Export of a hyperexpressed mammalian globular cytochrome b₅ precursor in *Escherichia coli* is dramatically affected by the nature of the amino acid flanking the secretory signal sequence cleavage bond

Naheed N. Kaderbhai,* Khalil Ahmed, and Mustak A. Kaderbhai

William Davies Laboratories, Institute of Biological, Environmental and Rural Sciences (IBERS), Plas Gogerddan Campus, Aberystwyth University Aberystwyth, Ceredigion, SY23 3EB, United Kingdom

Received 3 February 2010; Revised 23 April 2010; Accepted 28 April 2010

DOI: 10.1002/pro.411

Published online 5 May 2010 proteinscience.org

Abstract: A chimeric mammalian globular cytochrome b₅ fused to *Escherichia coli* alkaline phosphatase signal sequence (SS) was used as a model probe to investigate the influence of substituting each one of the standard 20 amino acids at its N-terminus on the Sec-dependent export of the precursor to the periplasmic space of *E. coli*. Substituting the native Met⁺¹ of the passenger protein flanking the SS with any one of the remaining 19 amino acids introduced significant changes in the export of cytochrome b₅ without jamming the Sec-dependent translocon. Acidic and hydrophilic residues proved to be the most efficient promoters of export. Small, nonbulky and basic residues yielded intermediate levels of the hemoprotein export. Replacement with a Cys⁺¹ residue generated significant quantities of both monomeric and disulfide-linked dimeric forms. However, bulky, aromatic and hydrophobic residues caused a significant decline in the rates of secretion. In expectation with their absences in the natural periplasmically secreted proteins, Pro and Ile-tagged cytochrome b₅ precursors failed to generate any detectable secreted recombinant products. Although Ala, amongst the native *E. coli* periplasmic proteins, is the preferred X⁺¹ residue with an occurrence of 50% frequency, it proved half as effective in promoting export when inserted proximally to the SS of cytochrome b₅. The mechanisms involved for these export variations are discussed. The findings will prove beneficial for high-level generation of recombinant proteins by secretory means for pharmaceutical and related biotechnological applications.

Keywords: protein export; recombinant protein; cytochrome b₅; signal sequence; preprotein processing; periplasmic space; *Escherichia coli*

Abbreviations: Cyt, globular domain of rat liver endoplasmic reticulum cytochrome b₅; SPase, signal peptidase-I; SS, signal sequence; X, any one of the standard 20 amino acids.

Additional Supporting Information may be found in the online version of this article

Khalil Ahmed's address is Karakoram International University Gilgit, Baltistan, Pakistan.

Mustak A. Kaderbhai is deceased.

*Correspondence to: Naheed N. Kaderbhai, William Davies Laboratories, Institute of Biological, Environmental and Rural Sciences (IBERS), Plas Gogerddan Campus, Aberystwyth University Aberystwyth, Ceredigion, SY23 3EB, United Kingdom. E-mail: nnk@aber.ac.uk

Introduction

For structural and functional investigations, there is now continued and increasing demand for hyperexpression of novel cloned genes to manufacture recombinant proteins in *Escherichia coli*. Biotechnologically, the process is highly optimized through (i) efficient transcription involving powerful, regulated promoters and (ii) translational signals, (iii) compatible sets of codon usage,¹ and (iv) post-translational modifications.² We have been interested in understanding the basic principles governing the high-level production of recombinant proteins by promoting their

secretion into the periplasmic space. Unlike cytoplasmic production, this zone is advantageous in that the slower rate of polypeptide threading through the Sec translocon evades the formation of inclusion bodies and offers the biochemical machinery for numerous essential post-translational modifications such as disulfide bonding, covalent heme linkage,² phosphorylation,³ glycosylation,⁴ hydroxylation, pro-peptide processing, etc.⁵ More importantly, through cleavage of the signal sequence (SS) the exported mature protein can be designed to acquire the desired N-terminus sequence.

Despite significant advances in our ability to generate successful high-yielding production of recombinant proteins in *E. coli* by (non-)secretory means, the process remains unpredictable and the rules governing secretion consequently remain empirical. Secretory proteins are initially synthesized as precursors carrying transient amino-terminal extensions SSs⁶ that are exquisitely designed peptides mediating the translocation of the precursor across the cytoplasmic membrane typically through a sec-dependent translocon.⁷ Following translocation of the precursor into the periplasmic zone, the SS is cleaved by an exocytosomally exposed signal peptidase-I (SPase). Although the SSs vary in both their length (from 15 to 40 amino acids) and primary structure⁸ they display a consensus tripartite structure. The central hydrophobic core (*h*), which is of a length sufficient to span the lipid bilayer in an α -helical conformation, is flanked at the N-terminus by positive charged, polar residues. The inter- and intraspecies variation of the length of hydrophobic core may be an important determinant of the rate and extent of secretion.^{9,10} The polar C-terminal segment (*c*), composed of 3–7 residues, has the cleavage site for SPase which recognizes residues at the –1 and –3 (small, uncharged residues are favored at P1 and P3 positions) position relative to the scission bond. Further upstream SS segments often display residues such as Gly, Ser, and Pro with a propensity to adopt an extended conformation to facilitate recognition and cleavage by SPase. Apart from promoting preprotein engagement with the translocon and controlled translocation, SSs also prevent premature (mis)folding of the precursor. SS-like peptides, liberated from dormant, membrane-bound precursor proteins by the activities of intramembrane-cleaving proteases such as SS peptidase, are thought to play downstream functions in cell signaling and regulation, immune surveillance, and intercellular communication.¹¹ An antimicrobial peptide derivative of flesh fruit fly hemolymph mimics secretory SS and inhibited SPase in the export pathway.¹²

In addition to the core hydrophobic region, the functionality of SS is critically determined by juxtaposition of the charge to the cleavage bond.¹³ Both the amino acid composition and the sequence of the

early mature region adjoining the SS are important determinants for protein secretion.¹⁴ This is probably due to the fact that this upstream passenger portion flanking the cleavage bond is tightly lodged and interactive within the active site of SPase.¹⁵ “Peppering” with negatively charged residues enhances secretion whereas positively charged residues retard export.^{16,17} A novel simulated evolutionary strategy that randomly searches for peptide sequences at the secretory signal-target protein junction to promote protein export demonstrated that functional, charge-dependent peptides ranging in size from 1 to 13 residues displayed an inverse relationship between the export rates and in-transit membrane-bound precursor pools. The preprotein export rates within these clusters were related to the magnitude and polarity of the membrane energization states.^{18,19}

Numerous studies have investigated the influence of SSs either alone or through gene chimeric approaches on the levels of expression, translocation, and secretion. Single point mutations particularly those involving insertion of a charged residue(s) in the hydrophobic core prove to be most deleterious.^{20,21} A more recent proteomic study of the human genome has also demonstrated that single-nucleotide polymorphism in SSs correlate with pathogenesis accruing from cessation of secretion of the proteins.²² While SSs have a variable primary structure, the analysis of their role in secretion is compounded by the fact that each class of the secretory protein also has a different N-terminal mature protein sequence. The contribution of these repertoires on the export rates *in vivo* have remained indistinct, the more so because many previous studies have been conducted using radiolabeled pulse-chase methods involving miniscule levels of export studied over a period of just a few seconds. Many of these were conducted with constructs containing motifs around the cleavage site that differed markedly in their primary structure, appended to distinct passenger molecules.^{1,23}

A mammalian globular cytochrome *b*₅ (Cyt) tagged with(out) the alkaline phosphatase SS of *E. coli* has been exploited to investigate many diverse molecular features of high-level periplasmic protein secretion.^{18,24–28} It has proved an invaluable tool in reporting both the *in vitro* and *in vivo* sec-dependent status of recombinant protein generation by visual and spectroscopic means. It was shown that the amino acid sequence, composition, and charge distribution in the early mature region were important parameters for the proton motive force dependent export of evolutionarily generated Cyt isoforms through the use of a common SS. Until recently, the influence on high-level protein secretion in *E. coli* as a function of substituting each of the standard 20 amino acids at the +1 flanking positions in fusions

Figure 1. Scheme showing construction of plasmids (pM-SS- X^{+1} Cyt), derived from the progenitor plasmid pM-SSnCyt, encoding secretory forms of Cyt, each coding for one of the 20 amino acids following the cleavage bond of SS. The native Met codon is in bold. The italicized bases mark the endonuclease restriction sites used in the construction. *PphoA*, alkaline phosphatase promoter; S/D, Shine Dalgarno; SS and signal sequence.

composed of identical SS and the passenger moiety has remained unknown. In the present study, employing a chimeric precursor composed of *E. coli* alkaline phosphatase SS (21 residues) appended to the mammalian globular Cyt (99 residues), we demonstrate that alteration of any one of the standard 20 amino acids at the N-terminus of the passenger polypeptide dramatically alters the rate of the hemoprotein secretion. The likely biochemical mechanisms for the variations in the export rates are discussed.

Results and Discussion

Construction of secretory vectors encoding variant X^{+1} amino acid in cytochrome b_5 precursors

A plasmid cassette (pM-SSnCyt) coding for a secretory form of a Cyt, appended to *E. coli* alkaline phosphatase SS and expressible under the transcriptional control of the native *Pho A* promoter, was constructed. In this plasmid, an engineered *Hind* III site together with an in-built Amber codon juxtapositioning the SS aborts translation of the passenger cytochrome b_5 . This strategy facilitated subsequent dual cleavage of the vector with *Hind*III and the upstream *Bcl*I to incorporate a synthetic double stranded “stuffer” oligonucleotide. This in turn resets a contiguous reading frame and concomitantly facilitates incorporation of a codon for any one of the standard 20 amino acids as the first amino acid (+1) of Cyt, flanking the SS cleavage bond. The cloning strategy is depicted in Figure 1. Plasmid constructs were transformed into the TB1 strain and following their culturing in the phosphate-limited MOPS medium positive clones were selected using the visual PINK detection system,^{24,29} except for those encoding Ile⁺¹ and Pro⁺¹ (see below). The sequences of the potentially positive clones were further verified by DNA sequencing, thus permitting iso-

lation of all 20 variants of the Cyt precursor expressing strains (see supporting information).

Characterization of exported cytochrome b_5 expressed with X^{+1} residue proximal to SS

The growth patterns of the clones expressing Cyt addressed with alkaline phosphatase SS interposed with each one of the standard 20 amino acids were similar following 8 h of culturing in a phosphate-limited MOPS medium. The periplasmic secretion of the hemoproteins was initially examined qualitatively and subsequently quantitatively. All of the exported hemoproteins, with the exception of Pro⁺¹ and Ile⁺¹, displayed the characteristic absorption spectrum of the native globular Cyt with a prominent Soret peak for the reduced form at 423 nm and a drop at 409 nm, indicating that the secreted proteins were correctly folded. To eliminate the likelihood that the engineered sequence(s) of the mature protein might be at variance from the predicted DNA sequences, the N-terminal sequence of the gel-eluted, exported Cyt isoforms were also determined. The Cyt derivatives, being relatively small (~12,000 Da), globular, compact, and acidic proteins with a predicted *pI* ~ 4.5, resolve as the fastest-migrating major species amongst the Coomassie blue-detectable periplasmic proteins following electrophoretic separation under nondenaturing conditions (Figs. 2 and 3). Before their staining, the holohemoprotein bands were distinctly identifiable as visible red bands, which facilitated their recovery for the determination of the first five N-terminal amino acid sequencing by Edman degradation.

Export features of SS- X^{+1} -cytochrome b_5 isoforms

Qualitative analyses of the Cyt isoforms, recovered in the periplasmic extracts of all of the strains of *E.*

Figure 2. Nondenaturing gel electrophoretic analysis of periplasmic extracts from *E. coli* pM-SS-(X⁺)Cyt expressing Cyt variants. Before electrophoresis, the samples were treated with 10 μ M bovine heme to convert the apo form pools to holo cytochromes. The white dots mark the positions of holo Cyt visually detectable and identifiable by their pink coloration. Protein loadings were 20 μ g per lane.

coli, except pM-SS-Pro⁺Cyt, pM-SS-Ile⁺Cyt, and pM-SSnCyt, simultaneously cultured under identical growth conditions, displayed small but significant changes in their electrophoretic mobilities, most likely attributable to nature of the X⁺ of the mature, exported Cyt (Figs. 2 and 3). Indeed, single amino acid changes have been known to introduce altered electrophoretic mobilities.^{30–33} Clearly, a change in the +1 residue of the passenger protein annexing the same SS invoked significant variations in the hemoprotein content deposited in the periplasm. The profiles were reproducibly verified in three independent expression trials. Substitutions with Pro⁺ or Ile⁺ yielded undetectable export. Notably, a negative charge introduced through placement of acidic residues increased the hemoprotein mobilities and conversely the basic and bulky residues retarded the migration toward the cathode. In contrast to the basic X⁺ residue, the higher export rates of the acidic residues must have been facili-

Figure 3. Nondenaturing gel electrophoresis analysis of periplasmic extracts from *E. coli* pM-SS-(X⁺)Cyt expressing X⁺ variants of Cyt. Samples were analyzed following treatment with 10 μ M heme and reduction with 50 mM dithiothreitol. The white dots mark the holo Cyt visually identified by their pink coloration.

tated by the protonophore-dependent translocation of the early segment of pre-Cyt across the inner membrane (IM) of *E. coli*.^{18,34} Qualitatively, the specific content of the periplasmic pool of the Cyt derivatives was highly variable (Table I). In contrast, the electrophoretically characterizable “in-house” periplasmic proteins appear to remain unchanged (Figs. 1 and 2). Moreover, enzymatic measure of the coinduced periplasmic alkaline phosphatase activities in the periplasmic extracts of the various strains showed insignificant variation (3.5 ± 0.2 units per milligram of protein).

A number of possibilities could account for the differences in the export rates exhibited by the amino acid substitutions at +1 of the globular Cyt. Changes in the amino acid at this particular position could have introduced variations in the rate of precursor synthesis, turnover, translocation, and/or processing, possibly due to small but significant structural changes invoked beyond the cleavage site. Since the expression of all the variant Cyt was controlled through the identical promoter and translation elements, the first factor seems unlikely. To test whether the half-lives of the precursors or final secreted products was influenced by the nature of the early mature region,³⁵ a cocktail of protease inhibitors³⁶ was added to the growth medium at defined intervals during the induction regime of a selection of recombinant cell lines, under conditions that did not affect their cell growth. However, this did not significantly influence the changes in the

Table I. The Effect of Varying the Amino Acid Flanking the SS and Cytochrome *b*₅ Junction on the Export of the Hemoprotein to the Periplasm

(Amino acid) ⁺	Export rate ^a
S	1.17
T	1.11
Q	1.01
C	1.00
D	0.93
E	0.88
H	0.81
N	0.79
M	0.78
K	0.73
R	0.67
V	0.62
G	0.62
A	0.60
F	0.50
W	0.31
Y	0.25
L	0.19
P	0.02
I	0.01

Results are averages of three experimental values with the SD of $\pm 5\%$.

^a Milligrams of Cyt (reduced) exported/Liter/hour/culture OD₆₀₀.

export rates of the cytochrome isoforms. Indeed, translocation and exposure of the unprocessed precursor at the periplasmic surface of the inner membrane can occur despite a block in SPase activity.³⁶ The decreased secretion rates of the variant Cyt forms could also arise through either weaker engagement of the precursors with the translocation machinery/channel or a diminished extent of processing by SPase. The likelihood of the former is discussed below. To assess whether the substitution of the amino acid had affected their SPase-cleavability, the isolated precursor proteins³⁶ were kinetically tested for their *in vitro* processivity.³⁶ The isolatable precursor proteins were efficiently processed to their mature counterparts *in vitro* (data not shown), thus suggesting that precursor processing, at least *in vitro* was an unlikely limiting factor in influencing export rates. Nevertheless, *in vivo* SS has a temporal role in activation and maintenance of the polypeptide translocation.²⁸ SS also confers processing reaction specificity based on residues juxtapositioning the cleavage bond and thus precursor stability may have been compromised (see below).

Bulkiness of the X^{+1} is a major determinant of the export rate

For a more precise determination of the secreted rates, the periplasmic Cyt isoforms were tabulated as mass of hemoprotein exported/hour/OD₆₀₀/liter of culture (Table I). With X^{+1} variations, the secretion rates can be conveniently classified into five groups based on the ascending yields of periplasmic Cyt isoforms: (i) Ile and Pro (nil production); (ii) Leu, Tyr, Trp, and Phe (0.2–0.5); (iii) Ala, Gly, Val, Arg, and Lys (0.6–0.7); (iv) Met, Asn, His, Glu, and Asp, (0.8–1.0); and (v) Cys Gln, Thr, and Ser (1.05–1.2).

Although these export variations could not be correlated with the corresponding molecular weights of X^{+1} , a trend was apparent based on the bulkiness of the side chains. The latter is the simplest and most fundamental property of amino acids that determine steric constraints at higher protein structural levels, defining the nature of buried and accessible surfaces of the side chains in proteins. Bulkiness defines the ratio of the side chain volume to length and was assessed employing published indices.^{37,38}

Amino acid scales of the engineered variable X^{+1} residue at position 8 in the sequence of SS^{-7→-1} (LFTPVK^A) bridging the mature^{1→6} Cyt (X^{+1} AEQSD) were computed using a moving window setting of three residues with a score weighting for the residue at position 8, eventually edging to 50% (Fig. 4). This was in line with the length of peptide region accommodated at the active site/groove of SPase.³⁹ With a few exceptions, the data clearly reflect an inverse relationship between the bulkiness of the introduced amino acid and export rates. The hydro-

Figure 4. Correlations between rates of secretion of SS- X^{+1} Cyt, the bulkiness and hydrophobicity of the adjoining amino acid. The indices for hydrophobicity and bulkiness were those of Refs. 38 and 37, respectively. The details of the derivation of the algorithms are fully described in the text. All scales were normalized from 0 to 1.

phobicity, though not as significant as the bulkiness, demonstrated a mild trend with the more polar residues being secreted slightly better. Linear regression between export rate and bulkiness index was highly significant ($r = -0.57$; $P = 0.009$). Accordingly, no apparent correlation was noted between export rate and hydrophobicity ($r = -0.315$; $P = 0.177$).

Incorporation of Ile as X^{+1} , which is the bulkiest and most hydrophobic amino acid, completely prevents Cyt secretion (Figs. 2 and 3). Interestingly, this amino acid at the X^{+1} position also appears to be absent amongst the native periplasmic proteins of *E. coli* (see Fig. 5). Unlike Ile, its isomeric counterpart, Leu, which possesses a slightly increased polarity while having an identical bulkiness index, displays a marginally functional secreting activity. It is perplexing that Ile⁺¹ is neither functional in the present model nor found flanking the SS of the resident periplasmic proteins of *E. coli*. Although in theory Ala⁻¹-Ile⁺¹ bond should be cleavable by SPase, it is possible that the incision may be sterically tempered. The bulkiness and hydrophobicity of Ile may not readily accommodate this residue at the active site of SPase.

The amino acids Tyr, Trp, Phe, and Ala, which are of similar bulkiness, gradually display an increased secretion that appears to be more

Figure 5. Time-course processing of pre-Cyt containing Cys⁺¹ at the SS junction. The precursor was isolated from the membranes from a freshly induced culture by extraction with 20% (v/v) acetonitrile. The clarified extract (15,000g × 15min) was applied onto a FPLC Mono-Q column (Pharmacia-LKB, UK) in TE buffer and eluted by applying a gradient of NaCl ranging from 0 to 0.5M. The precursor was then incubated with isolated SPase as described previously.³⁶

significantly related to their increasing polarity. The residue Gly shows a marked increase in export while it is the least bulky. Gly has an exceptionally high polarity but being the least bulky, may have a loose fit at the active site but nevertheless be sufficiently susceptible to catalytic cleavage.

A few less hydrophilic neutral amino acids such as Val, Met, and Cys with above average core sizes were also exported exceptionally well. We speculate that the significantly high-export rate of Cys may be attributed to its prior modification by a polar moiety (unpublished finding). Substitution with basic residues also generates high-export rates but this is further substantiated with the acidic residues. Ser and Thr, which have small, uncharged, and polar residues, yielded the highest export.

More recently, Karla *et al.*⁴⁰ have identified residues at the active site of SPase which determine the basis of substrate specificity and binding. A large hydrophobic patch accommodates the bound SS (−1 to −3 positions relative to the cleavage site) at the active site, which bridges the remaining N-terminal substrate in a helical form traversing the membrane core. The crystal structure of SPase with a bound model SS⁴⁰ has allowed the identification of residues at the active site which determine the basis of substrate specificity and binding. A groove “peppered” by side chains furnishing a wave of uncharged and charged surface accommodates 13 residues of the precursor extending from −7 to +6 of the precursor in an extended conformation with distinctive twist of the backbone in the cleavage region at −3 to +1. The narrow cleft in this zone appears to play a crucial role in the high specificity of SPase cleavage. Although the specificity of the cleavage bond is defined by the −3, −1, amino acids with small

uncharged side groups, this nevertheless extends into the mature region up to +6 residues. Computational approaches to the structural modeling of *E. coli* SPase-substrate complex suggests that the nature, size, and charge of the amino acid residues accommodated in the cleft at the flanking X⁺¹ play a significant role in the export rates observed in this study.

Conformation of the trans-membrane SS predicted by Multivariate Linear Regression Combination (MLRC) was strikingly affected by a change in the X⁺¹ residue adjoining the alkaline phosphatase SS and the globular Cyt.⁴¹ A correlation was apparent between export rates and the extent of conformational “stretch” of SS induced by changes in the X⁺¹ residue. A “tighter” helical conformation induced by bulky, hydrophobic residues such as Ile may have buried the cleavage site within the lipid matrix thus restraining it from gaining access to the exocytosmic active site of SPase. These findings are in agreement with previous reports^{42–44} that overall length of the central core and sequence hydrophobicity determine the *in vivo* translocation efficiency of export rates. These observations provide the basis for the poorer export rates of the bulky X⁺¹ and for the peculiar behavior of other less bulky but hydrophilic residues such as Gly. Thus, it is possible SPase *per se* is part of the translocase and if so the translocated precursor with poor SS accessibility to the active site may be a major determinant of export through susceptibility to degradation by SS protease.

Cys⁺¹-containing cytochrome b₅ precursor is exported as both monomeric and dimeric SPase-processed forms

Electrophoretic analysis of the *E. coli* pM-SS-Cys⁺¹ Cyt resulted in production of two forms of hemoproteins (Fig. 2). Both the isolated forms were spectrally identifiable as Cyt forms. MALDI-TOF-MS analysis of this processed electro-eluted band gave [M + H]⁺ = 22,626 compared with the theoretical or predicted value of 22,625. On reduction with excess dithiothreitol, the slower migrating hemoprotein disappears and concomitantly comigrates with the faster migrating band (compare lanes 2 in Figs. 2 and 3). Thus, the Cys⁺¹ form secreted Cyt occurred in the periplasm as a monomer as well as a disulfide crosslinked dimer. The latter is probably generated in the more oxidizing periplasmic space by oxidation of the monomers. In our survey of the +1 residue proximal to the SS cleavage site amongst the *E. coli* periplasmically localized proteins, Cys does not appear to occur at position +1 (see Fig. 5). Indeed, lipoprotein(s) targeted to the outer membrane, synthesized as presecretory proteins with a Cys residue immediately after the SS cleavage site, were reported to be selectively processed by SPase II.⁴⁵ A

Figure 6. Comparison of the rates of secretion of SS- X^{+1} -Cyt b_5 and probabilities of the occurrence N-terminal residues in the native secreted proteins of *E. coli*.

prerequisite for precursor cleavage by SPase II was that the Cys residue was modified to glycerylcysteine for subsequent esterification with two fatty acyl chains. SPase II, which is a structurally distinct enzyme from SPase and unlike the latter, is sensitive to inhibition by globomycin.^{46,47} Surprisingly, the isolated Cys⁺¹-containing Cyt precursor on incubation with purified SPase was efficiently processed (Fig. 6).

Pro-containing cytochrome b_5 isoform is export incompetent

Inception of Pro⁺¹ following the terminal Ala⁻¹ of SS completely annulled export of Cyt (see Figs. 2 and 3). This phenomenon appears to be unrelated to the bulkiness of the residue. Interestingly, nor does this residue appear to occur in the corresponding position of the natural precursor of all of the *E. coli* periplasmic proteins analyzed (Fig. 5). Why is Pro⁺¹ flanking the SS in the Cyt precursor as well as in natural precursors unacceptable? With the exception of some retroviral proteases, it is extremely uncommon for a protease or a peptidase to cleave an X-Pro bond.⁴⁸ Pro is an imino acid whose R group is covalently bonded to the α -amino group forming a cyclic structure that is incorporated into the peptide chain. The peptide imine is not readily available to accept a proton, in general, acid-base catalysis, a common mechanism for protease cleavage. In another study, we constructed the precursor SS-Ala-Pro-Cyt with placement of a Pro residue following the Ala⁺¹. This approach would be expected to overcome the peptide cleavage in turn revealing the true nature of processing by the SPase. The bacterial expression of SS-Ala⁺¹-Pro-Cyt precursor resulted in the export of hemoprotein at a level comparable to SS-Ala⁺¹-Cyt.³² The sequence of the exported protein at the N-terminus occurred predominantly as Ala⁺¹Pro-Cyt. The time-course profile of the exported hemoprotein from SS-Ala⁺¹-Pro-Cyt displayed a similar trend to that of SS-Ala⁺¹-Cyt but in the second phase (past mid-stationary phase) displayed poorer stability. These findings, in general, imply that the composition of the X_n sequence in SS- X_n -Cyt constructs have varying effects at more than one stage of the export and assembly processes and influence the rate of proteolysis of the secreted proteins. It

has been reported that a SS with a Pro next to the cleavage site inhibited SPase when present in a sec-independent protein.⁴⁹ Pro residues rarely occur in any three (-3 to -1) C-terminal positions of bacterial SSs, and have never been found in position +1 immediately following the cleavage site. Impairment of SPase ability to cleave SS-Pro bond must have similarly placed an evolutionary constraint to exclude it at this position in the natural sequences of all organisms. We thus conclude that SPase lacks the ability to process SS-Ala⁻¹-Pro⁺¹ bond.

Natural and simulated X^{+1} residues juxtapositioning the SS cleavage site

We retrieved 155 amino acid sequences of periplasmic precursor proteins of *E. coli* MG1655 KEGG (Kyoto Encyclopedia of Genes and Genomes, Tokyo, Japan) repository. The probabilities of the SSs at the amino terminals were probed using the Neural Network Prediction Program, SignalP V3.0. The mature N-terminal residue in the SS-processed protein residing in the periplasm was established from the definitive cleavage sites, predicted employing Neural networks and Hidden Markov models⁸ for Gram-negative bacteria. The SS probability which defines the fitness of the secretory signal character of peptide sequences in Gram negative organisms was >0.9 of the analyzed sequences.

Since the rate of export of the engineered forms in the SS- X^{+1} -Cyt were clearly a function of the nature of amino acid juxtapositioning the cleavage site, it was of interest to compare these with the frequencies of the occurrence of X^{+1} in the naturally secreted periplasmic proteins (Fig. 5). Although there is no apparent direct relationship between these two parameters, some links are obvious. Ile and Pro which are nonfunctional in the present study correspond with their absences from the natural pool. Second, the frequencies of the bulky residues in the natural pool are also significantly low. Paradoxically, Ala⁺¹, the most frequently (50%) employed amino acid in the *E. coli* secreted protein (and higher organisms), is only half as efficient as the artificially engineered chimeric protein. Obviously, the remaining pool must have a need for the natural variation in the biodiversity of the residue. We do not exclude the possibility that our model may be at variance with the natural canonical system. Indeed, some N-terminal residues may be essential for post-translational modifications such as acetylation or further proteolytic processing. Others may also be essential to provide the secretory machinery with the flexibility, versatility, and perhaps accuracy to enact the chaperoning functions, proper SS recognition, processing, and finally proper folding to gain of biological activity. It is possible that residues beyond the natural SSs may have evolved to fit the structure and function, and possibly the rate of

export and also the half life of the natural protein in the periplasm. The present study provides an unambiguous solution to achieving high-level export of artificially engineered or recombinant secretory proteins.

Materials and Methods

Cloning, screening, and bacterial culturing

The standard operations involving DNA manipulations and sequencing were performed as described in the Current Protocols.⁵⁰ New England Biolabs, UK supplied DNA modifying enzymes, which were used as recommended by the manufacturer. Custom-synthesized oligonucleotides were supplied by Invitrogen (Paisley, UK). *E. coli* strain TB-1 [F *ara* Δ (*lac-proAB*) *rps* ϕ 80d *lacZ* Δ M15 *thi* *hsdR*17 *rpsL* (*Str*^r)] was used as the host throughout this study. The bacteria were cultured at 30°C in Luria-Bertani broth (1% (w/v) Tryptone, 0.5% (w/v) yeast extract and 1% (w/v) NaCl) containing 100 μ g ampicillin per milliliter. A saturated culture served as a 2% (v/v) starter inoculum for seeding a phosphate-limited (0.1 mM) MOPS medium containing 100 μ g ampicillin per milliliter and growth at 35°C. Potentially positive clones expressing Cyt isoforms were detected by the PINK Reporter expression system.²⁴

Expression, detection, and quantification of proteins

E. coli TB1 harboring appropriate plasmids were cultured in 100 mL MOPS media for 8 h at 35°C with orbital agitation at 125 rpm. The harvested cells (centrifuged at 5,000g) were resuspended in 5 mL 20% sucrose (w/v), 0.33M Tris-HCl (pH 8), and 1 mM EDTA (STE). After leaving at room temperature for 10 min, the recovered cells were resuspended in the residual STE medium (~100 μ L) and were osmotically shocked by rapid mixing with 1 mL of ice-cold 0.5 mM MgCl₂. After leaving on ice for 10 min, the cell suspension was centrifuged at 5000g for 10 min and the supernatant fraction representing the periplasmic extract was retained.

All of the Cyt forms described in this study were spectrally identical to the native form⁵¹ and were quantified from the Soret absorption peaks using absorbance coefficients of 185 mM⁻¹ cm⁻¹ for the dithiothreitol-reduced form.⁵² Complete absorption spectra of the solution appropriately diluted in 10 mM Tris-acetate pH 8 buffer were monitored by scanning from 350 to 450 nm in a 1-cm light path cuvette. When heme was included to convert the apo-Cyt to holo-Cyt, it was added as a stock 1 mM solution in 0.1M Tris-HCl (pH 8.2), 80% (v/v) ethylene glycol to both the test and the reference cuvettes so as to give final concentrations ranging from 1 to 5 μ M heme.

The periplasmic fractionated proteins were separated in discontinuous, nondenaturing 15% poly-

acrylamide gels⁵⁰ with gel buffer at pH 8.2 and detected by staining with Coomassie blue R-250. The relative amounts of protein in the bands, detected using Biorad GS-800 calibrated Densitometer, were quantified using Phoretix 1D software (version 5.1) operating under MS Windows XPTM. Protein content was estimated from the absorbance rise at 595 nm using the Coomassie Blue G-250 dye-binding procedure of Bradford,⁵³ employing bovine serum albumin as the standard.

The processing of isolated pre-Cyt proteins³⁶ by purified SPase was performed as described previously.¹⁸ Protein sequencing of Cyt electroeluted from nondenaturing gels was performed on an Applied Biosystem 473A sequencer.

Alkaline phosphatase activity was assayed using the substrate *p*-nitrophenylphosphate (1 mM) as described previously^{25,54} in 0.1M Tris-HCl (pH 8.0) at 30°C and the rate of hydrolysis to yellow *p*-nitrophenolate (absorption coefficient of 17×10^{-3} L mol⁻¹ cm⁻¹) was monitored at 400 nm.

Acknowledgments

Khalil Ahmed was supported by funding from Karakoram International University, Gilgit, Pakistan.

References

1. Vaillancourt PE, editor (2003) *E. coli* gene expression. New Jersey, USA: Humana Press.
2. Stampolidis P, Kaderbhai NN, Kaderbhai MA (2009) Periplasmically-exported lupanine hydroxylase undergoes transition from soluble to functional inclusion bodies in *Escherichia coli*. Arch Biochem Biophys 484:8–15.
3. Deutscher J, Saier MH, Jr (2005) Ser/Thr/Tyr protein phosphorylation in bacteria—for long time neglected, now well established. J Mol Microbiol Biotechnol 9: 125–131.
4. Langdon RH, Cuccui J, Wren BW (2009) N-linked glycosylation in bacteria: an unexpected application. Future Microbiol 4:401–412.
5. Hopper DJ, Kaderbhai MA, Marriott SA, Young M, Rogozinski J (2002) Cloning, sequencing and heterologous expression of the gene for lupanine hydroxylase, a quinocytocrome c from a *Pseudomonas* sp. Biochem J 367:483–489.
6. von Heijne G (1990) The signal peptide. The limits of variation. J Membr Biol 115:195–201.
7. de Keyser J, van der Does C, Driessen AJ (2003) The bacterial translocase: a dynamic protein channel complex. Cell Mol Life Sci 60:2034–2052.
8. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1–6.
9. Laforet GA, Kendall DA (1991) Functional limits of conformation, hydrophobicity, and steric constraints in prokaryotic signal peptide cleavage regions. Wild type transport by a simple polymeric signal sequence. J Biol Chem 266:1326–1334.
10. Wang L, Miller A, Kendall DA (2000) Signal peptide determinants of SecA binding and stimulation of ATPase activity. J Biol Chem 275:10154–10159.
11. Friedmann E, Lemberg MK, Weihofen A, Dev KK, Dengler U, Rovelli G, Martoglio B (2004) Consensus

- analysis of signal peptide peptidase and homologous human aspartic proteases reveals opposite topology of catalytic domains compared with presenilins. *J Biol Chem* 279:50790–50798.
12. Kaderbhai N, Khan T, Kaderbhai MA (2008) An antimicrobial peptide derivative of flesh fruit fly mimics secretory signal sequence and inhibits signal peptidase-I in the export pathway. *Int J Pept Res Ther* 14: 173–181.
 13. Rusch SL, Mascolo CL, Kebir MO, Kendall DA (2002) Juxtaposition of signal-peptide charge and core region hydrophobicity is critical for functional signal peptides. *Arch Microbiol* 178:306–310.
 14. Li P, Beckwith J, Inouye H (1988) Alteration of the amino terminus of the mature sequence of a periplasmic protein can severely affect protein export in *Escherichia coli*. *Proc Natl Acad Sci USA* 85:7685–7689.
 15. Paetzel M, Goodall JJ, Kania M, Dalbey RE, Page MG (2004) Crystallographic and biophysical analysis of a bacterial signal peptidase in complex with a lipopeptide-based inhibitor. *J Biol Chem* 279:30781–30790.
 16. Struyve M, Bosch D, Visser J, Tommassen J (1993) Effect of different positively charged amino acids, C-terminally of the signal peptidase cleavage site, on the translocation kinetics of a precursor protein in *Escherichia coli* K-12. *FEMS Microbiol Lett* 109:173–178.
 17. Summers RG, Harris CR, Knowles JR (1989) A conservative amino acid substitution, arginine for lysine, abolishes export of a hybrid protein in *Escherichia coli*. Implications for the mechanism of protein secretion. *J Biol Chem* 264:20082–20088.
 18. Kaderbhai MA, Davey HM, Kaderbhai NN (2004) A directed evolution strategy for optimized export of recombinant proteins reveals critical determinants for preprotein discharge. *Protein Sci* 13:2458–2469.
 19. Kajava AV, Zolov SN, Kalinin AE, Nesmeyanova MA (2000) The net charge of the first 18 residues of the mature sequence affects protein translocation across the cytoplasmic membrane of Gram-negative bacteria. *J Bacteriol* 182:2163–2169.
 20. Gennity J, Goldstein J, Inouye M (1990) Signal peptide mutants of *Escherichia coli*. *J Bioenerg Biomembr* 22: 233–269.
 21. Sung CY, Gennity JM, Pollitt NS, Inouye M (1992) A positive residue in the hydrophobic core of the *Escherichia coli* lipoprotein signal peptide suppresses the secretion defect caused by an acidic amino terminus. *J Biol Chem* 267:997–1000.
 22. Jarjanazi H, Savas S, Pabalan N, Dennis JW, Ozcelik H (2008) Biological implications of SNPs in signal peptide domains of human proteins. *Proteins* 70:394–403.
 23. McKnight CJ, Stradley SJ, Jones JD, Gierasch LM (1991) Conformational and membrane-binding properties of a signal sequence are largely unaltered by its adjacent mature region. *Proc Natl Acad Sci USA* 88: 5799–5803.
 24. Kaderbhai N, Gallagher J, He M, Kaderbhai MA (1992) A pink bacterium as a reporter system signaling expression of a recombinant protein. *DNA Cell Biol* 11: 567–577.
 25. Karim A, Kaderbhai N, Evans A, Harding V, Kaderbhai MA (1993) Efficient bacterial export of a eukaryotic cytoplasmic cytochrome. *Biotechnology (NY)* 11: 612–618.
 26. Kaderbhai NN, Harding VJ, Kaderbhai MA (1995) Over-production and isolation of a precursor protein in a form ideal as a substrate for leader peptidase. *Biochem Soc Trans* 23:58S.
 27. Kaderbhai N, Karim A, Hankey W, Jenkins G, Venning J, Kaderbhai MA (1997) Glycine-induced extracellular secretion of a recombinant cytochrome expressed in *Escherichia coli*. *Biotechnol Appl Biochem* 25 (Pt 1): 53–61.
 28. Kaderbhai NN, Harding V, Kaderbhai MA (2008) Signal peptidase I-mediated processing of an engineered mammalian cytochrome b(5) precursor is an exocytoplasmic post-translocational event in *Escherichia coli*. *Mol Membr Biol* 25:388–399.
 29. Kaderbhai MA, Sligar SG, Barnfield T, Reames T, Gallagher J, He M, Mercer EI, Kaderbhai N (1990) A novel series of pEX-PINK expression vectors for screening high-level production of (un)fused foreign proteins by rapid visual detection of PINK *Escherichia coli* clones. *Nucleic Acids Res* 18:4629–4630.
 30. Noel D, Nikaido K, Ames GF (1979) A single amino acid substitution in a histidine-transport protein drastically alters its mobility in sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Biochemistry* 18: 4159–4165.
 31. Obaru K, DeClue JE, Haseltine WA (1992) An amino-terminal amino acid affects the electrophoretic mobility of the HIV-1 nef protein. *J Acquir Immune Defic Syndr* 5:308–312.
 32. Harding V, Karim A, Kaderbhai N, Jones A, Evans A, Kaderbhai MA (1993) Processing of chimeric mammalian cytochrome b5 precursors in *Escherichia coli*: reaction specificity of signal peptidase and identification of an aminopeptidase in post-translocational processing. *Biochem J* 293:751–756.
 33. LaBranche CC, Sauter MM, Haggarty BS, Vance PJ, Romano J, Hart TK, Bugelski PJ, Marsh M, Hoxie JA (1995) A single amino acid change in the cytoplasmic domain of the simian immunodeficiency virus transmembrane molecule increases envelope glycoprotein expression on infected cells. *J Virol* 69:5217–5227.
 34. Geller B, Zhu HY, Cheng S, Kuhn A, Dalbey RE (1993) Charged residues render pro-OmpA potential dependent for initiation of membrane translocation. *J Biol Chem* 268:9442–9447.
 35. Bachmair A, Finley D, Varshavsky A (1986) *In vivo* half-life of a protein is a function of its amino-terminus residue. *Science* 234:179–186.
 36. Kaderbhai N, Kaderbhai MA (1996) Expression, isolation, and characterization of a signal sequence-appended chimeric precursor protein. *Protein Expr Purif* 7:237–246.
 37. Zimmerman JM, Eliezer N, Simha R (1968) The characterization of amino acid sequences in proteins by statistical methods. *J Theor Biol* 21:170–201.
 38. Chothia C (1976) The nature of the accessible and buried surfaces in proteins. *J Mol Biol* 105:1–12.
 39. Choo KH, Tong JC, Ranganathan S (2008) Modeling *Escherichia coli* signal peptidase complex with bound substrate: determinants in the mature peptide influencing signal peptide cleavage. *BMC Bioinform Proc Open Access* 9:S15.
 40. Karla A, Lively MO, Paetzel M, Dalbey R (2005) The identification of residues that control signal peptidase cleavage fidelity and substrate specificity. *J Biol Chem* 280:6731–6741.
 41. Guermeur Y, Geourjon C, Gallinari P, Deleage G (1999) Improved performance in protein secondary structure prediction by inhomogeneous score combination. *Bioinformatics* 15:413–421.
 42. Ryan P, Robbins A, Whealy M, Enquist LW (1993) Overall signal sequence hydrophobicity determines the

- in vivo* translocation efficiency of a herpesvirus glycoprotein. *Virus Genes* 7:5–21
43. Izard JW, Kendall DA (1994) Signal peptides: exquisitely designed transport promoters. *Mol Microbiol* 13: 765–773.
 44. Ryan P, Edwards CO (1995) Systematic introduction of proline in a eukaryotic signal sequence suggests asymmetry within the hydrophobic core. *J Biol Chem* 270: 27876–27879.
 45. Pugsley AP (1993) The complete general secretory pathway in Gram-negative bacteria. *Microbiol Rev* 57:50–108
 46. Ichihara S, Hussain M, Mizushima S (1981) Characterization of new membrane lipoproteins and their precursors of *Escherichia coli*. *J Biol Chem* 256:3125–3129.
 47. Dev IK, Harvey RJ, Ray PH (1985) Inhibition of pro-lipoprotein signal peptidase by globomycin. *J Biol Chem* 260:5891–5894.
 48. Oroszlan S, Luftig RB (1990) Retroviral proteinases. *Curr Top Microbiol Immunol* 157:153–185.
 49. Nilsson I, von Heijne G (1992) A signal peptide with a proline next to the cleavage site inhibits leader peptidase when present in a sec-independent protein. *FEBS Lett* 299:243–246.
 50. Ausubel FM, Brent R, Kingston RE, Moore DE, Seidman JG, Smith JA, Struhl KA (2001) *Current protocols in molecular biology*. New York: John Wiley & Sons.
 51. Strittmatter P, Fleming P, Connors M, Corcoran D (1978) Purification of cytochrome b5. *Methods Enzymol* 52:97–101.
 52. Burchell A (1985) A simple method for purification of rat hepatic microsomal cytochrome b5. *Biochem J* 226: 339–341.
 53. Bradford M (1976) A rapid and sensitive method for the quantitation of protein utilising the principle of protein-dye binding. *Anal Biochem* 72:248–254.
 54. Torriani A, editor (1967) *Alkaline phosphatase from Escherichia coli*. New York: Harper and Row Publishers.