Remnant signal peptides on non-exported enzymes: implications for the evolution of prokaryotic respiratory chains

Bérengère Ize,† Sarah J. Coulthurst,‡ Kostas Hatzixanthis,† Isabelle Cadelari,§ Grant Buchanan,‡ Elaine C. Barclay,‡ David J. Richardson,† Tracy Palmer‡ and Frank Sargent‡

1 School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK
2 Division of Molecular Microbiology, College of Life Sciences, University of Dundee, Dundee DD1 5EH, UK
3 Department of Biological Chemistry, John Innes Centre, Norwich NR4 7UH, UK

Correspondence
Bérengère Ize
Berengere.Ize@ifr88.cnrs-mrs.fr

Received 10 August 2009
Revised 19 September 2009
Accepted 21 September 2009

The twin-arginine translocation (Tat) pathway is a prokaryotic protein targeting system dedicated to the transmembrane translocation of folded proteins. Substrate proteins are directed to the Tat translocase by signal peptides bearing a conserved SRRxFLK ‘twin-arginine’ motif. In Escherichia coli, most of the 27 periplasmically located Tat substrates are cofactor-containing respiratory enzymes, and many of these harbour a molybdenum cofactor at their active site. Molybdenum cofactor-containing proteins are not exclusively located in the periplasm, however, with the major respiratory nitrate reductase (NarG) and the biotin sulfoxide reductase (BisC), for example, being located at the cytoplasmic side of the membrane. Interestingly, both NarG and BisC contain ‘N-tail’ regions that bear some sequence similarity to twin-arginine signal peptides. In this work, we have examined the relationship between the non-exported N-tails and the Tat system. Using a sensitive genetic screen for Tat transport, variant N-tails were identified that displayed Tat transport activity. For the NarG 36-residue N-tail, six amino acid changes were needed to induce transport activity. However, these changes interfered with binding by the NarJ biosynthetic chaperone and impaired biosynthesis of the native enzyme. For the BisC 36-residue N-tail, only five amino acid substitutions were needed to restore Tat transport activity. These modifications also impaired in vivo BisC activity, but it was not possible to identify a biosynthetic chaperone for this enzyme. These data highlight an intimate genetic and evolutionary link between some non-exported redox enzymes and those transported across membranes by the Tat translocation system.

INTRODUCTION

Targeting of proteins to their sites of physiological function is an essential feature of all biological systems. Many eubacteria and archaea have the ability to transport pre-folded, often oligomeric, proteins across ionically sealed membranes (Berks et al., 2003). The proteins so transported are synthesized with, or associated with, folded partner proteins that possess distinctive N-terminal signal peptides that bear a common ‘twin-arginine’ SRRxFLK amino acid sequence motif (Berks, 1996). Twin-arginine signal peptides have a common tripartite structure that includes a polar N-terminal (n-) region of variable length, a moderately hydrophobic (h-) region of 12–20 amino acids, and a C-terminal (c-) region that often contains basic residues (Berks, 1996; Cristobal et al., 1999). The consensus motif is always located at the junction between the n- and h-regions (Berks, 1996; Cristobal et al., 1999). All proteins bearing twin-arginine signal peptides are transported across a lipid bilayer by the so-called twin-arginine translocation (Tat) system (Berks et al., 2003).
In prokaryotes, generation of energy by respiratory electron-transfer chains involves the plasma membrane. Respiratory enzymes can be located on either side of, or embedded within, this lipid bilayer, and protein-targeting pathways therefore play central roles in the biogenesis of respiratory chains. Nitrate is commonly exploited as a respiratory electron acceptor across the entire spectrum of prokaryotes (Richardson et al., 2001), and the electrogenic membrane-bound respiratory nitrate reductases catalyse the reduction of nitrate to nitrite using electrons usually supplied from the quinol pool. The well-characterized respiratory nitrate reductase from *Escherichia coli* is a heterotrimeric enzyme comprising a *bis*-molybdopterin guanine dinucleotide (*bis*-MGD)- and [Fe–S]-cluster-containing catalytic subunit (NarG), an [Fe–S]-cluster-containing subunit (NarH), and a integral membrane cytochrome b subunit (NarI) (Bertero et al., 2003). NarGH forms a stable dimer (Jormakka et al., 2004) that is attached to NarI at the cytoplasmic side of the membrane, and this type of architecture and topology is typical for eubacterial enzymes (Martinez-Espinosa et al., 2007). Homologues of these enzymes are found throughout the prokaryotic domains; however the *Haloflexus mediterranei* NarG, for example, is synthesized with a Tat signal peptide (Martinez-Espinosa et al., 2007; Sargent, 2007b). Moreover, the NarG subunits of the archaea *Pyrobaculum aerophilum* and *Archaeoglobus fulgidus* are also predicted to have N-terminal twin-arginine signal peptides (Martinez-Espinosa et al., 2007; Sargent, 2007b).

Recently, sequence analysis has suggested that the non-exported eubacterial NarG proteins may have amino acid signatures that resemble the canonical SRxRFXK twin-arginine motif (Sargent, 2007b; Turner et al., 2004). In *E. coli* NarG this region is part of an ‘N-tail’ that extends away from the cofactor-containing core and forms part of the binding interface with the NarI subunit (Bertero et al., 2005; Vergnes et al., 2006). Prior to the interaction with NarI, the NarG N-tail serves as a binding site for the NarI biosynthetic chaperone (Vergnes et al., 2006). NarI is a protein of the TatD family of peptide-binding proteins (Sargent, 2007b; Turner et al., 2004) and coordinates cofactor-loading and assembly of the nitrate reductase (Lanciano et al., 2007). Together, these observations led to the hypothesis that *E. coli* NarG displays a ‘vestige’ (Turner et al., 2004) or ‘remnant’ (Martinez-Espinosa et al., 2007; Sargent, 2007b) twin-arginine signal peptide at its N-terminus.

Sequence analysis suggests that other molybdenum-dependent enzymes may display similar remnant signal peptide ‘N-tails’. *E. coli* produces a non-exported biotin sulfoxide/methionine sulfoxide reductase encoded by the *bisC* gene (Ezraty et al., 2005; Pierson & Campbell, 1990). BisC is a homologue of the well-characterized *E. coli* Tat substrate trimethylamine *N*-oxide (TMAO) reductase (Tora) (Mejean et al., 1994), but is a cytoplasmic enzyme with an N-tail analogous to that identified on NarG (Ezraty et al., 2005; Sargent, 2007b; Turner et al., 2004).

In this work, the interrelationships between molybdoprotein N-tails, biosynthetic machineries and protein-targeting systems were explored. It is demonstrated that the normally transport-inactive N-tails of both NarG and BisC from *E. coli* can be converted into active Tat signal peptides through substitution of key amino acid residues. Further investigations reveal that biosynthesis of the native reductases is disrupted upon conversion of the N-tails to active signal peptides. It is proposed that N-tails are evolutionary remnants of twin-arginine signal peptides that retain essential biological functions in the biosynthesis of cytoplasmic molybdoenzymes.

**METHODS**

**Strain construction and bacterial growth.** The *E. coli* strains utilized in this work are listed in Supplementary Table S1, available with the online version of this paper. Chromosomal gene deletions, such as the complete deletion of *bioA* and *bisC* in strain BIF029, were performed using a PCR-based method (Datsenko & Wanner, 2000). Deletion alleles were first designed to incorporate an apramycin-resistance cassette flanked by FLP recognition target (FRT) sites on plasmid pLJ773 and then integrated into the chromosome of strain BW25113 (Datsenko & Wanner, 2000). Marked deletion alleles were transferred to various genetic backgrounds by P1 transduction (Sambrook & Russell, 2001). The resistance cassettes were subsequently excised (leaving an 81 bp scar) by induction of the yeast FLP recombinase encoded on pCP20 at 42 °C (Cherepanov & Wackernagel, 1995).

Incorporation of point mutations at precise chromosomal loci, such as inclusion of the *narG*-6' coding sequence at the *narG* locus in strain FTD901, and the BisC-5' coding sequence at the *bisC* locus in strain BIF032, were performed as described by Hamilton et al. (1989).

Bacterial growth tests were performed in LB or M9 medium supplemented with carbon sources and/or antibiotics as necessary (Sambrook & Russell, 2001). Growth rates in liquid culture were monitored in 96-well plates in a temperature-controlled, shaking plate reader (Biotek). The preparation of d-biotin-β-sulfoxide was exactly as described by Melville (1954); when required, this supplement was added to growth media at a final concentration of 5 μg ml⁻¹.

**Plasmid construction.** The vectors utilized in this work are listed in Supplementary Table S1. For constitutive production of proteins, including fusions to the mature domain of AmiA or full-length BisC, plasmids pSUPROM (Kanβ) and pUNIPROM (Ampβ) (Jack et al., 2004) were chosen.

For sequential peptide affinity tagging (SPA) of proteins, plasmid pKS-SPA (Ampβ) was constructed following PCR amplification of the SPA cassette from plL148 (Zehgout et al., 2004) with primers SPAHind3Sp2 and SPASaldown, digestion with HindIII and Sall, and cloning into pBluescript. The *bisC* and *ycdY* genes were amplified and cloned as EcoRI and HindIII fragments into pKS-SPA to give pKSbisC-SPA and pKSYcdY-SPA, respectively. Note that for bisC a mutation was introduced in the primer to replace the native UUG initiation codon with a common AUG sequence.

For the bacterial two-hybrid system, the *E. coli* narI gene was amplified by PCR and cloned into plasmid pT25 (Karimova et al., 1998), while sequences encoding the NarG N-tail, and the NarG-6' derivative, were cloned into the pUT18 vector (Karimova et al., 2001).

Site-directed mutagenesis experiments were performed according to the Quikchange system (Stratagene), and all PCR-based constructs were sequenced on both strands.
Construction and screening of mutant libraries. The NarG and BisC N-tail random mutant libraries were generated by error-prone PCR using defined concentrations of each dNTP to allow an error rate of approximately 1.5% (Fromant et al., 1995). DNA templates used were pBisC-AmiA, pBisC-2-AmiA or pBisC-4-AmiA, and pNarG-AmiA, pNarG-1-AmiA or pNarG-5-AmiA. The PCR products corresponding to the mutated genes were then purified and digested with XbaI and BamHI before being cloned into pAmiA. Each library contained approximately 15 000 independent clones. The libraries were then used to transform MCDSSAC before plating onto LB agar containing SDS (1%, w/v). In total, approximately 30 000 different clones were screened per library. Individual clones that supported growth in the initial screen were patchied first onto media without SDS and then onto LB containing SDS at a higher concentration (3%, w/v). Each potential clone was also grown in the presence of SDS in liquid medium as previously described (Ize et al., 2003). Clones found to consistently confer resistance to exogenous SDS were isolated and sequenced.

Isothermal titration calorimetry. This was performed in 50 mM sodium phosphate buffer (pH 7.5) at 28 °C in a VP-ITC microcalorimeter (MicroCal). Typically, a 1 mM solution of a synthetic peptide was loaded into the syringe. Two synthetic peptides were prepared (Severn Biotech), one corresponding to residues 2–30 of E. coli NarG and another bearing the six amino acid changes displayed by the NarG-6* sequence. The sample cell contained 100 μM of a recombinant, purified NarJiso protein. Titrations consisted of 35 × 8 μl injections of the synthetic peptides into buffer alone (as control), or the NarJiso protein. Data analysis was performed with the Origin software supplied by MicroCal.

Protein methods and enzyme assays. Overproduction and purification of the NarJiso protein were developed especially for squalistically affinity-tagged proteins (Hatzixanthis et al., 2005b). Isolation and identification of SPA-tagged proteins followed the published protocols (Butland et al., 2005). Subcellular fractionation was performed as previously described (Hatzixanthis et al., 2003), as were SDS-PAGE and Western immunoblotting (Laemmli, 1970; Towbin et al., 1979). Briefly, washed cell pellets were resuspended in 50 mM Tris/HCl (pH 7.5), 40% (w/v) sucrose at 10 ml per g cells (wet wt). EDTA was then added to 5 mM (final concentration), and fresh lysozyme was added to 0.6 mg ml⁻¹, before incubation at 37 °C for 30 min. Spheroplasts were harvested by centrifugation and taken up in an equivalent volume of 50 mM Tris/HCl (pH 7.5) before French pressure treatment. Following ultracentrifugation of the resultant crude extract, isolated membranes were again taken up in an equivalent volume of 50 mM Tris/HCl (pH 7.5). This protocol ensures that equivalent proportions of each cell fraction are assayed and thus allows a direct comparison of the subcellular localization of enzyme activities. As a result all enzyme assays are given as units per g cells. Nitrate: benzyl viologen oxidoreductase and β-galactosidase assays were performed as previously described (Karimova et al., 1998; Sargent et al., 1998).

RESULTS

NarG and BisC contain remnant signal peptides

The hypothesis that E. coli NarG and BisC display remnant twin-arginine signal peptides (Martinez-Espinosa et al., 2007; Sargent, 2007b; Turner et al., 2004) was explored using a molecular-genetic approach. A reliable and facile reporter system for positive Tat translocation was developed based around the SDS-sensitivity phenotype of an E. coli tat mutant that arises from non-export of the Tat-dependent amidases AmiA and AmiC (Ize et al., 2003). Strain MCDSSAC is an E. coli amiA/amiC double mutant, which shows identical SDS sensitivity to a tat mutant (Ize et al., 2003). Supply of amiA in multicopy restores growth of the amiA/amiC double mutant on SDS (Fig. 1b, c) in a Tat-dependent manner. A reporter system was therefore designed encoding signal-less AmiA on a multi-copy vector with a constitutive tat promoter. Using this system, a fusion to the Tat signal peptide from E. coli TorA resulted in translocation of the AmiA reporter protein and rescue of the amiA/amiC mutant phenotype (Fig. 1b, c). Without its signal peptide AmiA is not competent for transport and the amiA/amiC double mutant does not grow in the presence of SDS (Fig. 1b, c). The AmiA mature domain is a robust Tat reporter protein that is not compatible for transport by the Sec pathway, since a fusion to the MalE Sec signal peptide could not rescue growth of the reporter strain on SDS (data not shown). Thus the ability of peptide fusions to the AmiA mature domain to rescue growth of the MCDSSAC reporter strain on SDS is a rapid and reliable genetic system for Tat transport.

Next, AmiA fusions to the N-tails of E. coli NarG and BisC were constructed. Neither N-tail could direct export of AmiA in this assay (Fig. 1b, c) and both are therefore transport inactive, confirming the known subcellular localizations of NarG and BisC. To explore the key molecular features responsible for the translocation inactivity of the NarG and BisC N-tails a series of mutations were introduced using both site-directed and random mutagenesis approaches. The NarG and BisC N-tail primary structures broadly follow the tripartite structure of signal peptides (Fig. 1a). A pseudo-Tat motif can be identified in which an aromatic side-chain occupies the position of the second arginine (Fig. 1a). First, NarG F8 and BisC Y8 were each substituted by arginine and export of the AmiA reporter was tested. Neither single mutation was able to restore growth of the reporter strain on SDS (data not shown). Attention next turned to the pseudo h-regions of the N-tails. Polar or charged side-chains are very unlikely to be present in signal peptide h-regions (Cristobal et al., 1999); thus these side-chains in the pseudo h-regions of both NarG and BisC N-tails were targeted for mutagenesis. By first imposing a strict rule of allowing only one nucleotide change per codon, a further five amino acid substitutions were sequentially added to the NarG F8R N-tail: first E16A, followed by D20A, H22L, Q24L, and finally K14I. After the incorporation of each additional amino acid substitution the ability of the new NarG-AmiA fusion to rescue the ability of the reporter strain to grow on SDS was tested. Following the addition of the sixth successive mutation, growth of the reporter strain on SDS was restored and the variant NarG N-tail responsible was designated NarG-6* (Fig. 1).

The most likely explanation for the above result is that NarG-6* has Tat transport activity. This was corroborated by mutagenesis of the new synthetic Tat motif in NarG-6*. A NarG-6* R8F variant, where the original phenylalanine
side chain was restored to position 8, was unable to rescue growth of the reporter strain on SDS (data not shown). Thus the phenotype conferred by expression of the NarG-6*-AmiA fusion is dependent upon an intact twin-arginine motif, and the insertion of phenyalanine within this motif is sufficient to block transport. In addition, incorporation of K14I as the fourth (instead of H22L) or fifth (instead of Q24L) successive mutation, instead of the sixth, did not lead to a variant NarG N-tail-AmiA fusion capable of rescuing growth of the reporter strain on SDS (data not shown). Other combinations of the site-directed substitutions were not constructed and a random mutagenesis approach was employed instead (below) to explore the minimum changes required to induce transport activity in the NarG N-tail.

The pseudo h-region of the BisC N-tail-AmiA fusion (Fig. 1a) could also be modified by site-directed mutagenesis to

![Figure 1](http://mic.sgmjournals.org)

**Fig. 1.** Reactivation of protein transport activity in remnant signal peptides. (a) Similarity between N-tails and twin-arginine signal peptides: sequence alignment of the *E. coli* BisC N-tail (Ec-BisC), the *E. coli* NarG N-tail (Ec-NarG), the *P. pantotrophus* NarG N-tail (Pp-NarG), the *E. coli* TorA twin-arginine signal peptide (Ec-TorA), and the *Haloferax mediterranii* NarG twin-arginine signal peptide (Hm-NarG). Also included are the reactivated NarG-6* and BisC-5* peptides identified in this work (key amino acid changes highlighted in red). The figure was compiled with ClustalW2 and Boxshade. (b) Reactivation of Tat transport activity in the NarG N-tail. *E. coli* strain MCDSSAC (deleted for regions encoding the signal peptides of AmiA and AmiC) was transformed with an empty control plasmid (ΔamiAΔamiC), plasmid pAmiA, which produces only mature AmiA with no signal peptide (pAmiA(-sp)), pTorA-AmiA, which produces a fusion between AmiA and the TorA signal peptide (pTorA AmiA), pNarG-AmiA, which produces a fusion between AmiA and the NarG N-tail (pNarG-AmiA), and pNarG-6*-AmiA, which produces a fusion between AmiA and the modified NarG-6* N-tail (pNarG-6*-AmiA). The parent strain MC4100 was also transformed with a control plasmid. All were plated onto LB + kanamycin, or LB + SDS (1 %, w/v) + kanamycin, plates and incubated aerobically at 37 °C overnight. (c) Reactivation of Tat transport activity in the BisC N-tail. MCDSSAC was transformed with an empty control plasmid (ΔamiAΔamiC), plasmid pAmiA (pAmiA(-sp)), pTorA-AmiA (pTorA AmiA), pBisC-AmiA, which produces a fusion between AmiA and the native BisC N-tail (pBisC-AmiA), and pBisC-5*-AmiA, which produces a fusion between AmiA and the modified BisC-5* N-tail (pBisC-5*-AmiA). All were plated onto LB + kanamycin, or LB + SDS (1 %, w/v) + kanamycin, plates and incubated aerobically at 37 °C overnight.
activate Tat transport activity. A total of four amino acid substitutions were sequentially added on top of the original BisC Y8R change: first E22A, followed by D24A, E26A, and finally a P18L substitution was added. The resultant BisC-5* N-tail was able to rescue growth of the reporter strain on SDS (Fig. 1c). Incorporation of P18L as the third change (rather than D24A) did not generate an export activity in the AmiA-linked BisC N-tail (data not shown).

Taking these results together, it is clear that limited mutagenesis of both the NarG and BisC N-tails can activate a Tat transport function within these peptides.

**Random mutagenesis identifies transport-active variant N-tails**

In a bid to identify the minimum number of mutations required to induce Tat transport activity in the N-tails an error-prone PCR-based random mutagenesis approach was taken. First, DNA encoding the native NarG N-tail was amplified by error-prone PCR and cloned into the AmiA reporter system. This approach did not lead to the isolation of any clones conferring SDS resistance to the reporter strain. Similarly, the use of the NarG F8R N-tail as PCR template did not lead to the identification of any Tat transport active, as PCR template for the random mutagenesis experiment. In this case, from a library containing over 15000 clones, seven different NarG N-tail variants were isolated. Interestingly, the NarG-5.1* variant was taken. First, DNA encoding the native NarG N-tail was required to induce Tat transport activity in the N-tails an random mutagenesis of both the NarG and BisC N-tails can activate

**The role of the NarG N-tail in nitrate reductase assembly and activity was next investigated.** The NarG N-tail has been identified as an important binding site for the NarJ accessory protein (Chan et al., 2006; Lanciano et al., 2007). Here, the ability of NarJ to interact with the Tat-active NarG-6* N-tail was first investigated with a bacterial two-hybrid system (Karimova et al., 1998). The native NarG N-tail and the reactivated NarG-6* N-tail were genetically fused to the N-terminus of the T18 fragment of the adenylate cyclase domain of Bordetella pertussis CyaA. NarJ itself was fused to the T25 fragment of the adenylate

| Table 1. Transport-active variants of the NarG N-tail isolated by random mutagenesis |

<table>
<thead>
<tr>
<th>Primary sequence†</th>
<th>Changes</th>
<th>Transport activity‡</th>
<th>Times isolated§</th>
</tr>
</thead>
<tbody>
<tr>
<td>NarG</td>
<td>MSKFLDRFRYFKQKGETFADGHGQLINTNRDWEDGY</td>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>NarG-6*</td>
<td>MSKFLDRRYFKQKGFATFAAGLGILINTNRDWEDGY</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>NarG-5</td>
<td>MSKFLDRRRYFKQKGFATFAAGLGILINTNRDWEDGY</td>
<td>5</td>
<td>No</td>
</tr>
<tr>
<td>NarG-5.1*</td>
<td>MSKFLDRRRYFKQKGFATFAAGLGILINTNRDWEDGY</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>NarG-5.2*</td>
<td>MSKFLDRRRYFKQKGFATFAAGLGILINTNRDWEDGY</td>
<td>7</td>
<td>Yes</td>
</tr>
<tr>
<td>NarG-5.3*</td>
<td>MSKFLDRRRYFKQKGFATFAAGLGILINTNRDWEDGY</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>NarG-5.4*</td>
<td>MSKFLDRRRYFKQKGFATFAAGLGILINTNRDWEDGY</td>
<td>7</td>
<td>Yes</td>
</tr>
<tr>
<td>NarG-5.5*</td>
<td>MSKFLDRRRYFKQKGFATFAAGLGILINTNRDWEDGY</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>NarG-5.6*</td>
<td>MSKFLDRRRYFKQKGFATFAAGLGILINTNRDWEDGY</td>
<td>8</td>
<td>Yes</td>
</tr>
<tr>
<td>NarG-5.7*</td>
<td>MSKFLDRRRYFKQKGFATFAAGLGILINTNRDWEDGY</td>
<td>8</td>
<td>Yes</td>
</tr>
</tbody>
</table>

†Pseudo twin-arginine motifs are underlined and pseudo h-regions are italicized. Amino acid changes introduced by site-directed mutagenesis are shown in **bold**, and those introduced by random mutagenesis are shown in *double-underlined bold*.

‡Transport activity was judged by the ability of an AmiA fusion protein to rescue growth of the MCDSSAC reporter strain on SDS-containing solid medium.

§NA, Not applicable.

The error-prone PCR random mutagenesis approach was also taken to further analyse the BisC N-tail. In this case the Tat-inactive BisC-2 (carrying two codon changes: Y8R and E22A) and BisC-4 (carrying four codon changes: Y8R, E22A, D24A and E26A) were used as PCR templates. Only experiments with BisC-4 resulted in the isolation of variant N-tails that showed apparent Tat transport activity (Table 2). A total of 11 different, apparently Tat-active, BisC N-tail variants were identified (Table 2). The most frequently isolated variants, BisC-4.2* and BisC-4.3*, carried two additional mutations, resulting in the incorporation of H15L/T27I and H15L/T27A changes, respectively (Table 2). Interestingly, an apparently Tat-active clone designated BisC-4.1* was isolated three times that carries a P18L substitution (Table 2), an identical amino acid change as that deliberately imposed by site-directed mutagenesis (Fig. 1).

**The native NarG remnant signal peptide is required for nitrate reductase biosynthesis**

The role of the NarG N-tail in nitrate reductase assembly and activity was next investigated. The NarG N-tail has been identified as an important binding site for the NarJ accessory protein (Chan et al., 2006; Lanciano et al., 2007). Here, the ability of NarJ to interact with the Tat-active NarG-6* N-tail in vivo was first investigated with a bacterial two-hybrid system (Karimova et al., 1998). The native NarG N-tail and the reactivated NarG-6* N-tail were genetically fused to the N-terminus of the T18 fragment of the adenylate cyclase domain of Bordetella pertussis CyaA. NarJ itself was fused to the T25 fragment of the adenylate
Table 2. Transport-active variants of the BisC N-tail isolated by random mutagenesis

<table>
<thead>
<tr>
<th>Primary sequence†</th>
<th>Changes</th>
<th>Transport activity‡</th>
<th>Times isolated§</th>
</tr>
</thead>
<tbody>
<tr>
<td>BisC</td>
<td>MANSSRSVLAAHWGMLVETDGETVPSSRGALA</td>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>BisC-5*</td>
<td>MANSSRSVLAAHWGMLVATAAGTVFVSSRGALA</td>
<td>5</td>
<td>Yes</td>
</tr>
<tr>
<td>BisC-4</td>
<td>MANSSRSLTTAHWGMLVATAAGTVFVSSRGALA</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>BisC-4.1*</td>
<td>MANSSRSLTTAHWGMLVATAAGTVFVSSRGALA</td>
<td>5</td>
<td>Yes</td>
</tr>
<tr>
<td>BisC-4.2*</td>
<td>MANSSRSLTTAHWGMLVATAAGTVFVSSRGALA</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>BisC-4.3*</td>
<td>MANSSRSLTTAHWGMLVATAAGTVFVSSRGALA</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>BisC-4.4*</td>
<td>MANSSRSLTTAHWGMLVATAAGTVFVSSRGALA</td>
<td>5</td>
<td>Yes</td>
</tr>
<tr>
<td>BisC-4.5*</td>
<td>MANSSRSLTTAHWGMLVATAAGTVFVSSRGALA</td>
<td>7</td>
<td>Yes</td>
</tr>
<tr>
<td>BisC-4.6*</td>
<td>MANSSRSLTTAHWGMLVATAAGTVFVSSRGALA</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>BisC-4.7*</td>
<td>MANSSRSLTTAHWGMLVATAAGTVFVSSRGALA</td>
<td>5</td>
<td>Yes</td>
</tr>
<tr>
<td>BisC-4.8*</td>
<td>MANSSRSLTTAHWGMLVATAAGTVFVSSRGALA</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>BisC-4.9*</td>
<td>MANSSRSLTTAHWGMLVATAAGTVFVSSRGALA</td>
<td>7</td>
<td>Yes</td>
</tr>
<tr>
<td>BisC-4.10*</td>
<td>MANSSRSLTTAHWGMLVATAAGTVFVSSRGALA</td>
<td>9</td>
<td>Yes</td>
</tr>
</tbody>
</table>

†Pseudo twin-arginine motifs are underlined and pseudo h-regions are italicized. Amino acid changes introduced by site-directed mutagenesis are shown in bold, and those introduced by random mutagenesis are shown in double-underlined bold.
‡Transport activity was judged by the ability of an AmiA fusion protein to rescue growth of the MCDSSAC reporter strain on SDS-containing solid medium.
§NA, Not applicable.

cyclase domain. The strong red colour associated with the strain expressing NarJ together with the native NarG N-tail (Fig. 2a) on MacConkey medium is indicative of an interaction and this was confirmed by quantitative β-galactosidase assays (Fig. 2a). The in vivo two-hybrid assay was able to detect an interaction between NarJ and NarG-6* N-tail on the indicator plate (Fig. 2a); however, subsequent quantification of the binding suggested that the interaction was impaired (Fig. 2a).

In order to corroborate the data generated in vivo, attention next turned to a biophysical in vitro technique that would allow quantification of the protein–peptide interactions. Isothermal titration calorimetry is a sensitive technique that has proven useful in studying interactions between peptides and TorD-family chaperones (Buchanan et al., 2006; Chan et al., 2008; Hatzixanthis et al., 2005a). Here, synthetic peptides were prepared based on the native NarG and NarG-6* sequences. Titration of purified affinity-tagged NarJ with the NarG peptide resulted in a sigmoidal binding curve that reached a clear saturation point (Fig. 3a). Analysis of the binding isotherm gave an apparent $K_D$ of 0.7 μM for this interaction. Titration of NarJ with the modified NarG-6* peptide also showed that these proteins can interact in vitro (Fig. 3b). However, the binding curve is markedly different from that generated by the interaction between the native NarG peptide and NarJ, indicative of changes in the thermodynamics of the interaction (Fig. 3a). The best fit to the NarJ–NarG-6* data gave an apparent $K_D$ of 1.1 μM (Fig. 3b). These in vitro binding data corroborate the in vivo two-hybrid results and demonstrate that recognition of NarG-6* by NarJ is impaired by the six amino acid substitutions, but not abolished.

To assess the physiological role of the N-tail in assembly and activity of the nitrate reductase itself, a molecular-genetic approach was taken. A strain was constructed (FTD900) that was genetically devoid of nitrate reductase Z and the periplasmic nitrate reductase (NapA). In addition, the narl gene, encoding the transmembrane cytochrome b subunit of nitrate reductase A, was deleted from this strain, leaving only the NarGH active soluble dimer. Next, DNA encoding the Tat-active NarG-6* N-tail replaced exactly that encoding the native NarG N-tail at the native chromosomal locus to give strain FTD901. These strains thus enabled an analysis of the effect of the NarG-6* changes on the assembly of the core nitrate reductase only, without possible complications from binding to the Narl subunit in the membrane.

The benzyl-viologen-linked nitrate reductase activity was determined for strains FTD900 and FTD901 (Fig. 2b). Following subcellular fractionation, the majority of enzymatic activity was located in the cytoplasm for the strain producing the native NarGH dimer (Fig. 2b). However, the strain producing the modified NarGH carrying the NarG-6* N-tail (FTD901) displayed only very low levels of nitrate reductase (Fig. 2b). In an attempt to visualize the NarGH polypeptides in both strains, the subcellular fractions were subjected to Western analysis (Fig. 2c). Using a polyclonal antibody raised against NarGH it was clear that the enzyme bearing the modified NarG-6* N-tail was completely destabilized, and probably rapidly degraded (Fig. 2c).

Taking these results together, it is clear that while the NarG-6* N-tail exhibits Tat targeting activity (Fig. 1), this new function is at the expense of NarJ binding (Figs 2a and
3a, b), and that this in turn leads to the incomplete assembly of NarGH and its rapid degradation in the cytoplasm (Fig. 2b, c).

The role of the remnant signal peptide in biotin sulfoxide reductase activity

To examine the role of the BisC N-tail in biotin sulfoxide reductase activity a molecular-genetic approach was taken. In the absence of a functional biotin synthesis system (for example, in a bioA mutant) E. coli cannot grow on minimal media with no exogenous biotin source (Cleary & Campbell, 1972). In the presence of biotin sulfoxide, however, the BisC molybdoenzyme generates biotin, which will rescue growth of a bioA strain on minimal media (Pierson & Campbell, 1990). This early experiment was partially repeated here by the construction of strain BIF029 (as MC4100, ΔbioA ΔbisC), which, as expected, could grow aerobically on rich medium but was unable to grow on minimal medium even when supplemented with biotin sulfoxide (Fig. 4a).

To examine more closely the role and activity of BisC in E. coli, a ΔbioA strain was constructed that was also deleted for genes encoding BisC homologues TorA and DmsA (BIF023: as MC4100, ΔbioA ΔtorA ΔdmsABC::KanR). BIF023 remains bisC+, however, and so was able to grow on minimal medium in the presence of biotin sulfoxide (Fig. 4a). The bisC coding sequence carries an unusual UUG start codon, which would be expected to limit cellular levels of the BisC protein. A derivative of BIF023 was therefore constructed that contained a more common

![Fig. 2.](image-url) The NarG remnant signal peptide has a role in enzyme biosynthesis. (a) The NarG-6* peptide is impaired in NarJ binding in vivo. The BTH101 reporter strain was transformed with combinations of the empty pT25 and pUT18 vectors, a pT25 derivative encoding a Nar fusion (pT25-NarJ), a pUT18 derivative encoding a NarG N-tail fusion (pT18-NarG), and a pUT18 derivative encoding the reactivated NarG-6* peptide (pT18-NarG-6*). The transformants were plated onto MacConkey-maltose indicator plates (a red colour is indicative of an interaction), and were also grown aerobically in liquid culture (LB medium) before being subjected to whole-cell β-galactosidase assays. (b) NarG-dependent nitrate reductase activity requires the native N-tail. A strain expressing NarGH only (FTD900: ΔnarI ΔnapA ΔnarZYWV), and an identical strain carrying the NarG-6* reactivated signal (FTD901: narG6* ΔnarI ΔnapA ΔnarZYWV), were grown in LB + 1% (w/v) sodium nitrate + 0.5% (v/v) glycerol, harvested and separated into periplasmic (P), total membranes (M) and cytoplasmic (C) fractions before assaying benzyl-viologen-linked nitrate reductase activity. Units are μmol BV reduced per min per g cells and error bars represent SEM (n=3). (c) The NarG-6* peptide destabilizes nitrate reductase A. The same subcellular fractions as prepared for (d) were separated by SDS-PAGE (10%, w/v, acrylamide), blotted onto nitrocellulose, and challenged with a polyclonal anti-NarGH serum (1:5000 dilution). The positions of the NarG and NarH proteins are indicated. To allow direct comparison, protein loadings were matched so that identical proportions of proteins between different fractions, and between different strains, were analysed.
AUG start codon for bisC (BIF024: as MC4100, ΔbioA ΔtorA ΔdmsABC::KanR ΔbioA AUG). This BIF024 strain showed enhanced growth on minimal medium supplemented with biotin sulfoxide (Fig. 4a). Finally, to investigate the effect of the BisC-5* N-tail on in vivo BisC activity the sequence encoding this reactivated signal peptide (including the AUG start codon) was incorporated into the native bisC locus in strain BIF023 to give strain BIF032 (as MC4100, ΔbioA ΔtorA ΔdmsABC::KanR bisC5*). However, the inclusion of the variant N-tail on BisC severely impaired the ability of the mutant strain to grow with biotin sulfoxide as the sole biotin source (Fig. 4a).

There are two possible explanations for the apparent loss of in vivo biotin sulfoxide reductase activity in the strain producing the modified BisC-5* enzyme (BIF032). First, the N-tail may be a binding site for a biosynthetic chaperone of the TorD family (Sargent, 2007b), and so deleting the cytoplasm of biotin sulfoxide reductase activity. To test the latter hypothesis, a strain carrying the bisC5* allele was further modified by the disruption of the genes encoding the Tat translocase

The search for a biosynthetic chaperone for biotin sulfoxide reductase

It remains a possibility from the data presented here, as well as previous research (Ezraty et al., 2005) and sequence analysis (Sargent, 2007b; Turner et al., 2004), that the BisC N-tail plays a key role in the biosynthesis of the enzyme. By analogy with homologous systems, including NarG and TorA, it is possible that the N-tail acts as a docking site for a biosynthetic chaperone of the TorD family (Sargent, 2007b; Turner et al., 2004). Thus, a series of experiments were designed in an attempt to identify a chaperone for the BisC protein.

First, a mutant strain was constructed (BIF013) that was devoid of bioA as well as genes encoding three TorD-like proteins in E. coli (torD, dmsD and ycdY). The ability of the BIF013 strain to grow in minimal medium was then tested in liquid culture. The parent strain (MC4100) is bioA+ and was able to grow in minimal medium without additional biotin or biotin sulfoxide (Fig. 4c). Deletion of the bioA gene alone (BIF010), or in combination with the torD-like genes (BIF013), prevented growth of both of these strains in minimal medium (Fig. 4c). Supplementation of the growth medium with biotin sulfoxide, however, restored growth of both mutant strains to similar levels (Fig. 4c). This suggests that genetic removal of TorD, DmsD and YcdY does not sufficiently impact on BisC activity to cause an obvious defect in biotin sulfoxide metabolism.

Next, both biochemical and genetic experiments were performed in attempts to identify direct binding-partners for BisC. The SPA affinity-tagging system has recently proved useful in identifying protein complexes, including molybdoenzymes, in E. coli (Butland et al., 2005; Luke et al., 2008). The SPA system comprises the genetic fusion of
Fig. 4. The BisC remnant signal peptide has a role in enzyme biosynthesis. (a) The native BisC N-tail is important for full biotin sulfoxide reductase activity in vivo. E. coli strains MC4100 (parent strain), BIF023 (ΔbioA ΔtorA ΔdmsABC:::Kan<sup>5</sup>), BIF024 (bisC<sub>AUG</sub> ΔbioA ΔtorA ΔdmsABC:::Kan<sup>5</sup>), BIF029 (ΔbisC ΔbioA) and BIF032 (bisC<sup>5*</sup> ΔbioA ΔtorA ΔdmsABC:::Kan<sup>5</sup>) were grown aerobically on LB rich medium, M9 minimal medium and M9 + biotin sulfoxide (BSO) at 5 µg ml<sup>-1</sup>. (b) Increasing expression levels of BisC-5* restores biotin sulfoxide reductase activity. Strain BIF029 (ΔbisC ΔbioA) was transformed with an empty control plasmid (BIF029), a second control of an expression plasmid with no bisC gene (pUNIPROM), a vector encoding full-length BisC enzyme (pBisC), and a plasmid encoding the full-length BisC enzyme carrying the reactivated BisC-5*N-tail (pBisC*). Each transformant was grown aerobically on LB rich medium, M9 minimal medium, and M9 + biotin sulfoxide (BSO) at 5 µg ml<sup>-1</sup>. (c) TorD-like chaperones are not essential for biotin sulfoxide metabolism. E. coli strains MC4100 (parent strain), BIF010 (ΔbioA) and BIF013 (ΔbioA ΔtorD ΔdmsD ΔycdY) were grown in liquid culture in a 96-well plate format in a shaking-incubator plate reader at 37 °C. The strains were grown in M9 minimal media or M9 + biotin sulfoxide (BSO) at 5 µg ml<sup>-1</sup>. (d) A TorD-SPA fusion protein identifies TorA as a binding partner. A derivative of E. coli strain DY330 carrying a torD-SPA allele at the native tor locus (Butland et al., 2005) was cultured anaerobically in rich medium in the presence of 0.4 % (w/v) TMAO. Following two rounds of affinity chromatography, directed against the FLAG tags and the calmodulin-binding peptide, isolated proteins were separated by SDS-PAGE, stained with Coomassie blue G-250, and identified by tryptic peptide mass fingerprinting. Protein bands identified as TorD and TorA are indicated. (e) SPA-tagging fails to identify binding partners for BisC or YcdY. E. coli strain DY330 was transformed separately with pKSycdY-SPA and pKSBisC-SPA and grown aerobically in rich medium. Following two rounds of affinity chromatography, isolated proteins were separated by SDS-PAGE.
three FLAG tags and a calmodulin-binding peptide, separated by a TEV protease site, to the C-terminus of a protein of interest. Two rounds of affinity chromatography are then performed, allowing co-purifying binding partners to be identified following SDS-PAGE and tryptic peptide mass fingerprinting. As an initial control, a modified TorD chaperone was isolated that carried a SPA tag covalently fused to its C-terminus. This approach correctly identified the known binding partner of TorD, the BisC homologue TorA (Fig. 4d). A similar experiment in which BisC itself was SPA-tagged successfully isolated BisC, but the enzyme did not co-purify with any other polypeptides (Fig. 4e). It had previously been suggested that the ‘orphan’ YcdY chaperone may be responsible for BisC maturation (Turner et al., 2004); however, isolation of a SPA-tagged version of this protein did not identify any binding partners (Fig. 4e).

The results of the biochemical experiments were reinforced by genetic studies. A bacterial two-hybrid system was used to screen for both BisC- and YcdY-binding proteins. Genetic fusions to each protein were screened against an extensive library with coverage of all six reading frames of the entire E. coli genome (Handford et al., 2009). No interacting partners were identified by this method (data not shown).

**DISCUSSION**

This work has examined the relationship between the N-tails of two non-exported molybdoenzymes and the Tat translocation pathway. The N-tails of NarG and BisC have the appearance of degenerate signal peptides and here a modest number of amino acid changes were found to induce Tat transport activity in both. For the NarG N-tail, six amino acid changes (from seven base-pair changes) in a 36-residue peptide (108 bp) was the minimum required. Fewer alterations were necessary to restore transport activity to the BisC N-tail: five amino acid changes (six base pair changes) in a 36-residue peptide (108 bp). These experiments demonstrate a very clear and intimate relationship between the native N-tails of these proteins and bona fide twin-arginine signal peptides. The low number of substitutions required to generate an active signal peptide must point to an evolutionary link between the N-tails and Tat signals. Indeed, given that non-native, but real, twin-arginine signal peptides are usually non-functional when heterologously expressed in but real, twin-arginine signal peptides are usually non-

The remnant signal peptide hypothesis

The phenotypic growth test described here is sensitive and powerful, however, and the random mutagenesis experiments outlined in this work can shed light on what the structural and functional requirements for a generic synthetic twin-arginine signal peptide might be. Not surprisingly, polar residues such as lysine (NarG K14), threonine (NarG T17) and histidine (BisC H15) are not tolerated at the N-terminal end of the h-region. In the case of NarG K14, however, the negative effect of this positive charge can be nullified by the inclusion of a branched-chain hydrophobic side-chain (isoleucine or valine) nearby (compare NarG-5.2*, NarG-5.3*, NarG-5.4* and NarG-5.5* in Table 1), possibly on the same face of an α-helix. For BisC, both site-directed and random mutagenesis identified P18 as being important in inhibiting Tat transport activity. This proline is unlikely to be behaving as a classic ‘helix breaker’ and so destroying signal peptide structure since a proline at position 18 can be tolerated in an active signal so long as H15 is replaced by leucine (compare BisC-4.1* and BisC-4.4* in Table 2). Finally, random mutagenesis of the NarG N-tail identified a Tat-active variant with an RL dipeptide within the Tat motif (NarG-5.7*, Table 1). This was interesting since the non-exported Paracoccus pantotrophus NarG homologue naturally carries a leucine at this position (Fig. 1a). This raises the possibility that, outwith the E. coli model, the remnant signal peptides from other biological systems may require fewer genetic changes in order to reactivates protein transport activity.

Taken together, these data strongly support a hypothesis that the NarG and BisC N-tails are ‘remnants’ of twin-arginine signal peptides that were once active in a common ancestor. These findings raise questions surrounding the molecular evolution of the Tat pathway and of components of bacterial respiratory chains in general. What came first, the Tat pathway or the non-exported NarG system? To form an opinion on this, BLAST searches (Altschul et al., 1990) of either E. coli NarG or E. coli TatC were performed against completed prokaryotic genome databases. The two sets of results were then compared against each other to identify the genomes that encode a NarG but not a TatC protein, the logic behind this being that if non-exported NarG-like enzymes preceded (or evolved independently from) the Tat transport system then there would be no obvious correlation between the two. This bioinformatic analysis identified only five bacterial species that harboured a gene encoding a NarG-like nitrate reductase but no Tat translocase – Staphylococcus epidermidis RP62A and ATCC 12228, Lactobacillus plantarum WCFS1, Lactobacillus reuteri JCM1112 and Lactobacillus fermentum IFO3956. Closer inspection of the genomic regions surrounding the
**narGHJI** operons in these organisms suggest that horizontal acquisition is at play. The *S. epidermidis narGHJI* operon is near genes involved in nitrite metabolism and some transposase-like genes; in the lactobacilli the *narGHJI* operon is next to molybdenum cofactor biosynthesis genes, and in one case (*L. reuteri*) it is adjacent to multiple integrase and transposase genes. The presence of nearby mobile genetic element genes implicates horizontal gene transfer events in their acquisition by these biological systems. In the overwhelming majority of cases (1673 accession numbers belonging to completed genomes were searched), genes encoding NarG-like proteins are found in genomes that also encode a Tat translocase. This, together with the very compelling observation that many archael NarG homologues bear intact twin-arginine signal peptides (Martinez-Espinosa *et al.*, 2007) that are transport-active (Widdick *et al.*, 2008), suggests that the cytoplasmic NarG enzyme evolved from a Tat-exported progenitor.

The apparent connection between the Tat pathway and the NarG N-tail is strengthened considerably by the fact that this peptide is a binding site for the NarJ protein. *E. coli* NarJ is a member of a huge family of TorD-like chaperones, almost all of which are known or predicted to establish. The Tat pathway and the NarG N-tail is strengthened considerably by the fact that many archaeal NarG homologues bear intact twin-arginine signal peptides (Sargent, 2007a, b; Turner *et al.*, 2004). NarJ homologues are encoded by all *nar* operons so far identified in both bacteria and archaea (Martinez-Espinosa *et al.*, 2007), highlighting the central importance of NarJ in the biosynthesis of NarG-like nitrate reductases. Very interestingly, however, the conversion of the NarG N-tail to an active signal peptide described here upsets the recognition of this peptide by NarJ *in vivo* and *in vitro*. Moreover, although the interaction between NarG-6* and NarJ is not completely abolished *in vitro*, the assembly of the native enzyme is severely compromised *in vivo* and the protein is apparently rapidly degraded. The reason for this effect is not clear. Given that the Tat transport activity of NarG-6* is very low (see above), it is probably not the case that interactions with the Tat translocase, or even efficient transport to the periplasm, are the primary cause of this degradation. It seems more likely that the amino acid sequence of NarG-6* either precludes efficient binding of NarJ, or impedes the biological function of NarJ once bound to the peptide (a process which remains to be clearly defined). Note also that the overlapping (remnant) transport and chaperone-binding activities displayed by the NarG N-tail help to reinforce the theory that chaperone binding is required to suppress Tat targeting of bona fide signal peptides (Jack *et al.*, 2004; Sargent, 2007a). A second evolutionary question is therefore: what came first, the Tat proofreading function, or the biosynthetic function, of the peptide-binding proteins? Is there any tangible difference between these two processes? To answer this fully, it is important to reach a consensus as to what the chaperones are actually doing when bound to their peptide partners, and this will require future multi-disciplinary approaches to establish.

In comparison to the NarG family, non-exported BisC-like biotin sulfoxide reductases are very rare. However, the similarity of the BisC N-tail to Tat signal peptides, the fact that this remnant signal can be reactivated for transport, the central role the N-tail plays in enzyme biosynthesis, and the relatedness of BisC to the TorA-type of molybdoenzymes, together all point to the BisC N-tail being the docking site for a biosynthetic chaperone of the TorD/NarJ type. Indeed, deletion of this N-tail led to a complete loss of *in vivo* enzymic activity (Ezraty *et al.*, 2005). However, despite the use of multiple complementary experimental approaches in a concerted effort to identify such a chaperone, none has been forthcoming. It may be that the thioredoxin-type protein implicated in electron transfer to BisC (del Campillo-Campbell & Campbell, 1982) doubles as a biosynthetic chaperone. Certainly thioredoxin-like proteins take the role of TorD-family proteins in some [NiFe] hydrogenase assembly networks and are believed to be involved in Tat signal peptide recognition (Parish *et al.*, 2008; Schubert *et al.*, 2007). However, if a protein is interacting with the BisC N-tail, its binding affinity is not sufficiently strong to be identified by the techniques described here. The possibility should, of course, be considered that no chaperone binds to the BisC N-tail at all, and that no chaperone is required to insert the molybdenum cofactor into biotin sulfoxide reductase. While it is certainly true that some molybdo-proteins can be produced in a cofactor-containing form when heterologously expressed in *E. coli* without their signal peptides (e.g. Hilton *et al.*, 1999), there is as yet no correlation between signal peptide binding by a specific chaperone and the cofactor loading event (e.g. Jack *et al.*, 2004). Indeed, there is very often a second chaperone-binding site within the mature region of the enzyme that may be the key to cofactor loading (e.g. Jack *et al.*, 2004; Vergnes *et al.*, 2006), while some enzymes even retain significant activity (up to 50 %) in strains completely deleted for chaperone genes (Pommier *et al.*, 1998; Jack *et al.*, 2004). It is possible, therefore, to assemble a molybdoprotein in the complete absence of biosynthetic chaperones. Thus it is conceivable that BisC has lost all requirements for chaperone activity in both proofreading and cofactor-loading roles, since no binding partners could be unearthed when using remnant signal alone, or full-length BisC, as probes in this work. Why, then, is the BisC N-tail so important for enzymic activity? Alternative possibilities are that the N-tail acts as an ‘intramolecular chaperone’, and is therefore itself intimately involved in enzyme assembly (Chen & Inouye, 2008). Or perhaps the structure of the mRNA transcript in this region is important, with its disruption by mutagenesis subsequently affecting enzyme activity. Clearly, further experimentation is required to fully understand all the molecular processes at work during the assembly of biotin sulfoxide reductase.

Is there any evidence for other degenerate or remnant twin-arginine signal peptides? There are certainly some native, active Tat signal peptides that deviate slightly from the
SRRxFLK consensus motif (Hinsley et al., 2001; Ignatova et al., 2002). The Salmonella TtrB signal peptide, which carries an unusual SKRqFLQ motif, may be beginning to degenerate, as the mature protein is thought to form a tight complex with TtrA, which has a Tat signal peptide of its own (Hensel et al., 1999). More extreme cases of degenerate Tat signal peptides can be identified, however. Also in Salmonella, STM0612 encodes an [Fe–S] protein in an operon with genes encoding a Tat-dependent molybdoprotein (STM0611) and a TordD-like chaperone (STM0610). The STM0612 gene product bears an MRRaFLV motif at its extreme N terminus, but no subsequent h-region that would constitute a true signal peptide. As for TtrAB, the genetics suggest strongly that STM0612 should form a functional dimer with STM0611 and be exported with the single signal peptide on that subunit. It is tempting to speculate that STM0612 carries the ‘scar’ of a remnant signal peptide, the function of which has long since been rendered redundant.

In conclusion, these genetic relics associated with molybdenum and [Fe–S] proteins illustrate the efficiency, intricacy and evolutionary paths of complex enzyme biosynthesis in prokaryotes. Some exported enzymes utilize their N-terminal peptides both as export signals and as binding sites for biosynthetic chaperones. The central importance of this biosynthetic synergy is highlighted by the hypothesis documented here that some non-exported proteins retain a remnant N-terminal peptide as a key site for chaperone binding. Although remnant signal peptides are inactive for Tat transport, they can be reactivated through relatively few genetic modifications.

ACKNOWLEDGEMENTS

We thank G. Chandra (JIC, Norwich) for performing bioinformatic analyses, G. Butland (LBNL, Berkeley) for providing the TordD-SPA strain, and D. H. Boxer (IFR, Norwich) for supplying an anti-NarGH serum. This research was funded in the UK through BBSRC awards BB/D000386/1 (to D. J. R.), and BB/D000386/1 (to T. P.). D. J. R. is the recipient of a Royal Society Wolfson Merit Award and T. P. is an MRC Senior Non-Clinical Research Fellow.

REFERENCES


Edited by: J. Mair