Research article

THE PREDICTION OF NOVEL MULTIPLE LIPID-BINDING REGIONS IN PROTEIN TRANSLOCATION MOTOR PROTEINS: A POSSIBLE GENERAL FEATURE

ROB C. A. KELLER*

Section Chemistry, Charlemagne College, Wilhelminastraat 13-15,
Nijmegen 3913 NH, The Netherlands

Abstract: Protein translocation is an important cellular process. SecA is an essential protein component in the Sec system, as it contains the molecular motor that facilitates protein translocation. In this study, a bioinformatics approach was applied in the search for possible lipid-binding helix regions in protein translocation motor proteins. Novel lipid-binding regions in *Escherichia coli* SecA were identified. Remarkably, multiple lipid-binding sites were also identified in other motor proteins such as BiP, which is involved in ER protein translocation. The prokaryotic signal recognition particle receptor FtsY, though not a motor protein, is in many ways related to SecA, and was therefore included in this study. The results demonstrate a possible general feature for motor proteins involved in protein translocation.

Key words: Lipid-binding regions, Motor proteins, Protein-lipid interactions, Protein translocation, SecA

INTRODUCTION

Protein translocation is an important cellular process. Thanks to a combination of genetics, molecular biology, biochemistry and biophysics studies, all of the factors involved in protein translocation in *E. coli* have basically been identified [1]. These factors include the so-called Sec-proteins [2] and the phospholipids [3, 4].

* Author for correspondence. e-mail: r.keller@kgcnijmegen.nl

Both anionic and non-bilayer-forming phospholipids were found to be essential for efficient protein translocation [5, 6]. Sources of energy, ATP and a proton motive force are also essential [2]. In the last decade, an overwhelming amount of the three-dimensional structures of several Sec-proteins have been resolved (see [2, 7] for recent overviews). The stream of publications resulting from this newly obtained structural information mainly focused on protein-protein interactions.

However, understanding the role of protein-lipid interactions in the protein translocation process remains important, and interesting results have already been obtained in this line of research. For example, the insertion ability of SecA has been demonstrated in model membrane systems [8] with a preference for anionic phospholipids [9]. This is intriguing because the primary structure of SecA reveals it to be an overall negatively charged protein [10]. Based on an experimentally observed vesicle aggregation phenomenon and the use of some SecA deletion mutants [11, 12], two lipid-binding sides were proposed. *In vitro* studies led to the idea of an insertion and deinsertion cycle for SecA in the membrane [13], similar to what was found in the nucleotide-dependent SecA-phospholipid interaction studies using model systems [9, 11, 14]. The SecA protein even traverses the inner membrane towards the periplasmic side [15], as was found in a pure lipid model membrane system [16]. Furthermore, a number of studies described the role of different ligands in SecA membrane binding [17, 18], which not only explained some apparent contradictions in the literature, but also provided interesting mechanistic details of SecA functioning in the protein translocation process. Recently, an allosteric regulation of SecA by Mg^{2+} was found, with a special role for the anionic phospholipid, cardiolipin [19].

Not only can crystal structures serve as inspiration for further research, but bioinformatics can also play a prominent role in biochemistry and the life sciences. A recently developed bioinformatics approach led to the creation of the web-server Heliquest [20]. This web-based software includes a feature to identify (potential) lipid-binding helix regions in proteins. Stretches of amino acids can be investigated, and the use of a discrimination factor enables discrimination between the lipid-binding and non-lipid binding regions of proteins and peptides. The Heliquest program has two options: a screening option and an analysis option. The screening option includes an automatic screening for possible transmembrane regions that utilizes an external program specialized in identifying such regions.

The indicated discrimination factor used in this study utilized the analysis mode. Here, the classical transmembrane regions cannot be automatically identified. This issue will be discussed in detail in relation to protein translocation motor proteins in the Results and Discussion section of this paper. Based on this bioinformatics approach, the *E. coli* SecA protein was thoroughly investigated, and multiple lipid-binding sites were identified, divided over several lipid-binding domains (LBD). Additionally, other motor proteins found in chloroplasts (cpSecA), ER (BiP) and mitochondria (mtHsp70) were screened.

The prokaryotic signal recognition particle receptor FtsY, though not a motor protein, is included in this study due to its close resemblance to SecA. Surprisingly similar patterns in terms of multiple lipid-binding regions and organization into different lipid-binding domains were found for all the studied motor proteins. Some regions were further specified and characterized, and the possible implications for protein translocation are discussed herein.

MATERIALS AND METHODS

Primary and secondary structure identification

The primary structure of *E. coli* SecA (P10408) was obtained from the Swiss-Prot sequence database and the primary structures of the corresponding regions, identified as the lipid-binding helix, were collected (Tab. 1). In a similar way, detailed descriptions of the primary structure of BiP (P16474-1), cpSecA (Q41062), mtHsp70 (P48721) and FtsY (P10121) were obtained (see Tabs 2 and 3 for details). The included regions were checked for the extent of helicity either using the available crystal structural data and/or via secondary structure analysis using the program SOPMA [21], available at http://npsa-pbil.ibcp.fr/. In the case of conflicting results, this is indicated in the table S1 (see supplemental data in Tab. S1 at http://dx.doi.org/10.2478/s11658-010-0036-y for a detailed comparison of the experimental data extracted from PDB and the data from the secondary structure predictions for *E. coli* SecA).

Determination of the lipid-binding potential

Using the program Heliquest [20], available at http://heliquest.ipmc.cnrs.fr/, the mean hydrophobicity ($<H>$), the hydrophobic moment (μH) and the net charge (z) were calculated. In the analysis, 18-residue windows were used, and for each sequence under investigation the window with the highest discrimination factor was selected. In essence, a stepwise discriminant analysis module was implemented in TSAR 3.3 (Oxford Molecular) with segment length, $<H>$, $<\mu H>$ and z as explanatory variables (for additional information see the website http://heliquest.ipmc.cnrs.fr/TablePeptide.htm).

A stepwise procedure was used to select a subset of the explanatory variables and optimize the classification rule which appeared to be the discrimination factor (D):

$$D = 0.944 \ (<\mu H>) + 0.33 \ (z)$$

When this discrimination factor is above 0.68, the corresponding region can be considered to be a (potential) lipid-binding helix. See [20] or the website http://heliquest.ipmc.cnrs.fr/HelpProcedure.htm for additional information. The help page indicated gives detailed information about the way the discrimination factor is defined and how for example the screening mode can be utilized for certain purposes.

Helical wheel plot
The helical wheel representations were produced using the Heliquest software [20]. The obtained helical wheel plots were subsequently redrawn and customized.

RESULTS AND DISCUSSION

Identification of the lipid-binding regions of *Escherichia coli* SecA
The complete sequence of SecA was run through the Heliquest program. Only the mostly helical regions were selected for further analysis. Based on the discrimination factor (D), which comprises both the hydrophobic moment (μH)

Tab. 1. Various identified lipid-binding regions of the motor protein *E. coli* SecA and some examples of known lipid-binding regions from other proteins or peptides.

Name	Sequence	z	<H>	<μH>	Confirmed[a]
SecA(1-21)	MLIKLLTKVFGSRNDRTLRRM	3	0.442	0.303	[11, 22]
SecA(14-33)	NDRTLRRMRKVVNIINAME	4	0.176	0.471	[11]
SecA(43-60)	LKGKTAEFRARLEKGEVL	2	0.116	0.207	
SecA(66-90)	AFAVVREASKRVFGMRHFDVQLLG	3	0.368	0.209	
SecA(108-125)	KTLTATLPAYLNALTGKG	2	0.437	0.352	[23]
SecA(370-395)	QTLASITFQNYFRLYEKLAG MTGTAD	1	0.558	0.345[c]	
SecA(466-488)	SNELTKAGIKHNVLNAKFHANEA	3	0.272	0.126	
SecA(593-614)	ALMRIFASDRVSGMMRKLGMKP	3	0.425	0.131	[22]
SecA(635-660)	ESRNFDIRKQLLEYDDVANDQRRAIY	3	0.023	0.292	
SecA(804-822)	KRESFSMFAAMLESLKYEV	1	0.422	0.339	
SecA(865-882)	AAAAALAAQTGERKVGRN	2	0.049	0.088	[11, 12]
SecA(877-895)	RKVGRNDPCPCGSGKKYKQ	5	-0.038	0.195	[11, 12]
SecA(489-508)	AIVAQAGYPAAVTIATNMAG	0	0.584	0.041[b]	-
PhoE SP	MKKSTLALVVMGIVASASVQA	2	0.558	0.045	[26]
Apocyt c (2-21)	DVEKGKKIFVQKCAQCHTVE	3	0.333	0.341	[27]
Apocyt c (80-101)	MIPAGIKKKTEREDLIAYLKKA	3	0.046	0.129	[27]
Kes 1p (7-29)	SSSWTSFLKSIASFNGDLSSLSAP	0	0.500	0.523	[28]
Glycophorin A (92-114)	ITLIIFGVMAGVIGTILLISYGI	0	1.133	0.213	[29]

[a]The reported evidence and indications in the literature confirm the predicted behaviour. [b]According to the Heliquest program calculation and its prediction, this is a non-lipid binding region. However, the frequently used TMpred program indicates that this is an alpha helical transmembrane segment (TMS). Furthermore, the H and μH values are remarkably similar to those found for the PhoE signal sequence, which is well-known for its lipid-binding and inserting abilities [26]. Due to these noteworthy findings, this sequence is included in this table. [c]According to the extracted PDB data, this region is primarily beta sheet, but the secondary structure prediction indicates a stretch of 15 AA in a helical conformation, and it is therefore included (see Tab. S1 for details).

and the net charge (z), a number of potential lipid-binding regions were identified (Tab. 1). For a number of the identified regions, direct or indirect experimental or theoretical evidence already exists.

For example, the SecA regions 1-21 and/or 14-33 found with Heliquest correspond well with the earlier proposed lipid-binding site that comprises the first 25 AA of the N-terminus [11]. The existence of this region was found to be a reasonable explanation (combined with the proposed C-terminal lipid-binding site) for the observed vesicle aggregation phenomena found when SecA is added to DOPG liposomes [11]. In addition, a recently performed ESR approach looking at mobility changes within SecA upon binding to different ligands [22] mentioned residue G11 as being part of a possible lipid-binding site, which is obviously right in the middle of the here identified SecA(1-21) region. This ESR approach also pointed to the special position of K609, which corresponds well with the region SecA(593-614) as identified here. Furthermore, it is interesting to see that with a recently developed prediction program for identifying transmembrane sequences [23], a 12-residue long helix was found in SecA proteins originated from different organisms. In *E. coli*, this appeared to be the region 108-120, which corresponds remarkably well to the predicted potential lipid-binding region SecA(108-125) as found by the Heliquest program.

The previously proposed C- and N-terminal lipid-binding sites were postulated to be different in character [11]. Indeed, the corresponding higher hydrophobic moment for the N-terminus (Tab. 1) indicates a more hydrophobic component in the protein-lipid interaction, while the low hydrophobic moment and low mean hydrophobicity for the C-terminal region suggest the more electrostatic nature of the protein-lipid interaction. This is confirmed by looking at the helical wheel plots for the two regions. It is clear that the sequence corresponding to region SecA(1-21) is an amphiphilic helix (Fig. 1A), perfectly suited to bind to lipid membranes, both electrostatic and hydrophobic in nature, while in the SecA(877-895) region, the alignment of positive charges is clearly indicated in the helical wheel plot (Fig. 1B). More SecA regions with a relatively high mean hydrophobicity have been identified (Tab. 1). There are a number of interesting earlier observations in relation to this. First of all, there is the observation that some C-terminal SecA deletion mutants fully abolished aggregation when added to DOPG vesicles, which was explained by the disappearance of one lipid-binding site [12]. Second, when SecA is added to DOPC first, and then DOPG vesicles are added, the wild-type SecA and mutant SecA with a C-terminal deletion of 64 amino acids both induce vesicle aggregation at comparable levels [24]. Third, C-terminal deletion mutants of SecA gave almost wild-type monolayer insertion levels in DOPG [12] and in DOPC [24].

This leads to the idea that N- or C-terminal deletions within SecA lead to the exposure of other more hydrophobic or amphiphilic regions of SecA. These exposed regions might very well be one or more of the lipid-binding helix regions identified here.
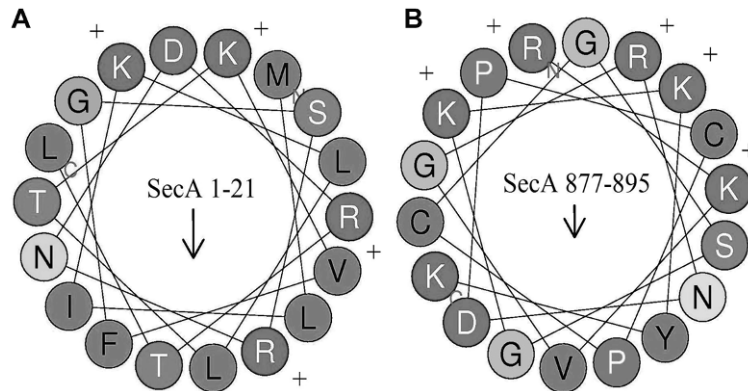
Fig. 1. Helical wheel representations of SecA(1-21) (A) and SecA(877-895) (B). The best 18-residue window result is depicted (see experimental procedures for details).

A detailed description of the identified regions

The positions of the identified regions of *E. coli* SecA in the overall sequence have been aligned to the primary sequence (Fig. 2). The lipid-binding regions are divided over four lipid-binding domains (LBD). The LBD is a proposed functional domain as an indicative attempt to group lipid-binding regions in such away that they presumably belong together on the basis of their close distance in the aligned structure.
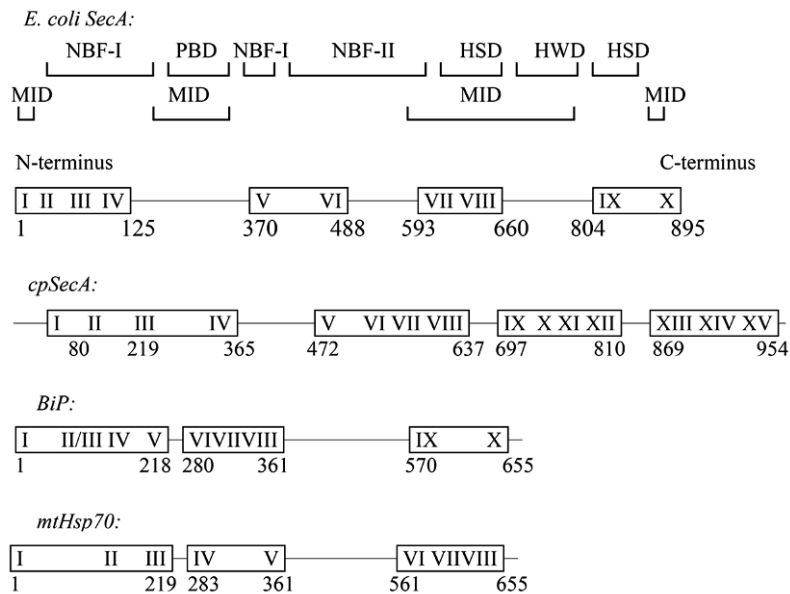

Fig. 2. Schematic representations of the motor proteins. The newly identified lipid-binding regions are depicted in roman numerals and the lipid-binding domains (LBD) are indicated in bars with grouped lipid-binding regions. For comparison purposes, the different functional sites of *Escherichia coli* SecA are indicated at the top of the figure.

Tab. 2. The various identified lipid-binding regions in cpSecA, BiP and mtHsp70.

Name	Sequence	z	\<H\>	\<μH\>
cpSecA(80-97)	RKQYAAIVNTINGLEPKI	2	0.324	0.453[a]
cpSecA(125-144)	LLPEAFAVVREASKRVLGLR	2	0.287	0.275
cpSecA(194-219)	YLARRARRDCEWVGQVPRFLGMKV	2	0.402	0.340
cpSecA(348-365)	QWASFVINAIKELFLRDV	2	0.583	0.332
cpSecA(472-493)	RATTGKWRAVVVEISRMNKTGRP	4	0.179	0.343
cpSecA(566-583)	FMARLKLREIMMPRVVKL	4	0.567	0.175
cpSecA(587-606)	GEFVSVKKPPPSKTWKVNEK	4	0.214	0.235
cpSecA(617-637)	TELAEKAVQLAVKTWGKRSLT	3	0.259	0.230
cpSecA(697-714)	RHESRRIDNQLRGRSGRQ	4	-0.276	0.243
cpSecA(731-748)	FRIFGGDRIQGLMRAFRV	3	0.467	0.281
cpSecA(766-783)	QKKVENYFFDIRKQLFEY	1	0.274	0.477
cpSecA(790-810)	QRDRVYTERRRALQSVNLQSL	3	0.182	0.172
cpSecA(869-880)	GLRSYLRLRGKEAYLQKR	5	0.117	0.215
cpSecA(909-928)	IDRLWKEHLQALKFVQQAVG	2	0.447	0.461
cpSecA(954-974)	AQIRRNVIYSIYQFKPVLLKQ	4	0.488	0.265
BiP(1-22)	MFFNRLSAGKLLVPLSVVLYAL	2	0.782	0.184
BiP(102-119)	KNQVAANPQNTIFDIKRL	2	0.217	0.290
BiP(117-134)	KRLIGLKYNDRSVQKDIK	4	0.044	0.244
BiP(157-174)	KKVFTPEEISGMILGKMK	2	0.359	0.234
BiP(201-218)	RQATKDAGTIAGLNVLRI	2	0.282	0.107
BiP(280-297)	YKIVRQLIKAFKKKHGID	5	0.253	0.422
BiP(302-319)	NKALAKLKREAEKARRAL	5	-0.122	0.490
BiP(344-361)	RAKFEELNLDLFKKTLKP	2	0.225	0.332
BiP(570-593)	SEDASIKAKVESRNKLENYAHSLK	3	-0.067	0.275
BiP(638-655)	KFESLSKVAYPITSKLYG	2	0.427	0.378
mtHsp70(1-18)	MISASRAAAARLVGTAAS	2	0.347	0.250
mtHsp70(158-175)	GKLYSPSQIGAFVLMKMK	3	0.522	0.118
mtHsp70(202-219)	RQATKDAGQISGLNVLRV	2	0.203	0.090
mtHsp70(283-306)	LRHIVKEFKRETGVDLTKDNMALQR	2	0.169	0.366
mtHsp70(339-361)	PKHLNMKLTRAQFEGIVTDLIKR	2	0.358	0.359
mtHsp70(561-578)	MVKNAEKYAEEDRRKKER	2	-0.383	0.310
mtHsp70(608-625)	CNKLKEEISKMRALLAGK	3	0.189	0.420
mtHsp70(638-655)	SSLQQASLKLFEMAYKKM	2	0.376	0.365

[a]Values are based on utilizing the 18-residue window result with the highest discrimination factor (see the Materials and Methods section for details).

It is remarkable that the majority of the lipid-binding regions are found in the N- and C-terminal ends of SecA (respectively 1-100 and 600-900), with two exceptions: SecA(370-395) and SecA(466-488). Interestingly enough, these latter two regions overlap with the two nucleotide-binding folds (NBF). This might mechanistically explain the nucleotide dependency of the SecA-lipid insertion [11, 12, 14]. Depending on how strictly one considers the positions within the primary sequence of already assigned and identified functional regions in SecA (NBF, HWD etc.), a certain degree of overlap between these regions and the identified lipid-binding regions is apparent. This could mean that some of these lipid-binding regions will either function secondarily and/or certain regions might be multi-functional, since SecA is a key protein which interacts with different ligands, and is a well-known multi-functional protein. A completely different approach characterizing the protease-protected fragments of SecA using immunodetection [25] proposed two new membrane-interacting domains (MID). These two MID, SecA(240-330) and SecA(610-820), are remarkably close to or even overlap two of the lipid-binding domains proposed here (Fig. 2). This is interesting since Chen *et al.* [25] suggested that these regions are part of an integral membrane form of SecA. Since lipid binding could serve as anchoring for the contact of an adjacent part of the SecA protein with a membrane protein, this seems to indicate one of the possible roles for a number of the identified lipid-binding regions during protein translocation.

In order to emphasize the reliability of the method, examples of well-know lipid-binding regions of another protein and peptide were included in Tab. 1. The signal peptide of PhoE is an example of a well-studied peptide in terms of peptide-lipid interactions, and its ability to interact specifically with anionic phospholipids has been experimentally demonstrated [26]. Indeed the Heliquest program perfectly identifies the sequence of this signal peptide as a lipid-binding helix, corresponding to the hydrophobic core and the positively charged N-terminus of the signal peptide. Apocytochrome c is a frequently studied protein when it comes to protein-lipid interactions, and it is clear (Tab. 1) that the Heliquest program unquestionably identifies the two lipid-binding helix regions which were found experimentally [27]. It is interesting to note that the N- and C-terminal regions of Apocytochrome c show a similar duality as found for *E. coli* SecA when it comes to its lipid-binding sites, with one site being more hydrophobic and the other being more electrostatic in nature. These similarities between the proteins seem to underline the earlier hypothesis that *E. coli* SecA and proteins like Apocytochrome c are members of the same family of proteins [14], a class of proteins which can be found either in a soluble or in a membrane-bound state. As indicated in the introduction, the used analysis cannot identify (classical) transmembrane regions. Two additional examples are included in Tab 1. According to the discrimination factor, the known transmembrane region of the *H. sapiens* glycophorin A [28] and the surface membrane-binding segment of the *S. cerevisiae* Kes1p sterol transporter [29] would be considered non-lipid binding (D < 0.68) in this Heliquest approach.

The segment AA 7-29 in Kes1p appeared to be a member of a new class of membrane-binding amphipathic helices especially suited for curved membrane surfaces [29]. The identification of transmembrane regions can be performed either automatically using the screening mode of the Heliquest program, which utilized the specialized program TMHMM [30], or manually before or after a Heliquest analysis. Thus, the manual analysis option of Heliquest as described in this study cannot identify possible transmembrane regions. However, there are indications that it is possible to extract the identification of transmembrane regions directly from the Heliquest data (Keller, manuscript in preparation). In the case of protein translocation motor proteins, this aspect is not required in the analysis because it is known that for example SecA does not contain 'pure' transmembrane regions (see [23] for a more elaborate discussion on this).

The identification of the lipid-binding regions of protein translocation motor proteins

Motor proteins are not only found across *E. coli* inner membranes, but also in other protein translocation systems, e.g. cpSecA is found in the chloroplasts, mtHsp70 in the mitochondria, and BiP in the ER [31]. The results found for these motor proteins (Fig. 2 and see Tab. 2 for a detailed description) demonstrate a common feature for all motor proteins: multiple lipid-binding regions. These novel lipid-binding regions are divided over several lipid-binding domains. It is interesting to note that an apparent correlation between the anionic phospholipid content of the membrane and the number of lipid-binding regions in the corresponding motor proteins can be observed. The anionic phospholipid content in the corresponding membranes is different: 20-25% in the *E. coli* inner membrane [32], 14-23% in the mitochondrial membranes [33, 34], 11-20% in the thylakoid chloroplast membranes [35, 36], and 10-20% in the ER membranes [37]. It is tempting to see an increase in the extent of lipid-binding regions with a decrease in the anionic phospholipid content in the corresponding membrane. Provided that one takes 20 AA as the average length for each region, the percentages of the helical lipid-binding region extents does indeed increase, to 22%, 23.5%, 29.5% and 29.5%, respectively. In this respect, it is of interest to note that two reported photo-cross-linking studies found *E. coli* SecA-lipid labelling with 5-10% efficiency [38, 39]. Since it is unlikely that all of the regions identified in this study are involved in protein-lipid interactions at the same time, a quarter (5.5%) or third (7.5%) seems more reasonable. This would correspond well with the experimentally observed efficiency. Unlike *E. coli* SecA, little is known about any possible involvement of protein-lipid interactions in other motor proteins. However, for cpSecA, some indications have been found for specific lipid and signal peptide preferences [36]. Additionally, a recent experimental approach developed to screen a large number of different proteins for their lipid-binding capability identified three proteins belonging to the heat shock protein 70 family [40]. In relation to this, obviously *E. coli* SecA and cpSecA are related proteins (in two different protein

translocation systems). BiP and mtHsp70 both belong to the Hsp70 family (again in two different protein translocation systems). The fact that BiP and mtHsp70 are closely related is apparent in Fig. 2: the lipid-binding domains are almost identical (although on the primary structural level there is no resemblance, see Tab. 2).

FtsY, a membrane-associated protein involved in protein translocation
Of all the motor proteins, *E. coli* SecA is the best characterized protein in terms of protein-lipid interactions. FtsY is a membrane-associated homologue of the eukaryotic SRP receptor and is not a motor protein. However, FtsY and SecA are remarkably similar at a number of levels. For example, both are nucleotide-binding proteins with a low basal hydrolysis activity *in vitro* that is stimulated by anionic phospholipids [41, 42]. Both show an unusual subcellular distribution *in vivo*: partly soluble and partly firmly associated with the inner membrane [43, 44]. Both interact directly with phospholipids, with a preference for anionic phospholipids despite their overall acidic nature (pI SecA = 5.50, pI FtsY = 4.49), and can give rise to anionic phospholipid vesicle aggregation [11, 42].

Tab. 3. Identified lipid-binding regions in *E. coli* FtsY.

Name	Sequence	z	<H>	<μH>	Confirmed[a]
FtsY(1-18)	MAKEKKRGFFSWLGFGQK	4	0.277	0.332	[45, 46]
FtsY(188-208)	EKPTKEGFFARLKRSLLKTKE	5	0.198	0.254	[46, 47]
FtsY(245-262)	TRKIITNLTEGASRKQLR	4	0.088	0.290	
FtsY(387-407)	LQNKSHLMEELKKIVRVMKL	4	0.258	0.595	
FtsY(226-243)	DLFEELEEQLLIADVGVE	0	0.584	0.041[b]	-
FtsY(265-282)	EALYGLLKEEMGEILAKV	-2	0.449	0.427[b]	-

[a]The reported evidence and indications in the literature confirm the predicted behaviour. [b]According to the Heliquest program calculation and its prediction, this is a non-lipid binding region. However the helices indicated here have a hydrophobic face, which could be involved in FtsY interactions with PE (see text for discussion).

Because of this large number of similarities, the protein FtsY was included in this study. Indeed, multiple lipid-binding regions have been identified for *E.coli* FtsY (Tab. 3). It is interesting to note that the two most described and characterized lipid-binding regions in FtsY can be clearly identified with the Heliquest approach. The region FtsY(1-18) corresponds well with the earlier identified 14 AA long N-terminal region [45, 46], while the region FtsY(188-208) corresponds remarkably well with the earlier identified essential and autonomous membrane-binding amphipathic helix in *E. coli* FtsY, which was found to be region 197-211 [46, 47].
Besides these described regions, a number of novel lipid-binding regions have been identified. The protein FtsY consists of three domains: A(1-196), N(197-280), and G(281-497). Each has a distinct function (see [47] and references therein). One of the now identified novel lipid-binding regions FtsY(245-262) is situated

in the middle of the N-domain, while the other lipid-binding region FtsY(387-407) is found near the C-terminal side of the G-domain. Additionally, two more regions are included: FtsY(226-243) and FtsY(265-282) (Tab. 3). Although according to the corresponding discrimination factor, these regions were not identified as anionic phospholipid-binding regions, both regions exhibit two interesting features. Both were found to be helical and include in their sequence a strong hydrophobic face in the corresponding helical wheel plot (data not shown). A protein-lipid interaction based primarily on the hydrophobic nature of the regions FtsY(226-243) and FtsY(265-282) is postulated based on these features. This might be of particular interest since a specific FtsY-lipid interaction has been demonstrated for PE [48].

Concluding remarks
The data presented here indicates multiple lipid-binding sites in *E. coli* SecA. The number of binding sites is surprisingly high. A closer look at the distribution of the lipid-binding sites clearly demonstrates an N- and C-terminal domain distribution, with the exception of regions V and VI, which seems to overlap with (parts of) the NBF-domains. The results might inspire a closer look at these regions during protein translocation. Possible experimental confirmation opens the possibility to determine which regions are actively involved at the different stages of protein translocation. This could provide important mechanistic information. The finding that multiple lipid-binding regions were found not only in *E. coli* SecA, but also in other motor proteins, is intriguing. Basically two groups of protein translocation motor proteins have been studied, members of the SecA and the Hsp70 family in four different protein translocation systems. The finding that FtsY, which is in many aspects closely related to SecA, also contains multiple lipid-binding sites might emphasize the importance of this apparent common feature of all these proteins involved in protein translocation. The general feature that multiple lipid helical binding regions can be identified in protein translocation motor proteins might indicate an underlying functional aspect of this family of proteins.
Additionally the results may yield important and distinctive characteristics of the N- and C-terminal domain of some members of the amphitrophic protein family [49]. Apparently the presence of lipid-binding helical regions, and the absence of transmembrane regions, enables in this typical distribution both a soluble and a membrane-bound form. The presented bioinformatics approach might serve as a starting point for studying proteins which have not yet been characterized in detail when it comes to protein-lipid interactions.

REFERENCES

1. Arkowitz, R.A. and Bassilana, M. Protein translocation in Escherichia coli. **Biochim. Biophys. Acta** <u>1197</u> (1994) 311-343.
2. Driessen, A.J.M. and Nouwen, N. Protein translocation across the bacterial cytoplasmic membrane. **Annu. Rev. Biochem.** <u>77</u> (2008) 643-667.
3. de Kruijff, B., Breukink, E., Demel, R.A., van 't Hoff, R., de Jong, H.H.J., Jordi, W., Keller, R.C.A., Killian, J.A., de Kroon, A.I.M.P., Kusters, R. and Pilon, M. Lipid involvement in protein translocation. In: Membrane Biogenesis and Protein Targeting, **New Comprehensive Biochemistry**, Vol. <u>22</u>, (Neupert, W. & Lill, R. Eds.), Elsevier, Amsterdam, 1992, 85-100.
4. van Klompenburg, W. and de Kruijff, B. The role of anionic phospholipids in protein insertion and translocation in bacterial membranes. **J. Membr. Biol.** <u>162</u> (1998) 1-7.
5. de Vrije, T., de Swart, R.I., Dowhan, W. and de Kruijff, B. Phosphatidylglycerol is involved in protein translocation across Escherichia coli inner membranes. **Nature** <u>334</u> (1988) 173-175.
6. Rietveld, A.G., Koorengevel, M.C. and de Kruijff, B. Non-bilayer lipids are required for efficient protein transport across the plasma membrane of Escherichia coli. **EMBO J.** <u>14</u> (1995) 5506-5513.
7. Flower, A.M. The SecY translocation complex: convergence of genetics and structure. **Trends Microbiol.** <u>15</u> (2007) 203-210.
8. Ulbrandt, N.D., London, E.L. and Oliver, D.B. Deep penetration of a portion of Escherichia coli SecA protein into model membranes is promoted by anionic phospholipids and by partial unfolding. **J. Biol. Chem.** <u>267</u> (1992) 15184-15192.
9. Breukink, E., Demel, R.A., de Korte-Kool, G. and de Kruijff, B. SecA insertion into phospholipids is stimulated by negatively charged lipids and inhibited by ATP: A monolayer study. **Biochemistry** <u>31</u> (1992) 1119-1124.
10. Schmidt, M.G., Rollo, E.E., Grodberg, J. and Oliver, D.B. Nucleotide sequence of the secA gene and secA(Ts) mutants preventing protein export in Escherichia coli. **J. Bacteriol.** <u>170</u> (1988) 3404-3414.
11. Breukink, E., Keller, R.C.A. and de Kruijff, B. Nucleotide and negatively charged lipid-dependent vesicle aggregation caused by SecA. **FEBS Lett.** <u>331</u> (1993) 19-24.
12. Breukink, E., Nouwen, N., van Raalte, A., Mizushima, S., Tommassen, J. and de Kruijff, B. The C terminus of SecA is involved in both lipid binding and SecB binding. **J. Biol. Chem.** <u>270</u> (1995) 7902-7907.
13. Economou, A. and Wickner, W. SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. **Cell** <u>78</u> (1994) 835-843.
14. Keller, R.C.A., Snel, M.M.E., de Kruijff, B. and Marsh, D. SecA restricts in a nucleotide-dependent manner acyl chain mobility up to the center of a phospholipid bilayer. **FEBS Lett.** <u>358</u> (1995) 251-254.

15. Kim, Y.J., Rajapandi, T. and Oliver, D. SecA protein is exposed to the periplasmic surface of the E. Coli inner membrane in its active state. **Cell** 78 (1994) 845-853.
16. Ahn, T. and Kim, H. SecA of Escherichia coli traverses lipid bilayer of phospholipid vesicles. **Biochem. Biophys. Res. Commun.** 203 (1994) 326-330.
17. Benach, J., Chou, Y.-T., Fak, J.J., Itkin, A. Nicolae, D.D., Smith, P.C., Wittrock, G., Floyd, D.L., Golsaz, C.M., Gierasch, L.M. and Hunt J.F. Phospholipid-induced monomerization and signal-peptide-induced oligomerization of SecA. **J. Biol. Chem.** 278 (2003) 3628-3638.
18. Shin, J.-Y, Kim, M. and Ahn, T. Effect of signal peptide and adenylate on the oligomerization and membrane binding of soluble SecA. **J. Biochem. Mol. Biol.** 3 (2006) 319-328.
19. Gold, V.A.M., Robson, A., Clarke, A.R. and Collinson, I. Allosteric regulation of SecA. **J. Biol. Chem.** 282 (2007) 17424-17432.
20. Gautier, R., Douguet, D., Anthonny, B. and Drin, G. Heliquest: a web-server to screen sequences with specific α–helical properties. **Bioinformatics** 24 (2008) 2101-2102.
21. Geourjon, C. and Deleage, G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. **Comput. Appl. Biosci.** 11 (1995) 681-684.
22. Cooper, D.B., Smith, V.F., Crane, J.M., Roth, H.C., Lilly, A.A. and Randall, L.L. SecA, the motor of the secretion machine, binds diverse partners on one interactive surface. **J. Mol. Biol.** 382 (2008) 74-87.
23. Hu, H.-J., Holley, J., He, J., Harrison, R.W., Yang, H., Tai, P.C. and Pan, Y. To be or not to be: Predicting soluble SecAs as membrane proteins. **IEEE Trans. NanoBioscience** 6 (2007) 168-179.
24. Keller, R.C.A. Interactions between lipids and protein components of the prokaryotic secretion pathway. **PhD thesis**, University of Utrecht, The Netherlands, 1995.
25. Chen, X., Brown, T. and Tai, P.C. Identification and characterization of protease-resistant SecA fragments: secA has two membrane-integral forms. **J. Bacteriol.** 180 (1998) 527-537.
26. Keller, R.C.A., Killian, J.A. and de Kruijff, B. Anionic phospholipids are essential for α-helix formation of the signal peptide of prePhoE upon interaction with phospholipid vesicles. **Biochemistry** 31 (1992) 1672-1677.
27. Jordi, W., de Kruijff, B. and Marsh, D. Specificity of the interaction of amino- and carboxy-terminal fragments of the mitochondrial precursor protein apocytochrome c with negatively charged phospholipids. A spin-label electron spin resonance study. **Biochemistry** 28 (1989) 8998-9005.
28. Treutlein, H.R., Lemmon, M.A., Engelman, D.M. and Brunger, A.T. The glycophorin A transmembrane domain dimer: sequence-specific propensity for a right-handed supercoil of helices. **Biochemistry** 31 (1992) 12726-12732.

29. Drin, G., Casella, J-F., Gautier, R., Boehmer, T., Schwartz, T.U. and Antonny, B. A general amphipathic α–helical motif for sensing membrane curvature. **Nat. Struct. Mol. Biol.** 14 (2007) 138-146.

30. Krogh, A, Larsson, B, von Heijne, G, Sonnhammer, E.L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. **J. Mol. Biol.** 305 (2001) 567-80.

31. Tomkiewicz, D., Nouwen, N. and Driessen, A.J.M. Pushing, pulling and trapping − Modes of protein supported protein translocation. **FEBS Lett.** 581 (2007) 2820-2828.

32. Dowhan, W. Molecular basis for membrane phospholipid diversity: Why are there so many phospholipids. **Annu. Rev. Biochem.** 66 (1997) 199-232.

33. Hovius, H., Lambrechts, H., Nicolay, K. and de Kruijff, B. Improved methods to isolate and subfractionate rat liver mitochondria. Lipid composition of the inner and outer membrane. **Biochim. Biophys. Acta** 1021 (1990) 217-226.

34. Ardail, D., Privat, J.-P., Egret-Charlier, M., Levrat, C., Lerme, F. and Louisot, P. Mitochondrial contact sites. **J. Biol. Chem.** 265 (1990) 18797-18802.

35. Chapman, D.J., De-Felice, J. and Barber, J. D. Growth temperature effects on thylakoid membrane lipid and protein content of pea chloroplasts. **Plant Physiol.** 72 (1983) 225-228.

36. Sun, C., Rusch, S.L., Kim, J. and Kendall, D.A. Chloroplast SecA and *Escherichia coli* SecA have distinct lipid and signal peptide preferences. **J. Bacteriol.** 189 (2007) 1171-1175.

37. Yeung, T., Gilbert, G.E., Shi, J., Silvius, J. Kapus, A. and Grinstein, S. Membrane phosphatidylserine regulates surface charge and protein localization. **Science** 11 (2008) 210-213.

38. Eichler, J. Brunner, J. and Wickner, W. The protease-protected 30 kDa domain of SecA is largely inaccessible to the membrane lipid phase. **EMBO J.** 16 (1997) 2188-2196.

39. van Voorst, F., van der Does, C., Brunner, J., Driessen, A. J. M. and de Kruijff, B. Translocase-bound SecA is largely shielded from the phospholipid acyl chains. **Biochemistry** 37 (1998) 12261-12268.

40. Fernández-Murray, J.P. and McMaster, C.R. Identification of novel phospholipid binding proteins in *Saccharomyces cerevisiae*. **FEBS Lett.** 580 (2006) 82-86.

41. Lill, R., Dowhan, W. and Wickner W. The ATPase activity of SecA is regulated by acidic phospholipids, SecY, and the leader and mature domains of precursor proteins. **Cell** 26 (1990) 271-280.

42. de Leeuw, E., te Kaat, K., Moser, C., Menestrina, G., Demel, R. and de Kruijff, B. Anionic phospholipids are involved in membrane association of FtsY and stimulate its GTPase activity. **EMBO J.** 19 (2000) 531-541.

43. Cabelli, R.J., Dolan, K.M., Qian, L. and Oliver, D.B. Characterization of membrane-associated and soluble states of SecA from wild-type and *SecA51* (TS) mutant strains of *Escherichia*. **J. Biol. Chem.** 266 (1991) 24420-24427.

44. Luirink, J., ten Hagen-Jongman, C.M., van der Weijden, C.C., Oudega, B., High, S., Dobberstein, B. and Kusters, R. An alternative protein targeting pathway in *Escherichia coli*: studies on the role of FtsY. **EMBO J.** 13 (1994) 2289-2296.

45. Weiche, B., Bürk, J., Angelini, S., Schiltz, E., Thumfart, J-O and Koch, H-G. A cleavable N-terminal membrane anchor is involved in membrane binding of the Escherichia coli SRP receptor. **J. Mol. Biol.** 28 (2008) 761-773.

46. Braig, D., Bär, C., Thumfart, J-O. and Koch, H-G. Two cooperating helices constitute the lipid-binding domain of the bacterial SRP receptor. **J. Mol. Biol.** 390 (2009) 401-413.

47. Parlitz, R., Eitan, A., Stjepanovic, G., Bahari, L., Bange, G., Bibi, E. and Sinning, I. *Escherichia coli* signal recognition particle receptor FtsY contains an essential and autonomous membrane-binding amphipathic helix. **J. Biol. Chem.** 44 (2007) 32176-32321.

48. Millman, J.S., Qi, H-Y., Vulcu, F., Bernstein, H.D. and Andrews, D.W. FtsY binds to the Escherichia coli inner membrane via interactions with phosphatidylethanolamine and membrane proteins. **J. Biol. Chem.** 276 (2001) 25982-25989.

49. Halskau, Ø., Muga, A. and Martínez, A. Linking new paradigms in protein chemistry to reversible membrane-protein interactions. **Curr. Prot. Pept. Sci.** 10 (2009) 339-359.