Research Article

Journal of PeptideScience

Received: 8 June 2010 Revised: 4 October 2010 Accepted: 8 October 2010 Published online in Wiley Online Library: 18 February 2011

(wileyonlinelibrary.com) DOI 10.1002/psc.1324

Structural determinants of protein translocation in bacteria: conformational flexibility of SecA IRA1 loop region

Pasquale Palladino,[a]* Gabriella Saviano,[b] Teodorico Tancredi,[c] Ettore Benedetti,[a] Filomena Rossi[a] and Raffaele Ragone[a]

Bacteria employ the SecA motor protein to push unfolded proteins across the cytoplasmic membrane through the SecY protein-conducting channel complex. The crystal structure of the SecA–SecY complex shows that the intramolecular regulator of ATPase1 (IRA1) SecA domain, made up of two helices and the loop between them, is partly inserted into the SecY conducting channel, with the loop between the helices as the main functional region. A computational analysis suggested that the entire IRA1 domain is structurally autonomous, and was the basis to synthesize peptide analogs of the SecA IRA1 loop region, to the aim of investigating its conformational preferences. Our study indicates that the loop region populates a predominantly flexible state, even in the presence of structuring agent. This provides indirect evidence that the SecA loop–SecY receptor docking involves loop-mediated opening of the SecY channel. Copyright © 2011 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: translocase; SecA; SecY; IRA1; motor protein; protein-conducting channel; intramolecular regulator of ATPase1

Introduction

Proteins synthesized in cytosol reach a functional location by crossing, partially or completely, the cellular membrane. This process is mediated by translocases, which consist of a protein-conducting channel and an associated motor protein, usually an ATPase, at either the cis-side of the prokaryotic membrane or the trans-side of the eukaryotic endoplasmic reticulum membrane [1].

Among translocation systems, Sec-translocases present a conducting channel well conserved throughout all kingdoms of life, but differ for motor proteins. In particular, eukaryotes employ the trans-acting binding protein, whereas bacterial Sec translocase utilizes cis-side SecA motor protein to push unfolded protein across the cytoplasmic membrane through the SecYEG protein-conducting channel trimeric complex [2]. Two domains can be identified in SecA, namely, the N-terminal DEAD motor domain (roughly, residues 1–609), and the C-terminal region (roughly, residues 610–901), including the helix1(H1)-loop-helix2(H2) domain (roughly residues 767–790, 791–799 and 800–818, respectively), better known as intramolecular regulator of ATPase1 (IRA1) [3]. The crystal structure of the hetero-trimeric receptor SecYEG with SecA [4] shows that the IRA1 loop region, is inserted into a narrow conducting channel of the SecY receptor subunit, playing a functional role (Figure 1) [5]. Actually, it has been reported that SecA binds SecYEG mainly through the SecA DEAD domain (K_d = 110 nM) [3]. The SecA C-terminal region alone seems unable to bind the receptor, but the decrease in the complex binding affinity caused by amino acid substitution or truncation in the IRA1 functional motif suggests that it could optimize the interaction stabilizing the main binding site located in the DEAD domain, and/or creating a secondary binding site located in the IRA1 region. Indeed, truncation in the IRA1 domain (Δ783–795) was shown to cause about fourfold decrease in the binding affinity [3].

We have focused our attention on the loop connecting the two IRA1 helices, because there is no evidence that binding and/or activity are to be ascribed to helical residues in the IRA1 domain, but a few data on helix-2 [3]. We have collected several hints that the function of the SecA–SecYEG system could be determined by the presence of a few residues in the IRA1 loop. Thus, this restricted region can be considered as governing affinity and activity in SecA IRA1. On the basis of this evidence, we have performed an analysis by up-to-date computational tools [6], which suggested that the entire IRA1 domain is structurally autonomous.

Although the range of conformations adopted by the IRA1 loop is likely affected by the two long helices, investigation of intrinsic structural properties of loop-derived peptides does not require the presence of the whole IRA1 domain. On this basis, we synthesized wild-type and single-mutant analogs of the *Escherichia coli* SecA IRA1 loop region, to the aim of investigating how the loop conformation is affected by substitution of the Y794 residue, which has been proposed as governing the functional activity of SecA [5]. In particular, we have performed a CD and NMR analysis in

* *Correspondence to: Pasquale Palladino, Dipartimento delle Scienze Biologiche and CIRPeB, Università degli Studi Federico II, Naples, Italy. E-mail: pasquale.palladino@unina.it*

a *Dipartimento delle Scienze Biologiche and CIRPeB, Università degli Studi Federico II, Naples, Italy*

b *Dipartimento di Scienze e Tecnologie per l'Ambiente e il Territorio, Università degli Studi del Molise, Pesche, Isernia, Italy*

c *Istituto di Chimica Biomolecolare, CNR, Pozzuoli, Italy*

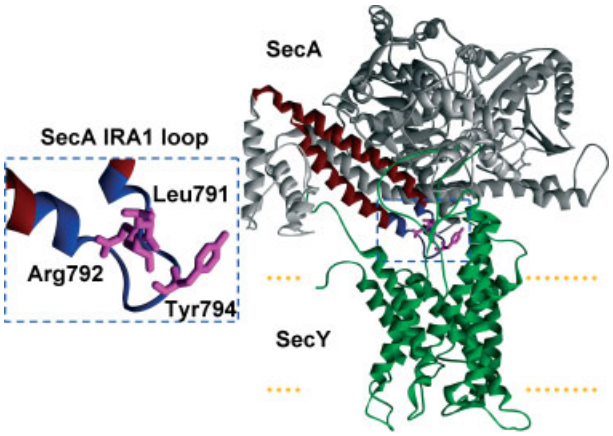

Figure 1. SecA–SecY complex, as derived from crystallographic data [4]. The SecA C-terminal IRA1 helix–loop–helix domain (red–blue–red) is inserted into the channel of the SecY receptor (green). The functional residues L791, R792, and Y794 [5] are highlighted in pink.

water and in H_2O/TFE mixtures on the synthetic wild-type peptide, SecA[788–804], and its analog, SecA[788–804]Y794A, which is derived from a completely inactive SecA mutant, both in the acetylated and amidated form to mimic the protein environment. Although TFE is known to stabilize secondary structure, our study suggests a predominant flexible state for both peptides, even in presence of high TFE percentage, except for peptide regions corresponding to SecA helices. These results, together with literature data, underline that the high flexibility of this SecA domain is independent of the presence of the functional loop tyrosine, but could be fundamental for SecA loop docking in SecY cavity and subsequent complex activation.

Materials and Methods

Chemicals

Chemicals were purchased from Sigma-Aldrich (Milan, Italy). Columns for peptide purification and characterization were from Phenomenex (Torrance, CA, USA).

Peptide Synthesis

The peptides *E. coli* SecA[788–804] Ac-GIHLRGYAQKDPKQEYK-NH_2 and SecA[788–804]Y794A Ac-GIHLRGAAQKDPKQEYK-NH2, N- and C-blocked, corresponding to the 788–804 segment in the SecA IRA1 domain, were synthesized in batches by a standard 9-fluorenylmethyl carbamate chemistry protocol using Rink-amide MBHA resin and purified by RP-HPLC using a C18 Jupiter (250 Å $\sim$ 22 mm) column. Peptide purity and integrity were confirmed by LC-MS technique (Finnigan Surveyor, Thermo Electron Corporation).

Structural Predictions

Predictions of secondary structure and structural disorder were performed by the MeDor (Metaserver of Disorder), freely available at http://www.vazymolo.org/MeDor/index.html [6], as reported in Figure 2:

- Secondary structure prediction is obtained by using Pred2ary software [7] based on StrBioLib java library [8].

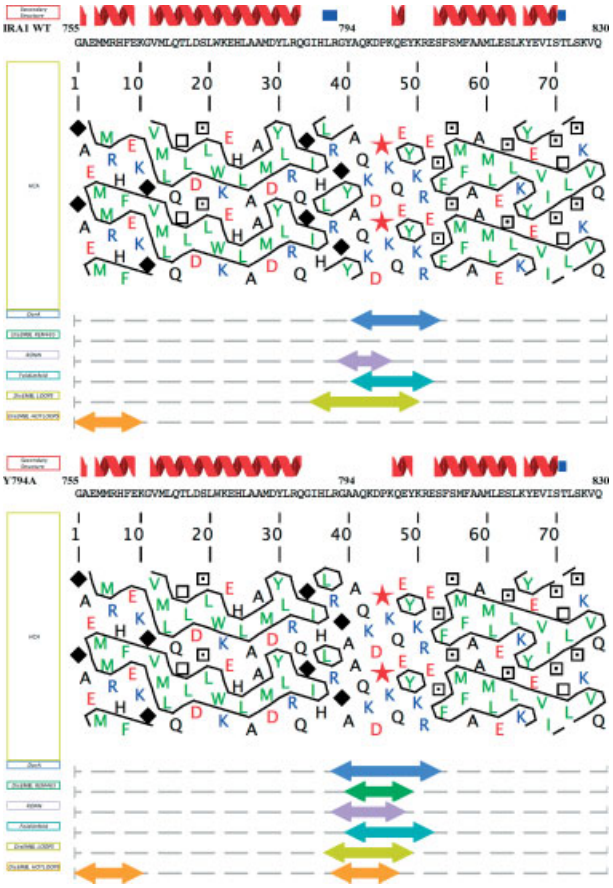

Figure 2. Graphical output of the MeDor metaserver [6] for *E. coli* SecA IRA1 (sequence 755–830). IRA1WT for wild type (upper) and Y794A for mutant (lower).
Secondary structure prediction is obtained by using Pred2ary software [8]. It reports α-helices as red ribbons and β-strands as blue strips.
HCA plots the protein sequence in a bidimensional diagram obtained unrolling a theoretical α-helix amino acid distribution [9,10]. V, I, L, F, M, Y, and W are in green. A black line highlights clusters that correspond to regular secondary structures. In particular, α-helices and β-strands appear as horizontal and vertical clusters, respectively. K and R, D and E, and A, C, Q, H, and N, are in blue, red, and black, respectively. Red stars (★), diamonds (♦), dotted (⊡), and empty squares (□) indicate P, G, S, and T, respectively. Disorder predictors employed are indicated on the left column: DorA; DisEMBL LOOPS, DisEMBL HOTLOOPS and DisEMBL REM465 [11]; RONN [12]; FoldUnfold [13]. The output is represented on the right, under the HCA plot, as colored left–right double arrows (⇿) spanning on possible disordered amino acids.

- HCA stands for hydrophobic cluster analysis, which uses a graphical representation of the sequence to recognize unstructured regions and identify secondary structure elements. In this representation, α-helices and β-strands appear as horizontal and vertical clusters, respectively [9,10].
- DorA (Disorder Analyzer) is a predictor based on the combined use of a disorder scoring matrix and HCA. This scoring matrix is derived from a training set issued from the DISPROT database and the CAZy database (http://www.cazy.org/).
- The DISEMBL predictors of disorder (http://dis.embl.de/) [11] are as follows:
  I. DisEMBL LOOPS predictor recognizes regions devoid of regular secondary structure (α-helix, 3_{10}-helix or β-strand) as loops/coils. Loop assignment is a necessary but not sufficient requirement for protein disorder.

Journal of PeptideScience

II. DisEMBL HOTLOOPS predictor recognizes a subset of loops with a high degree of mobility, as determined from C-α temperature factors (B-factors), as indicative of protein disorder.
III. DisEMBL REM465 predictor identifies protein regions called 'Remark 465', which miss assignment of electron density in the protein data bank (PDB).

- RONN (Regional Order Neural Network) detects natively disordered regions in proteins by sequence alignment without consideration of amino acid hydrophobicity or charge (http://www.strubi.ox.ac.uk/RONN) [12].
- FoldUnfold calculates the expected average number of contacts (mean packing density) per residue from the protein sequence. Regions with weak expected packing density would be responsible for the appearance of disordered regions [13].

To assess a comparison, the AGADIR software gave less than 2% helical content for the 788–804 loop region and more than 10% for 774–785 and 810–819 helices in Ac-755–830-Am (http://agadir.crg.es).

CD Measurements

Far UV CD spectra were recorded from 190 to 260 nm on a Jasco J-810 spectropolarimeter at 20 °C, using 1-mm quartz cell containing 5.25-μM peptide dissolved in H_2O/TFE solutions. Each spectrum was obtained subtracting contributions from other species, and converting the signal to mean residue ellipticity in units of deg cm^2 $dmol^{-1}$ per residue. Other experimental settings were 20 nm min^{-1}, scan speed; 1.0 nm, bandwidth; 0.2 nm, resolution; 50 mdeg, sensitivity; and 4 s, response.

NMR Measurements

Both samples were prepared by dissolving each peptide in water (90/10 v/v H_2O/D_2O) or in H_2O/TFE_{d2}-OH (50/50 v/v) up to a concentration of approximately 1 mM. NMR spectra were acquired at 300 K using a 600 MHz Bruker Avance spectrometer equipped with a cryoprobe. Natural abundance $^1H-^{15}N$ HSQC, TOCSY, NOESY, and double-quantum-filtered COSY spectra were used for resonance assignments. NOESY mixing times were set at 200 and 300 ms to follow the NOE build-up rates. Two-dimensional TOCSY experiments were recorded with a mixing time of 70 ms. NMRPipe and NMRView programs were used for data processing and spectral analysis, respectively. Spin system identification and assignment of individual resonances of SecA[788–804] and SecA[788–804]Y794A, in both solvents, were carried out by using a combination of TOCSY, NOESY, and DQF-COSY spectra, according to the standard procedure [14,15]. $^3J_{NH-CH}$ coupling constant values were measured for resolved NH amidic protons from 1D or DQF-COSY spectra [16].

Structure Calculations

Peak integrals were evaluated by NMRView, transferred to the program package DYANA 1.0.6, and converted to upper distance limits by using the CALIBA module of DYANA. Distance constraints were then worked out by the GRIDSEARCH module to generate a set of allowed dihedral angles. Structure calculation was carried out with the macro ANNEAL module by torsion angle dynamics. Eighty structures were calculated by Torsion Space Simulated Annealing (TSSA), starting with a total of 10 000 MD steps and a default value of maximum temperature. The 30 best structures in terms of target function were subjected to cluster analysis by the best fitting of backbone atoms of residues from Gly793 to Lys797 for both peptides with the program MOLecule analysis and MOLecule display (MOLMOL) [17].

Results

Structural Predictions

The graphical output of secondary structure and disorder predictions based on the MeDor metaserver [6] for the *E. coli* SecA IRA1 wild-type domain (residues 755–830) is shown in Figure 2. This result almost exactly matches the X-ray structure, suggesting that the secondary organization of IRA1 is nearly independent of tertiary interactions, which implies the structural autonomy of IRA1 helices and flexibility of the connecting loop [18]. MeDor predictors locate disorder in the well-conserved region A795-K800, and indicate the presence of a strand centered on residues L791 and R792 (Figure 1). Moreover, we have performed comparative simulations on IRA1 single-point mutants. Table 1 summarizes the predictions on mutants for which SecA affinity and/or translocation activity data were available [3,5]. Intriguingly, as compared to the wild-type mutant, most significant differences occur in the region where residues important for binding and/or function are located. Therefore, it seems reasonable to infer that structure–activity relationships involve the preservation of L791 and R792 residues. In fact, the prediction of strand conformation for these residues seems to be associated with retained activity and binding ability, but for a few H1 and H2 mutants, which could also be important for binding and/or function. As an example, Figure 2 shows predictions concerning the inactive Y794A mutant. The disorder consensus for this mutant is higher than the wild type, also affecting L791 and R792 conformation. On this basis, we have synthesized the wild-type peptide, SecA[788–804], and its analog, SecA[788–804]Y794A, to limit structural investigations to conformational differences possibly occurring in this critical region.

Circular Dichroism

The CD analysis of both *E. coli* SecA[788–804] and SecA[788–804] Y794A in H_2O showed spectral features typical of unstructured peptides, with a negative band centered on 195 nm (Figure 3A and B, respectively). The addition of TFE induced a conformational rearrangement that resulted in increased helical content for both peptides, consistently with the organization of the corresponding SecA helix–loop–helix tip region, as observed by the inspection of available structural data. Helix formation is better appreciated by difference spectra, which show the growth of a positive and two negative bands centered around 195 nm, and 208 nm and 222 nm, respectively, as typical of α-helical conformation. However, the low helical content gained by both peptides, even in TFE, suggests high flexibility of this SecA model region.

Nuclear Magnetic Resonance

E. coli SecA[788–804] and SecA[788–804]Y794A spectra recorded in water (90/10 v/v H_2O/D_2O) are in substantial agreement with the CD analyses for both peptides. Indeed, NOESY spectra enhance only $d_{\alpha N}(i,i+1)$ sequential effects, as typical of linear or unstructured peptides. In addition, the $^3J_{NH-CH}$ coupling constant

Table 1. *E. coli* SecA IRA1 mutants:.binding affinity, translocation activity, and disorder predictors

Mutant*	Region	K_d (nM)	Activity	Strand prediction	DorA	REM465	RONN	Fold/Unfold	Loops	HOT LOOPS
wt	–	30	100%	L^{791}-R^{792}	A^{795}-S^{807}	–	G^{793}-K^{800}	A^{795}-E^{806}	I^{789}-K^{804}	G^{755}-K^{764}
W775A	H1	50	yes	L^{791}-R^{792}	A^{795}-S^{807}	–	H^{761}-M^{767}, G^{793}-K^{800}	A^{795}-E^{806}	I^{789}-K^{804}	–
L785R	H1	n.d.	no	L^{791}-R^{792}	A^{795}-S^{807}	–	Y^{794}-K^{800}	A^{795}-E^{806}	I^{789}-K^{804}	G^{755}-K^{764}
Δ783-795	H1	130	no	–	–	–	W^{775}-Q^{801}	M^{782}-E^{806}	K^{797}-A^{804}	G^{755}-K^{764}
G788A	H1	n.d.	~60%	–	A^{795}-E^{806}	–	–	A^{795}-E^{806}	R^{792}-K^{804}	G^{755}-K^{764}
I789A	H1	n.d.	~20%	–	R^{786}-S^{807}	–	R^{792}-Q^{801}	A^{795}-E^{806}	A^{789}-K^{804}	G^{755}-K^{764}
I789R	H1	170	no	–	R^{786}-S^{807}	–	H^{790}-Q^{801}	A^{795}-E^{806}	G^{788}-K^{804}	G^{755}-K^{764}, H^{790}-E^{802}
H790A	H1	n.d.	~60%	–	A^{795}-S^{807}	–	Y^{794}-P^{799}	A^{795}-E^{806}	R^{792}-K^{804}	G^{755}-K^{764}
L791A	loop	n.d.	~0%	–	A^{795}-S^{807}	Y^{794}-E^{802}	H^{790}-E^{802}	A^{795}-E^{806}	H^{790}-K^{804}	G^{755}-K^{764}, R^{792}-Q^{801}
R792A	loop	180	no	–	A^{795}-S^{807}	–	–	A^{795}-E^{806}	H^{790}-K^{804}	G^{755}-K^{764}
G793A	loop	n.d.	~130%	L^{791}-R^{792}	A^{795}-S^{807}	–	–	Q^{796}-E^{806}	A^{793}-A^{804}	G^{755}-K^{764}
Y794A	loop	n.d.	<10%	–	R^{792}-S^{807}	Y^{794}-Y^{803}	R^{792}-E^{802}	A^{794}-E^{806}	L^{791}-Y^{803}	G^{755}-K^{764}, R^{792}-Q^{801}
Q796A	loop	n.d.	~130%	L^{791}-R^{792}	A^{795}-S^{807}	–	–	A^{795}-E^{806}	H^{790}-K^{804}	G^{755}-K^{764}
K797A	loop	n.d.	~130%	L^{791}	A^{795}-S^{807}	–	Q^{796}-P^{799}	–	I^{789}-Y^{803}	G^{755}-K^{764}
D798A	loop	n.d.	~60%	–	A^{795}-S^{807}	–	–	–	H^{790}-K^{804}	G^{755}-K^{764}
P799A	loop	n.d.	<10%	–	A^{795}-S^{807}	–	–	A^{795}-E^{806}	–	G^{755}-K^{764}
K800A	H2	n.d.	~130%	L^{791}	A^{795}-S^{807}	–	R^{792}-Q^{801}	A^{795}-E^{806}	I^{789}-K^{804}	G^{755}-K^{764}
Q801A	H2	n.d.	~150%	L^{791}-R^{792}	A^{795}-S^{807}	–	Q^{796}-P^{799}	A^{795}-E^{806}	I^{789}-K^{804}	G^{755}-K^{764}
E802A	H2	n.d.	<20%	L^{791}-R^{792}	A^{795}-S^{807}	–	–	–	I^{789}-Y^{803}	G^{755}-K^{764}
Y803A	H2	n.d.	<20%	L^{791}-R^{792}	A^{795}-S^{807}	Q^{796}-K^{804}	R^{792}-K^{804}	A^{795}-E^{806}	I^{789}-Y^{803}	G^{755}-K^{764}
E806A	H2	n.d.	no	L^{791}-R^{792}	A^{795}-S^{807}	–	G^{793}-K^{800}	–	I^{789}-Y^{803}	G^{755}-K^{764}
F811A	H2	n.d.	no	L^{791}-R^{792}	A^{795}-S^{809}	–	G^{793}-K^{804}	A^{795}-E^{806}	I^{789}-K^{804}	G^{755}-K^{764}

* Conserved residues are in bold. Regions to which residues belong are indicated according to Ref. [5]. K_d and activity data are from Refs [3,5], respectively. Disorder predictors are from the MeDor metaserver [6]. n.d. = not determined.

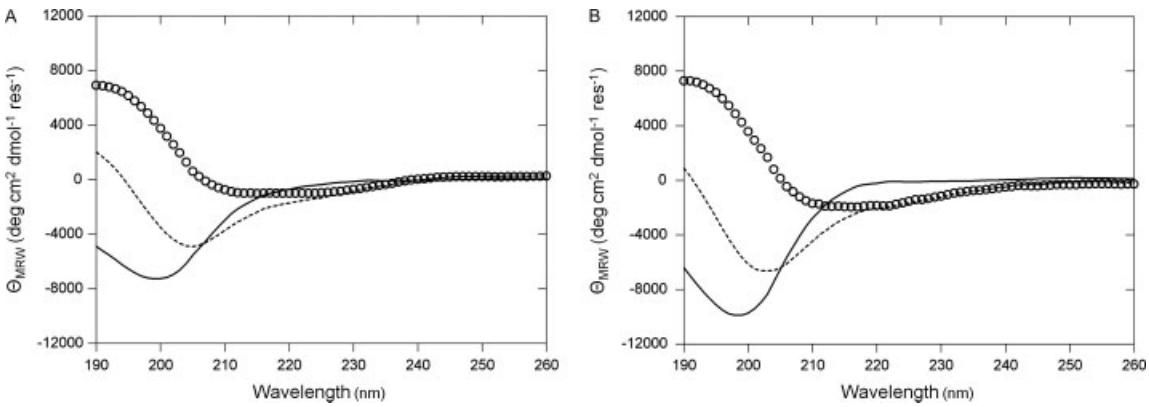

Figure 3. Far UV CD spectra of SecA[788–804] (A) and SecA[788–804]Y794A (B). Open circles (○) correspond to the difference between spectra in 1 : 1 v/v H_2O/TFE (- - - -) and H_2O (———).

values fall in the range of multiple φ-angle values (6 ÷ 8 Hz) [14]. The lack of strong evidence for secondary structure did not allow performing any conformational calculation by energy minimization (EM).

The addition of TFE up to 50/50 (v/v) H_2O/TFE_{d2}-OH induces some changes in the chemical shifts of amidic protons for both peptides, increasing the number of $d_{NN}(i,i+1)$ NOEs, consistently with the CD analysis. The spectra recorded in this solvent system show well-resolved resonances for almost all residues. The Asp798–Pro799 peptide bond was found in trans conformation, as in the aqueous solution, but the $^3J_{NH-CH}$ coupling constant values were, however, in the range of multiple angle values for both peptides as in H_2O. NOESY spectra show $d_{NN}(i,i+1)$ sequential effects in the amino terminal region, spanning residues I789-H790-L791 and R792-G793-Y794 (or A794), suggestive of folding. Moreover, the presence of a short $d_{NN}(i,i+1)$ between K800-Q801 in the C-terminal region could be indicative of incipient helical fold (Figure 4A and B). However, from a preliminary analysis of all spectral parameters, the presence of $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$ effects for both peptides suggests the concomitance of both extended and folded regions in the backbone [19]. Sequential and medium-range NOEs are shown in Figure 5A and B.

Structure Calculations and Analysis

A total of 248 observed NOEs (obtained at m_t = 300 ms) were used for structure calculations on SecA[788–804]. About a hundred distance restraints derived from intraresidue, sequential and medium-range NOEs were introduced in TSSA calculation

Journal of PeptideScience

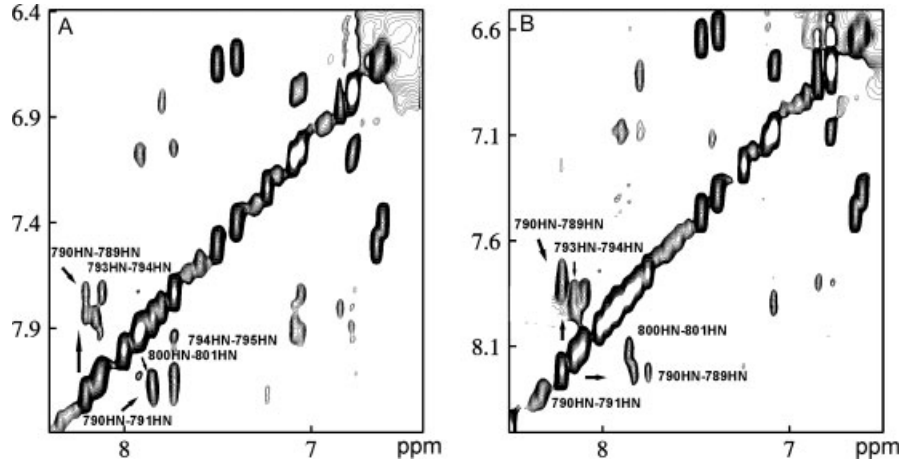

Figure 4. Regions of 300 ms NOESY spectra in 1 : 1 v/v H_2O/TFE. Panels (A) and (B) show the $d_{NN}[i,i+1]$ connectivities of SecA[788–804] and SecA[788–804]Y794A, respectively.

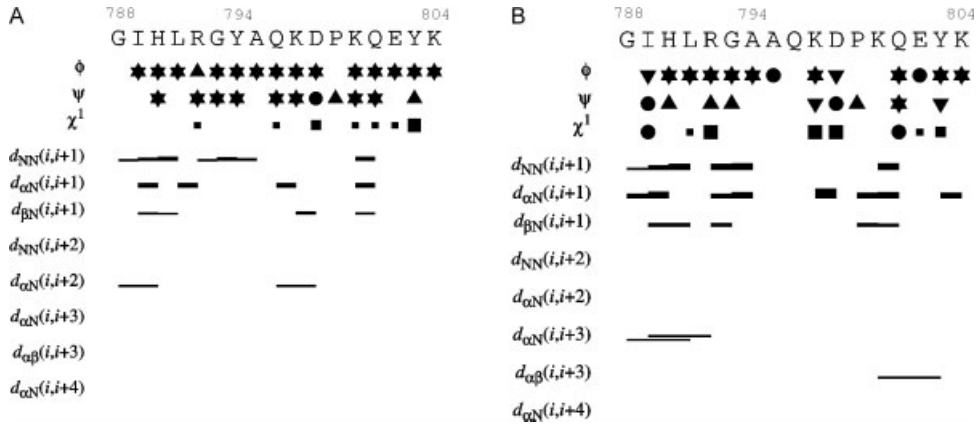

Figure 5. Sequential and medium-range NOE connectivities for SecA[788–804] (A) and SecA[788–804]Y794A (B) in 1 : 1 v/v H_2O/TFE. Connectivities were derived from NOESY spectra at 300 ms mixing time. Backbone NOE connectivities are indicated by horizontal lines between residues, with thickness indicating their relative magnitude. The first three lines below the amino acid sequence represent torsion angle restraints for the backbone torsion angles ϕ and ψ, and for the side-chain torsion angle χ^1. For ϕ and ψ, a ▲ symbol indicates compatibility with an ideal α-helix or 3_{10}-helix; a ▼ symbol indicates compatibility with an ideal parallel or antiparallel β-strand; a ✱ symbol encloses conformation of both α and β secondary structure types; and a ● symbol marks a restraint that excludes the torsion angle values of these regular secondary structure elements. Torsion angle restraints for χ^1 are depicted by filled squares ■ of three different decreasing sizes, depending on whether they allow for one, two, or all three of the staggered rotamer positions.

performed by DYANA package. The best 30 structures in terms of target function (0.45 and 0.00213 for wild type and mutant, respectively) were selected from 80 structures sampled in TSSA calculations. A similar procedure was employed for structure calculations on SecA[788–804]Y794A, using a total of 256 observed NOEs (Figure 5A and B). Backbone clustering analysis for SecA[788–804] in the region 6–10 led to the identification of five structural families. The most populated one contains nine structures with a backbone RMSD of 0.57 ± 0.23 Å for residues 6–10. The same cluster analysis was carried out for SecA[788–804]Y794A, leading to the identification of seven structural families. The most populated one contains 13 structures with a backbone RMSD of 0.63±0.14 Å for residues 6–10. However, given the high flexibility of both peptides, we do not report the average structures.

Discussion

The SecY-assisted translocation of bacterial proteins across the cytoplasmic membrane takes place via the SecA motor protein. SecA is a flexible cytosolic protein, as inferred by the variability of its conformation, both in solution and in the solid state [4,18,20,21], which undergoes ionic strength-dependent monomer–dimer self-association, owing to electrostatic complementarities at the dimer interface [22]. However, it has been proposed that, in the course of protein translocation, monomeric SecA interacts with dimeric SecY [20]. In particular, the SecA monomer binds a first copy of SecY by its N-terminal domain, and then activates the translocation pore of a second copy of SecY through its C-terminal helix–loop–helix IRA1 domain [23,24]. SecA binds SecY mainly through the SecA DEAD domain [3]. Thus, the SecA IRA1 domain could optimize binding both protecting the main binding site and anchoring SecA to SecY, as suggested by the affinity decrease caused by residue substitution or truncation in the IRA1 loop region [3,5]. This view is reinforced by structural information on both SecA and the SecA–SecYEG complexes (Figure 1) [5,23,24]. Mutagenesis has confirmed that SecA IRA1 helix–loop–helix residues are important for translocation activity [3,5]. In particular, single alanine substitution of Y794 (*E. coli* IRA1 loop) reduces the translocation activity at least by 80% compared to wild-type SecA.

Moreover, 40–80% of the translocation activity can be retained by substitution of Y794 with other bulky hydrophobic residues. Finally, cysteine cross-linking experiments have shown that several *E. coli* IRA1 loop residues are in contact with the SecY-complexed translocating protein, but their single mutation to alanine does not decrease the SecA activity [5].

Some authors [4,5] suggested that SecA IRA1 loop Y794 pulls proteins across the SecY channel by direct contact with translocating polypeptide side chains. This hypothesis is based on the consideration of known ATPase mechanisms, which also employ loop tyrosine or other bulky residues to drag proteins, but is hardly suitable for the SecA-SecY system. Indeed, most ATPases are hexameric, and the translocating protein is forced to cross a gate formed by six identical exposed loops with tyrosine on the tips [25–28], whereas SecA is monomeric, and its IRA1 loop is buried in the SecY cavity [4].

Secondary structure predictions here reported suggest that the SecA IRA1 domain (residues 755–830) is structurally autonomous. In particular, MeDor disorder predictors for IRA1 mutants with known binding and/or translocation activity [3,5] agree in locating disorder in the well-conserved region A795-K800, and indicate the presence of a strand centered on residues L791 and R792, which is seemingly associated with SecA decreased activity and binding affinity (Figure 2 and Table 1). As a consequence, we have focused our attention on the H1–H2 connecting loop, which is the only part of IRA1 that contains well-conserved residues responsible for SecA-SecY binding and activation [3,5]. It seems fully representative of the functional IRA1 region and suitable for peptide synthesis and conformational analyses by CD and NMR in solution.

Considering that SecA IRA1 interacts with the interior of the membrane receptor SecY (Figure 1), we adopted TFE to mimic the water–membrane interface-mediated polypeptide folding [29]. It is known that this cosolvent is able to enhance naturally occurring structural propensities, including loop conformation [29–32].

Results here reported indicate that peptides derived from the *E. coli* SecA IRA1 helix–loop–helix region are highly flexible in solution, even in the presence of TFE, which is barely able to induce into a helical conformation a few residues belonging to the IRA1 helices, as expected, but seems unable to stabilize any kind of rigid loop conformation. Alanine substitution of Y794 has a reduced effect on the SecA loop plasticity. According to a mechanism proposed by other authors, SecA IRA1 loop is involved in substrate protein dragging across the SecY channel mainly by Y794. However, as a result of our data, the large decrease of translocation activity for the Y794A mutant, as previously reported, does not appear to be related to changes in the rigidity of this protein region. Furthermore, the alleged Y794 involvement with SecY-crossing polypeptide side chains is not experimentally supported. Instead, we suggest that Y794 is involved in SecA–SecY docking prior to loop-mediated opening of the SecY channel.

Conclusion

On the basis of the crystal structure of SecA–SecY complex [4], it was suggested that the functional SecA IRA1 loop tyrosine pulls proteins across the SecY channel by direct contact with translocating polypeptide side chains [5], somehow resembling the mechanism of multimeric ATPases. However, the alleged mechanism is hardly suitable for monomeric SecA, whose loop is buried into the SecY cavity [4], whereas hexameric ATPases drag proteins exposing six loops crowding around the channel entrance [23–26]. Considering that the SecA IRA1 loop Y794 seems to point away from the SecY translocating channel in the X-ray structure of the complex (Figure 1) [4], and that the translocation activity for the Y794F mutant is almost retained [5], it appears more reasonable that the loop dynamics permits opening of the SecY channel after the tyrosine mediated docking of the SecA IRA1 loop region in between SecY helices.

The present study aims to verify if the loss of activity of the Y794A mutant might be ascribed to severe changes of the SecA IRA1 loop flexibility. We have performed a conformational analysis of the SecA IRA1 loop, which is the main functional region of SecA motor protein, using computational tools for investigating the entire IRA1 domain, and thereafter analyzing two synthetic analogs of IRA1 loop region in solution, by both CD and NMR. Results suggest that the conformational flexibility of these peptides is not affected by Y794A substitution. It seems reasonable to infer that Y794 binding to a SecY hydrophobic site drives docking of the SecA loop, whose dynamics underlies unlocking and activation of the SecY channel.

Acknowledgements

Pasquale Palladino thanks the Centro di Competenza in Diagnostica e Farmaceutica Molecolari (CRdC DFM) for funds (grant number 198/1).

References

1. Tomkiewicz D, Nouwen N, Driessen AJM. Pushing, pulling and trapping-modes of motor protein supported protein translocation. *FEBS Lett.* 2007; **581**: 2820–2828.
2. Driessen AJM, Nouwen N. Protein translocation across the bacterial cytoplasmic membrane. *Annu. Rev. Biochem.* 2008; **77**: 643–667.
3. Vrontou E, Karamanou S, Baud C, Sianidis G, Economou A. Global Coordination of Protein Translocation by the SecA IRA1 Switch. *J. Biol. Chem.* 2004; **279**: 22490–22497.
4. Zimmer J, Nam Y, Rapoport TA. Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature* 2008; **455**: 936–943.
5. Erlandson KJ, Miller SBM, Nam Y, Osborne AR, Zimmer J, Rapoport TA. A role for the two-helix finger of the SecA ATPase in protein translocation. *Nature* 2008; **455**: 984–987.
6. Lieutaud P, Canard B, Longhi S. MeDor: a metaserver for predicting protein disorder. *BMC Genomics* 2008; **9**(Suppl. 2): S25.
7. Chandonia JM, Karplus M. New methods for accurate prediction of protein secondary structure. *Proteins* 1999; **35**: 293–306.
8. Chandonia JM. StrBioLib: a Java library for development of custom computational structural biology applications. *Bioinformatics* 2007; **23**: 2018–2020.
9. Callebaut I, Labesse G, Durand P, Poupon A, Canard L, Chomilier J, Henrissat B, Mornon JP. Deciphering protein sequence information though hydrophobic cluster analysis. Current status and perspectives. *Cell. Mol. Life Sci.* 1997; **53**: 621–645.
10. Woodcock S, Mornon JP, Henrissat B. Detection of secondary structure elements in proteins by hydrophobic cluster analysis. *Protein Eng.* 1992; **5**: 629–635.
11. Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB. Protein disorder prediction: implications for structural proteomics. *Structure* 2003; **11**: 3792–3794.
12. Yang ZR, Thomson R, McMeil P, Esnouf RM. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics* 2005; **21**: 3369–3376.
13. Galzitskaya OV, Garbuzynskiy SO, Lobanov MY. FoldUnfold: web server for the prediction of disordered regions in protein chain. *Bioinformatics* 2006; **1**: 2948–2949.
14. Wüthrich K. *NMR of Proteins and Nucleic Acids*. Wiley: New York; 1986.
15. Palladino P, Tizzano B, Pedone C, Ragone R, Rossi F, Saviano G, Tancredi T, Benedetti E. Structural determinants of unexpected

agonist activity in a retro-peptide analogue of the SDF-1α N-terminus. *FEBS Lett.* 2005; **579**: 5293–5298.
16. Kim Y, Prestegard JH. Measurement of vicinal couplings from cross peaks in COSY spectra. *J. Magn. Reson.* 1989; **84**: 9–13.
17. Koradi R, Billeter M, Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* 1996; **14**: 51–55.
18. Papanikolau Y, Papadovasilaki M, Ravelli RGG, McCarthy AA, Cusack S, Economou A, Petratos K. Structure of dimeric SecA, the *Escherichia coli* preprotein translocase motor. *J. Mol. Biol.* 2007; **366**: 1545–1557.
19. Dyson HJ, Wright PE. Defining solution conformations of small linear peptides. *Ann. Rev. Biophys. Biophys Chem.* 1991; **20**: 519–538.
20. Gelis I, Bonvin AMJJ, Keramisanou D, Koulaki M, Gouridis G, Karamanau S, Economou A, Kalodimos CG. Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* 2007; **131**: 756–769.
21. Chen Y, Pan X, Tang Y, Quan S, Tai PC, Sui S-F. Full-length *Escherichia coli* SecA dimerizes in a closed conformation in solution as determined by cryo-electron microscopy. *J. Biol. Chem.* 2008; **283**: 28783–28787.
22. Or E, Rapoport T. Cross-linked SecA dimers are not functional in protein translocation. *FEBS Lett.* 2007; **581**: 2616–2620.
23. Osborne AR, Rapoport TA. Protein translocation is mediated by oligomers of the SecY complex with one SecY copy forming the channel. *Cell* 2007; **129**: 97–110.
24. Tsukazaki T, Mori H, Fukai S, Ishitani R, Mori T, Dohmae N, Perederina A, Sugita Y, Vassylyev DG, Ito K, Nureki O. Conformational transition of Sec machinery inferred from bacterial SecYE structures. *Nature* 2008; **455**: 988–992.
25. Park E, Rho YM, Koh O-j, Ahn SW, Seong IS, Song J-J, Bang O, Seol JH, Wang J, Eom SH, Chung CH. Role of the GYVG Pore Motif of HslU ATPase in protein unfolding and translocation for degradation by HslV peptidase. *J. Biol. Chem.* 2005; **280**: 22892–22898.
26. Yamada-Inagawa T, Okuno T, Karata K, Yamanaka K, Ogura T. Conserved pore residues in the AAA protease FtsH are important for proteolysis and its coupling to ATP hydrolysis. *J. Chem. Biol.* 2003; **278**: 50182–50187.
27. DeLaBarre B, Christianson JC, Kopito RR, Brunger AT. Central pore residues mediate the p97/VCP activity required for ERAD. *Mol. Cells* 2006; **22**: 451–462.
28. Martin A, Baker TA, Sauer RT. Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Mol. Cells* 2008; **29**: 441–450.
29. Buck M. Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Q. Rev. Biophys.* 1998; **31**: 297–355.
30. Reisen H, Rees AR. Trifluoroethanol may form a solvent matrix for assisted hydrophobic interactions between peptide side chains. *Protein Eng.* 2000; **13**: 739–743.
31. Roccatano D, Colombo G, Fioroni M, Mark AE. Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: a molecular dynamics study. *Proc. Natl. Acad. Sci. U.S.A.* 2002; **99**: 12179–12184.
32. Salinas RK, Shida CS, Pertinez TA, Spisni A, Nakaie CR, Paiva ACM, Schreier S. Trifluoroethanol and binding to model membranes stabilizes a predicted turn in a peptide corresponding to the first extracellular loop of the angiotensin II AT_{1A} receptor. *Biopolymers* 2002; **65**: 21–31.

