

High Incidence of Disorder-Promoting Amino Acids in the Amino Terminus of Mature Proteins in *Bacillus subtilis*

¹Guangqiang Wang, ¹Yining Dong, ²Haiqin Chen,
³Heping Zhang, ²Yuanda Song, ²Hao Zhang and ¹Wei Chen

¹State Key Laboratory of Food Science and Technology,
²Centre for Food Biotechnology, School of Food Science and Technology,
Jiangnan University, Wuxi 214122, P.R. China

³Department of Food Science and Engineering, College of Food Science and Engineering,
Inner Mongolia Agricultural University, Hohhot 010018, P.R. China

Abstract: Problem statement: Exported polypeptides are usually synthesized as preproteins which are composed with mature domains and the signal peptides. The foreign proteins are usually inefficiently secreted by the aid of signal peptides. Although we know that the mature protein played an important role in the protein export process, we do not know what properties of the mature domains are critical for protein export? This study explores the influence of the property of the mature domains on protein export in *Bacillus subtilis*. **Approach:** We analyzed the amino acid composition of 241 predicted exported proteins and 457 cytoplasmic proteins of *Bacillus subtilis* that have been previously reviewed using the program DAMBE, Codeprotein and Codeprotein Win. **Results:** Disorder-promoting amino acids in the first 14 residues of the mature domains were overrepresented in comparison to the residues in the NH₂-terminal region of cytoplasmic proteins or the cytoplasmic proteins. Further, the NH₂ terminus of the mature domain had more negatively charged residues and fewer positively charged residues than the NH₂ terminus of the signal peptide. **Conclusion/Recommendations:** The presence of disordered domains in the NH₂-terminal regions of mature domains may provide the binding site and maintain the exported proteins in a secretion-competent conformation and the contrary charged residues distribution in the two ends of hydrophobic core may help form the hairpin-like structure of the signal peptide. These findings may have important implications for the understanding of the contribution of mature domains to membrane targeting and to translocation.

Key words: Protein export, disorder-promoting amino acid, mature domain, signal peptide, *Bacillus subtilis*, hydrophobic core, charged residues, hairpin-like structure, disordered domains

INTRODUCTION

Exported proteins are initially synthesized as precursors with an amino-terminal signal peptide. These precursors are subsequently processed by a leader peptidase to yield the mature domains. The signal peptide can be divided into 3 distinct domains: a positively charged N-terminus, which has been proposed to provide stable interactions with the negatively charged inner membrane phospholipids; a central hydrophobic core, which plays an important role in binding to the secretory components; and a cleavage site, which is processed by the respective signal peptidase (Fekkes and Driessen, 1999; Tjalsma *et al.*, 2000).

Although the importance and roles of the signal peptides are generally accepted, only a signal peptide is

not sufficient for the export of any fused protein. The foreign proteins are usually inefficiently secreted by the aid of signal peptides (Schein *et al.*, 1986; Tommassen and Kroon, 1987), despite the fact that many of them are secretory proteins by nature (Simonen *et al.*, 1992). Chen and Nagarajan (1993) have shown that the rate of secretion of barnase was improved by replacement of the barnase signal peptide with a heterologous signal peptide and the barnase signal peptide exported *Escherichia coli* alkaline phosphatase faster than mature barnase (Chen and Nagarajan, 1993). These facts suggested that some features of the mature protein played an important role in the protein export process. Indeed, some experimental work has shown that the membrane anchor sequences can block protein export (Davis and Model, 1985) and that the net charge of the

Corresponding Author: Wei Chen, State Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi 214122, P.R. China

NH₂ terminus of the mature sequence affected protein translocation (Li *et al.*, 1988; Kajava *et al.*, 2000). However, the latter rule was restricted to gram-negative bacteria, with no effect on gram-positive bacteria (Kajava *et al.*, 2000). Furthermore, Gouridis have demonstrated that mature domains contained prominent targeting determinants and the targeting signals required for docking at the membrane were only presented on 'non-native' preproteins (Gouridis *et al.*, 2009). Nevertheless, neither of those experiments answered the following question: what properties of the mature domains are critical for protein export? Currently, sufficient amino acid sequences are available to conduct detailed statistical studies on the property of the mature domains of exported proteins in *Bacillus subtilis*.

Although disordered regions lack rigid 3-D structures, they carry out various biological functions (Dunker *et al.*, 2002; Patil and Nakamura, 2006; Salma *et al.*, 2009). Dunker *et al.* (2002) identified twenty-eight separate functions for 98 out of 115 disordered regions that included molecular recognition via binding to other proteins. Patil and Nakamura (2006) reported that disordered domains and high surface charge played a significant role in the binding ability of hubs. Hence disordered regions is important for protein-protein binding. It was not clear whether the binding site of mature proteins was related to the disordered domains during protein targeting. In this study, the amino acid composition of 241 predicted exported proteins and 457 cytoplasmic proteins of *Bacillus subtilis* was explored and revealed preference for disorder-promoting amino acids in the NH₂-terminal regions of mature domains that may help understand the contribution of mature domains to membrane targeting.

MATERIALS AND METHODS

Generation of data sets: The sequence of every Open Reading Frame (ORF) of the genome of *B. subtilis* 168 was obtained via the National Center for Biotechnology Information (accession NC_000964). The exported proteins were taken from the previous survey of the secretome (Tjalsma *et al.*, 2000). To reduce the number of false-positives, PhoD and YwbN, that were exported via the Tat pathway, anomalous exported proteins (those whose lengths of mature proteins were less 60 residues) and proteins with membrane-spanning domains annotated in the UniProtKB database were excluded. The proteins located in the cytoplasm and reviewed in the UniProtKB database were considered as cytoplasmic proteins. This resulted in a total of 241 exported proteins and 457 cytoplasmic proteins in *B. subtilis*. The folded proteins data set was compiled by Prilusky *et al.* (2005).

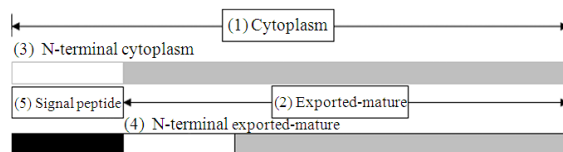


Fig. 1: Derivation of protein sequences comprising data sets 1-5. The top bar depicts cytoplasmic proteins and the bottom bar depicts exported proteins. The regions of sequence depicted are not shown to scale

The amino acid sequences from the exported and cytoplasmic categories were obtained and used to derive 5 data sets (Fig. 1). Data set 1 (cytoplasm) was the amino acid sequences of cytoplasmic proteins. Data set 2 (exported-mature) was the mature domains of exported proteins. Data set 3 (N-terminal cytoplasm) was derived by copying an excerpt of the N-terminal end of data set 1. Data set 4 (N-terminal exported-mature) was the N-terminal end of data set 2. Data set 5 (signal peptide) was the signal peptide of the exported proteins.

A more detailed analysis has provided information on the difference in the composition of ordered and disordered proteins and has enabled the classification of amino acids into 3 classes: order-promoting amino acids (Trp, Tyr, Phe, Ile, Leu, Val, Cys and Asn), disorder-promoting amino acids (Ala, Arg, Gly, Gln, Ser, Glu, Lys and Pro) and uncertain amino acids (His, Asp, Met and Thr) (Uversky and Dunker, 2010).

Analysis of data sets and statistical methods: For each data set, the mean percentage of amino acids was calculated using the program DAMBE (Xia and Xie, 2001), Codeprotein and CodeproteinWin (written in the Python programming language). For analysis, the first amino acid Met of all proteins (except proteins in the folded proteins data set) and the first amino acid Cys at the mature proteins of lipoproteins has been removed in the analysis process. The results were analyzed by one-way ANOVA. R version 2.7.1 was used for all statistical analysis.

RESULTS

Proportion of disordered residues in mature domains and cytoplasmic proteins: Table 1 shows the differences in the amino acid composition of the concatenated data sets. On average, the first 10 residues of data set 4 had 32.05% more disorder-promoting amino acids than the first 10 residues of data set 3 ($p < 2.2e-16$), with 22.66% more compared to data set 1 ($p < 2.2e-16$). But the proportion of disorder-promoting amino acids in data set 2 and data set 1 were virtually the same (48.90 and 48.81%, respectively, $P = 0.1945$).

Table 1: Amino acid composition in the data sets

Data set	Order-promoting acid (%)	Disorder-promoting amino acids (%)	Uncertain amino amino acids (%)
(1) Cytoplasm ^a	35.10	48.81	16.06
(2) Exported-mature ^a	33.84	48.90	17.23
(3) N-terminal cytoplasm (the first ten) ^b	37.64	45.34	17.02
(3) N-terminal cytoplasm (the first twenty) ^b	36.41	47.88	15.70
(4) N-terminal exported-mature (the first ten) ^b	24.70	59.87	15.44
(4) N-terminal exported-mature (the first twenty) ^b	27.11	55.62	17.27
Genomic mean ^a	37.66	46.95	15.38

^a: The percent of order-promoting amino acids were calculated using the the program DAMBE; ^b: The percent of order-promoting amino acids were calculated using the software Codeprotein

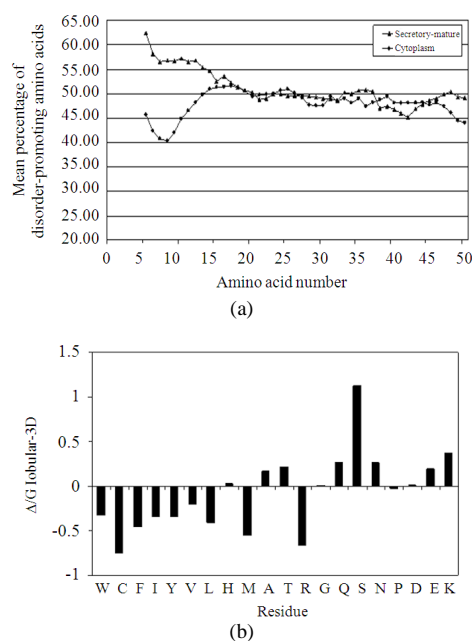


Fig. 2: (A) Mean percentage of disorder-promoting amino acids found in cytoplasmic proteins and exported mature domains. CodeproteinWin analysis of the first 50 amino acids of the sequences in the cytoplasmic proteins and exported mature domains, using a window size of 5. The amino acid number (terminal of window) is indicated on the x-axis and the mean percentage of disordered residues are plotted on the y-axis. (B) Order/the first 14 residues of mature domains composition profile. Comparisons of amino acid compositions of ordered protein with the first 14 residues of mature domains. The ordinates are (% amino acid in the first 14 residues of mature domains-% amino acid in ordered dataset) /(% amino acid in ordered dataset) = Δ /Globular-3D

The results indicated that the NH₂-terminal regions of the mature domains had more disordered residues.

However, the trend seen in the comparison of data set 4 and data set 3 or data set 1 becomes weaker as the length of data set 4 and data set 3 becomes longer. For example, the first 20 residues of data set 4 have only 16.17% more disordered residues than data set 3 (Table 1, $p < 2.2 \times 10^{-16}$). To explore whether the significance resulted from the difference in the early NH₂ terminus of proteins and where these disordered residues are located within the sequences examined, the small software named CodeproteinWin was written in Python to analysis of the first 100 residues of the sequences of secretory protein and cytoplasmic proteins, using a window size of 5. The overall distributions of disordered residues in the cytoplasmic and exported groups of sequences are shown in Fig. 2A. The proportion of disordered residues is relatively higher at the NH₂ terminus of the mature domains and the level of disordered residues rapidly declines from left to right. At a position number 11 (from residue 7 to residue 11), the NH₂ terminus of the mature domains had a massive bias for disordered residues compared to the cytoplasmic protein ($p = 1.356 \times 10^{-7}$). At number 12 and number 13, disordered residues were also preferred ($p = 1.065 \times 10^{-5}$, $p = 0.003945$), as the length becomes longer, the probability of difference between mature proteins and cytoplasmic proteins is equal to or more than the significance level (at number 14, $p = 0.03726$, at number 15, $p = 0.5772$, at number 17, $p = 0.7045$). After number 14, the mean percentage of disordered residues in the cytoplasmic proteins and secretory proteins is virtually the same over the remainder of the plotted sequence (48.88 and 48.47%, respectively, $p = 0.3483$). Therefore, the first 14 residues of the mature domains were mainly composed of disordered residues.

The difference in the amino acid propensities of the first 14 residues of the mature domains and ordered proteins is shown in Fig. 2B.

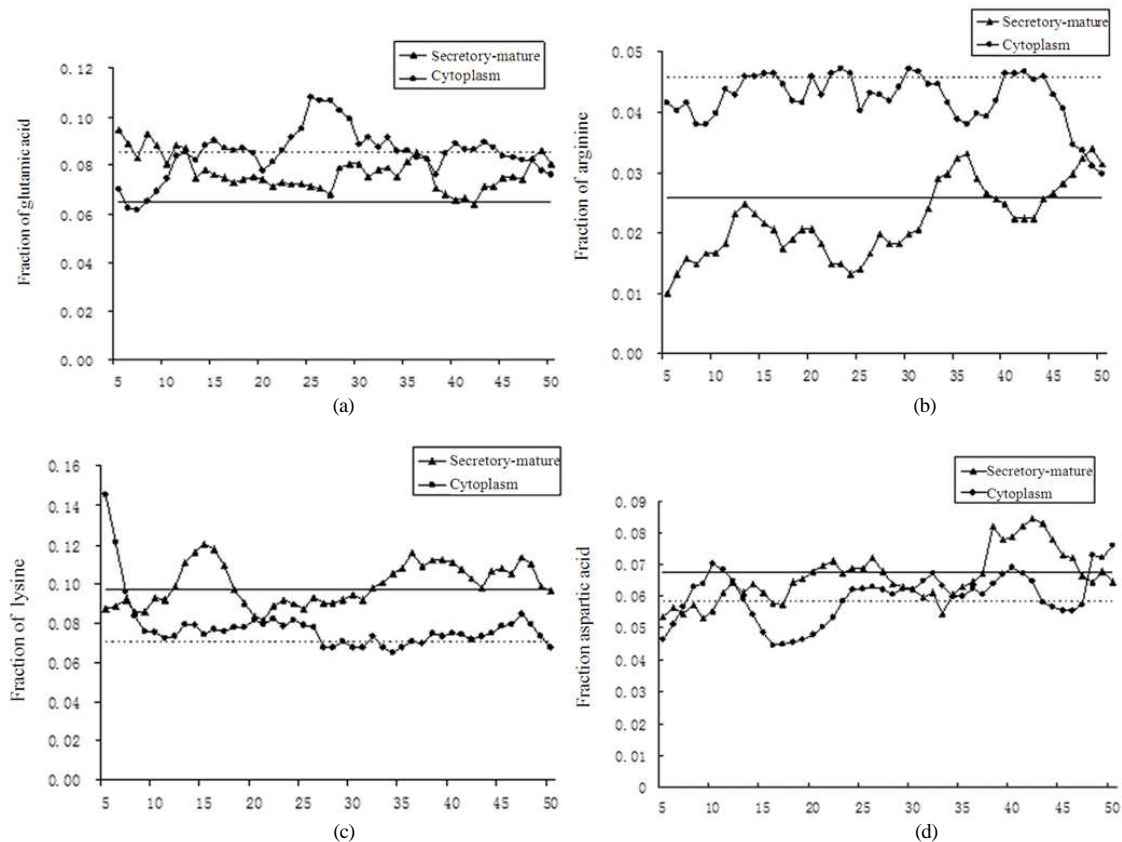


Fig. 3: CodeproteinWin analysis of the first 50 residues of the mature sequences and cytoplasmic proteins, using a window size of 5. The x-axis reports the amino acid number of the terminal of the window and the mean percentage of the amino acid are plotted on the y-axis. The solid line and dashed line respectively represents the charged amino acid's percent composition in the mature domains and cytoplasmic proteins. A, mean percentage of glutamic acid in the first 50 residues of the mature sequences and cytoplasmic proteins; B, mean percentage of arginine; C, mean percentage of lysine; D, mean percentage of aspartic acid

Except for Met, Arg and Pro, the others showed the same trends in the order/disorder composition profile (Uversky and Dunker, 2010). The first 14 residues of the mature domains had more charged and polar residues like Glu, Ser, which were enriched in disordered proteins, but had less residues which were commonly found in ordered proteins, such as Trp, Cys, Phe (Dunker *et al.*, 2001). These facts showed that the first 14 residues of the mature domains were difficult to fold into an ordered conformation and easily formed the disordered domains.

Analysis of distribution of charged residues in the NH₂-terminal regions of mature domains: As shown in Fig. 3, the percentage of Glu in data set 4 was higher than that in data set 3 or data set 2 (for example, in the first 10 residues, 21.33 and 35.19% higher, $p = 0.002044$ and $p = 0.02495$, respectively) and the

number of Arg in data set 4 was markedly lower than in data set 3 or data set 2 (for example, in the first 10 residues, 67.32 and 48.65% lower, $p = 5.266e-10$ and $p = 4.747e-07$, respectively). Despite the abundance of Lys residues in data set 3, especially in the first 5 residues, the number of Lys residues in data set 4 was not significantly different from that in data set 2 (for example, in the first 10 residues, $p = 0.1441$). The number of Asp residues in data set 4 was not significantly different from that in data set 3 ($p = 0.5134$). These observations may implicate that the charged residues in the NH₂-terminal regions of the mature domains are factors affecting protein export.

Moreover, as shown in Fig. 4, we compared the charged N-terminus of the signal peptide (with an average of 5 residues (Tjalsma *et al.*, 2004) with the first 5 residues of the mature domains.

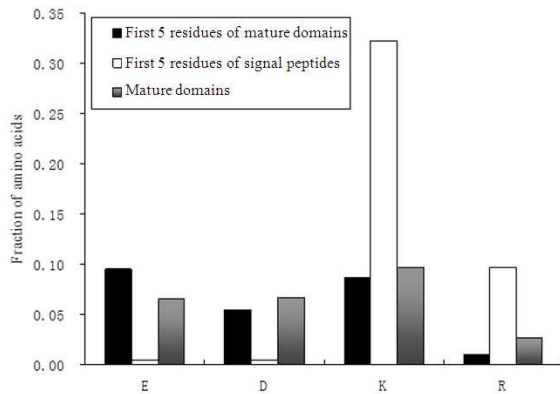


Fig. 4: Mean percentage of glutamic acid (E), aspartic acid (D), lysine (K) and arginine (R) found in the first 5 residues of mature domains (black), the first 5 residues of signal peptides (white) and mature domains (gray)

As confirmed by Heijne and Abrahmsen (1989), the NH₂-terminus of the signal peptide contains more positively charged residues, especially Lys, but only few negatively charged residues.

However, the first 5 residues of the mature domains contained 18.1 times more negatively charged residues and 4.3 times less positively charged residues than the positively charged N-domain. In addition, compared with the mature domains, the NH₂-terminal regions of the mature domains had a tendency to contain more Glu and less Arg (Fig. 4), which mean the two ends of hydrophobic core had contrary charged residues distribution.

DISCUSSION

Although several experiments have suggested that the mature proteins are critical for protein export (Tomassen and Kroon, 1987; Simonen *et al.*, 1992; Chen and Nagarajan, 1993; Davis and Model, 1985; Li *et al.*, 1988; Kajava *et al.*, 2000; Gouridis *et al.*, 2009), the precise roles and the property of mature domains are unknown. In this study, we identified the amino acid composition of the mature domains in *B. subtilis* to study the property of the mature domains that facilitates their export.

The study has showed that the first 14 residues of the mature domains contained more disordered residues than the corresponding residues in the cytoplasmic proteins and mature domains. This was further corroborated in the comparison to the order/disorder composition profile (Uversky and Dunker, 2010). Surprisingly, Met, Arg and Pro were not as prevalent in

these regions as in the disordered proteins. The depletion of Met may be caused by the analyzed mature proteins that did not contain the first Met (the folded proteins data set contained the first Met) and the roles of the N-terminus of mature domains may be related to the depletion of Arg and Pro.

It has been proved that the mature domain was the primary binding determinant in the targeting event (Gouridis *et al.*, 2009). The presence of disordered domains in the NH₂-terminal regions of mature domains can provide the binding site and some advantages for binding, such as fast speed of interaction (Shoemaker *et al.*, 2000) or interaction with multiple proteins (Patil and Nakamura, 2006). Because Pro may hinder the disorder-to-order transition that many protein-protein binding arose from (Patil and Nakamura, 2006; Uversky and Dunker, 2010), the binding function of mature domains may result in the depletion of Pro. The other role of the disordered domains in the NH₂-terminal regions of the mature domains may be the maintenance of exported proteins in a secretion-competent conformation for a long time to facilitate the interaction with the secretory components and protein export. Comparison to the NH₂-terminal regions of the mature domains revealed that the NH₂-terminal regions of the cytoplasm proteins have more order-promoting amino acids, so they can fold to the ordered globular structure quickly and then will not be recognized by protein export pathways. This feature may underly that besides the signal hypothesis (Manson, 1974), there is another mechanism that ensures that the cell efficiently recognizes the secretory proteins and cytoplasm proteins, which is the NH₂-terminal regions of the mature domains enriching disorder-promoting amino acids while the NH₂-terminal regions of the cytoplasm proteins containing a large number of order-promoting amino acids.

Apart from the disordered domains, we have shown that the NH₂-terminal regions of the mature domains had more Glu and less Arg. Several different experiments have shown that the net charge of the amino terminus of the mature domains had either a neutral or negative net charge (Li *et al.*, 1988; Kajava *et al.*, 2000; Heijne, 1986), which only applied to gram-negative bacteria; neither eukaryotic nor gram-positive bacteria had this charge bias (Kajava *et al.*, 2000). However, some experiments proved that a negatively charged N-terminus in the mature protein increased the secretion efficiency in *Lactococcus lactis* (Loir *et al.*, 2001; Dieye *et al.*, 2001). In our study, the NH₂-terminal regions of the mature domains had more Glu; thus, the presence of charged residues appeared to be more crucial than the net charge.

Mutations in the positively charged NH₂ region of the levansucrase signal peptide have shown that the secretion efficiency became slow when the 3 positive charges were reduced to 2 and the signal peptide was no longer processed when the charges were reduced to zero (Borchert and Nagarajan, 1991) and we showed the NH₂-terminal regions of the mature domains had contrary charged residues to NH₂-terminus of the signal peptide. These observations may implicate that the negatively charged residues in the NH₂-terminal regions of the mature domains interacted with the positively charged N-terminus of the signal peptide. It was proposed that the signal peptide formed a hairpin-like structure that can insert into the membrane and unlooping of this hairpin can create a complete signal peptide that can insert into the membrane (Vrije *et al.*, 1990). Our study suggested that the interaction between the NH₂ terminal of the signal peptide and the NH₂ terminal of the mature domains may help form the hairpin-like structure of the signal peptide. However, because the charged residues were not necessary for protein export in some secretory proteins (Chen and Nagarajan, 1994), it may be complementary to the role of the central hydrophobic core. The effect of the central hydrophobic core is the dominant factor in the formation of the hairpin-like structure.

By the help of the negatively charged residues in the NH₂-terminal regions of the mature domains, the signal peptides form the hairpin-like structure that can insert into the membrane. The disordered domains of the mature domains may provide the primary binding site in targeting and docking events and can maintain the secretory proteins in a secretion-competent conformation for a long time that facilitates preproteins to interact with the secretory components. These features ensure efficient sorting of secretory proteins from cytoplasmic residents in the cells.

CONCLUSION

The amino acid composition of 241 predicted exported proteins and 457 cytoplasmic proteins of *Bacillus subtilis* was explored and we found that disorder-promoting amino acids in the first 14 residues of the mature domains were overrepresented in comparison to the amino acids in the cytoplasmic proteins or the NH₂-terminal region. The finding may have important implications for the understanding of the contribution of mature domains to membrane targeting and to translocation.

ACKNOWLEDGEMENT

This study was supported by the National Nature Science Foundation of China (No 20836003, 30871952), the Science and Nature Foundation in Jiangsu Province (Grant No. BK2008103), the Fundamental Research Funds for the Central Universities JUSRP 11017 and the 111 Project B07029.

REFERENCES

- Borchert, T.V. and V. Nagarajan, 1991. Effect of signal sequence alterations on export of levansucrase in *Bacillus subtilis*. *J. Bacteriol.*, 173: 276-282.
- Chen, M. and V. Nagarajan, 1993. The roles of signal peptide and mature protein in RNase (barnase) export from *Bacillus subtilis*. *Mol. Gen. Genet.*, 239: 409-415. DOI: 10.1007/BF00276939
- Chen, M. and V. Nagarajan, 1994. Effect of alteration of charged residues at the N termini of signal peptides on protein export in *Bacillus subtilis*. *J. Bacteriol.*, 176: 5796-5801.
- Davis, N.G. and P. Model, 1985. An artificial anchor domain: Hydrophobicity suffices to stop transfer. *Cell*, 41: 607-614. DOI: 10.1016/S0092-8674(85)80033-7
- Dieye, Y., S. Usai, F. Clier, A. Gruss and J.C. Piard, 2001. Design of a protein-targeting system for lactic acid bacteria. *J. Bacteriol.*, 183: 4157-4166. DOI: 10.1128/JB.183.14.4157-4166.2001
- Dunker, A.K., J.D. Lawson, C.J. Brown, R.M. Williams and P. Romero *et al.*, 2001. Intrinsically disordered protein. *J. Mol. Graph. Model.*, 19: 26-59. DOI: 10.1016/S1093-3263(00)00138-8
- Dunker, A.K., C.J. Brown, J.D. Lawson, L.M. Iakoucheva and Z. Obradovic, 2002. Intrinsic disorder and protein function. *Biochemistry*, 41: 6573-6582. PMID: 12022860
- Fekkes, P. and A.J. Driessen, 1999. Protein targeting to the bacterial cytoplasmic membrane. *Microbiol. Mol. Biol. Rev.*, 63: 161-173. PMID: 10066835
- Gouridis, G., S. Karamanou, I. Gelis, C.G. Kalodimos and A. Economou, 2009. Signal peptides are allosteric activators of the protein translocase. *Nature*, 462: 363-367. DOI: 10.1038/nature08559
- Heijne, G.V. and L. Abrahmsen, 1989. Species-specific variation in signal peptide design Implications for protein secretion in foreign hosts. *FEBS Lett.*, 244: 439-446. DOI: 10.1016/0014-5793(89)80579-4
- Heijne, G.V., 1986. Net N-C charge imbalance may be important for signal sequence function in bacteria. *J. Mol. Biol.*, 192: 287-290. DOI: 10.1016/0022-2836(86)90365-7

- Kajava, A.V., S.N. Zolov, A.E. Kalinin and M.A. Nesmeyanova, 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. DOI: 10.1128/JB.182.8.2163-2169.2000
- Li, P., J. Beckwith and H. Inouye, 1988. Alteration of the amino terminus of the mature sequence of a periplasmic protein can severely affect protein export in *Escherichia coli*. *Proc. Nat. Acad. Sci. USA.*, 85: 7685-7689. PMID: 3051001
- Loir, Y.L., S. Nouaille, J. Commissaire, L. Bretigny and A. Gruss *et al.*, 2001. Signal peptide and propeptide optimization for heterologous protein secretion in *Lactococcus lactis*. *Applied Environ. Microbiol.*, 67: 4119-4127. DOI: 10.1128/AEM.67.9.4119-4127.2001
- Manson, L.A., 1974. *Biomembranes*. 1st Edn., Plenum Press, New York, ISBN: 0306398915.
- Patil, A. and H. Nakamura, 2006. Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS Lett.* 580: 2041-2045. DOI: 10.1016/j.febslet.2006.03.003
- Prilusky, J., C.E. Felder, T. Zeev-Ben-Mordehai, E.H. Rydberg and O. Man *et al.*, 2005. FoldIndex©: A simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics*, 21: 3435-3438. DOI: 10.1093/bioinformatics/bti537
- Salma, P., C. Chhatbar and S. Seshadri, 2009. Intrinsically unstructured proteins: Potential targets for drug discovery. *Am. J. Infect. Dis.*, 5: 126-134. DOI: 10.3844/ajidsp.2009.126.134
- Schein, C.H., K. Kashiwagi, A. Fujisawa and C. Weissmann, 1986. Secretion of mature IFN- α 2 and accumulation of uncleaved precursor by *Bacillus subtilis* transformed with a hybrid α -amylase signal sequence-IFN- α 2 gene. *Nat. Biotechnol.*, 4: 719-725. DOI: 10.1038/nbt0886-719
- Shoemaker, B.A., J.J. Portman and P.G. Wolynes, 2000. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc. Natl. Acad. Sci. USA.*, 97: 8868-8873. DOI: 10.1073/pnas.160259697
- Simonen, M., E. Tarkka, R. Puohiniemi and M. Sarvas, 1992. Incompatibility of outer membrane proteins OmpA and OmpF of *Escherichia coli* with secretion in *Bacillus subtilis*: Fusions with secretable peptides. *FEMS Microbiol. Lett.*, 79: 233-241. DOI: 10.1016/0378-1097(92)90215-A
- Tjalsma, H., A. Bolhuis, J.D.H. Jongbloed, S. Bron and J.M.V. Dijk, 2000. Signal peptide-dependent protein transport in *Bacillus subtilis*: A genome-based survey of the secretome. *Microbiol. Mol. Biol. Rev.*, 64: 515-547. DOI: 10.1128/MMBR.64.3.515-547.2000
- Tjalsma, H., H. Antelmann, J.D.H. Jongbloed, P.G. Braun and E. Darmon *et al.*, 2004. Proteomics of protein secretion by *Bacillus subtilis*: Separating the "Secrets" of the secretome. *Microbiol. Mol. Biol. Rev.*, 68: 207-233. DOI: 10.1128/MMBR.68.2.207-233.2004
- Tommassen, J. and T.D. Kroon, 1987. Subcellular localization of a PhoE-LacZ fusion protein in *E. coli* by protease accessibility experiments reveals an inner-membrane-spanning form of the protein. *FEBS Lett.*, 221: 226-230. DOI: 10.1016/0014-5793(87)80930-4
- Uversky, V.N. and A.K. Dunker, 2010. Understanding protein non-folding. *Biochim. Biophys. Acta (BBA)- Proteins Proteomics*, 1804: 1231-1264. DOI: 10.1016/j.bbapap.2010.01.017
- Vrije, G.J.D., A.M. Batenburg, J.A. Killian and B.D. Kruijff, 1990. Lipid involvement in protein translocation in *Escherichia coli*. *Mol. Microbiol.*, 4: 143-150. DOI: 10.1111/j.1365-2958.1990.tb02024.x
- Xia, X. and Z. Xie, 2001. DAMBE: Software package for data analysis in molecular biology and evolution. *J. Hered.*, 92: 371-373. DOI: 10.1093/jhered/92.4.371