Volume 244, number 2, 439-446 FEB 06856 February 1989

Species-specific variation in signal peptide design

Implications for protein secretion in foreign hosts

Gunnar von Heijne* and Lars Abrahmsén

*Department of Molecular Biology, Center for Biotechnology, Karolinska Institutet, Huddinge Hospital-K87, S-141 86 Huddinge and Department of Biochemistry, Royal Institute of Technology, S-100 44 Stockholm, Sweden

Received 3 January 1989

Secretory signal peptides from individual prokaryotic and eukaryotic species have been analyzed, and the lengths and amino acid compositions of the positively charged amino-terminal region, the central hydrophobic region, and the carboxy-terminal cleavage-region have been compared. We find distinct differences between species in all three regions. Implications for protein secretion in foreign hosts are discussed.

Protein secretion; Signal peptide; Leader peptide

1. INTRODUCTION

The mechanism of protein secretion is highly conserved throughout the living world. In the great majority of cases, secretory proteins are made with an amino-terminal extension, a signal peptide, that targets the precursor to translocation sites on the appropriate membrane. During or shortly after translocation, the signal peptide is removed by a signal peptidase.

Signal peptides from bacteria to plants and mammals share a common design [1]. The 'canonical' signal peptide is characterized by a short, positively charged amino-terminal region (n-region) followed by a central hydrophobic region (h-region) and a more polar carboxy-terminal region that contains the cleavage site (c-region). The signal peptidase apparently recognizes a '(−3, −1)-pattern' with small, uncharged residues in positions −3 and −1 relative to the cleavage site [2].

Correspondence address: G. von Heijne, Department of Molecular Biology, Center for Biotechnology, Karolinska Institutet, Huddinge Hospital-K87, S-141 86 Huddinge, Sweden

Despite these conserved features, previous comparisons have shown that eukaryotic signal peptides differ in detail from prokaryotic ones [3], and that signal peptides from Gram-positive bacteria tend to be longer than those from Gram negative species [4]; also, signal peptides from one organism do not always function efficiently when expressed in foreign hosts [5,6]. With the much larger database now available, we have undertaken a thorough comparative study of signal peptides from different organisms in an attempt to define more precisely the species-specific variations in signal peptide design.

2. METHODS

All signal peptides were selected from the current version of the SIGPEP database [7], which holds a total of about 200 prokaryotic and 900 eukaryotic sequences. Analysis of the h- and c-regions was carried out for signal peptides with known cleavage sites from *E. coli* (28 sequences), *Bacillus* (16 sequences), *Staphylococcus* (6 sequences), *Streptomyces* (6 sequences), *Homo sapiens* (147 sequences), and a collection of plant signal peptides (22 sequences). In the analysis of the n-regions, additional sequences where the signal peptidase cleavage site is not known were included (31 from *E. coli*, 13

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50 © 1989 Federation of European Biochemical Societies

Table 1

Bacterial and yeast signal peptides analyzed in this paper

Signal peptide		Ref.
Bacillus		
α-Amylase	MFAKRFKTSLLPLFAGFLLLFYLVLAGPAAASA:ETANKSNELT	20
α-Amylase	MIQKRKRTVSFRLVLMCTLLFVSLPITKTSA:VNGTLMQFEW	21
β-Lactamase II	MKKNTLLKVGLCVGLLGTIQFVSTISSVQA:SQKVEKTVIK	22
Middle wall protein	MKKVVNSVLASALALTVAPMAFA:AEEAATTTAP	23
Extracellular amylase	MKMRTGKKGFLSILLAFLLVITSIPFTLVDVEA:HHNGTNGTMM	24
α-Amylase	MKQHKRLYARLLPLLFALIFLLPHSAAAA:ANLNGTLMQY	25
Cyclodextrin glucanotransferase	MKRFMKLTAVWTLWLSLTLGLLSPVHA:APDTSVSNKQ	26
Cyclodextrin glucanotransferase	MKSRYKRLTSLALSLSMALGISLPAWA:SPDTSVDNKV	27
Alkaline cellulase	MLRKKTKQLISSILILVLLLSLFPTALAA:EGNTREDNFK	28
α-Amylase	MLTFHRIIRKGWMFLLAFLLTALLFCPTGQPAKA:AAPFNGTMMQ	29
Levansucrase	MNIKKFAKQATVLTFTTALLAGGATQAFA:KETNQKPYKE	30
Outer wall protein	MNKKVVLSVLSTTLVASVAASAFA:APKDGIYIGG	31
Xylanase	MNLRKLRLLFVMCIGLTLILTAVPAHA:RTITNNEMGN	32
Subtilisin E	MRSKKLWISLLFALTLIFTMAFSNMSAQA:AGKSSTEKKY	33,34
β-Amylase	MTLYRSLWKKGCMLLLSLVLSLTAFIGSPSNTASA:AVADDFQASV	35
Sphingomyelinase	MKGKLLKGVLSLGVGLGALYSGTSAQA:EASTNQNDTL	36
β-Lactamase II	MFVLNKFFTNSHYKKIVPVVLLSCATLIGCSNSNTQSES	37
Neutral protease	MGLGKKLSSAVAASFMSLTISLPGVQAAENPQLKENL	38
Amylase	MKGKKWTALALTLPLAASLSTGVDAETVHKGKAPT	39
Levanase	MKKKVLALAAAITVVAPLQSVAFAHENDGGSKIK	40
β-Lactamase	MKKRLIQVMIMFTLLLTMAFSADAADSSYYDEDY	41
β-Lactamase	MKLWFSTLKLKKAAAVLLFSCVALAGCANNQTNASQ	41
β-Lactamase I	MKNKKMLKIGMCVGILGLSITSLVTFTGGALQVEAKEKTG	42
β-Lactamase II	MKNTLLKLGVCVSLLGITPFVSTISSVQAERTVEHKVIK	43
Endo-β-1,4-glucanase	MKRSISIFITCLLITLLTMGGMIASPASAAGTKTPVAKN	44
Ribonuclease	MMKMEGIALKKRLSWISVCLLVLVSAAGMLFSTAAKTETSSHKAE	45
Subtilisin Carlsberg	MMRKKSFWLGMLTAFMLVFTMAFSDSASAAQPAKNVEKD	46
Neutral protease	MNKRAMLGAIGLAFGLLAAPIGASAKGESIVWNEQ	47
β-Glucanase	MPYLKRVLLLLVTGLFMSLFAVTATASAKTGGSFFDPF	48
Staphylococcus		
Nuclease	MAISNVSKGQYAKRFFFFATSCLVLTLVVVSSLSSSANA:SQTDNGVNRS	49
Exofoliative toxin B	MDKNMFKKIILAASIFTISLPVIPFESTLQA:KEYSAEEIRK	50
Protein A	MKKKNIYSIRKLGVGIASVTLGTLLISGGVTPAANA:AQHDEAQQNA	51
Enterotoxin A	MKKTAFTLLLFIALTLTTSPLVNG:SEKSEEINEK	52
α-Toxin	MKTRIVSSVTTTLLLGSILMNPVAGA:ADSDINIKTG	53
Enterotoxin C1	MNKSRFISCVILIFALILVLFTPNVLA:ESQPDPTPDE	54
Lipase	MKETKHQHTFSIRKSAYGAASFMVASCIFVIGGGVAEANDSTTQTT	55
Serine protease	MKGKFLKVSSLFVATLTTATLVSSPAANALSSKAMDNHP	56
Lysostaphin	MKKTKNNYYTRPLAIGLSTFALASIVYGGIQNE	57
Staphylokinase	MLKRSLLFLTVLLLLFSFSSITNEVSASSSFDKGKYK	58
Lipase	MLRGQEERKYSIRKYSIGVVSVLAATMFVVSSHEAQASEKTSTNAAA	59
Exofoliative toxin A	MNNSKIISKVLLSLSLFTVGASAFVIQDELMEK	60
Streptococcus		
Scarlet fever toxin	MENNKEVLKKMVFFVLMKFLGLTILPKGIC:STRPKPSQLQ	61
Protein G	MEKEKKVKYFLRKSAFGLASVSAFLVGSTVFA:VDSPIEDTPI	62,63
Type 6 M protein	MAKNNTNRHYSLRKLKKGTASVAVALSVIGAGLVVNTNEVSA:RVFPRGTVEN	64
Streptokinase	MKNYLSFGMFALLFALTFGTVNSVQA:IAGPEWLLDR	65
Fructosyltransferase	METKVRKKMYKKGKFWVVATITTAMLTGIGLSSVQADEANSTQVSS	66
Streptolysin O	MSNKKTFKKYSRVAGLLTAALIIGNLVTANAESNKQNTAST	67
Streptomyces		
α-Amylase	MARRLATASLAVLAAAATALTAPTPAAA:APPGAKDVTA	68

Table 1 (continued)

Signal peptide		Ref.
SapA	MKRSMQAVGATLTAVGAIGAGLLVTAPAAGAATAGAT:ASYNGVCGSG	69
Streptavidin	MRKIVVAAIAVSLTTVSITASASA:DPSKDSKAQV	70
Agarase	MVNRRDLIKWSAVALGAGAGLAGPAPAAHA:ADLEWEQYPV	71
DD-peptidase	MVSGTVGRGTALGAVLLALLAVPAQAGTAAA:ADLPAPDDTG	72
α-Amylase	MQQRSRVLGGTLAGIVAAAAATVAPWPSQA:TPPGQKTVTA	73
Cellulase	MENPRTTPTPTPLRRRSERRARGGRVLTALTGVTLLAGLAIAPAATGASPSPAPPASP	74
EndoH	MFTPVRRRVRTAALALSAAAALVLGSTAASGASATPSPAPA	75
β-Lactamase	MHPSTSRPSRRTLLTATAGAALAAATLVPGTAHASSGGRGHGSGSVSDAERRLAGL	76
β-Galactosidase	MPHSPVSPAESPAPQPGRPRPVVSRRRLLEGGAAVLGALALSASPLTAQAAVRRAAADEPPEWNDF	77
Protease A	MTFKRFSPLSSTSRYARLLAVASGLVAAAALATPSAVAAPEAESKATV	78
Protease B	MRIKRTSNRSNAARRVRTTAVLAGLAAVAALAVPTANAETPRTFSANQ	78
XP55	MTARRTRWTRRTDRSLPIRSAAAAVAFAAGATACSAPTGGGGDGGTEAAE	79
Yeast		
α-Galactosidase	MFAFYFLTACISLKGVFG:VSPSYNGLGL	80
Acid phosphatase	MFKSVVYSILAASLANA:GTIPLGKLAD	81
Carboxypeptidase Y	MKAFTSLLCGLGLSTTLAKA:ISLQRPL...	82
28 kDa killertoxin	MKIYHIFSVCYLITLCA:AATTAREEFF	83
Invertase	MLLQAFLFLLAGFAAKISA:SMTNETSDRP	84
Mating factor α-1	MRFPSIFTAVLFAASSALA:APVNTTTEDE	85,86
PEP4	MFSLKALLPLALLLVSANQVAAKVHKAKIYKH	87
Mating factor α-2	MKFISTFLTFILAAVSVTASSDEDIAQVPA	88
BAR1	MSAINHLCLKLILASFAIINTITALTNDGTGHLE	89
K1 toxin	MTKPTQVLVRSVSILFFITLLHLVVALNDVAGPAET	90
Glucoamylase	MVGLKNPYTHTMQRPFLLAYLVLSLLFNSALGFPTALVPRGS	91
Killer plasmid ORF2	MNIFYIFLFLLSFVQGLEHTHRRGSL	92

Known signal peptidase I cleavage sites are indicated by :

from *Bacillus*, 6 from *Staphylococcus*, 6 from *Streptococcus*, 7 from *Streptomyces*, 12 from yeast, and 9 from plants). The sequences from Gram-positive bacteria and from yeast are listed in table 1. Listings of the other samples can be obtained from G.v.H., who also distributes the SIGPEP database in Macintosh format.

Statistical significance was assessed by χ^2 and two-sided *t*-test analysis.

3. RESULTS

Cumulative distributions of the overall lengths of the different signal peptide samples are presented in fig.1. There is a clear gradation from the short *Homo* and plant sequences (mean length = 22.5 and 23.9 residues) and the *E. coli* sequences (mean length = 24.1) to the much longer signal peptides from Gram-positive bacteria (mean length = 29–31). The differences in mean length between these two groups are statistically significant ($p < 0.005$ by two-sided *t*-test).

3.1. *The c-region*

Fig.2 shows plots of the incidence of hydrophobic (ACFILMV) and turn-promoting (DGNPS) [8] residues when the different samples are aligned with coincident cleavage sites. The mean c-region length, as read off from the figure, is five residues for the *Homo* and plant samples, six residues for the *E. coli*, and eight residues for the *Bacillus* samples. For *Staphylococcus*, *Streptococcus*, and *Streptomyces*, the number of signal peptides with known cleavage sites is too small to make reliable estimates of the c-region length.

As for the cleavage site, the $(-3,-1)$-rule is faithfully obeyed in all samples (not shown).

3.2. *The h-region*

We define the border between the n- and h-regions as being located immediately after the most C-terminal charged residue in the n-region. In some cases, this residue is followed by one or more uncharged but polar residues before the first strongly hydrophobic residue is encountered, but for simplicity we adhere to the above definition in these cases as well. In practice, this will not matter

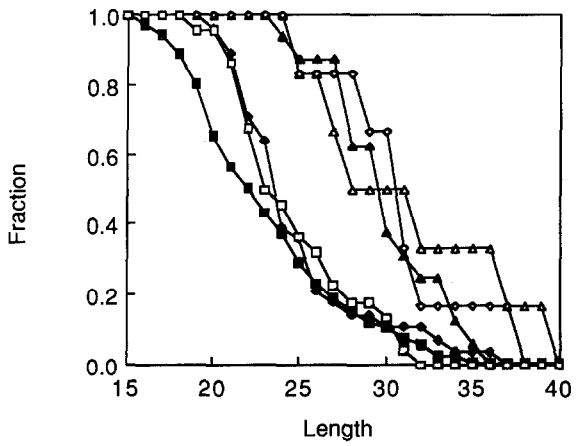

Fig.1. Cumulative distributions of the overall lengths of signal peptides with known cleavage sites from different species [*Homo* (□), plants (■), *Bacillus* (▲), *Staphylococcus* (△), *Streptomyces* (◇), *E. coli* (◆)].

for our qualitative conclusions and will have only a minor influence on the mean lengths and amino acid compositions calculated below.

For the combined h- and c-regions, we find that the total mean lengths are 17.6 for *Homo*, 20.0 for the plant sample, 18.2 for *E. coli*, 21.8 for *Bacillus*, 23.0 for *Staphylococcus*, and 24.5 for *Streptomyces*. Subtracting the c-region lengths found above, we thus calculate mean h-region lengths ranging from around 12 residues (*Homo* and *E. coli*) to 15 or more residues for the Gram-positive signal peptides. In terms of overall amino acid composition, the eukaryotic h-regions are relatively rich in Leu (40%) and contain less Ala (10%); h-regions from *E. coli* and all the Gram-positive bacteria contain 25–35% Leu and 10% (*Bacillus, Staphylococcus*) or 30% (*E. coli, Streptomyces*) Ala.

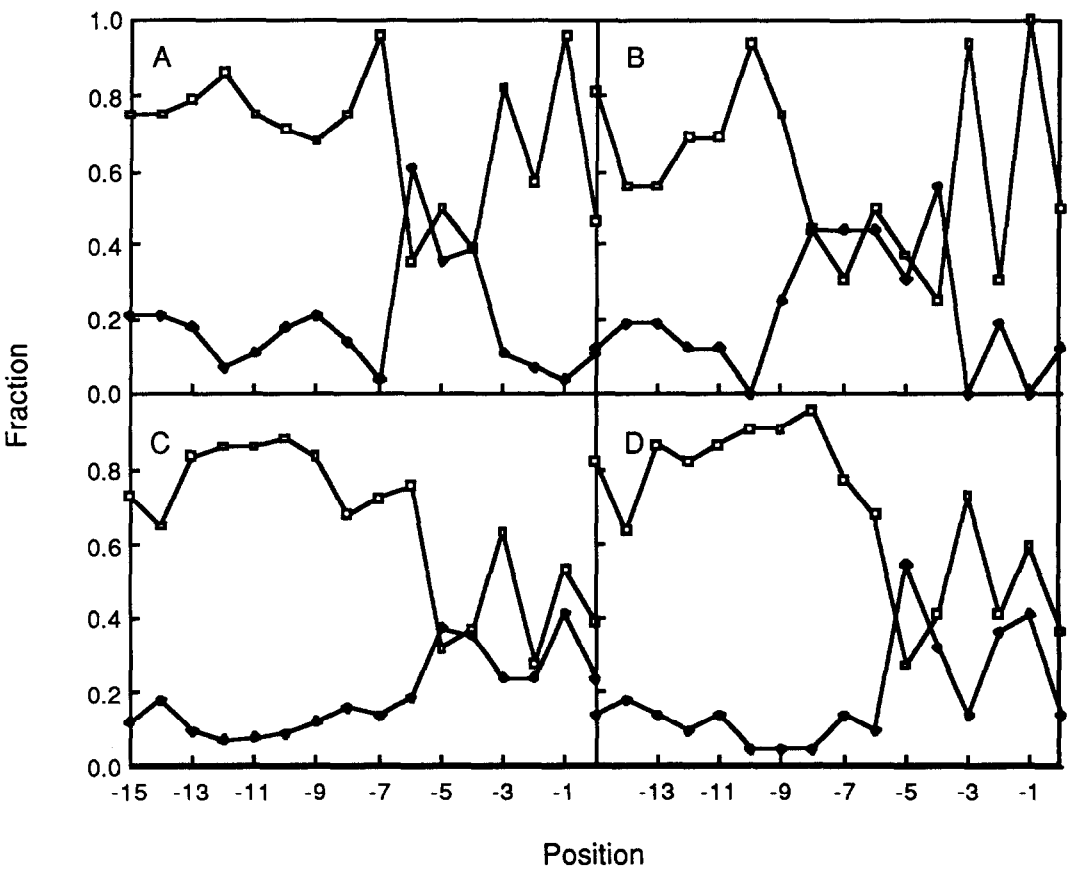

Fig.2. Distribution of hydrophobic [Ala + Cys + Phe + Ile + Leu + Met + Val (□)] and turn-promoting [Asp + Gly + Asn + Pro + Ser (■)] residues relative to the signal peptidase I cleavage site (between position −1 and +1) for different species [(A) *E. coli*, (B) *Bacillus*, (C) *Homo*, (D) plants].

3.3. *The n-region*

The length and net charge distributions for the n-region follow similar trends, with the *Homo*, yeast, and plant samples being shortest (mean length 4–5 residues) and carrying the smallest net charge (mean net charge around +0.8, not counting the amino group on the initiator Met), the *E. coli* sample being intermediate (5.5 residues, +2.0), and the *Bacillus* and *Staphylococcus* samples being longer (7–8 residues) and more highly charged (+3.0 and +2.8). The n-regions from *Streptococcus* and *Streptomyces* are particularly long (12 residues) and have mean net charges of +4.3 and +3.5, respectively.

In terms of overall amino acid composition, the n-regions from *Streptomyces* stand out from the other samples by virtue of their very high Arg (30%) and low Lys (3%) content. The other bacterial samples have around 10% Arg and 35% Lys in this region.

4. DISCUSSION

As demonstrated above, there are indeed clear-cut differences between signal peptides from various species. In particular, eukaryotic signal peptides tend to have shorter n-, h- and c-regions than bacterial signal peptides, and among the latter, those from Gram-positive bacteria generally have longer n-, h- and c-regions than those from the Gram-negative *E. coli*. The net charge of the n-region is higher for the Gram-positive signal peptides.

Differences in the c-region may well be related to slightly different signal peptidase specificities in the different species. It is less obvious what may cause the differences in the n- and h-regions. A simple explanation would be that differences in the lipid composition of membranes play a role [9]. Thus, *E. coli* membranes contain more of the zwitterionic phosphatidylethanolamine and less of the negatively charged cardiolipin than, for example, *Bacillus* and *Streptomyces* [10,11].

Eukaryotic signal peptides have more hydrophobic h-regions than those of bacteria (more Leu and less Ala). The n-regions of the eukaryotic sequences contain less Lys (10–20% vs 30%) but approximately the same percentage of Arg (10%) as the prokaryotic ones.

Signal peptides from the Gram-positive *Streptomyces* are distinct from all other signal peptides in that their n-regions are much longer and contain much Arg (~30%) but almost no Lys (3%) residues ($p < 10^{-4}$). It is likely that these peculiarities result, at least in part, from the high G+C content of these bacteria (~70–75%) and the fact that lysine codons are GC-poor (AAA, AAG) whereas arginine codons are GC-rich (CGN, AGA, AGG). However, mature cytoplasmic and extracellular proteins from *Streptomyces* contain about the same percentages of Lys (3%) but only 6–7% Arg (not shown). Possibly, Arg can functionally replace Lys in signal peptides but not with equal ease in the mature parts of proteins.

Nevertheless, the excessive lengths of the Arg-rich n-regions from *Streptomyces* could be an indication that Arg may be in some sense less 'efficient' than Lys in signal peptide n-regions, making it necessary for these regions to be longer and/or more highly charged. We have tested this idea by comparing n-regions from eukaryotic signal peptides that contain only Arg but no Lys (108 sequences) with those containing only Lys but no Arg (100 sequences). Indeed, the former group tends to have longer (mean length 5.7 residues, $p < 10^{-4}$) but not more charged (mean net charge +1.2) n-regions than the latter (3.2 residues, +1.0) (fig.3). The Arg-containing n-regions also have a higher content of Pro than the Lys-containing ones (10% vs 1%, $p < 10^{-4}$); the same is true for n-regions from *Streptomyces* compared to n-regions from other bacterial species (12% vs 1%, $p < 10^{-3}$). Thus, when Arg is the only positively charged residue, it tends to be part of a rather long, possibly unstructured, n-region. Signal peptides from bacteria other than *Streptomyces* rarely have n-regions totally lacking Lys residues (none of these sequences listed in table 1 lack Lys in this region; also, out of the 59 *E. coli* sequences analyzed here, only 7 lack Lys in their n-region. Conversely, 8 out of 13 *Streptomyces* signal peptides contain Arg but no Lys; none contains Lys but no Arg).

A clear understanding of the differences in the c-regions is especially important when secretory proteins are expressed in foreign hosts. Although the (−3, −1)-rule seems to be valid for all organisms, more than one site compatible with this rule often exists in the vicinity of the normal cleavage site. It is therefore interesting to note that 'aberrant'

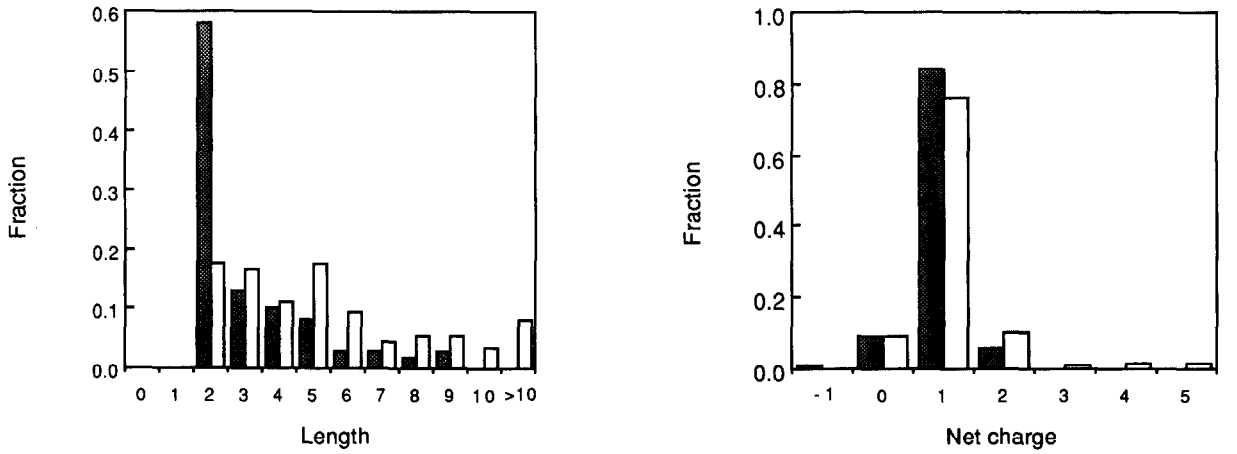

Fig.3. Length and net charge ($= n_{Arg} + n_{Lys} - n_{Asp} - n_{Glu}$) distributions for the n-regions of eukaryotic signal peptides that contain lysine (shaded bars) or arginine (unfilled bars) as the only positively charged residue.

cleavages have sometimes been found when signal peptides from Gram-positive bacteria have been expressed in *E. coli*. Thus, when α-amylase from *Bacillus stearothermophilus* was expressed and secreted by *E. coli*, a large fraction (40%) was incorrectly processed [12]. The aberrant cleavage took place three residues to the amino-terminal side of the normal cleavage site. This is what we would expect from the data presented here, namely that *E. coli* signal peptidase I seems to cleave preferentially six residues after the end of the h-region, whereas the corresponding proteases in Gram-positive bacteria seem to prefer a slightly longer distance from the h-region (7–9 residues).

Another example is provided by the major outer membrane lipoprotein from *E. coli*, which is normally cleaved by signal peptidase II. A mutant lipoprotein signal peptide that is not recognized by this peptidase is nevertheless cleaved by signal peptidase I when expressed in *E. coli* and yeast; however, the cleavage in yeast takes place four residues upstream of the *E. coli* cleavage site [13]. Generalizing from these examples, we anticipate that some eukaryotic signal peptides will be found to be cleaved downstream of their natural cleavage site when expressed in bacterial, and in particular Gram-positive, hosts.

A phenomenon of considerable biotechnological interest has recently been shown to be caused by the use of a heterologous signal peptide [19]. When the signal peptide from staphylococcal protein A was used to secrete proteins in *E. coli*, the morphology of the host cell was affected and periplasmic proteins leaked out to the growth medium [15,19]. A similar 'leaky' phenotype of the host cell has also been observed when other Gram-positive proteins have been expressed and secreted in *E. coli* ([14,16–18]; Nygren, P.-Å., personal communication). Whether this effect was in all cases caused by the foreign signal peptide is unknown. In the light of the results presented above it is, however, tempting to suggest that this phenomenon is caused by a 'mismatch' between the signal peptide and the *E. coli* translocation machinery, resulting in a pleiotropic secretion defect with secondary effects on the structure of the outer membrane.

Acknowledgement: This work was supported by grants from the Swedish Natural Sciences Research Council.

REFERENCES

[1] Von Heijne, G. (1985) J. Mol. Biol. 184, 99–105.
[2] Von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4690.
[3] Von Heijne, G. (1984) EMBO J. 3, 2315–2318.
[4] Abrahmsén, L., Moks, T., Nilsson, B., Hellman, U. and Uhlén, M. (1985) EMBO J. 4, 3901–3906.
[5] Smith, R.A., Duncan, M.J. and Moir, D.T. (1985) Science 229, 1219–1224.
[6] Chang, C.N., Rey, M., Bochner, B., Heynecker, H. and Gray, G. (1987) Gene 55, 189–196.
[7] Von Heijne, G. (1987) Protein Seq. Data Anal. 1, 41–42.

[8] Levitt, M. (1978) Biochemistry 17, 4277–4285.
[9] De Vrije, T., De Swart, R.L., Dowhan, W., Tommassen, J. and De Kruijff, B. (1988) Nature 334, 173–175.
[10] Datta, D.B. (1987) A Comprehensive Introduction to Membrane Biochemistry, pp.72, Floral Publishing, Madison.
[11] Sternberg, B., Gumpert, J., Reinhardt, G. and Gawrisch, K. (1987) Biochim. Biophys. Acta 898, 223–230.
[12] Suominen, I., Karp, M., Lautamo, J., Knowles, J. and Mantsälä, P. (1987) Extracellular Enzymes of Microorganisms (Chaloupka, J. and Krumphanz, V. eds) pp.129–137, Plenum, New York.
[13] Pines, O., Lunn, C.A. and Inouye, M. (1988) Mol. Microbiol. 2, 209–217.
[14] Götz, F., Popp, F., Korn, E. and Schleifer, K.H. (1985) Nucleic Acids Res. 13, 5895–5906.
[15] Abrahmsén, L., Moks, T., Nilsson, B. and Uhlén, M. (1986) Nucleic Acids Res. 14, 7487–7500.
[16] Malke, H. and Ferretti, J.J. (1984) Proc. Natl. Acad. Sci. USA 81, 3557–3561.
[17] Recsei, P.A., Gruss, A.D. and Novick, R.P. (1987) Proc. Natl. Acad. Sci. USA 84, 1127–1131.
[18] Murphy, N., McConell, D.J. and Cantwell, B.A. (1984) Nucleic Acids Res. 12, 5355–5367.
[19] Abrahmsén, L., Nygren, P.-Å., Uhlén, M., Anderson, S. and Nilsson, B. (1988) submitted.
[20] Yamazaki, H., Ohmura, K., Nakayama, A., Takeichi, Y., Otozai, K., Yamasaki, M., Tamura, G. and Yamane, K. (1983) J. Bacteriol. 156, 327–337.
[21] Takkinen, K., Pettersson, R.F., Kalkkinen, N., Palva, I., Söderlund, H. and Kääriäinen, L. (1983) J. Biol. Chem. 258, 1007–1013.
[22] Hussain, M., Carlino, A., Madonna, M.J. and Lampen, J.O. (1985) J. Bacteriol. 164, 223–229.
[23] Tsuboi, A., Uchichi, R., Adachi, T., Sasaki, T., Hayakawa, S., Yamagata, Y., Tsukagoshi, N. and Udaka, S. (1988) J. Bacteriol. 170, 935–945.
[24] Tsukamoto, A., Kimura, K., Ishii, Y., Takano, T. and Yamane, K. (1988) Biochem. Biophys. Res. Commun. 151, 25–31.
[25] Sibakov, M. and Palva, I. (1984) Eur. J. Biochem. 145, 567–572.
[26] Kimura, K., Kataoka, S., Ishii, Y., Takano, T. and Yamane, K. (1987) J. Bacteriol. 169, 4399–4402.
[27] Takano, T., Fukuda, M., Monma, M., Kobayashi, S., Kainuma, K. and Yamane, K. (1986) J. Bacteriol. 166, 1118–1122.
[28] Fukumori, F., Kudo, T., Narahashi, Y. and Horikoshi, K. (1986) J. Gen. Microbiol. 132, 2329–2335.
[29] Nakajima, R., Imanaka, T. and Aiba, S. (1985) J. Bacteriol. 163, 401–406.
[30] Steinmetz, M., Le Coq, D., Aymerich, S., Gonzy-Treboul, G. and Gay, P. (1985) Mol. Gen. Genet. 200, 220–228.
[31] Tsubio, A., Uchihi, R., Tabata, R., Takahashi, Y., Hashiba, H., Sasaki, T., Yamagata, H., Tsukagoshi, N. and Udaka, S. (1986) J. Bacteriol. 168, 365–373.
[32] Fukusaki, E., Panbangred, W., Shinmyo, A. and Okada, H. (1984) FEBS Lett. 171, 197–201.
[33] Wong, S.-L., Price, C.W., Goldfarb, D.S. and Doi, R.H. (1984) Proc. Natl. Acad. Sci. USA 81, 1184–1188.
[34] Wong, S.-L. and Doi, R.H. (1986) J. Biol. Chem. 261, 10176–10181.
[35] Kawazu, T., Nakanishi, Y., Uozumi, N., Sasaki, T., Yamagata, H., Tsukagoshi, N. and Udaka, S. (1987) J. Bacteriol. 169, 1564–1570.
[36] Yamada, A., Tsukagoshi, N., Udaka, S., Sasaki, T., Nakamura, S., Little, C., Tomita, M. and Ikezawa, H. (1988) Eur. J. Biochem. 175, 213.
[37] Hussain, M., Pastor, F.I.J. and Lampen, J.O. (1987) J. Bacteriol. 169, 579–586.
[38] Shimada, H., Honjo, M., Mita, I., Nakayama, A., Akaoka, A., Manabe, K. and Furutani, Y. (1985) J. Biotechnol. 2, 75–85.
[39] Metz, R.J., Allen, L.N., Cao, T.M. and Zeman, N.W. (1988) Nucleic Acids Res. 16, 5203.
[40] Schörgendorfer, K., Schwab, H. and Lafferty, R.M. (1987) Nucleic Acids Res. 15, 9606.
[41] Nielsen, J.B.K. and Lampen, J.O. (1982) J. Biol. Chem. 257, 4490–4495.
[42] Wang, W., Mezes, P.S.F., Yang, Y.Q., Blacher, R.W. and Lampen, J.O. (1985) J. Bacteriol. 163, 487–492.
[43] Lim, H.M., Pène, J.J. and Shaw, R.W. (1988) J. Bacteriol. 170, 2873–2878.
[44] MacKay, R.M., Lo, A., Willick, G., Zuker, M., Baird, S., Dove, M., Moranelli, F. and Seligy, V. (1986) Nucleic Acids Res. 14, 9159–9170.
[45] Paddon, C.J. and Hartley, R.W. (1986) Gene 40, 231–239.
[46] Jacobs, M., Eliasson, M., Uhlén, M. and Flock, J.-I. (1985) Nucleic Acids Res. 13, 8913–8926.
[47] Takagi, M., Imanaka, T. and Aiba, S. (1985) J. Bacteriol. 163, 824–831.
[48] Murphy, N., McConnell, D.J. and Cantwell, B.A. (1984) Nucleic Acids Res. 12, 5355–5367.
[49] Shortle, D. (1983) Gene 22, 181–189.
[50] Jackson, M.P. and Iandolo, J.J. (1986) J. Bacteriol. 167, 726–728.
[51] Löfdahl, S., Guss, B., Uhlén, M., Philipson, L. and Lindberg, M. (1983) Proc. Natl. Acad. Sci. USA 80, 697–701.
[52] Betley, M.J. and Mekalanos, J.J. (1988) J. Bacteriol. 170, 34–41.
[53] Gray, G.S. and Kehoe, M. (1984) Infect. Immun. 46, 615–618.
[54] Bohach, G.A. and Schlievert, P.M. (1987) Mol. Gen. Genet. 209, 15–20.
[55] Liebl, W. and Götz, F. (1986) Mol. Gen. Genet. 204, 166–173.
[56] Carmona, C. and Gray, G.L. (1987) Nucleic Acids Res. 15, 6757.
[57] Recsei, P.A., Gruss, A.D. and Novick, R. (1987) Proc. Natl. Acad. Sci. USA 84, 1127–1131.
[58] Sako, T. and Tsuchida, N. (1983) Nucleic Acids Res. 11, 7679–7693.
[59] Lee, C.Y. and Iandolo, J.J. (1986) J. Bacteriol. 166, 385–391.
[60] Sakurai, S., Suzuki, H. and Kondo, I. (1988) J. Gen. Microbiol. 134, 711–714.
[61] Johnson, L.P., L'Italien, J.J. and Schlievert, P.M. (1986) Mol. Gen. Genet. 203, 354–356.

[62] Fahnestock, S.R., Alexander, P., Nagle, J. and Filpula, D. (1986) J. Bacteriol. 167, 870–880.
[63] Olsson, A., Eliasson, M., Guss, B., Nilsson, B., Hellman, U., Lindberg, M. and Uhlén, M. (1987) Eur. J. Biochem. 168, 319–324.
[64] Hollingshead, S.K., Fischetti, V.A. and Scott, J.R. (1986) J. Biol. Chem. 261, 1677–1686.
[65] Malke et al. (1985) Gene 34, 357–362.
[66] Shiroza, T. and Kuramitsu, H.K. (1988) J. Bacteriol. 170, 810–816.
[67] Kehoe, M.A., Miller, L., Walker, J.A. and Boulnois, G.J. (1987) Infect. Immun. 55, 3228–3232.
[68] Long, C.M., Virolle, M.-J., Chang, S.-Y., Chang, S. and Bibb, M.J. (1987) J. Bacteriol. 169, 5745–5754.
[69] Giujarro, J., Santamaria, R., Schauer, A. and Losick, R. (1988) J. Bacteriol. 170, 1895–1901.
[70] Argarana, C.E., Kuntz, I.D., Birken, S., Axel, R. and Cantor, C.R. (1986) Nucleic Acids Res. 14, 1871–1882.
[71] Buttner, M.J., Fearnley, I.M. and Bibb, M.J. (1987) Mol. Gen. Genet. 209, 101–109.
[72] Duez, C., Piron-Fraipont, C., Joris, B., Dusart, J., Urdea, M.S., Martial, J.A., Frére, J.-M. and Ghuysen, J.-M. (1987) Eur. J. Biochem. 162, 509–518.
[73] Hoshiko, S., Makabe, O., Nojiri, C., Katsumuta, K., Satoh, E. and Nagaoka, K. (1987) J. Bacteriol. 196, 1029–1036.
[74] Nakai, R., Horinuchi, S. and Beppu, T. (1988) Gene 65, 229–238.
[75] Robbins, P.W., Trimble, R.B., Wirth, D.F., Hering, C., Maley, F., Maley, G.F., Das, R., Gibson, B.W., Royal, N. and Biemann, K. (1984) J. Biol. Chem. 259, 7577–7583.
[76] Dehottay, P., Dusart, J., De Meester, F., Joris, B., Van Beeumen, J., Epricum, T., Frére, J.-M. and Ghuysen, J.-M. (1987) Eur. J. Biochem. 166, 345–350.
[77] Eckhardt, T., Strickler, J., Gorniak, L., Burnett, W.V. and Fare, L.R. (1987) J. Bacteriol. 169, 4249–4256.
[78] Henderson, G., Krygsman, P., Liu, C.J., Davey, C.C. and Malek, L.T. (1987) J. Bacteriol. 169, 3778–3784.
[79] Burnett, W.V., Henner, J. and Eckhardt, T. (1987) Nucleic Acids Res. 15, 3926.
[80] Liljeström, P.L. (1985) Nucleic Acids Res. 13, 7257–7268.
[81] Arima, K., Oshima, T., Kubota, I., Nakamura, N., Mizunaga, T. and Toh-e, A. (1983) Nucleic Acids Res. 11, 1657–1672.
[82] Bird, P., Gething, M.-J. and Sambrook, J. (1987) J. Cell Biol. 105, 2905–2914.
[83] Stark, M.J.R. and Boyd, A. (1986) EMBO J. 5, 1995–2002.
[84] Taussig, R. and Carlson, M. (1983) Nucleic Acids Res. 11, 1943–1954.
[85] Kurjan, J. and Herskowitz, I. (1982) Cell 30, 933–943.
[86] Waters, M.G., Evans, E.A. and Blobel, G. (1988) J. Biol. Chem. 263, 6209–6214.
[87] Ammerer, G., Hunter, C.P., Rothman, J.H., Saari, G.C., Valls, L.A. and Stevens, T.H. (1986) Mol. Cell. Biol. 6, 2490–2499.
[88] Singh, A., Chen, E.Y., Lugovoy, J.M., Chang, C.N., Hitzeman, R.A. and Seeburg, P.H. (1983) Nucleic Acids Res. 11, 4049–4063.
[89] MacKay, V.L., Welch, S.K., Insley, M.Y., Manney, T.R., Holly, J., Saari, G.C. and Parker, M.L. (1988) Proc. Natl. Acad. Sci. USA 85, 55–59.
[90] Skipper, N., Thomas, D.Y. and Lau, P.C.K. (1984) EMBO J. 3, 107–111.
[91] Yamashita, I., Suzuki, K. and Fukui, S. (1985) J. Bacteriol. 161, 567–573.
[92] Stark, M.J.R., Mileham, A.J., Romanos, M.A. and Boyd, A. (1987) Nucleic Acids Res. 15, 6011–6030.