Structural diversity in twin-arginine signal peptide-binding proteins


*Centre for Metalloprotein Spectroscopy and Biology, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom; ‡Center for Molecular and Biomolecular Informatics, Nijmegen Center for Molecular Life Sciences, and **Protein Biophysics Group, Institute for Molecules and Materials, Radboud University, 6525 ED, Nijmegen, The Netherlands; and †Department of Molecular Microbiology, John Innes Centre, Norwich NR4 7UH, United Kingdom

Edited by William M. Clemens, California Institute of Technology, Pasadena, CA, and accepted by the Editorial Board August 15, 2007 (received for review April 30, 2007)

The twin-arginine transport (Tat) system is dedicated to the translocation of folded proteins across the bacterial cytoplasmic membrane. Proteins are targeted to the Tat system by signal peptides containing a twin-arginine motif. In Escherichia coli, many Tat substrates bind redox-active cofactors in the cytoplasm before transport. Coordination of cofactor insertion with protein export involves a “Tat proofreading” process in which chaperones bind twin-arginine signal peptides, thus preventing premature export. The initial Tat signal-binding proteins described belonged to the TorD family, which are required for assembly of N- and O-side reductases. Here, we report that E. coli NapD is a Tat signal peptide-binding chaperone involved in biosynthesis of the Tat-dependent nitrate reductase NapA. NapD binds tightly and specifically to the NapA twin-arginine signal peptide and suppresses signal peptide translocation activity such that transport via the Tat pathway is retarded. High-resolution, heteronuclear, multidimensional NMR spectroscopy reveals the 3D solution structure of NapD. The chaperone adopts a ferredoxin-type fold, which is completely distinct from the TorD family. Thus, NapD represents a new family of twin-arginine signal-peptide-binding proteins.

The Tat (twin-arginine translocation) system is a protein-targeting pathway found in the cytoplasmic membranes of many prokaryotes (1). Tat-targeted proteins are synthesized as precursors with N-terminal signal peptides that contain distinctive SRRxFLK “twin-arginine” amino acid motifs (2). Twin-arginine signal peptides have a common tripartite structure that includes a polar N-terminal (n) region, a moderately hydrophobic (h) region of 12–20 aa, and a C-terminal (c) region that often contains basic residues (2, 3). The consensus motif is always located at the junction between the n and h regions (2). Precursors bearing Tat signal peptides are transported by a membrane embedded export apparatus comprising a signal peptide recognition complex (4) and a protein-conducting channel (5). The physiological role of the Tat system is to transport fully folded proteins (1, 6, 7), and pathogenic bacteria deficient in Tat transport demonstrate reduced virulence, which has led to consideration of this system as a target for novel antimicrobials (8). The model eubacterium Escherichia coli produces 27 Tat-targeted proteins (9). The majority of these are multisubunit, cofactor-containing respiratory enzymes that must bind their redox cofactors, fold, and often associate with signal-less partner subunits before membrane transport can proceed (10, 11). Coordination of these different processes has been the focus of recent research. For instance, the E. coli trimethylamine N-oxide reductase TorA is subject to a chaperone-mediated quality control process known as “Tat proofreading,” which prevents premature targeting of TorA until all biosynthetic processes are complete (11). The TorD chaperone recognizes the TorA signal peptide and binds directly to it, thus shielding it from the Tat transporter (11, 12). E. coli TorD is a member of a family of peptide-binding proteins that share a common fold comprising almost entirely of α-helices and completely devoid of β-strand (13, 14).

The E. coli napFDAGHBC operon encodes a periplasmic dissimilatory nitrate reductase [supporting information (SI) Fig. 4] (15). Periplasmic nitrate reductases are found in a broad spectrum of proteobacteria, including pathogens, and play a central role in the global nitrogen cycle (16, 17). The nap gene encodes the catalytic subunit of the nitrate reductase, a periplasmic protein that binds a [4Fe–4S] cluster and bis-molybdopterin guanine dinucleotide as redox cofactors (18) and is targeted by a Tat signal peptide (Fig. 1B; ref. 19). The other proteins encoded by the nap operon have roles either in electron transfer to NapA (20) or in the biosynthesis of NapA (21, 22). NapD has been shown to be essential for NapA-dependent periplasmic nitrate reductase activity in E. coli (23). The similar cofactor requirements of NapA and TorA (24) suggest that NapA should also be subject to Tat proofreading during assembly. However, detailed sequence analysis suggests NapA and TorA followed different evolutionary lines (25, 26), consistent with this torD genes never colocalize with nap operon, suggesting that the respective biosynthetic mechanisms are also distinct.

In this work, E. coli NapD was structurally and functionally defined as the paradigm representative of a second family of twin-arginine signal-peptide-binding proteins. The solution structure of NapD showed a βαββα/βαβα/βαβα “ferredoxin-type” fold. Genetic and biophysical techniques demonstrate that NapD binds directly and specifically to the NapA twin-arginine signal peptide. NMR was used to probe the peptide-binding site on NapD.

Author contributions: J.M. and C.A.E.M.S. contributed equally to this work; J.M., C.A.E.M.S., G.W.V., and F.S. designed research; J.M., C.A.E.M.S., G.W.V., T.P., and F.S. performed research; D.J.R., T.P., G.W.V., and F.S. analyzed data; and J.M., C.A.E.M.S., D.J.R., T.P., G.W.V., and F.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. W.M.C. is a guest editor invited by the Editorial Board.

Abbreviations: HSQC, heteronuclear single quantum correlation; ITC, isothermal titration calorimetry; Tat, twin-arginine translocation.

Data deposition: The resonance assignments have been deposited in the Biological Magnetic Resonance Data Bank, www.bmrbr.wisc.edu (BMRB accession no. 15381), and the atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2jsx).

1Present address: Environment Naturel, Architectural et Construit–Institut des Sciences et Technologies de l’Environnement/Laboratoire de Biotechnologie Environnementale, École Polytechnique Fédérale Lausanne, Bâtiment Chimie-B Ecublens, CH-1015 Lausanne, Switzerland.

2Present address: Spronk NMR Consultancy, Pylimo Gatve 6, LT-01117 Vilnius, Lithuania.

3Present address: College of Life Sciences, University of Dundee, Dundee DD1 5EH, Scotland.

1To whom correspondence should be addressed. E-mail: f.sargent@dundee.ac.uk.

This article contains supporting information online at www.pnas.org/cgi/content/full/0703967104/DC1.

© 2007 by The National Academy of Sciences of the USA
NapD, and a transport assay suggested that the physiological role of NapD was to prevent premature interaction between the Tat translocase and the NapA signal peptide.

**Results**

**NapD Interacts with the NapA Twin-Arginine Signal Peptide.** The *E. coli* *napD* gene was predicted to encode a protein of 87 aa (9,462 Da). Western blot analysis demonstrated NapD was a water-soluble cytoplasmic protein (Fig. 1A) and thus located on the opposite side of the membrane from the mature NapA protein (Fig. 1A). In the absence of the *napD* gene, *E. coli* lacks all periplasmic nitrate reductase activity (23). Western blot analysis demonstrated the NapA polypeptide was completely destabilized and degraded in a *napD*− background (Fig. 1A). Taken together, these data point to a role for NapD in the biosynthesis of NapA. The ability of NapD to recognize the NapA Tat signal peptide was tested by using a bacterial two-hybrid system based on reconstitution of *Bordetella pertussis* adenylate cyclase in an *E. coli* *cya* mutant (27). NapD was fused to the C terminus of the T25 fragment, and the NapA signal was fused to the N terminus of the T18 fragment of *B. pertussis* adenylate cyclase. In this system, the efficacy of protein–protein interactions was reflected by β-galactosidase activity in a reporter strain. Coproduction of the fusion proteins induced high β-galactosidase activity (6,220 ± 474 Miller units), indicating that NapD recognized the NapA signal peptide *in vivo*

The ability of NapD to bind directly to the NapA signal peptide *in vitro* was tested by using isothermal titration calorimetry (ITC). The NapA twin-arginine signal peptide contained a long stretch of consecutive alanine and valine side-chains (Fig. 1B), which precluded efficient *in vitro* synthesis. Instead, the peptide was synthesized *in vivo* as a C-terminal fusion to maltose-binding protein MalE and purified via a C-terminal His tag, thus ensuring only full-length chimera was included in ITC experiments. Titration of φMalE::NapAHis with purified NapDHis generated negative enthalpy values and a sigmoidal binding curve that reached saturation (Fig. 1C). Analysis of the binding isotherm gave a best fit to a noncooperative model with a binding stoichiometry of 0.59 and an apparent *Kd* of 7 nM (Fig. 1C). Control titrations failed to detect any interactions between NapDHis and MalEHis lacking the signal peptide (SI Fig. 5A).

The specific, tight binding of the NapA signal peptide by NapD was corroborated by size exclusion chromatography (SI Fig. 5C). The φMalE::NapAHis chimera was premixed with NapDHis and loaded onto a molecular exclusion column. The φMalE::NapAHis::NapDHis complex elutes as a higher-molecular-mass species than either φMalE::NapAHis or NapDHis alone (SI Fig. 5C). No binding of NapDHis to φMalE::TorAHis was observed by this technique (SI Fig. 5C), confirming the specific nature of the NapD::NapA interaction.

In ITC, the experimentally observed enthalpy change (ΔHobs) upon peptide binding by NapD is the sum of the true enthalpy change upon binding (ΔHbinding) plus the enthalpy change contributed by the number of protonation/deprotonation events occurring in the buffering system (ΔHionization). To determine the true binding enthalpy of the protein–peptide interaction, identical ITC experiments in a range of buffers were carried out (Fig. 1D). The gradient of the line-of-best-fit indicated that a net release of one proton (nH+ = 0.88) from the protein–peptide complex occurred during the binding event (Fig. 1D).

As a result, the true enthalpy change (ΔHbinding) of the NapDHis::φMalE::NapAHis interaction could be determined by linear regression (Fig. 1D) and was −4.92 kcal/mol (1 kcal = 4.18 kJ). The true entropy change for the interaction (TΔSbinding) could then be calculated as +6.28 kcal/mol. Thus, the binding thermodynamics point to a mechanism driven almost equally through favorable enthalpy and entropy changes (28).

Fig. 1. NapD binds directly to the NapA signal peptide. (A) *E. coli* K-12 strains R611-HN (MC4100 parent strain, nap+), R611-HND (as R611-HN, ΔnapD), and FTD140 (as MC4100, ΔnapA::AprA) were cultured anaerobically in LB medium supplemented with 0.5% (vol/vol) glycerol and 0.4% (wt/vol) fumarate. Whole cells were harvested, washed, and either taken up directly in disaggregating buffer (ΔnapD and ΔnapA lanes) or fractionated into periplasmic (P), total membranes (M), and cytoplasmic (C) proteins before being subjected to SDS/PAGE, blotted, and challenged with an anti-NapD serum (Upper) or an anti-NapA serum (Lower). The asterisk denotes the position of a nonspecific immunoreactive band. (B) The amino acid sequence of the NapA twin-arginine signal peptide with the consensus motif outlined in gray and with the hydrophobic region underscored. (C) Calorimetric titration of a φMalE::NapAHis fusion protein with NapDHis in Tris-HCl buffer (pH 7.5). (Upper) The raw data for the heat effect during the titration. (Lower) The binding isotherm. (D) ITC experiments performed by using identical concentrations of NapDHis and φMalE::NapAHis but utilizing a range of buffers with differing specific heats of ionization. The experimentally observed enthalpy change (ΔHobs) for each experiment was plotted against the ionization (deprotonation) enthalpy of the buffer used (ΔHionization): sodium phosphate, 1.22 kcal/mol; Pipes, 2.74 kcal/mol; Mops, 4.90 kcal/mol; Aces, 7.51 kcal/mol; and Tris-HCl, 11.49 kcal/mol.
The Solution Structure of NapD. The solution structure of recombinant NapD<sub>Hi</sub> (95 residues; 10,528 Da) was determined by using high-resolution, heteronuclear, multidimensional NMR spectroscopy on the basis of 1,340 distance and 119 dihedral restraints (Table 1). No resonance assignments were obtained for residues 1, 2, 79, 80, and the C-terminal His tag residues 89–95 because of chemical exchange and flexibility of the N- and C-terminal parts of the NapD structure. The 20 conformers with the lowest energy values were chosen to represent the most probable structures from 100 randomized starting models and represented a well defined ensemble with good structural characteristics (Table 1). Residue-specific validation (29) revealed a well defined core and no regions in the protein with consistent characteristics (Table 1). Residue-specific validation (29) revealed a probable structures from 100 randomized starting models and the lowest energy values were chosen to represent the most C-terminal parts of the NapD structure. The 20 conformers with residues 1, 2, 79, 80, and the C-terminal His tag residues 89–95 constraints (Table 1). No resonance assignments were obtained for high-resolution, heteronuclear, multidimensional NMR spectra clearly demonstrated have a cylindrical shape of approximate dimensions of 40 Å by 24 Å by 22 Å, and a <sup>15</sup>N relaxation study clearly demonstrated NapD was a monomer in solution (SI Fig. 6). The polypeptide adopted a mixed α/β fold with four β-strands and two α-helices in a β-α-β-β-α-ββ-αββ-fold, where the antiparallel β-strands comprise residues 6–14, 32–38, 53–50, and 69–75 and the two helices comprise residues 19–26 and 53–64, respectively. (B) Residues most affected by NapA signal peptide binding are shown in surface representation, whereas conserved residues E33, E49, and Q43 are shown in Corey–Pauling–Koltun representation (color coding as in A). (C) Secondary structure, solvent accessibility (blue bars), and PROCHECK NMR G factors (colored squares representing δ<sub>B</sub>, 1/2, x1/2, x1 only, and overall from top to bottom, respectively) as a function of residue. Solvent accessibility is color-coded from light-blue (accessible) to dark-blue (buried). G factors are color coded from yellow (good) to red (poor).

NapD Has Antitransport Activity. An in vivo assay was developed to assess the effect of NapD on the Tat transport activity of the NapA signal peptide. The chloraphenicol acetyl transferase (Cat) protein is exported exclusively by the Tat pathway when fused to a twin-arginine signal peptide (30). A genetic fusion of the NapA signal peptide to Cat (pNapA::Cat) resulted in a host strain sensitive to high chloraphenicol concentrations (Fig. 3A) because the reporter protein was targeted to the periplasm (Fig. 3C). However, coexpression of the napD gene with the pnapA::cat chimera prevented both signal peptide processing and transport of the reporter protein to the periplasm (Fig. 3C) and thus conferred chloraphenicol resistance to the host strain (Fig. 3B). This effect was entirely specific to NapD because coexpression of torD in this assay did not result in increased chloraphenicol resistance (SI Fig. 9). Furthermore, the level of chloraphenicol resistance induced by excess NapD in this assay was comparable to that caused by complete inactivation of the Tat translocase itself (SI Fig. 9). Thus, it must be concluded that NapD binding to the NapA signal peptide completely blocks transport of pNapA::Cat on the Tat pathway. This assay suggests NapD has a strong “antitransport” activity such that tight distinct patch on the β-sheet face of NapD (Fig. 2B). The residues that were least affected upon peptide binding were, in general, localized within the α-helical and turn regions (residues 53–69 and 17–31) on the opposite side of the molecule, or within the flexible C terminus which consisted of the binding site being located within the β-sheet.

NapD residues Y76 and S9, together with the conserved side chains E33, Q43, E49, and H77, were chosen for site-directed mutagenesis (SI Fig. 8). Separate substitutions of each of these NapD residues with Ala did not significantly impair either in vitro signal peptide-binding activity or NapD-dependent nitrate reductase biosynthesis in vivo (SI Fig. 8).

Table 1. Structural statistics for NapD

<table>
<thead>
<tr>
<th>Experimental restraints</th>
<th>Number of nonredundant distance restraints</th>
<th>Number of dihedral angle restraints</th>
<th>Structure Z-scores</th>
<th>χ1/2 rotamer normality</th>
<th>Backbone conformation</th>
<th>rmsd (backbone/heavy atom), Å</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1,340 (0.014 ± 0.001)</td>
<td>119 (0.22 ± 0.06)</td>
<td></td>
<td>2.89</td>
<td>0.87</td>
<td>0.63/1.18</td>
</tr>
</tbody>
</table>

The rmsd values of the nonredundant distance and dihedral angle restraints are given in parentheses. The statistics for Z scores and rmsd (backbone/heavy atom) were calculated for the well defined core residues 6–75. For the Ramachandran plot appearance, the values in parentheses are the percentages of residues in the favored, additionally allowed, generously allowed, and disallowed regions, respectively, of the PROCHECK Ramachandran plot.

Maillard et al.  
PNAS | October 2, 2007 | vol. 104 | no. 40 | 15643

BIOCHEMISTRY
binding to the NapA signal peptide can, in some circumstances, out compete its interaction with the Tat translocase itself.

**Discussion**

NapA is a member of the DMSO reductase family of molybdenum-containing enzymes (24), which can be further sorted into three evolutionarily distinct groups (25, 26): types I, II, and III. NapA is a type I enzyme and the well studied Tat substrate TorA belongs to the type III group. The napD gene is the most common nonstructural gene found in prokaryotic nap clusters and is never found colocalized with genes encoding other molybdoenzymes. Indeed, sequence analysis suggests the major differences between the NapA nitrate reductases and the non-exported type I nitrate reductases (NarB-like) are the presence of a Tat signal on NapA and the requirement for a NapD accessory protein in NapA biosynthesis (16). Here, we demonstrate that the genetic link between napD and a structural gene encoding a Tat-dependent nitrate reductase reflects a functional relationship. Tat transport activity of the NapA signal peptide can be repressed upon binding by NapD; however, this anti-transport effect shows the interaction is difficult to reverse in nonnative systems. This is not an issue for the native NapA–NapD system in which NapA is clearly efficiently assembled and exported. It is possible that signal-bound NapD can “sense” the folded state of the native NapA protein, which in turn triggers NapD release but not the folded state of the Cat reporter protein. A possible mechanism is not immediately obvious; however, this would require only a transient and unstable “interaction” between mature NapA and NapD, and, consistent with this hypothesis, no interaction between NapA and NapD, the purified, fully mature and active NapA enzyme could be detected by ITC or analytical ultracentrifugation (T. A. Clarke and F. S., unpublished results). It is conceivable, however, that signal-bound NapD could adopt a different conformation from free NapD that would allow it to “talk” to the NapA mature enzyme or to a second accessory protein, for example, NapF.

The structure of NapD introduces a new family of twin-arginine signal peptide-binding proteins (SI Fig. 10). The ferredoxin-like fold is common and found in proteins with diverse biological functions (31). The data presented here point to the β-sheet face of NapD as harboring a single-peptide-binding site. Interestingly, the β-sheet face of ferredoxin-like proteins is so often involved in mediating interactions with other moieties that it has been described as a “supersite” for ligand binding (32). Some members of the ferredoxin-like family have been at least implicated in peptide binding and protein transport systems, for example, the EpsM protein of the bacterial type II secretion system (33), and others are chaperones involved in metalloprotein assembly processes, for example, CopZ (34), UreE (35), and HypF (36).

The ferredoxin-like fold of NapD is completely different from the all-helical fold exhibited by members of the previously described TorD family of twin-arginine signal-peptide-binding proteins (SI Fig. 11). Thus, consistent with the different evolutionary paths taken by NapA and TorA (25, 26), the concomitant mechanisms of signal-peptide recognition during biosynthesis are also evolutionarily distinct. In addition, FdhE and HyaE have been suggested as possible signal-peptide-binding chaperones for the Tat-dependent formate dehydrogenases and hydrogenses, respectively (1). Although detailed biochemical analyses of FdhE and HyaE have not been reported, structural genomics consortia have recently solved the 3D structures of these proteins, and they too adopt independent folds (SI Fig. 11). Thus, despite the Tat signal peptides from each of these complex enzymes having a final common purpose in interacting with the Tat translocase, the structures of the signal peptides must also be precision-tuned to fit only one type of binding protein. A future challenge will be to understand the molecular basis of how this exquisite chaperone specificity is balanced with a common transport function and, in some cases, the additional involvement of housekeeping chaperones (37, 38).

This work highlights the structural diversity of chaperone-mediated quality control systems operating on the Tat pathway. The solution structure of NapD presented here will enable future studies of this mode of regulation of enzyme biosynthesis and the ability of NapD to completely suppress the transport activity of a signal peptide may have future medical or biotechnological applications.

**Methods**

NapD Production and Purification. The pQE–NapD plasmid encodes NapD with a C-terminal hexahistidine tag. Production of isotopically labeled and unlabeled NapD was by well established protocols described in detail in SI Methods.

NMR Spectroscopy and Resonance Assignment. NMR samples were suspended in 20 mM deuterated Tris, pH 7.4 (Cambridge Isotope Labs)/1 mM DTT/5% (vol/vol) 2H2O. Three samples containing 100% 15N-labeled, 10% 13C-labeled/100% 15N-labeled, or 100% 13C/100% 15N-labeled NapD at 0.5 mM were used for resonance assignments and NOE data collection. Spec-
Assignments of backbone and side-chain NMR resonances were obtained from CBCA(CO)NNH, HBHA(CBCA-CO)NNH, HNCACB, HN(CA)HA, and H(CO)CH— and (HCO)CH—total correlation spectroscopy experiments recorded at 600 MHz by using standard assignment protocols. The assignment of aromatic side chain resonances was performed on a 13C-NOESY–HSQC spectrum optimized for the aromatic region by using aromatic side-chain-specific chemical shifts and expected intraresidual distances. Stereospecific assignments of the prochiral methyl groups of valines and leucines, except valines 7, 36, 69, 72, 75, were obtained by using a 10% 13C-labeled/100% 15N sample (39).

NOE assignments were performed by using AtnosCandid (40,41) for automated NOE assignment in combination with CYANA2.0 for intermediate structure calculations. The input for AtnosCandid consisted of three 3D NOESY spectra recorded at 800 MHz: an aliphatic 13C-NOE–HSQC [spectral tolerances: 13C, 0.5 ppm; 1H (acquisition), 0.035 ppm; 1H (NOE), 0.035 ppm], 15N-NOESY–HSQC [spectral tolerances: 14N, 0.4 ppm; 1H (acquisition), 0.035 ppm; 1H (NOE), 0.035 ppm], and an aromatic 13C-NOE–HSQC [spectral tolerances: 13C, 0.5 ppm; 1H (acquisition), 0.035 ppm; 1H (NOE), 0.035 ppm]. Before the start of the iterative procedure, all chemical shifts were matched to the NOESY–HSQC spectrum optimized for the aromatic region by using standard assignment protocols. The assignment of aromatic side chain resonances was performed on a 13C-NOESY–HSQC [spectral tolerances: 13C, 0.5 ppm; 1H (acquisition), 0.035 ppm; 1H (NOE), 0.035 ppm].

Dihedral angle restraints were derived from chemical shifts via iterative AtnosCandid procedure were used to calculate and refine a final set of structure models. To ultimately refine the structures in explicit solvent itself imposes limits on the proximity of atoms, van der Waal interactions are properly treated, and, therefore, all lower bounds were set to 0 Å. From this adapted set of distance restraints and the Talos-derived dihedral angle restraints, 100 structures were set to 0 Å. From this adapted set of distance restraints and the Talos-derived dihedral angle restraints, 100 structures were acquired at 25°C on Unity Inova 600- and 800-MHz spectrometers (Varian, Palo Alto, CA).

Structure Calculation. Structure Calculations, Statistics, and Validations. A total of 1,340 NOE-based distance restraints obtained in the sixth cycle of the iterative AtnosCandid procedure were used to calculate and refine a final set of structure models. To ultimately refine the structures in explicit solvent itself impose limits on the proximity of atoms, van der Waal interactions are properly treated, and, therefore, all lower bounds were set to 0 Å. From this adapted set of distance restraints and the Talos-derived dihedral angle restraints, 100 structures were calculated by using a standard simulated annealing protocol in X-PLOR–NIH with the OPLSX force field (43). Fifty-four energy structures were used for the final analysis.

Results and Discussion. We thank K. Hatziathanthis (University of East Anglia, Norwich, United Kingdom) for performing preliminary ITC experiments, V. Breukels (Radboud University) for analysis of the NapD relaxation data, and J. A. Cole (University of Birmingham, Birmingham, U.K.) for providing a strain MG1655 at 37°C harboring combinations of these plasmids were monitored in a Synergy 2 plate reader/incubator (BioTek, Winooski, VT).
SUPPORTING INFORMATION

Structural diversity in twin-arginine signal peptide binding proteins
Supporting Figure 4

Supporting Figure 4: The *E. coli* Nap system

Supporting Figure 5: NapD binds specifically to the NapA signal peptide

NapD binds the NapA signal peptide in vitro. (A) Calorimetric titration of MalE<sub>his</sub> only (no signal peptide fused) with NapD<sub>his</sub> in Tris buffer (pH 7.5). (B) For comparison the calorimetric titration of a φMalE::NapA<sub>his</sub> fusion protein with NapD<sub>his</sub> in Tris buffer (pH 7.5) is shown. The upper panels show the raw data for the heat effect during the titrations, and the lower panels are the binding isotherms. Both experiments were done under identical conditions with identical concentrations of proteins. There is no significant interaction between NapD and the MalE carrier protein. (C) NapD<sub>his</sub>, φMalE::NapA<sub>his</sub>, and φMalE::TorA<sub>his</sub> were purified. (i) A 10 times Molar excess of NapD<sub>his</sub> was mixed with either φMalE::NapA<sub>his</sub> (solid line) or φMalE::TorA<sub>his</sub> (dotted line) and run separately on a Superdex-75 size-exclusion column. The NapD<sub>his</sub>-φMalE::NapA<sub>his</sub> complex elutes with a relative mass of 69.1 kDa. (ii) Both free MalE fusion proteins elute with relative masses of 53.7 kDa, φMalE::NapA<sub>his</sub> is shown here. (iii) Free NapD<sub>his</sub> elutes with a relative molecular mass of 19.1 kDa by

www.pnas.org/cgi/doi/10.1073/pnas.0703967104
this technique. Thus, the \( \phi \text{MalE::NapA}^{\text{his}}::\text{NapD}^{\text{his}} \) complex elutes as a higher molecular mass species than either \( \phi \text{MalE::NapA}^{\text{his}} \) or \( \text{NapD}^{\text{his}} \) alone. No binding of \( \text{NapD}^{\text{his}} \) to \( \phi \text{MalE::TorA}^{\text{his}} \) was observed with this technique, confirming the specific nature of the NapD-NapA interaction.

**Supporting Figure 6**

Supporting Figure 6: **NapD is a monomer.**

Histogram of the residue-specific \( \tau_c \) values derived from the Fushman analysis (Fushman D, Weisemann R, Thuring H, Ruterjans H. (1994) *J. Biomol. NMR* 4:61-78) using the \(^{15}\text{N}-\text{R}_1\) and \(^{15}\text{N}-\text{R}_{1p}\) relaxation rates of 45 non-overlapping peaks, corresponding to well-structured NapD residues. The \( \tau_c \) values vary between 4.2 and 6 ns (isotropic value \( \sim 4.6 \) ns) and indicate an approximately axial tensor (\( D_{\text{parallel}}/D_{\text{perpendicular}} \sim 1.4 \)), consistent with the elongated shape of the NapD molecule. The values suggest a protein of \( \sim 8 \) kDa and compare well with for example the protein ubiquitin of 76 amino acids, which shows an isotropic \( \tau_c \sim 4.1 \) ns at 27 C (Tjandra N, Feller SE, Pastor RW, Bax A. (1995) *J. Am. Chem. Soc.* 117:12562-12566).
Supporting Figure 7: Spectral changes upon peptide binding by NapD.

Unlabelled φMalE::NapA$_{his}$ was titrated into a solution of $^{15}$N-labelled NapD$_{his}$ at sub-stochiometric NapA-to-NapD ratios. (A) An overlay of $^{15}$N-HSQC spectra of NapD$_{his}$ in unbound state (black) and in presence of sub-stochiometric amounts of φMalE::NapA$_{his}$ (red). A general broadening and only minor shifts of the resonances was observed. This allowed a direct transfer of the unbound NapD resonance assignments to the complex. The full spectral region with backbone assignments is shown. The black boxes denote regions enlarged in (B) and (C). Since crosspeak positions hardly change, the effect of the titration is only clearly visible in the relative peak heights and volumes. As two clear examples of isolated peaks, Val69 (C), which is the least affected crosspeak in the spectrum apart from Thr86, and Leu10 (B) were chosen. The spectral regions were enlarged and the $^1$H chemical shifts in 1-D shown.
Supporting Figure 8

The NapD family currently contains ~100 members. (A) A sequence alignment of a selection of NapD family proteins. Homologs from *E. coli* (Ec), *Agrobacterium tumefaciens* (Ag), *Rhodobacter sphaeroides* (Rh), *Shewanella oneidensis* (Sh), *Bradyrhizobium japonicum* (Br) and *Paracoccus pantotrophus* (Pa) are shown. The long N-terminal extensions of the *B. japonicum* and *P. pantotrophus* proteins are not shown. The *E. coli* NapD residues chosen for mutagenesis are marked with arrows. Many of the conserved residues are hydrophobic side-chains that form the β-sheet, however 2 ever-present acidic residues can be identified (E33 and E49 in *E. coli* NapD) solvent-exposed at the beginning of β-strand 2 and the end of β-strand 3.

**Table B**

<table>
<thead>
<tr>
<th>NapD&lt;sup&gt;ras&lt;/sup&gt;</th>
<th>N</th>
<th>K&lt;sub&gt;d&lt;/sub&gt; (nM)</th>
<th>ΔG (kcal/mol)</th>
<th>ΔH&lt;sub&gt;obs&lt;/sub&gt; (kcal/mol)</th>
<th>TΔS&lt;sub&gt;obs&lt;/sub&gt; (kcal/mol)</th>
<th>ΔΔG (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>native</td>
<td>N</td>
<td>0.59 ± 0.01</td>
<td>7.0 ± 1.4</td>
<td>-11.2</td>
<td>-15.9</td>
<td>0.00</td>
</tr>
<tr>
<td>S9A</td>
<td>N</td>
<td>0.67 ± 0.15</td>
<td>8.5 ± 2.3</td>
<td>-11.1</td>
<td>-14.1</td>
<td>-3.0</td>
</tr>
<tr>
<td>E33A</td>
<td>N</td>
<td>0.77 ± 0.05</td>
<td>19.8 ± 2.9</td>
<td>-10.6</td>
<td>-14.3</td>
<td>-3.7</td>
</tr>
<tr>
<td>Q43A</td>
<td>N</td>
<td>0.71 ± 0.01</td>
<td>14.0 ± 2.7</td>
<td>-10.8</td>
<td>-10.0</td>
<td>+0.8</td>
</tr>
<tr>
<td>E49A</td>
<td>N</td>
<td>0.92 ± 0.02</td>
<td>47.6 ± 2.8</td>
<td>-10.1</td>
<td>-11.4</td>
<td>-1.3</td>
</tr>
<tr>
<td>Y76A</td>
<td>N</td>
<td>0.63 ± 0.03</td>
<td>26.2 ± 3.2</td>
<td>-10.4</td>
<td>-13.6</td>
<td>-3.2</td>
</tr>
<tr>
<td>H77A</td>
<td>N</td>
<td>0.88 ± 0.01</td>
<td>15.9 ± 3.2</td>
<td>-10.7</td>
<td>-9.0</td>
<td>+1.7</td>
</tr>
</tbody>
</table>

**Supporting Figure 8: Mutagenesis of NapD**

The NapD family currently contains ~100 members. (A) A sequence alignment of a selection of NapD family proteins. Homologs from *E. coli* (Ec), *Agrobacterium tumefaciens* (Ag), *Rhodobacter sphaeroides* (Rh), *Shewanella oneidensis* (Sh), *Bradyrhizobium japonicum* (Br) and *Paracoccus pantotrophus* (Pa) are shown. The long N-terminal extensions of the *B. japonicum* and *P. pantotrophus* proteins are not shown. The *E. coli* NapD residues chosen for mutagenesis are marked with arrows. Many of the conserved residues are hydrophobic side-chains that form the β-sheet, however 2 ever-present acidic residues can be identified (E33 and E49 in *E. coli* NapD) solvent-exposed at the beginning of β-strand 2 and the end of β-strand 3.
respectively. A solvent-exposed basic residue at the beginning of β-strand 3 is also conserved, although *E. coli* NapD has Q43 at this position. In addition to these three residues the conserved H77 together with S9 and Y76, which were implicated in peptide binding, were chosen for mutagenesis. (B) The ability of each of the alanine-substituted NapD variants to interact with the NapA signal peptide was investigated by ITC (1 kcal = 4.18 kJ). All experiments were carried out in Tris buffer (pH 7.6) under identical conditions to allow direct comparisons. None of the substitutions resulted in significantly altered binding constants or relative changes in binding free energy [\(\Delta \Delta G = RT \ln \left( \frac{K_d \text{(variant)}}{K_d \text{(native)}} \right) \)]. (C) The physiological role of the conserved NapD residues was also investigated. The LP202 (ΔnapD) strain, which is devoid of NapA activity (Potter LC, Cole JA. (1999) Biochem J 344:69-76), was transformed with the original cloning vector (pUNI-PROM), the pUP-NapD, pUP-E49A, and pUP-NapD-ΔC plasmids. Cultures, including the nap' parent strain LCB2048, were grown anaerobically in LB supplemented with 0.5% (v/v) glycerol, 0.4% (w/v) fumarate, and 0.2% (w/v) nitrate. Aliquots were withdrawn form each culture at the time points indicated, OD600 noted, and nitrite concentration calculated. The relative nitrite production for each strain was then plotted against time. The E49A NapD variant tested here behaved identically to the five other point mutants constructed (not shown) and could rescue NapA activity to the same level as native NapD, indicating these side chains are not involved in a biochemical activity distinct from peptide binding. However, the C-terminally truncated NapD (‘ΔC-term’), which is lacking the final 10 very flexible residues, was partially compromised in its physiological activity, suggesting that the unstructured C-tail may have role in NapD function. (D) Western analysis of truncated NapD. LP202 (ΔnapD) harbouring pUNI-PROM (‘ΔnapD’), pUP-NapD (‘napD’), and pUP-NapD-ΔC (‘ΔC’) were harvested and fractionated into periplasm, membranes, and cytoplasm. Proteins from the soluble cytoplasmic fraction were separated by SDS-PAGE, blotted, and challenged with anti-NapD serum. The C-terminally truncated NapD is synthesized and soluble.
Supporting Figure 9: Export of \( \phi \text{NapA::Cat} \) is unaffected by TorD and is Tat-dependent.

(A) An *E. coli* strain MG1665 (tat\(^+\)) was transformed with pUNI-NapA encoding a fusion between the NapA signal peptide and chloramphenicol acetyl transferase (‘\( \phi \text{NapA::Cat} \)’). A plasmid (pSU-TorD; Jack RL, Buchanan G, Dubini A., Hatzixanthis K, Palmer T & Sargent F (2004) EMBO J 23:3962-3972) was then introduced and growth rate was monitored aerobically in LB medium in the presence (‘+ Chlor’) or absence (‘-Chlor’) of chloramphenicol (final concentration 200 \( \mu \text{g/ml} \)). The presence of excess TorD in this system was unable to retard export of the Cat reporter protein and therefore the strain remains chloramphenicol sensitive.

(B) An *E. coli* strain completely blocked for Tat transport, DADE (\( \Delta \text{tatABCD}, \Delta \text{tatE} \); Wexler M, Sargent F, Jack RL, Stanley NR, Bogsch EG, Robinson C, Berks BC & Palmer T. (2000) J Biol Chem 275:16717-16722), was transformed with pUNI-NapA encoding \( \phi \text{NapA::Cat} \). Growth rate was then monitored aerobically in LB medium in the presence (‘+ Chlor’) or absence (‘-Chlor’) of chloramphenicol (final concentration 200 \( \mu \text{g/ml} \)). The inability of the DADE mutant strain to export the \( \phi \text{NapA::Cat} \) fusion protein results in resistance to a high concentration of chloramphenicol.
**Supporting Figure 10**

Supporting Figure 11: The ferredoxin-like fold of *E. coli* NapD is unlike that of other known or predicted twin-arginine signal peptide binding proteins.

(A) The solution structure of *E. coli* NapD, this work, (PDB accession number 2JSX) is shown. This protein has a ferredoxin-like fold unlike any other type of Tat signal binding proteins. (B) The crystal structure of a TorD homodimer from *Shewanella massilia* (PDB number 1N1C) published by Tranier S, Iobbi-Nivol C, Birck C, Ilbert M, Mortier-Barriere I, Mejean V, Samama JP. (2003) *Structure* 11:165-174. This protein has an all-helical fold and the *E. coli* homolog of this protein binds directly to the twin-arginine signal peptide of the TorA protein (Hatzixanthis K, Clarke TA, Oubrie A, Richardson DJ, Turner RJ, Sargent, F. (2005) *Proc Natl Acad Sci USA* 102:8460-8465). (C) The crystal structure of the DmsD protein from *Salmonella typhimurium* LT2 (PDB number 1S9U). This structure is unpublished and was solved by Qiu, Y, Zhang R, Tereshko V, Kim Y, Collart F, Joachimiak A, Kossiakoff A. from the Midwest Center for Structural Genomics in 2004. DmsD proteins are closely-related to TorD (Turner RJ, Papiş AL, Sargent F. (2004) *Can J Microbiol* 50:225-238) and the *E. coli* homolog binds directly to the DmsA twin-arginine signal peptide.
Supporting Methods

**NapD production and purification**

The *E. coli* napD gene was PCR amplified and cloned as an *SphI/BglII* fragment into pQE-70 (AmpR; Qiagen) yielding pQE-NapD. S9A, E33A, Q43A, E49A, Y76A, and H77A versions were constructed using QuikChange (Stratagene). The host strain for NapD\(^{\text{his}}\) production was *E. coli* M15 (Qiagen) harboring pREP4 (*lacI\(^{+}\), Kan\(^{R}\); Roche). Protein production was performed at 37°C in 1 litre batches of LB supplemented with 1% (w/v) glucose, appropriate antibiotics and 2 mM IPTG (final). Isotopically-labelled NapD\(^{\text{his}}\) was prepared in M9 minimal medium supplemented with 0.02-0.2% (w/v) \(^{13}\)C-glucose and/or 0.1% (w/v) \(^{15}\)NH\(_4\)Cl. Expression was induced with 2 mM IPTG (final) at an OD\(_{600}\) of 0.4, and the cells were then incubated overnight at 37°C with shaking.

Cell pastes were resuspended at 10 ml/g in 20mM Tris (pH 7.5), 150 mM NaCl and a protease inhibitor cocktail (Calbiochem). Cells were broken by French Press and crude extracts loaded onto 5 ml HisTrap affinity columns (GE Healthcare) equilibrated in 20 mM Tris (pH 7.5), 150 mM NaCl, 1 mM DTT. Bound proteins were eluted with a 30 ml linear gradient to 500 mM imidazole. NapD fractions were pooled, diluted 10 × in 20 mM Tris (pH 7.5), 1 mM EDTA, 1 mM DTT, and loaded onto a 5 ml HiTrap Q HP anion exchange column (GE Healthcare). Proteins were eluted with a 150 ml linear gradient to 1 M NaCl. NapD fractions were pooled, concentrated, and protein concentration estimated by A\(_{280}\) (ε = 6970 M\(^{-1}\)). For isotopically-labeled NapD\(^{\text{his}}\) a size-exclusion chromatography step was added using Superdex-75 (GE Healthcare) equilibrated in 20 mM Tris (pH 7.5), 200 mM NaCl, 1 mM DTT.

**Production and purification of MalE fusion proteins**

The coding sequence for mature (signal-less) maltose binding-protein (MalE), together with a C-terminal poly-asparagine linker and a Factor Xa recognition sequence, was amplified by PCR using pMal-p2x (New England Biolabs) as template and cloned into pQE-60 (AmpR; Qiagen) to give pQM. The pQM construct thus encodes a MalE protein with a flexible C-terminal extension punctuated by a hexahistidine tag and using the unique in-frame *XmaI* and *NcoI* sites in this vector fusions are generated in which a peptide is located C-terminal to MalE and between the Factor Xa site and the hexahistidine tag. Coding sequences for the *E. coli* NapA signal peptide and the *E. coli* TorA signal peptide were PCR amplified and cloned into pQM resulting in the plasmids pQM-NapASP and pQM-TorASP. MalE fusions were produced in the same manner as NapD\(^{\text{his}}\) except *E. coli* strain C43(DE3) (Miroux B, Walker JE. (1996) *J Mol Biol* 260:289-980) replaced M15. MalE fusion proteins were purified via their C-terminal histags, as described for NapD\(^{\text{his}}\), to ensure only full-length chimeras were included in ITC experiments. MalE fusions were further purified by size-exclusion
chromatography as described for NapD<sup>his</sup>, and protein concentration estimated by measuring A<sub>280</sub> (e.g. φMalE::NapA<sup>his</sup> ε = 70,410 M<sup>-1</sup>).

**Bacterial two-hybrid system**


**In vivo nitrite production assay**

The napD gene was PCR amplified and cloned into pUNI-PROM (Amp<sup>R</sup>; Jack RL, Buchanan G, Dubini A, Hatzixanthis K, Palmer T, Sargent F. (2004) *EMBO J* 23:3962-3972) as a BamHI/HindIII fragment, yielding pUP-NapD. S9A, E33A, Q43A, E49A, Y76A, and H77A versions were constructed using QuikChange (Stratagene). A clone producing NapD lacking the final 10 amino acids was also constructed by PCR (pUP-NapD-ΔC). *E. coli* strains LCB2048 (narGHJI-, narZYWV-, narL-, nap<sup>+</sup>) and LP202 (as LCB2048, ΔnapD) (Potter LC, Cole JA. (1999) *Biochem J* 344:69-76) were transformed with the pUP-NapD series and grown at 37ºC under anaerobic conditions in LB medium supplemented with 0.5 % (v/v) glycerol, 40 mM fumarate, and 20 mM nitrate. Samples were taken for OD<sub>600</sub> measurement and nitrite quantification. The nitrite concentration in the extracellular medium was calculated by diluting 200 μl of cell-free culture supernatant in 600 μl 0.1 M Tris HCl (pH 7.5), and adding 400 μl of a 2:1 mixture of 4 % sulfanilamide (in 25 % (v/v) conc. HCl) and 0.08 % (w/v) N-(1-naphthyl)ethylenediamine. After 15 mins incubation the absorbance at 540 nm was measured and nitrite concentration calculated from a standard curve.

**General biochemical methods**

Subcellular fractionations were performed by a lysozyme/EDTA method (Hatzixanthis K, Palmer T, Sargent F (2003) *Mol Microbiol* 49:1377-1390). Antiserum to NapD<sup>his</sup> was raised by Davids Biotechnologie GmbH, Cat antiserum was from 5′→3′ inc., and NapA antiserum was a gift from Prof Jeff Cole (Birmingham, UK). SDS-PAGE and immunoblotting were by Lämmli UK (1970) *Nature* 277:680-685 and Towbin H, Staehelin T, Gordon J (1979) *Proc Natl Acad Sci USA* 76:4350-4354, respectively.

www.pnas.org/cgi/doi/10.1073/pnas.0703967104