

Processing of *Escherichia coli* Alkaline Phosphatase

SEQUENCE REQUIREMENTS AND POSSIBLE CONFORMATIONS OF THE –6 TO –4 REGION OF THE SIGNAL PEPTIDE*

Andrey V. Kajava^{‡§}, Sergey N. Zolov[¶], Konstantin I. Pyatkov^{||}, Andrey E. Kalinin^{¶**},
and Marina A. Nesmeyanova^{§¶}

From the [‡]Center for Molecular Modeling, CIT, National Institutes of Health, Bethesda, Maryland 20892, the [¶]Laboratory of Protein Secretion in Bacteria, Skryabin Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia, and the ^{||}Institute of Cellular Biophysics, 142290 Pushchino, Russia

Analysis of the precursors of bacterial exported proteins revealed that those having bulky hydrophobic residues at position –5 have a high incidence of Pro residues at positions –6 and –4, Val at position –3, and Ser at positions –4 and –2. This led to a hypothesis that the previously observed inhibition of processing by bulky residues at position –5 can be suppressed by introduction of Pro, Ser, or Val in the corresponding nearby positions. Subsequent mutational analysis of *Escherichia coli* alkaline phosphatase showed that, as it was predicted, Pro on either side of bulky hydrophobic –5 Leu, Ile, or Tyr completely restores efficiency of the maturation. Introduction of Val at position –3 also partially suppresses the inhibition imposed by –5 Leu, while a Ser residue at position –4 or –2 does not restore processing. In addition, effective maturation of a mutant with Pro residues at positions from –6 throughout –4 proved that polyproline conformation of this region is permissive for processing. To understand the effects of the mutations, we modeled a peptide substrate into the active site of the signal peptidase using the known position of the β -lactam inhibitor. The inhibitory effect of the –5 residue and its suppression by either Pro –6 or Pro –4 can be explained if we assume that Pro-containing –6 to –4 regions adopt a polyproline conformation whereas the region without Pro residues has a β -conformation. These results permit us to specify sequence requirements at –6, –5, and –4 positions for efficient processing and to improve the prediction of yet unknown cleavage sites.

In prokaryotes and eukaryotes, most exported proteins are synthesized as precursors with an amino-terminal extension called the leader or signal peptide. The signal peptide directs protein translocation across membranes and is removed by a membrane-bound peptidase after transition through the membrane (1). Despite their common purpose, signal peptides have

very little amino acid sequence similarity, although they do share general features. Typically 15–30 amino acids long, signal peptides of prokaryotic proteins consist of three distinct regions: a 1–5-residue amino-terminal positively charged segment, a 10–15-residue central hydrophobic core, and a more polar 5–7-residue carboxyl-terminal cleavage region (c-region)¹ (2, 3). In addition, most bacterial proteins have a 14–18-residue region in the mature part immediately downstream of the signal sequence, which has a negative or neutral net charge (4–6). As a result of extensive research over the last two decades, the role of each region of the exported proteins has been mainly elucidated (for review see Ref. 7). The export of proteins is initiated by interactions of the positively charged amino terminus with negatively charged phospholipid head-groups of the cytoplasmic membrane (8–12) and by insertion of the hydrophobic core of the signal peptide into the apolar environment of the membrane (3, 13). The insertion of the signal peptide into the lipid bilayer proceeds in association with proteins of the Sec translocation machinery (7, 14, 15). The positive charge of the amino terminus can also govern the N_{in}-C_{out} orientation of the signal peptide within the membrane (16). In this orientation, the c-region of the signal peptide is exposed on the periplasmic side where it can be recognized and cleaved by the signal peptidase (SPase) between positions –1 and +1 (1, 17, 18). A sequence motif with small residues at positions –3 and –1 defines the cleavage site (2, 3, 19). The conformational characteristics of the signal peptide are also mainly established. There is a consensus view based on several *in vitro* experimental studies (20, 21) that the region of the signal peptide inserted into the membrane adopts a α -helical conformation. It is now known that the –3 to –1 region has an extended β -structural conformation, which is recognized by SPase (19, 22). Despite this progress, the critical physical and structural characteristics of residues –6, –5, and –4 that delineate the hydrophobic core and peptidase recognition site of the signal peptide are still poorly understood. Bacterial signal peptides frequently have α -helix-breaking residues such as proline and glycine at –6 to –4 positions (16), and this suggested that the disruption of the helical conformation in this region is an important requirement for efficient processing. A number of experimental data supported this conclusion (23, 24). Based on the analysis of natural sequences (16) and experimental evidence, it was also proposed that the hydrophilicity of this region rather than its conformation may be important

* This study was supported in part by Grants 99-04-48153, 02-04-06304, and 02-04-49765 from the Russian Foundation for Basic Research. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence may be addressed: CRBM-CNRS UPR 1086, 1919, route de Mende, 34293 Montpellier, Cedex 5, France. Fax: 33-4-67521559; E-mail: kajava@crbm.cnrs-mop.fr or E-mail: aniram@ibpm.serpukhov.su (for biological samples).

** Present address: Laboratory of Skin Biology, NIAMS, National Institutes of Health, Bldg. 50, Rm. 1527, Bethesda, MD 20892.

¹ The abbreviations used are: c-region, cleavage region; PhoA, alkaline phosphatase; pre-PhoA, alkaline phosphatase precursor; SPase, bacterial type I signal peptidase; RMS, root-mean-square.

for the maturation (25). However, none of these rules has absolute support from the recent collection of natural sequences: there are exported proteins with c-regions consisting of only apolar or helix-fostering residues. The conformation of the -4 to -6 region is also unknown. The Pro, Gly, and Ser residues that frequently occupy -4, -5, and -6 positions (2, 16) are typical for β -turns of globular proteins (26). This observation resulted in a widely accepted opinion that this region has a β -turn conformation (2, 27). Furthermore, some mutagenesis studies of exported proteins showed that a decrease of the processing efficiency in mutant proteins correlates with a low probability of β -turn formation (24, 28). However, it was shown that when Pro residues are simultaneously present at both the -5 and -4 positions of alkaline phosphatase from *E. coli*, this protein is processed properly (29). The steric constraints of this Pro-Pro tandem allow only a β -conformation of the -5 residue, and, as a consequence, this result cast doubt on the presence of the β -turn in the -6 to -4 region. Rather, it was suggested that the c-region has an extended β -conformation (29). In this conformation, the -5 residue may have contact with SPase, and this can explain why the processing is sensitive to the size of the -5 residue (29). The determination of the three-dimensional structure of the bacterial type I SPase co-crystallized with its inhibitor (19) allows a final rejection of the β -turn hypothesis and favors the extended conformation of the c-region. However, despite the knowledge of the active site of the SPase and docking of the peptide substrate into its binding pocket, the exact conformation of the -6 to -4 region remains unknown. This could be considered a minor academic problem if it was not known that amino acid substitutions within this region can significantly diminish or even block the maturation of exported proteins (23, 29–31).

The goal of this work is to define the sequence requirements and conformation of the -6 to -4 region and its interactions with SPase during the processing. We approached this problem by using sequence analysis of exported proteins, mutational analysis, and molecular modeling.

EXPERIMENTAL PROCEDURES

Sequence Analysis—Sequences from Gram-negative bacteria were taken from SwissProt using Sequence Retrieval System software (www.ebi.ac.uk/srs/) and then checked manually. They were 110 proteins from *E. coli* (68 with known and 42 with well-predicted cleavage sites) and 81 proteins from other Gram-negative bacteria with known cleavage sites. The collection did not include highly homologous sequences with more than 80% identity. Anomalous signal sequences (those whose lengths of the hydrophobic core did not fall into the range between 7 and 17 residues), and proteins, secreted by other or modified secretion machineries (hydrogenases having a RR**F*K pattern within the signal sequence, where * denotes any residue Ref. 32; pili, Ref. 15; and lipoproteins) were also excluded. The collection of the 191 sequences is available over the World Wide Web (cmm.cit.nih.gov/kajava/gram-negat.dat). The data sets of 114 exported proteins from Gram-positive bacteria and 1011 human exported proteins have been taken from the SIGNALP data base www.cbs.dtu.dk/services/SignalP/sp_matrices.html (33).

Bacterial Strains and Plasmids—*E. coli* strain E15 (*Hfr* Δ phoA8 *fadL701 tonA22 garB10 ompF627 relA1 pit-10spoT1T2*) (34) was used as a host strain for the expression of wild-type and mutant *phoA* genes cloned in plasmids. *E. coli* strain Z85 (*thi* Δ (*lac-proAB*) Δ (*srl-recA*) *hsdR::Tn10* (*F'* *traD proAB lacI^qZM15*)) (35) was used to construct mutant *phoA* genes.

Wild-type alkaline phosphatase gene (*phoA*) was cloned into *Hind*III/*Bam*HI sites of vector p15SK(-) containing multiple cloning sites identical to pBluescript SK (Stratagene), p15A *ori* of replication and chloramphenicol-acetyltransferase gene.² The resulting phagemid was used to construct and express mutant *phoA* genes. Helper phage R408 was used to isolate single-strand recombinant phagemids. The plasmid har-

boring the gene of amber suppressor tRNA^{Ala} from *E. coli* in the vector pGFTB (36) was provided by Dr. J. Miller.

Media and Culture Conditions—Bacteria for cloning and oligonucleotide-directed mutagenesis were grown on LB or 2YT medium at 37 °C. All media were supplemented with 25 μ g/ml chloramphenicol to either select for or maintain *phoA*-containing plasmids. To screen for colonies expressing active alkaline phosphatase, *E. coli* cells were grown on agar plates made of LB medium free of inorganic phosphate and containing 40 μ g/ml 5-bromo-4-chloro-3-indolyl-phosphate (37). For alkaline phosphatase expression, cells were grown on minimal medium (38) with 1 mM K₂HPO₄ and 0.1% peptone to the mid-log phase and transferred to medium without orthophosphate and peptone.

Oligonucleotide-directed Mutagenesis—To generate mutant forms of *phoA*, we used a new two step method, which allowed us to omit hybridization with labeled nucleotides during selection of clones containing mutant genes (6, 39). Isolation of single-strand phagemid DNA and plasmid DNA, electrophoresis of DNA fragments in agar gels, phosphorylation of oligonucleotides, and transformation of *E. coli* cells were performed by standard procedures (40). Mutations (Table I) were confirmed by DNA sequencing (41).

Alkaline Phosphatase Maturation—Pulse-chase experiments were used to analyze the alkaline phosphatase maturation. *E. coli* cells grown to the mid-log phase in the minimal medium with 1 mM K₂HPO₄ were harvested, washed, and incubated for 10 min in the same medium without orthophosphate to induce alkaline phosphatase synthesis. The cells were labeled with 50 μ Ci/ml [³⁵S]methionine for 60 s and chased for 0.1, 1.0, 5.0, or 60.0 min by addition of unlabeled methionine to a final concentration of 0.05%. Proteins were precipitated with 10% trichloroacetic acid. Alkaline phosphatase and its precursor were immunoprecipitated with rabbit antibodies and separated by 10% SDS-PAGE followed by autoradiography. Proteins were quantified using a LKB UltraScan laser densitometer. The relative quantity of mature alkaline phosphatase and its precursor was calculated with adjustment for the difference in number of methionine residues between the precursor and mature form.

Alkaline Phosphatase Isoforms and Activity—Cells expressing alkaline phosphatase were harvested and converted to spheroplasts in 20 mM Tris-HCl, pH 7.5, 10 mM EDTA, 50 mM sucrose, and 1 mg/ml lysozyme for 15 min on ice. Periplasmic fraction was separated from the cell debris by centrifugation at 12,000 \times g for 5 min. The samples were analyzed by non-denaturing electrophoresis in 7.5% PAGE (42). Staining of the alkaline phosphatase isoforms was performed by incubation of the gel with α -naphthyl phosphate (Sigma, N-7255) and Fast Red Dye TR (Chemapol, Czech Republic) (43). The alkaline phosphatase activity was determined by measuring the rate of *p*-nitrophenylphosphate hydrolysis, taking the activity of hydrolysis of 1 μ mol of substrate per 1 min at 37 °C as a unit of enzymatic activity (unit). Total cell protein was assayed by the Lowry method (44).

Molecular Modeling—Initial docking of a peptide corresponding to the -3 to +1 region of the alkaline phosphatase into the active site of SPase was made manually based on the known position of the β -lactam inhibitor and using Insight II program (45). Possible conformations of the region -6 to -4 were selected based on two constraints: first, the absence of steric clashes within the peptide chain and between the peptide and SPase; second, direction of signal peptide α -helix (residues -21 to -7) into the cytoplasmic membrane. Then the complexes between SPase and alkaline phosphatase precursor (-21 to +2) were subjected to energy minimization using DISCOVER module of Insight II (300 steps of minimization based on the steepest descent algorithm and the next 500 steps using conjugate gradients algorithm). The CHARMM force field (46) and the distance-dependent dielectric constant were used for the energy calculations. During the minimization (i) the backbone atoms of SPase were tethered to their positions in the crystal structure, (ii) a carbonyl carbon atom in the -1 residue was covalently linked to the oxygen atom of the Ser-90 side chain forming a tetrahedral intermediate, and (iii) several hydrogen bonds (between the oxygen of the peptide group of -1 residue and hydroxyl group of Ser-88, between the backbone oxygen of -2 residue and NH group of Ile-144, between the backbone nitrogen of -2 residue and backbone oxygen of Asp-142) were enforced by setting the distance constraints with moderate force ($K = 50$), in order to improve their geometry. In addition, when the region -6 to -4 in the β -conformation was energy minimized the distance constraints were imposed on hydrogen bonds between the backbone CO group of Gln-85 and NH group of the -3 residue; the backbone NH group of Gln-85 and CO group of -4 residue; the CO group of Pro-83 and NH group of -5 residue. To allay the concern that these constraints generated significant tension in the minimized structure, the last calculation was performed without any restrictions to an

² R. Fischer and W. Hengstenberg, unpublished observations.

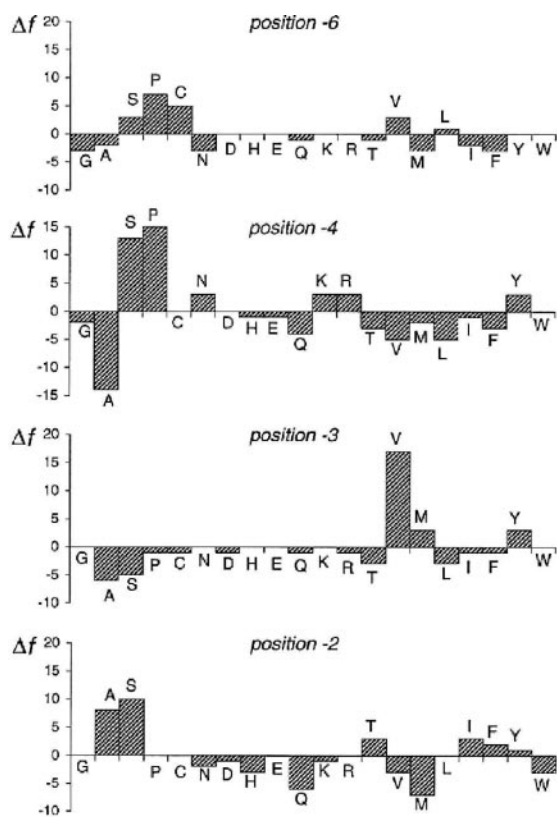


FIG. 1. The difference between frequencies of amino acids (Δf) calculated for all exported proteins of Gram-negative bacteria and a subset of these proteins having bulky hydrophobic residues (Trp, Ile, Phe, Leu, Met, and Tyr) at the -5 position. $\Delta f_{aa} = (f_{aa_bulky} - f_{aa_total}) \times 100$, where f_{aa_total} is a count of a given amino acid in a complete collection of the bacterial proteins divided by the number of proteins in this collection; f_{aa_bulky} is a count of a given amino acid among the bacterial proteins with -5 bulky hydrophobic residues divided by the number of such proteins.

RMS derivative of 0.4 kcal/(mol·Å). A module "Struct_Check" of Insight II program (45) was used to check the quality of the modeled complexes. The figures were generated with Molscript (47).

RESULTS

Rationale of the Selected Amino Acid Substitutions in the -6 to -4 Region—Our previous study showed that the introduction of bulky residues at position -5 of *E. coli* alkaline phosphatase causes a decrease in the efficiency of its maturation (29). In agreement with this result, the occurrence of large hydrophobic residues Trp, Ile, Phe, Leu, Met, and Tyr (written in the order of decreasing hydrophobicity, Ref. 48) at position -5 for the c-regions of eukaryotic and Gram-positive bacterial proteins is lower than at positions -6 and -4 (17% versus 45 and 27%, and 7% versus 20 and 18% correspondingly). Surprisingly, the exported proteins of Gram-negative bacteria have an opposite distribution: 21% of large residues at position -5 against 14 and 17% at positions -6 and -4 . We analyzed a subset of bacterial proteins with bulky hydrophobic residues at position -5 and found that they have a higher incidence of Pro residue at positions -6 and -4 , Val residue at position -3 , and a Ser residue at positions -4 and -2 compared with the complete collection of these proteins (Fig. 1). This observation led to the hypothesis that the inhibitory effect of bulky hydrophobic residue in position -5 can be suppressed by introducing Pro, Ser, or Val in the corresponding nearby positions. In accordance with this, a series of mutant *E. coli* alkaline phosphatases were obtained (Table I) to test the hypothesis. In addition, a mutant having Pro residues in all -6 , -5 , and -4 positions was also obtained. The -6 to -4 region of such a protein is

sterically constrained in the polyproline conformation, and it was of interest to determine whether it was processed. The fact that this tandem of three Pro was not found in natural sequences provided an additional motivation to study this mutant.

Effect of the Mutations on Translocation and Processing of the Alkaline Phosphatase—All mutant proteins were enzymatically active in cells (data not shown). This result implies that the mutants were translocated across the cytoplasmic membrane, because it is known that alkaline phosphatase becomes active only after translocation into the periplasm, where disulfide bond formation and enzyme dimerization take place (49).

The effect of the amino acid substitutions on alkaline phosphatase maturation was assessed by the rate of conversion of a pulse-labeled mutant protein precursor into the mature form *in vivo* using the standard pulse-chase method. As shown in Fig. 2, the presence of bulky Leu, Ile, or Tyr residue at position -5 (proteins L(-5), I(-5) and Y(-5)) notably impaired the maturation of the precursor in comparison with wild-type protein. Even after 60 min of chase almost half of the mutant protein precursor remained unprocessed. In agreement with our hypothesis, introduction of a Pro residue on either side of -5 Leu, Ile, or Tyr restored efficiency of maturation (proteins P(-6)-L(-5), L(-5)P(-4), P(-6)I(-5), I(-5)P(-4), P(-6)Y(-5), and Y(-5)P(-4)). Introduction of a Ser residue at position -4 or -2 or Val residue at position -3 also partially suppressed the effect of the -5 Leu mutation (proteins L(-5)S(-4), L(-5)S(-2) and L(-5)V(-3), correspondingly). Pre-PhoA with a stretch of Pro residues in positions from -6 to -4 (protein P(-6 , -5 , -4)), was converted into the mature form with almost the same efficiency as the wild-type precursor.

An unprocessed protein can reside in the cytoplasm or be translocated to the periplasmic side. It is known that the unprocessed but translocated precursor of *E. coli* alkaline phosphatase has an enzymatic activity, while precursor, which remains in the cytoplasm, does not (50). This property of alkaline phosphatase was used to distinguish which of these two situations is true for the unprocessed portions of the analyzed mutants. We visualized alkaline phosphatase isoforms (I, II, and III) in gel after electrophoresis under non-denaturing conditions. Active alkaline phosphatase can be stained in the gel by treatment with the enzyme substrate α -naphthyl phosphate and an appropriate dye. Furthermore, active mature protein and the translocated precursor can be distinguished on non-denaturing gel, since they have different electrophoretic mobilities (29, 51). The precursor translocated across the membrane can be found at the top of the gel, probably due to aggregation caused by the presence of hydrophobic signal peptide. Such active precursor was detected (Fig. 3, a series of mutants containing Leu in -5 position are shown) in all cases when significant amount of unprocessed pre-PhoA was present after 60 min of chase (Fig. 2, proteins L(-5), I(-5), Y(-5), L(-5)S(-4), and L(-5)S(-2)). This implies that these mutant precursors are located at the periplasmic side of the cytoplasmic membrane. Thus, we showed that inefficient processing of L(-5), I(-5), Y(-5), L(-5)S(-4), and L(-5)S(-2) proteins is due to the failure of their recognition or cleavage by SPase, but not translocation across the membrane.

Molecular Docking of the Peptide Substrate into the Binding Pocket of SPase—Docking of the signal peptide into the SPase active site was used to understand the molecular mechanism of the effects caused by the mutations. We modeled a peptide corresponding to the c-region of the alkaline phosphatase precursor into the active site of the enzyme by using the known position of the β -lactam inhibitor (19). The -3 to -1 region of the peptide substrate fits the binding pocket only when it has

TABLE I
 Mutant *E. coli* alkaline phosphatases

Protein	-6	-5	-4	-3	-2	-1	Mutagenic oligonucleotides
Wild-type	T	P	V	T	K	A	
L(-5)	T	L	V	T	K	A	5'-GGCTTTTGTCCACCAGGGTAAACAGTAACG-3'
P(-6)L(-5)	P	L	V	T	K	A	5'-GCTTTTGTCCACCAGCGGAAACAGTAACGGTAAG-3'
L(-5)P(-4)	T	L	P	T	K	A	5'-GGCTTTTGTCCGGCAGGGTAAACAGTAAC-3'
L(-5)S(-4)	T	L	S	T	K	A	5'-CGGGCTTTTGTGGACAGGGTAAACAGTAACG-3'
L(-5)V(-3)	T	L	V	V	K	A	5'-CCGGCTTTTAAACCACAGGGTAAACAGTAACG-3'
L(-5)S(-2)	T	L	V	T	S	A	5'-GTGTCCGGGCAGATGTCCACCAGGGTAAACAGTAAC-3'
I(-5)	T	I	V	T	K	A	5'-GGCTTTTGTCCACAATGGTAAACAGTAACG-3'
P(-6)I(-5)	P	I	V	T	K	A	5'-GGCTTTTGTCCACAATCGGAAACAGTAACGGTAAG-3'
I(-5)P(-4)	T	I	P	T	K	A	5'-CCGGCTTTTGTCCGGAATGGTAAACAGTAACGGTA-3'
Y(-5)	T	Y	V	T	K	A	5'-GGCTTTTGTCCACGTAGGTAAACAGTAACG-3'
P(-6)Y(-5)	P	Y	V	T	K	A	5'-GGCTTTTGTCCACGTACGGAAACAGTAACGGTAAG-3'
Y(-5)P(-4)	T	Y	P	T	K	A	5'-CCGGCTTTTGTCCGGTAGGTAAACAGTAACGGTA-3'
P(-6, -5, -4)	P	P	P	T	K	A	5'-CGGGCTTTTGTCCGGCGGAAACAGTAACGG-3'

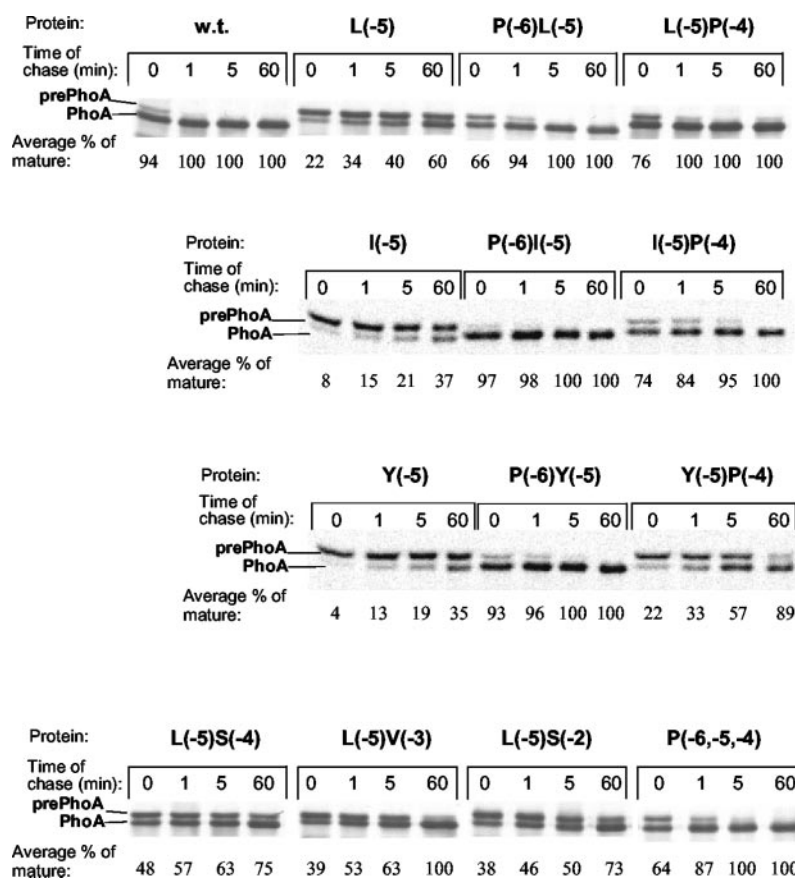


FIG. 2. The dynamics of maturation of wild-type and mutant alkaline phosphatases. Cells were pulse-labeled with [³⁵S]methionine for 60 s, and the radioactivity was chased for indicated periods of time (0, 1, 5, and 60 min). Pulse-labeled alkaline phosphatase and its precursor were immunoprecipitated using affinity-purified rabbit antibodies against the alkaline phosphatase and resolved on 10% SDS-PAGE followed by autoradiography.

an extended β -conformation (Fig. 4) similar to one already proposed in the previous works (19, 29, 52). In this arrangement, the oxygen of the peptide groups of the -1 residue forms hydrogen bonds with the NH group of Ser-90 and hydroxyl group of Ser-88, and backbone nitrogen and oxygen of the -2 residue interact with the CO group of Asp-142 and NH group of Ile-144, respectively. The model explains the cleavage-site Ala-X-Ala specificity: the side chain at -1 position points into the apolar pocket formed by residues Ile-86, Met-91, Leu-95, and Ile-144, the side chain at -2 position points out of the pocket, and the side chain at position -3 is located in the apolar pocket mostly formed by Phe-84, Ile-86, Ile-144, and Asp-142. The modeling also shows that the peptide can bind SPase only if +1 residue adopts a β -conformation with dihedral angle φ ranged between -170° and -115° . This constraint can explain the absence of the Pro residue (with its φ angle restricted between -90° and -50°) at +1 position of exported proteins. The extended -3 to +1 region bends at position -1.

The region -6 to -4 of the signal peptide points out of the active site and this makes it impossible to find its unique conformation based exclusively on the constraints imposed by the enzyme. Nevertheless, we succeeded in the identification of at least one conformation by using a P(-6, -5, -4) mutant. In a tandem of several Pro residues with *trans* peptide groups, the dihedral angles φ are fixed around $-70^\circ \pm 20^\circ$ by the pyrrolidine ring while dihedral angles ψ are restrained around $145^\circ \pm 20^\circ$ by steric interactions between Pro residues. As a result of these steric constraints, the region -6 to -4 of the P(-6, -5, -4) mutant can only have a polyproline conformation in the pocket of SPase (Fig. 4). Our experiments show that the P(-6, -5, -4) mutant can be processed by SPase (Fig. 2), and this result proves that the polyproline conformation of the -6 to -4 region is permissive for processing.

The location of the signal peptide within the membrane and flatness of the SPase surface surrounding the active site make it possible to imagine the overall spatial arrangement of the

membrane, SPase, and signal peptide (Fig. 5). The -6 to -4 region should have a conformation that orients the N-terminal α -helix of the signal peptide (-21 to -7 residue) out of SPase and into the membrane. This provides an additional constraint on possible conformation. The polyproline conformation is in agreement with this condition (N2 on Fig. 5). Previously, it was also suggested that the whole c-region has an extended β -conformation (N1 on Fig. 5) (20, 29). Indeed, if the region -6 to -4 does not have Pro residues, it can adopt a β -conformation and form several hydrogen bonds with the β -strand of SPase (Fig. 4). The backbone of Gln-85 is hydrogen-bonded to the NH group of the -3 residue and the CO group of the -4 residue, and the CO group of Pro-83 interacts with the NH group of the -5 residue. This hydrogen bonding can induce a β -conformation of residues -4 and -5 , while residue -6 needs to adopt a β -conformation to direct the α -helix of the signal peptide (-21 to -7 residue) out of the active site and into the membrane. This arrangement of peptide in the binding pocket is similar to the one that has been already suggested for the -6 to -4 region (19) based on the crystal structure of the analogous enzyme LexA with its bound cleavage site region (52). Our stereochemical analysis shows that there are at least two other conformations of the -6 , -5 , -4 region ($\alpha_6\alpha_5\beta_4$ and $\beta_6\beta_5\alpha_4$ where α and β denote α -helical and β -structural conformations of indi-

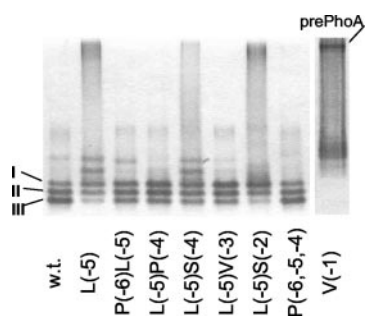
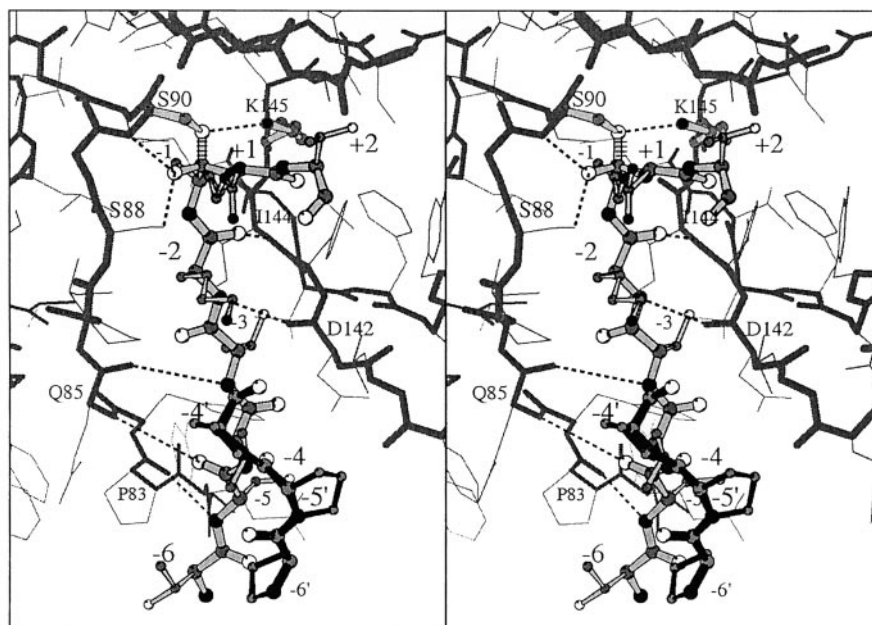


FIG. 3. **Isozymic spectra of wild-type and mutant alkaline phosphatases.** Samples were analyzed by electrophoresis (7.5% PAGE) under nondenaturing conditions, and the active enzyme was revealed by treatment of the gel with α -naphthyl phosphate as the alkaline phosphatase substrate and Fast Red Dye RR. Positions of alkaline phosphatase isoforms (I, II, III) and active precursor (pre-PhoA) are indicated. A mutant with Val in position -1 is shown as an example of the precursor translocated to the periplasmic side (51).

FIG. 4. **Stereoview of a model of the signal peptide bound to the active site of the *E. coli* SPase.** The -6 to $+2$ region of the signal peptide having a β -conformation and TSVTKART sequence corresponding to a well-processed alkaline phosphatase mutant (29) is shown by *ball-and-stick* representation. Sticks of this peptide are in *gray*. Another peptide fragment with *black sticks* represents the -6 to -4 region having three Pro residues in a polyproline conformation (residues -3 to $+2$ are not shown). Side chains Ser-90 and Lys-145 of SPase are also in *ball-and-stick* representation. Oxygen atoms are in *white*, nitrogen atoms are in *gray*, carbon atoms are in *black*. Potential hydrogen-bonding interactions between the signal peptide and SPase are shown by dotted lines. A temporary covalent bond between oxygen of Ser-90 and carbon of -1 residue in the tetrahedral intermediate is hatched.



vidual residues), which meet the basic constraints: absence of steric clashes with SPase and direction of signal peptide α -helix (residues -21 to -7) into the membrane (N3 and N4 on Fig. 5). The occurrence of these conformations cannot be completely ruled out.

DISCUSSION

Previously, based on the observed inhibitory effect of bulky residues at position -5 and molecular modeling, we suggested a β -conformation for the c-region of the signal peptide bound by SPase (29). Knowing that Pro residues can break not only α -helices but also β -structures, we assumed that proline-containing c-regions can have different conformation and sequence requirements compared with c-regions without prolines. The sequence analysis, performed in this work, suggested that the inhibitory effect of bulky residues at position -5 can be suppressed by the introduction of Pro at either position -6 or -4 . The subsequent mutational analysis of *E. coli* alkaline phosphatase strongly supported the predicted effect of the Pro residue. A search of the literature revealed additional proof of inhibition by bulky hydrophobic -5 residue and alleviation of this inhibition by -4 Pro residue. A β -lactamase from *Salmonella typhimurium* has a Leu at position -5 and Pro at position -4 . Substitutions of Pro -4 with the other residues (Ser, Phe, or Leu) lead to partial or complete block of its maturation (23, 31). Furthermore, randomization of the -4 to $+2$ positions (AAFCLXXX/XX, where X is a randomized position) and selection of functional signal peptides from a library of the random sequences, revealed that the c-region of the selected proteins had a consensus AAFCLPAXA/XX (30). This result shows that the preservation of Pro after Leu (-5) is essential for processing. It is worth mentioning that the original studies did not connect these results with the correlation between bulky residues at position -5 and Pro at position -4 .

Molecular docking of the signal peptide into the binding pocket of the SPase opens the possibility of better understanding the molecular mechanism of the effects revealed by our mutational analysis. Along this line, we designed a mutant that provides information about the conformation of the -6 to -4 region during processing. Indeed, the mutant with the -6 to -4 sequence that entirely consists of Pro residues can only have the conformation of a polyproline helix. It was shown that this mutant can be processed. Thus, to assume that the confor-

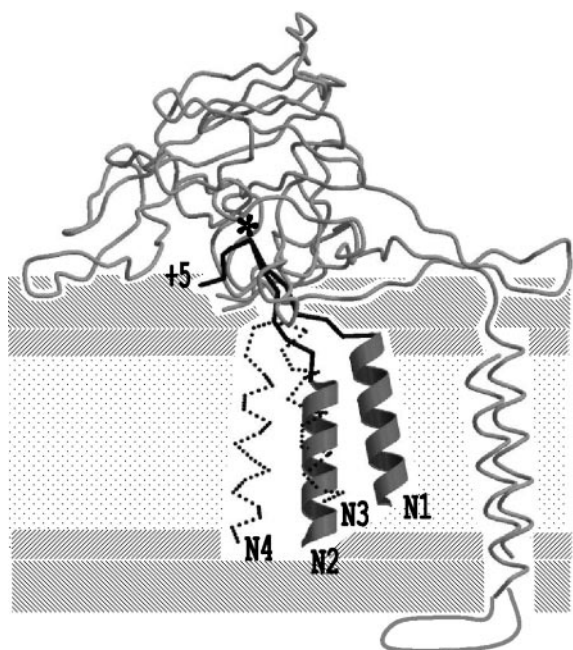


FIG. 5. A model of the overall spatial arrangement of signal peptide and SPase relative to the membrane. The backbone of the crystal structure of SPase is shown as a gray coiled rope. N1, N2, N3, and N4 denote signal peptides with -6 to -4 region in β -conformation, polyproline conformation, $\beta_6\beta_5\alpha_4$ and $\alpha_6\alpha_5\beta_4$ conformations, correspondingly. The N-terminal hydrophobic regions of signal peptides having the -6 to -4 region in β - and polyproline conformations are outlined by ribbon. Traces of two other signal peptides are shown by dotted lines. Asterisk marks the cleavage site of the signal peptide. The two polar regions of the membrane, constituted by phospholipid heads and by glycerol backbones, are shaded darker compared with the middle nonpolar layer. The dimensions of the membrane and the proteins are shown in scale, and the proportions of the layers are taken from Ref. 54. Two transmembrane α -helices of SPase (residues 1–22 and 58–77) were modeled.

mation of the -6 to -4 region is unique for all exported proteins, then our results indicate that it should be the polyproline conformation. On the other hand, the polyproline conformation alone cannot explain the inhibition by bulky -5 residues and its suppression by Pro at position -6 or -4 . Previously, it was suggested that the c-region has β -conformation, and the inhibition was explained by the inability of the corresponding pocket of SPase to accommodate the large -5 residue (29). However, the analysis of the newly available structure of SPase (19) shows that although the β -conformation of the c-region directs the -5 residue toward the enzyme (Fig. 4), there is enough room to accommodate such residues as Leu or Met. Knowing this data, we suggest the following explanation for the observed relationship between residues of the -6 to -4 region. Initially, an apolar membrane environment favors an α -helical conformation of the signal peptide, including the c-region. In order to be recognized by SPase, the α -helical conformation of the c-region needs to be unfolded. This unfolding can be facilitated by hydrogen bonding with the β -strand (residues 84–86) of SPase (Fig. 4). This unfolding into a β -conformation is accompanied by a transposition of the -5 residue from the apolar environment of the membrane to the more hydrophilic surface of SPase. In this situation, the inhibition by a large hydrophobic residue at position -5 can be explained by its reluctance to leave the membrane that hampers the unfolding. We were able to understand the suppression of this inhibition by Pro residues by suggesting that an introduction of Pro residues in the adjacent positions of residue -5 generates the polyproline conformation. In contrast to the β -conformation, the unfolding of the α -helical region -6 to -4 into the polypro-

line conformation is not accompanied by the transition of the -5 residue from the membrane to SPase (Fig. 4). The -5 hydrophobic side chain resides in the membrane and, therefore, will not hamper the unfolding and subsequent processing of the polypeptide. It is worth mentioning that the unfolding of the Pro-containing -6 to -4 region can be facilitated by the helix-breaking property of Pro residues rather than by formation of the hydrogen bonds with the β -strand (residues 84–86) of SPase. This explanation of our experimental results is based on molecular modeling and, therefore, needs further support from x-ray crystallography.

The sequence analysis also indicated that Val at position -3 or Ser at position -4 or -2 may lead to suppression of the inhibition by the -5 bulky residue. Our experiments show that Val at position -3 partially suppresses the inhibition while Ser at position -4 or -2 does not. Minor recovery of the processing for the L(-5)S(-4) mutant can be accounted for by the higher hydrophilicity and, therefore, an easier unfolding of its c-region compared with the L(-5) protein. The known structure of SPase, however, does not provide a simple explanation for a partial recovery of the L(-5) inhibition in the case of the L(-5)V(-3) mutant, and processing of this mutant requires further investigation. Another observation that will need further study is the presence of a bulky Leu at position -5 and non-proline residues at both adjacent positions in few exported bacterial proteins. Supposedly, they are efficiently processed despite the fact that we observed an inhibition of processing for mutants of alkaline phosphatase with similar c-regions (29).

It is worth mentioning that if our hypothesis about at least two possible conformations of the -6 to -4 region is true, the hydrophobic region of the signal peptide should not have a fixed position relative to the SPase (Fig. 5). This implies that SPase and the hydrophobic region of the signal peptide may not be connected with each other by rigid protein structures (e.g. translocase), and one or both of them float in the membrane at the moment of recognition. Experimental study supports this conclusion (53).

Acknowledgments—We thank Dr. A. L. Karamyshev for discussion and assistance and Dr. U. Blum-Tirouvanziam for critical reading of the article.

REFERENCES

- Dalbey, R. E., and Von Heijne, G. (1992) *Trends Biochem. Sci.* **17**, 474–478
- Perlman, D., and Halvorson, H. O. (1983) *J. Mol. Biol.* **167**, 391–409
- von Heijne, G. (1985) *J. Mol. Biol.* **184**, 99–105
- Summers, R. G., and Knowles, J. R. (1989) *J. Biol. Chem.* **264**, 20074–20081
- Kuhn, A., Kiefer, D., Kohne, C., Zhu, H. Y., Tschantz, W. R., and Dalbey, R. E. (1994) *Eur. J. Biochem.* **226**, 891–897
- Kajava, A. V., Zolov, S. N., Kalinin, A. E., and Nesmeyanova, M. A. (2000) *J. Bacteriol.* **182**, 2163–2169
- Fekkes, P., and Driessen, A. J. (1999) *Microbiol. Mol. Biol. Rev.* **63**, 161–173
- Nesmeyanova, M. A., and Bogdanov, M. V. (1989) *FEBS Lett.* **257**, 203–207
- de Vrije, T., de Swart, R. L., Dowhan, W., Tommassen, J., and de Kruijff, B. (1988) *Nature* **334**, 173–175
- Killian, J. A., de Jong, A. M., Bijvelt, J., Verkleij, A. J., and de Kruijff, B. (1990) *EMBO J.* **9**, 815–819
- Nesmeyanova, M. A., Karamyshev, A. L., Karamysheva, Z. N., Kalinin, A. E., Ksenzenko, V. N., and Kajava, A. V. (1997) *FEBS Lett.* **403**, 203–207
- Van Voorst, F., and De Kruijff, B. (2000) *Biochem. J.* **347**, 601–612
- Chou, M. M., and Kendall, D. A. (1990) *J. Biol. Chem.* **265**, 2873–2880
- Lill, R., Dowhan, W., and Wickner, W. (1990) *Cell* **60**, 271–280
- Pugsley, A. P. (1993) *Microbiol. Rev.* **57**, 50–108
- Von Heijne, G. (1986) *J. Mol. Biol.* **192**, 287–290
- Zimmermann, R., Watts, C., and Wickner, W. (1982) *J. Biol. Chem.* **257**, 6529–6536
- Dalbey, R. E., Lively, M. O., Bron, S., and van Dijk, J. M. (1997) *Protein Sci.* **6**, 1129–1138
- Paetzel, M., Dalbey, R. E., and Strynadka, N. C. (1998) *Nature* **396**, 186–190
- Izard, J. W., and Kendall, D. A. (1994) *Mol. Microbiol.* **13**, 765–773
- Plath, K., Mothes, W., Wilkinson, B. M., Stirling, C. J., and Rapoport, T. A. (1998) *Cell* **94**, 795–807
- Paetzel, M., Dalbey, R. E., and Strynadka, N. C. (2002) *J. Biol. Chem.* **277**, 9512–9519
- Koshland, D., Sauer, R. T., and Botstein, D. (1982) *Cell* **30**, 903–914
- Barkocy-Gallagher, G. A., Cannon, J. G., and Bassford, P. J., Jr. (1994) *J. Biol. Chem.* **269**, 13609–13613
- Laforet, G. A., and Kendall, D. A. (1991) *J. Biol. Chem.* **266**, 1326–1334

26. Chou, P. Y., and Fasman, G. D. (1979) *Biophys. J.* **26**, 367–373
27. Rosenblatt, M., Beaudette, N. V., and Fasman, G. D. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 3983–3987
28. Shen, L. M., Lee, J. I., Cheng, S. Y., Jutte, H., Kuhn, A., and Dalbey, R. E. (1991) *Biochemistry* **30**, 11775–11781
29. Karamyshev, A. L., Karamysheva, Z. N., Kajava, A. V., Ksenzenko, V. N., and Nesmeyanova, M. A. (1998) *J. Mol. Biol.* **277**, 859–870
30. Palzkill, T., Le, Q. Q., Wong, A., and Botstein, D. (1994) *J. Bacteriol.* **176**, 563–568
31. Kadonaga, J. T., Pluckthun, A., and Knowles, J. R. (1985) *J. Biol. Chem.* **260**, 16192–16199
32. Niviere, V., Wong, S. L., and Voordouw, G. (1992) *J. Gen. Microbiol.* **138**, 2173–2183
33. Nielsen, H., Engelbrecht, J., von Heijne, G., and Brunak, S. (1996) *Proteins* **24**, 165–177
34. Bachmann, B. J. (1987) in *Escherichia coli and Salmonella typhimurium: Cellular and molecular biology* (Neidhardt, F. C., ed), American Society for Microbiology, Washington, D. C.
35. Zaitsev, E. N., Zaitseva, E. M., Bakhlanova, I. V., Gorelov, V. N., and Kuz'min, N. P. (1986) *Genetika* **22**, 2721–2727
36. Kleina, L. G., Masson, J. M., Normanly, J., Abelson, J., and Miller, J. H. (1990) *J. Mol. Biol.* **213**, 705–717
37. Inouye, H., Michaelis, S., Wright, A., and Beckwith, J. (1981) *J. Bacteriol.* **146**, 668–675
38. Torriani, A. (1966) in *Procedures in nucleic acid research* (Cantoni, G. L., and Davis, R., eds), pp. 224–234, Harper and Row, New York
39. Kalinin, A. E., Mikhaleva, N. I., Karamyshev, A. L., Karamysheva, Z. N., and Nesmeyanova, M. A. (1999) *Biochemistry-Russia* **64**, 1021–1029
40. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning. A laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, New York
41. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
42. Davis, B. J. (1964) *Ann. N. Y. Acad. Sci.* **121**, 404–427
43. Lojda, Z., Grossrau, R., and Schibler, T. H. (1979) *Enzyme histochemistry: A laboratory manual.*, Springer-Verlag, Berlin, Germany
44. Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265–275
45. Dayring, H. E., Tramonato, A., Sprang, S. R., and Fletterick, R. J. (1986) *J. Mol. Graphics* **4**, 82–87
46. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Computat. Chem.* **4**, 187–217
47. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950
48. Fauchere, J., and Pliska, V. (1983) *Eur. J. Med. Chem.* **18**, 369–375
49. Michaelis, S., Inouye, H., Oliver, D., and Beckwith, J. (1983) *J. Bacteriol.* **154**, 366–374
50. Boyd, D., Guan, C.-D., Willard, S., Wright, W., Strauch, K., and Beckwith, J. (1987) in *Phosphate Metabolism and Cellular Regulation in Microorganisms* (Torriani-Gorini, A., Rothman, F. G., Silver, S., Wright, A. & Yagil, E., eds) pp. 89–93, American Society for Microbiology, Washington, D. C.
51. Karamyshev, A. L., Kalinin, A. E., Tsfasman, I. M., Ksenzenko, V. N., and Nesmeyanova, M. A. (1994) *Mol. Biol. (Mosk)* **28**, 362–373
52. Luo, Y., Pfuetzner, R. A., Mosimann, S., Paetzel, M., Frey, E. A., Cherney, M., Kim, B., Little, J. W., and Strynadka, N. C. (2001) *Cell* **106**, 585–594
53. Martoglio, B., Hofmann, M. W., Brunner, J., and Dobberstein, B. (1995) *Cell* **81**, 207–214
54. Venable, R. M., Zhang, Y., Hardy, B. J., and Pastor, R. W. (1993) *Science* **262**, 223–226