter half-life than ErmS and ErmE, respectively.

Bioinformatic Analysis of Unusually High Susceptibility of ErmN. Substitution with F at the position of 45 of ErmN does not seem enough to cause the unusually high susceptibility to chymotrypsin because a lot of preferential recognition sites are present without corresponding strong cleavage. Furthermore, F45 resides at the bottom of the cavity as a member of an α helix formed on the opposite side of the active site based on the structure of ErmC′, which does not appear to be favorable for the recognition and cleavage by chymotrypsin (Figure 4). If this is true for ErmN, it would be interesting to investigate the mechanisms by which chymotrypsin interacts with ErmN, recognizes the buried residue, unfolds it locally, and cleaves the peptide bond after phenylalanine rapidly in contrast to the established notion that burial and structural constraint from the interaction with other residues could hinder the proteolysis by proteases.[22] However, another possibility exists that should not be ruled out. In addition to the F45 substitution, ErmN harbors a loop connecting α1 and β1 that is four residues longer than the corresponding loop in the other four Erm proteins employed in this study (Figure 1). Hence, this longer loop could distort the local structure and expose the F45 residue, which would directly enable a reaction with chymotrypsin and/or increase the flexibility around F45 to facilitate easier recognition and cleavage. A multiple sequence alignment of the 37 Erm proteins known to date (Figure 6a) revealed that the position F45 of

DOI: 10.1021/acsomega.7b00446
ACS Omega 2017, 2, 8129−8140

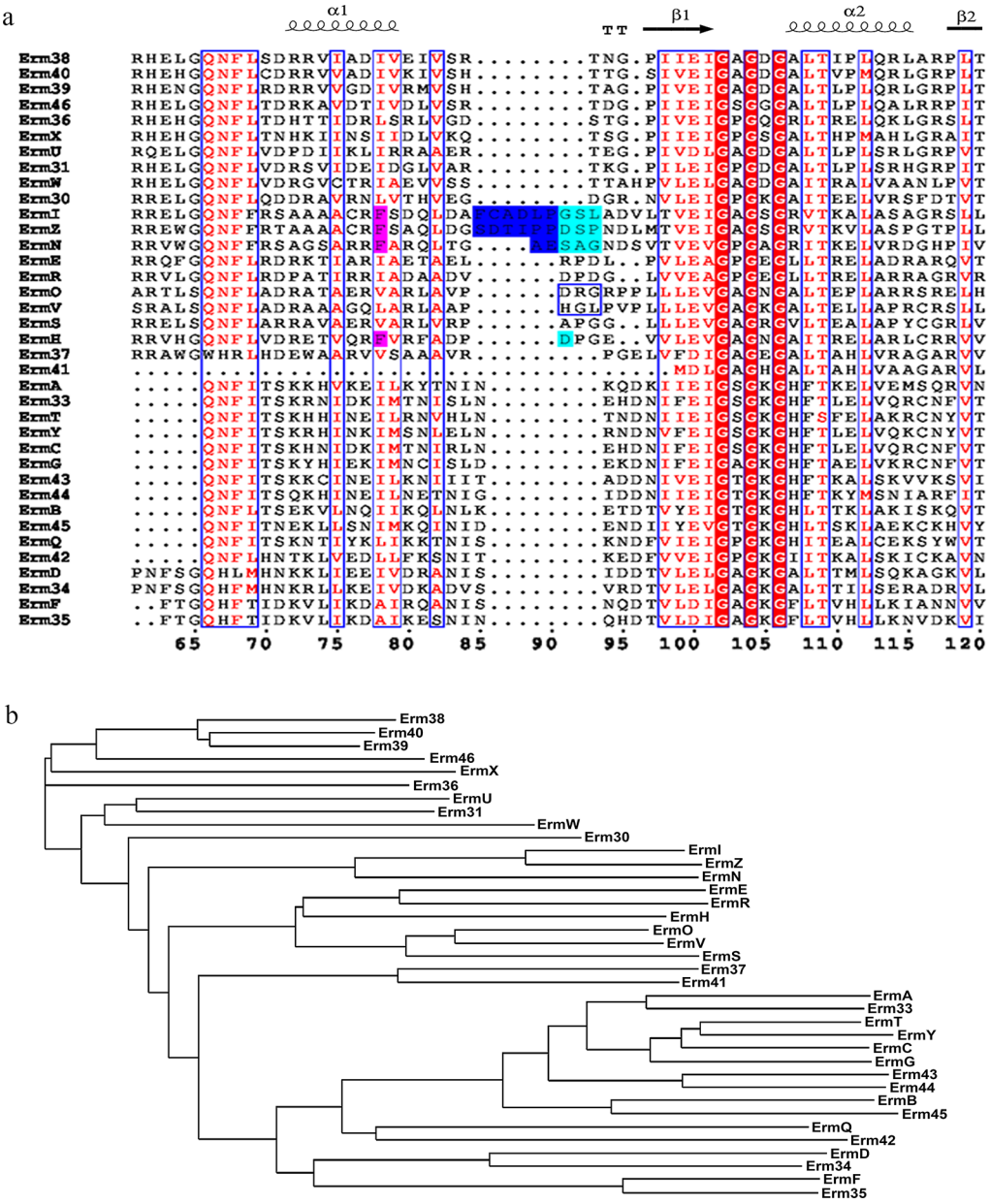


Figure 6. (a) Multiple sequence alignment of 37 Erm proteins known to date in the region encompassing α1 and β1 and the loop connecting two secondary structures. From the multiple sequence alignment, several Erm proteins were detected to harbor additional amino acids in the loop region connecting α1 and β1, which could put more flexibility to the regions around the chymotrypsin recognition site (F, colored pink) in ErmH, ErmI, ErmN, and ErmZ as well. Length of the added amino acids in ErmH, ErmI, ErmN, and ErmZ depends on which Erm sequence is compared to them. Among these Erm sequences in comparison, ErmO and ErmV, in which additional amino acids are boxed with a blue line, have the longest loop. On comparing with these two sequences, ErmN, ErmI, and ErmZ are found to harbor the extra amino acids (colored blue) of 2, 6, and 6, respectively. When the first 8 Erm sequences in a multiple sequence alignment are selected for comparison, which harbor the shortest loop, ErmH, ErmN, ErmI, and ErmZ contain 1, 3, 3, and 3 more amino acids (colored cyan), which might be added in the loop region, and the total extra amino acids in the loop become 1, 5, 9, and 9, respectively. On comparing with ErmC′, which gives a reference secondary structure on top of the multiple sequence alignment, ErmE, and ErmS, added extra amino acids are 4, 9, and 9 in ErmN, ErmI, and ErmZ, respectively. The substituted amino acid F in ErmH, ErmI, ErmN, and ErmZ is colored pink. Amino acid numbers are based on the column numbering. (b) Phylogenetic tree of 37 Erm proteins known to date. Erm proteins such as ErmI, ErmZ, and ErmN with a long loop region form one cluster, implying that they are originated from a common ancestor and ErmI and ErmZ with the longest loop are the closest members to each other.

ErmN is most often occupied by hydrophobic amino acids, such as isoleucine, leucine, and valine, and, in some cases, by alanine. In addition to ErmN, the substitution with phenylalanine at that position is found in ErmH (from *Streptomyces thermotolerans*),[23] ErmI (from *Streptomyces mycarofaciens*),[24] and ErmZ (from *Streptomyces ambofaciens*).[25] Although ErmH harbors a loop connecting α1 and β1 that has a length similar to most other Erm proteins in the multiple sequence alignment, ErmI and ErmZ contain a loop that is six to nine residues longer than the other Erm proteins (Figure 6a). In addition, ErmN harbors two to five longer loops than the other Erm proteins. This observation suggests that at least 3 (ErmI, ErmN, and ErmZ) out of the 37 known Erm proteins could exhibit the unusually high susceptibility to chymotrypsin digestion as well.

DOI: 10.1021/acsomega.7b00446
ACS Omega 2017, 2, 8129−8140

Furthermore, ErmI, ErmN, and ErmZ except ErmH form one cluster in the phylogenetic tree, implying that they originated from a common ancestor. In addition, ErmI and ErmZ with longer inserted amino acids are located close to each other in the phylogenetic tree[7] (Figure 6b).

DISCUSSION AND CONCLUSIONS

The previous report[13] showed that ErmN was apparently more susceptible than ErmS to the contaminating proteases contained in a commercially available DNase I, possibly indicating a structural difference between mono- and dimethyltransferases. This increased susceptibility to proteolysis could be caused by chymotrypsin and/or trypsin contamination. In regard to trypsin, there is greater variation in the number of susceptibility sites among ErmN, ErmS, and ErmE. However, the levels of chymotrypsin contamination in DNase I prepared from bovine pancreas are higher, and its preferential cleavage sites are similar in number among ErmN, ErmS, and ErmE. The fastest ErmN degradation rate could be ascribed to both the unusually rapid digestion at F45 and the resultant structural perturbation that accelerated the degradation of the resulting polypeptide. Even though there could be many other cleavage products that facilitate further degradation following the exposure of the hydrophobic core of other Erm proteins, the amounts of those products were notably less than band 1 of ErmN. The rapid formation of digestion products and their subsequent degradation could be major points that determine the protease degradation rate of the protein. In this sense, ErmN is quite remarkable because within 30 s of incubation, the first visible truncated product, ErmN band 1 (28.6 kDa), represented greater than 80% of the total band intensity and appeared to be digested at both C- and N-terminal ends as soon as it was formed to yield two degradation products that were approximately 16 and 27 kDa in size. In regard to the generation of the product that led to the exposure of the hydrophobic core and caused more rapid degradation, the half-life of native ErmN in relation to band 1 was more than 67-fold and 104-fold shorter than the half-lives of ErmS and ErmE, respectively, even though the comparison was drawn between one site in ErmN (F45) and several sites in ErmS and ErmE. After the ErmN band 1 was generated, its degradation rate (3.18 min half-life) was estimated to be approximately 1.3- and 4-fold shorter in terms of half-life than the degradation rate of ErmS and ErmE, respectively, in the region outside the core structure. Furthermore, the digestion rate at F45 of ErmN to form band 1 is faster than that of band 1 degradation by more than 15 times and is more than 60 and 20 times faster than the digestion rate of preferential recognition site on the N-terminal unstructured regions of ErmE and ErmS, respectively.

When considering the core structure, the number of preferential recognition sites in the core structure of each Erm protein investigated was close to each other: 18 for ErmE, 20 for ErmN (after F45), and 19 for ErmS. As well, at seven of these sites, the position and identity of the recognition amino acid were the same in the multiple sequence alignment. As well, the position was the same, but the preferential recognition amino acid varied at three other sites (hereafter, these 10 sites will be referred to as the preferential equivalent sites). Therefore, a time-dependent chymotrypsin digestion using standard conditions (an enzyme to substrate ratio of 1:50) could be employed to compare the conformational features surrounding the recognition sites and the chymotrypsin cleavage patterns, if the digestion occurs at the equivalent sites described above.[22,26] Over the course of the chymotrypsin reaction, ErmS exhibited many small bands in each lane in the course of incubation from 30 s to 30 min, with only two or three clearly noticeable bands that were derived from the N-terminal region. This suggests that chymotrypsin binds to many recognition sites and cleaves even without strongly preferred sites. In contrast to this observation, ErmE and ErmN, in particular, appeared to follow the major degradation pathway, displaying apparently discrete bands that resulted from the detection of a unique and distinct pattern of recognition sites during the course of incubation with chymotrypsin. ErmN yielded bands 1, 2, and 3. Band 1 resulted from a deviation (F45) from the consensus sequence that was not a member of the preferential equivalent sites. Likewise, band 2 was not from a preferential site. Band 3 originated from a preferential equivalent site, but only after digestion at F45 and the subsequent structural modification that presumably resulted in the exposure of the site for recognition by chymotrypsin. ErmE yielded bands 1, 2, 3, 4, and 5, all of which were from the native protein. Among these, bands 1 and 2 were derived from the digestion of the N-terminal region, whereas band 3 originated from the preferential equivalent sites [W277, ErmE numbering (W309, column numbering) or F281 (F313) or Y271 (Y303, nonequivalent but shared with ErmS)]. Bands 4 and 5 resulted from the cleavage at Y135 (Y165, one of the preferential equivalent sites) and W144 (174, nonequivalent site but shared with ErmS), respectively. Even though only clearly detectable bands were considered, no consensus cleavage site existed in either the native protein or the structurally perturbed protein after the digestion of ErmE, ErmN, or ErmS at their respective core structures. It is possible that preferential equivalent residues that matched with those from ErmE to generate bands 3, 4, and 5 existed among the digested but uninvestigated ErmS sites. Presumably, these probable bands were not detected because of either their very weak recognition and cleavage or their distinct structural environment, which subsequently resulted in no cleavage. However, after structural perturbation, the digestion of ErmN at F163 (ErmC′ numbering) could be considered unique because it is caused by exposure, which resulted in recognition and cleavage by chymotrypsin after the loss of a particular set of interactions, as described above. ErmE still clearly exhibited these intermediate degradation bands up until incubation with chymotrypsin for 30 min, whereas the ErmN intermediate degradation bands were not clearly visible on the SDS-PAGE gel, indicating that some of the ErmE intermediate states (especially, bands 4 and 5) were more stable than those of ErmN, although the explanation for this observation remains unclear at this time. It was apparent that ErmS exhibited a half-life of the core structure degradation that was more than 1.5-fold shorter than the corresponding half-life in ErmE, suggesting that many bands generated during the reaction with chymotrypsin might result from the digestion of the native protein. ErmN and ErmS both originated from *S. fradiae* and are homologous despite their functional differences as mono- and dimethyltransferases. ErmE and ErmS come from the same *Actinobacteria clade* and are located close to each other on the phylogenetic tree.[7] Despite this close relationship among the Erm proteins investigated, the microenvironments of each protein for chymotrypsin recognition are thought to be slightly different; thus, digestion with chymotrypsin cannot distinguish between the mono- and dimethyltransferase.

DOI: 10.1021/acsomega.7b00446
ACS Omega 2017, 2, 8129−8140

In this study, we intended to pay more attention to the preferential recognition sites (F, W, and Y) by employing sequencing grade chymotrypsin, which exhibited more specificity to these recognition sites than other chymotrypsin grades. Most of the cleavage sites investigated were observed to occur after preferable recognition sites, except D16−R17 or L27−G28, which generated band 1 in ErmE, and A52−E53, E53−S54, A55−G56, or D58−S59, which generated band 2 in ErmN. Although there are many other theoretical loop region positions matching less preferable recognition sites (L, M, A, D, and E), they were not detected in this study. Two less preferential cleavage recognition sites were located in the purportedly unstructured N-terminal region of the substrate polypeptide, which does not contain any preferential susceptibility sites near each site; thus, these two sites could presumably be recognized and truncated by chymotrypsin without any further need for the local unfolding of any secondary structures such as an α helix. The exemption from the need for local unfolding could outweigh the disadvantage associated with less preferable recognition sites, with regard to recognition and digestion by chymotrypsin, and enable faster digestion rates compared with other inefficient preferable sites. However, further investigations are necessary to identify the exact mechanism of preferable truncation at the less favorable recognition sites, especially because the observed speed at which band 2 was generated from ErmN band 1 was quite fast.

Generally, Erm proteins are believed to exhibit a high structural similarity. ErmN and ErmC′ share 20% sequence identity and 46% similarity. Furthermore, the two proteins have a paralogous relationship[7] and perform mono- or dimethylation, respectively. The KsgA/Dim1 family of proteins is the closest to the Erm family, sharing approximately 15−25% amino acid sequence identity. Despite their involvement in the distinct adenine dimethylation reaction, KsgA from *Thermus thermophilus* and ErmC′ exhibit similar structures and could be aligned with a root-mean-square deviation of 1.7 Å for 197 Cα atoms out of 236 residues, with the most notable divergence apparent in the C-terminal domains.[27] This similarity suggests the plasticity of the core structures of these two closely related adenine methyltransferases, which has allowed these orthologous and potentially paralogous proteins to fill multiple and distinct roles with only modest adaptive changes. The extensive phylogenetic analysis generated two clearly separate clusters of Erm proteins: those from *Actinobacteria*, which are composed of antibiotic producers (including ErmN), and those from *Firmicutes*, which represent pathogenic bacteria (containing ErmC′). Although these proteins could form different clusters in the phylogenetic tree, they are closer to each other than KsgA, suggesting that they could assume more similar structures than between the Erm and KsgA/Dim1 protein clusters. Furthermore, it is not possible to differentiate between a monomethyltransferase and a dimethyltransferase based on sequence comparison[28,29] or using a phylogenetic tree construction algorithm.[7] Therefore, the structural analysis of ErmN based on the structure of ErmC′ appeared to be reasonable in this study. In this regard, the location of F45 in ErmN (substituted with isoleucine in ErmC′) and the loop region around it, which existed in the more structurally conserved N-terminal domain, was determined based on the structure of ErmC′.

The unexpectedly high susceptibility of ErmN to chymotrypsin could not be predicted based on a typical or known protein degradation pathway, such as the intrinsically high susceptibility of sinemin to proteolysis,[1,30,31] the proline, glutamic acid, serine, and threonine protein sequence that targets a protein for rapid degradation,[32] or ubiquitin-dependent protein degradation,[33] to name a few. Further, the substitution of isoleucine, leucine, or valine with phenylalanine does not appear to occur by chance because the substitution is found in 4 out of 37 Erm proteins known to date. Even though the origin of the increased susceptibility conferred by the phenylalanine substitution, which likely occurs in conjunction with the juxtaposition of the longer loop region, could not be determined based on the biology of *Streptomyces* at this time and could be a serendipitous event, the fact that ErmI, ErmN, and ErmZ, which carry substitution with F and harbor longer loop region, originate from a common ancestor and ErmZ and ErmI containing the longest loop are the closest member to each other could be a starting point for investigation. The observations made in this study might enable us to define the novel recognition sites leading to unexpectedly high susceptibility to chymotrypsin, which include not only one amino acid (F, Y, and W) immediately before the scissile bond but also structural features around the recognition sites. In addition, hopefully, extremely high susceptibility observed in ErmN and other similar results, which are expected to be acquired, could contribute to elucidate the catalytic mechanism on how remote interaction from the active site could affect the catalytic activity and its regulation of serine protease, which lacks in our understanding of protease action mechanism at present. This understanding will be helpful to grasp the catalytic mechanism of related proteolytic enzymes including the most abundant chymotrypsin-like serine proteases with significant physiological roles and affect the ability to design effective drugs and new therapeutics. Considering the unexpectedly high susceptibility of ErmN to chymotrypsin, successful overexpression of ErmN might not correlate with the amount of purified protein and could result in little to no protein recovery because of the action of the chymotrypsin contaminant in the DNase I preparation used for the removal of DNA released during cell disruption.[13] To this end, the overexpressed ErmN in a prior study (approximately 41 mg/L of induced cell culture, unpublished result) was almost completely degraded by contaminating proteases in the DNase I product during purification (see lane 6 of Figure S1 in ref 13). Most of the commercially available DNase I endonuclease products that are currently marketed are not clearly labeled to indicate the extent of their protease contamination. In fact, 73% of the 101 products surveyed do not provide any information on protease contamination. Compared to the numbers obtained from a survey conducted 2 years ago, nothing appears to have changed with regard to indicating protease contamination on DNase I labels. Therefore, this unusual, unexpected, and thus unpredictable susceptibility to certain protease(s) could pose significant problems and could prevent accurate observations, interpretations, and conclusions during the performance of biological procedures including protein purification and other molecular and biochemical processes including proteases or reagents potentially contaminated with proteases.

MATERIALS AND METHODS

Materials. Luria−Bertani broth (cat. no. 3220142) was obtained from Difco Laboratories (Detroit, MI, USA). The inducer, isopropyl-β-D-thiogalactopyranoside (IPTG, cat. no. MB-I4385), was purchased from MB Cell (Los Angeles, CA, USA). His·Bind resin (cat. no. N71757-6) was obtained from

DOI: 10.1021/acsomega.7b00446
ACS Omega 2017, 2, 8129−8140

Novagen Inc. (Madison, WI, USA). PMSF (cat. no. PMS123.5) was purchased from BioShop Canada Inc. (Burlington, ON, Canada). The Pierce BCA protein assay kit (cat. no. 23225) was obtained from Thermo Fisher Scientific (Waltham, MA, USA). Sequencing grade chymotrypsin (cat. no. 11418467001) was purchased from Sigma-Aldrich (St. Louis, MO, USA). The protein marker (cat. no. EBM-1018) was purchased from Elpis Biotech (Daejeon, Korea). The polyvinylidene difluoride (PVDF) membrane (cat. no. IPVH 000 10) was purchased from Millipore (Billerica, MA, USA). All other reagents were purchased from Bio Basic (Markham, ON, Canada) or from Sigma-Aldrich.

Strains, Templates, and Plasmids. *E. coli* DH5α (Promega, Madison, WI, USA) and BL21(DE3) cells (Novagen Inc., Madison, WI, USA) were used for the general cloning and expression of His_6-tagged ErmE (UniProt accession ID is T2S3U9), ErmN (UniProt accession ID is P97178), and ErmS (UniProt accession ID is P45439). The total DNA from *S. erythraea* NRRL 2338 and *S. fradiae* NRRL 2702 was used as the DNA templates for the polymerase chain reaction (PCR) amplification of *ermE* and *ermN*.[34,35] The pET23b bacterial expression vector (Novagen Inc., Madison, WI, USA) was used to construct *ermE* and *ermN* expression systems.

Protein Sequence Alignment and Phylogenetic Tree Construction. The sequence alignment of several representative Erm proteins, including ErmB, ErmC′, ErmE, ErmN, and ErmS and all 37 Erm protein sequences, was generated using CLUSTALX2,[36] and the resulting alignment was visualized using ESPript,[17] with the secondary structure of ErmC′ (PDB entry 1QAM) on top of the alignment, as a reference. The residue locations were numbered according to the ErmN sequence (above the alignment) and column numbering (below the alignment). The relative degree of accessibility of each position is shown below the alignment according to the structure of ErmC′. The preferential chymotrypsin cleavage sites [tryptophan (W), tyrosine (Y), and phenylalanine (F)] were differentially marked depending on the degree of consensus among the aligned Erm proteins (see the Results and Figure 1 for more details). The phylogenetic tree of 37 Erm proteins known to date was constructed using CLUSTALX2.

Construction of the Expression Vector. The ErmS expression vector was constructed as previously described.[37] Expression vectors for ErmN and ErmE were constructed using the same strategy. Briefly, two oligonucleotides, 5′*GGAATTC*catatgAGCAGTTCGGACGAGCAGCCGCGCCCG3′ and 5′*CCG*ctcgagCCGCTGCCCGGGTCCGCC3′, were used as the forward and reverse primers of ErmE, respectively, and two oligonucleotides, 5′*GGAATTC*catatgCCGTCTCGTCCGCGTACCGATTCGCCCCACCGGCACGAGGG3′ and 5′*CCG*ctcgagGCGCCTCCGCTGCGGCGAGATGCG3′, were used as the forward and reverse primers of ErmN, respectively. The primer sequences were modified to include a restriction site for *Nde*I (5′catatg3′) overlapping the initiation methionine codon and a site for *Xho*I (5′ctcgag3′). In addition, the italicized sequence in each oligonucleotide was added at the end of DNA fragment for cleavage to occur more easily. The resulting PCR products were treated with the restriction enzymes and directly ligated and cloned into the pET23b*Nde*I−*Xho*I sites. Consequently, the overexpressed proteins in this study harbored six histidine residues at the C-terminus, which facilitated effective enzyme purification on a Ni^{2+} affinity column but had no discernible effect on the activity or specificity of the Erm proteins.[37,38] The resulting plasmids were transformed into *E. coli* DH5α cells to obtain a sufficient amount of the plasmids, which were then sequenced to confirm the sequence of the insert and frame. The plasmids were finally transformed into *E. coli* BL21(DE3) cells for the overexpression of the recombinant proteins.

Overexpression and Purification of Recombinant Proteins. Expression and purification of the His_6-tagged Erm proteins were conducted using a previously described strategy, with slight modifications.[13,37] Briefly, *E. coli* cells overexpressing the desired protein by culturing at 22 °C for 18−20 h after IPTG (1 mM) induction were harvested, and each 1 g of wet cell mass was resuspended in 15 mL of buffer A [500 mM KCl, 50 mM Tris-HCl (pH 7.5), and 10 mM $MgCl_2$] containing 5 mM imidazole and 1 mM PMSF. The disruption of the cells was accomplished by sonication on ice using a GEX-130 ultrasonic processor (130 W, 20 kHz) at 50% amplitude for 5 s pulses with 10 s intervals for cooling. The total sonication time was 5 min. The cell lysate was centrifuged at 11 000*g* at 4 °C for 15 min, and the clarified supernatant was applied to a column containing 2 mL of nickel (Ni^{2+})-charged affinity resin (Novagen, Madison, WI, USA) according to the manufacturer's instructions. The column was then washed with buffer A containing 25 mM imidazole, and the recombinant protein was eluted using buffer A containing 300 mM imidazole. The eluted fraction was dialyzed overnight at 4 °C in buffer B [50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10 mM $MgCl_2$, and 50% (v/v) glycerol]. The resulting protein concentrations were determined in triplicate using the Pierce BCA protein assay kit with bovine serum albumin as the protein standard (cat. no. 23225, Thermo Fisher Scientific, MA, USA).

Proteolytic Digestion Assay. The purified Erm proteins (10 μg) were treated with sequencing grade chymotrypsin (cat. no 11418467001, Sigma-Aldrich) at an enzyme/substrate ratio of 1:50 in 10 μL of 50 mM Tris-HCl buffer (pH 8.0) with 50 mM $CaCl_2$. After incubating at 37 °C for 0.5, 1, 2, 3, 5, 10, 15, and 30 min, the reaction was stopped by the addition of PMSF to a final concentration of 5 mM, and the mixture was boiled for 5 min. The samples were then subjected to 12% SDS-PAGE, and the resulting gel was stained with Coomassie brilliant blue R-250.[39]

Protein Blotting and N-Terminal Sequencing. After SDS-PAGE, the proteins/peptides were transferred to a 0.45 μm PVDF membrane using a TE 22 Mighty Small tank transfer system (Amersham Biosciences, Piscataway, NJ, USA). Briefly, the transfer was carried out using Towbin buffer (25 mM Tris, 192 mM glycine, and 20% v/v methanol, pH 8.3) at a constant current mode of 0.2 A in a cold chamber (Jisico, Seoul, Korea) for 6 h. The proteins were visualized by staining with 0.1% Coomassie brilliant blue R-250, followed by destaining in 50% methanol. The N-terminus of each protein was sequenced by Edman degradation using a Procise 491 HT protein sequencer (Applied Biosystems, Foster City, CA, USA) in accord with the experiments completed in a previous study.[13] In a separate experiment, the degraded products obtained in the prior study[13] and obtained in this study were subjected to SDS-PAGE to determine the identical peptide bands.

Estimating the Molecular Weights of the Proteins/Peptides. The molecular weight of the Erm proteins/degradation product peptides was deduced from the peptide sequence using the ExPASy ProtParam tool.[19] The calculation included the sequence incorporated at the C-terminus of each protein, including the hexa-His-tag, during the construction of the expression vector. This was also true for the digested

DOI: 10.1021/acsomega.7b00446
ACS Omega 2017, 2, 8129−8140

peptides retained on the Ni^{2+} affinity column during purification. As an alternative, the molecular weight of the Erm proteins/degradation product peptides was also estimated using a procedure recommended by G-Bioscience.[20] The sample was processed as described above and analyzed on a 12% SDS-PAGE gel using a set of commercially available protein markers (cat. no. EBM-1018, Elpis Biotech, Daejeon, Korea). The molecular weights of the Erm proteins/degradation product peptides were then determined based on a standard curve generated by plotting the log molecular weight versus the migration distance of the protein markers and the Erm proteins/their identified digestion peptides. The method typically yields an accuracy that deviates less than 5% for a band larger than 15 kDa.

Analyzing the Proteolytic Digestion Kinetics. Proteolytic degradation kinetics were determined by measuring the protein/peptide band intensities on an SDS-PAGE gel using the GelQuant.NET software provided by biochemlabsolutions.com. The intensity of the band corresponding to the untreated proteins served as a control for 100% protection against chymotrypsin action unless otherwise indicated (see Results). The band intensities were plotted against the incubation time to monitor the degradation kinetics, and the resultant curves appeared to follow the first-order degradation kinetics. The rate constant and half-life of the proteolytic degradation were determined by fitting one-phase decay to the above data using GraphPad Prism version 7.00 for Windows (exhibiting *R*-squared values ranging from 0.95 to 0.995, GraphPad software, La Jolla, CA, USA, www.graphpad.com).

AUTHOR INFORMATION

Corresponding Author

*E-mail: hjjin@suwon.ac.kr. Phone: 82-31-220-2290. Fax: 82-31-220-2519.

ORCID

Hyung Jong Jin: 0000-0002-8097-2250

Author Contributions

H.J.J. and H.J.L. conceived the experiment and together with T.L. carried it out; H.J.J. and T.L. designed and carried out the data analysis; and H.J.J. coordinated the study and wrote the paper. Figures 1, 3, 4, and 6a,b were prepared by H.J.J., and Figures 2 and 5 were prepared by T.L.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The project described was supported by grants from Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration (grants no. C0350515 to H.J.J.) in 2015 and the Basic Science Research Program (2010-0011442 to H.J.J.) through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology.

ABBREVIATIONS

Erm, erythromycin ribosome methylation; MLS_B antibiotics, macrolide−lincosamide−streptogramin B antibiotics; IPTG, isopropyl β-D-thiogalactopyranoside; PMSF, phenylmethylsulfonyl fluoride; PVDF, polyvinylidene fluoride

REFERENCES

(1) Bilak, S. R.; Sernett, S. W.; Bilak, M. M.; Bellin, R. M.; Stromer, M. H.; Huiatt, T. W.; Robson, R. M. Properties of the novel intermediate filament protein synemin and its identification in mammalian muscle. *Arch. Biochem. Biophys.* **1998**, *355*, 63−76.

(2) Taniguchi, T.; Kobayashi, T.; Kondo, J.; Takahashi, K.; Nakamura, H.; Suzuki, J.; Nagai, K.; Yamada, T.; Nakamura, S.; Yamamura, H. Molecular cloning of a porcine gene syk that encodes a 72-kDa protein-tyrosine kinase showing high susceptibility to proteolysis. *J. Biol. Chem.* **1991**, *266*, 15790−15796.

(3) Reyes-Vivas, H.; Martínez-Martínez, E.; Mendoza-Hernández, G.; López-Velázquez, G.; Pérez-Montfort, R.; de Gómez-Puyou, M. T.; Gómez-Puyou, A. Susceptibility to proteolysis of triosephosphate isomerase from two pathogenic parasites: characterization of an enzyme with an intact and a nicked monomer. *Proteins: Struct., Funct., Bioinf.* **2002**, *48*, 580−590.

(4) Rawlings, N. D. A large and accurate collection of peptidase cleavages in the MEROPS database. *Database* **2009**, *2009*, bap015.

(5) Lai, C. J.; Weisblum, B. Altered methylation of ribosomal RNA in an erythromycin-resistant strain of Staphylococcus aureus. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 856−860.

(6) Skinner, R.; Cundliffe, E.; Schmidt, F. J. Site of action of a ribosomal RNA methylase responsible for resistance to erythromycin and other antibiotics. *J. Biol. Chem.* **1983**, *258*, 12702−12706.

(7) Park, A. K.; Kim, H.; Jin, H. J. Phylogenetic analysis of rRNA methyltransferases, Erm and KsgA, as related to antibiotic resistance. *FEMS Microbiol. Lett.* **2010**, *309*, 151−162.

(8) Bussiere, D. E.; Muchmore, S. W.; Dealwis, C. G.; Schluckebier, G.; Nienaber, V. L.; Edalji, R. P.; Walter, K. A.; Ladror, U. S.; Holzman, T. F.; Abad-Zapatero, C. Crystal structure of ErmC′, an rRNA methyltransferase which mediates antibiotic resistance in bacteria. *Biochemistry* **1998**, *37*, 7103−7112.

(9) Schluckebier, G.; Zhong, P.; Stewart, K. D.; Kavanaugh, T. J.; Abad-Zapatero, C. The 2.2 Å structure of the rRNA methyltransferase ErmC′ and its complexes with cofactor and cofactor analogs: implications for the reaction mechanism. *J. Mol. Biol.* **1999**, *289*, 277−291.

(10) Yu, L.; Petros, A. M.; Schnuchel, A.; Zhong, P.; Severin, J. M.; Walter, K.; Holzman, T. F.; Fesik, S. W. Solution structure of an rRNA methyltransferase (ErmAM) that confers macrolide-lincosamide-streptogramin antibiotic resistance. *Nat. Struct. Biol.* **1997**, *4*, 483−489.

(11) Liu, M.; Douthwaite, S. Resistance to the macrolide antibiotic tylosin is conferred by single methylations at 23S rRNA nucleotides G748 and A2058 acting in synergy. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14658−14663.

(12) Weisblum, B. Erythromycin resistance by ribosome modification. *Antimicrob. Agents Chemother.* **1995**, *39*, 577−585.

(13) Le, T.; Lee, H. J.; Jin, H. J. An efficient method to eliminate the protease activity contaminating commercial bovine pancreatic DNase I. *Anal. Biochem.* **2015**, *483*, 4−6.

(14) Cundliffe, E. How antibiotic-producing organisms avoid suicide. *Annu. Rev. Microbiol.* **1989**, *43*, 207−233.

(15) Otsuka, A. S.; Price, P. A. Removal of proteases from DNase I by chromatography over agarose with covalently attached lima bean protease inhibitor. *Anal. Biochem.* **1974**, *62*, 180−187.

(16) Wang, D.; Moore, S. Preparation of protease-free and ribonuclease-free pancreatic deoxyribonuclease. *J. Biol. Chem.* **1978**, *253*, 7216−7219.

(17) Robert, X.; Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **2014**, *42*, W320−W324.

(18) Keil, B. *Specificity of Proteolysis*, 1st ed.; Springer-Verlag Berlin Heidelberg: NewYork, 1992; pp 206−214.

(19) Artimo, P.; Jonnalagedda, M.; Arnold, K.; Baratin, D.; Csardi, G.; de Castro, E.; Duvaud, S.; Flegel, V.; Fortier, A.; Gasteiger, E.; Grosdidier, A.; Hernandez, C.; Ioannidis, V.; Kuznetsov, D.; Liechti, R.; Moretti, S.; Mostaguir, K.; Redaschi, N.; Rossier, G.; Xenarios, I.; Stockinger, H. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* **2012**, *40*, W597−W603.

DOI: 10.1021/acsomega.7b00446
ACS Omega 2017, 2, 8129−8140

(20) Man, P. Determining Protein Molecular Weight with SDS-PAGE: An Overview of the Process. Web Blog Post. The Protein Man's Blog. G-Biosciences. http://info.gbiosciences.com/blog/bid/196974/Determining-Protein-Molecular-Weight-with-SDS-PAGE-An-Overview-of-the-Process, 2014.
(21) Tina, K. G.; Bhadra, R.; Srinivasan, N. PIC: Protein Interactions Calculator. *Nucleic Acids Res.* **2007**, *35*, W473−W476.
(22) Hubbard, S. J. The structural aspects of limited proteolysis of native proteins. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1998**, *1382*, 191−206.
(23) Epp, J. K.; Burgett, S. G.; Schoner, B. E. Cloning and nucleotide sequence of a carbomycin-resistance gene from Streptomyces thermotolerans. *Gene* **1987**, *53*, 73−83.
(24) Hara, O.; Hutchinson, C. R. Cloning of midecamycin (MLS)-resistance genes from Streptomyces mycarofacines, Streptomyces lividans and Streptomyces coelicolor A3(2). *J. Antibiot.* **1990**, *43*, 977−991.
(25) Karray, F.; Darbon, E.; Oestreicher, N.; Dominguez, H.; Tuphile, K.; Gagnat, J.; Blondelet-Rouault, M.-H.; Gerbaud, C.; Pernodet, J.-L. Organization of the biosynthetic gene cluster for the macrolide antibiotic spiramycin in Streptomyces ambofaciens. *Microbiology* **2007**, *153*, 4111−4122.
(26) Fontana, A.; de Laureto, P. P.; Spolaore, B.; Frare, E.; Picotti, P.; Zambonin, M. Probing protein structure by limited proteolysis. *Acta Biochim. Pol.* **2004**, *51*, 299−321.
(27) Demirci, H.; Belardinelli, R.; Seri, E.; Gregory, S. T.; Gualerzi, C.; Dahlberg, A. E.; Jogl, G. Structural rearrangements in the active site of the Thermus thermophilus 16S rRNA methyltransferase KsgA in a binary complex with 5′-methylthioadenosine. *J. Mol. Biol.* **2009**, *388*, 271−282.
(28) Pernodet, J. L.; Fish, S.; Blondelet-Rouault, M. H.; Cundliffe, E. The macrolide-lincosamide-streptogramin B resistance phenotypes characterized by using a specifically deleted, antibiotic-sensitive strain of Streptomyces lividans. *Antimicrob. Agents Chemother.* **1996**, *40*, 581−585.
(29) Pernodet, J.-L.; Cundliffe, E.; Gourmelen, A.; Blondelet-Rouault, M. H. Dispensable ribosomal resistance to spiramycin conferred by srmA in the spiramycin producer Streptomyces ambofaciens. *Microbiology* **1999**, *145*, 2355−2364.
(30) Granger, B. L.; Lazarides, E. Synemin: a new high molecular weight protein associated with desmin and vimentin filaments in muscle. *Cell* **1980**, *22*, 727−738.
(31) Sandoval, I. V.; Colaco, C. A.; Lazarides, E. Purification of the intermediate filament-associated protein, synemin, from chicken smooth muscle. Studies on its physicochemical properties, interaction with desmin, and phosphorylation. *J. Biol. Chem.* **1983**, *258*, 2568−2576.
(32) Rogers, S.; Wells, R.; Rechsteiner, M. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* **1986**, *234*, 364−368.
(33) Hochstrasser, M. Ubiquitin-dependent protein degradation. *Annu. Rev. Genet.* **1996**, *30*, 405−439.
(34) Labeda, D. P. Transfer of the Type Strain of Streptomyces erythraeus (Waksman 1923) Waksman and Henrici 1948 to the Genus Saccharopolyspora Lacey and Goodfellow 1975 as Saccharopolyspora erythraea sp. nov., and Designation of a Neotype Strain for Streptomyces erythraeus. *Int. J. Syst. Bacteriol.* **1987**, *37*, 19−22.
(35) Zalacain, M.; Cundliffe, E. Cloning of tlrD, a fourth resistance gene, from the tyiosin producer, Streptomyces fradiae. *Gene* **1991**, *97*, 137−142.
(36) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947−2948.
(37) Jin, H. J.; Yang, Y. D. Purification and biochemical characterization of the ErmSF macrolide−lincosamide−streptogramin B resistance factor protein expressed as a hexahistidine-tagged protein in Escherichia coli. *Protein Expression Purif.* **2002**, *25*, 149−159.
(38) Vester, B.; Nielsen, A. K.; Hansen, L. H.; Douthwaite, S. ErmE methyltransferase recognition elements in RNA substrates. *J. Mol. Biol.* **1998**, *282*, 255−264.
(39) Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *277*, 680−685.

DOI: 10.1021/acsomega.7b00446
ACS Omega 2017, 2, 8129−8140