Research background

My previous research experience was in the area of microbial bioenergetics, and I undertook my Ph.D research in the laboratory of Professor David Richardson at the University of East Anglia in Norwich, UK. My project goals were to initiate functional studies of the bacterial nitric oxide reductase from *Paracoccus denitrificans*, a gram-negative denitrifying bacterium. I expressed the membrane bound protein complex in *E. coli* and implemented a site-directed mutagenesis program targeting key conserved residues though to be involved in catalysis. I continued on to characterize wild-type, recombinant and mutant enzyme using a combination of optical and magneto-optical spectroscopy and redox potentiometry.

My current research focus, and that of the last four years, has been a functional proteomics project looking at protein-protein interactions in *E. coli*. Using a combination of novel microbiological techniques, modern affinity purification technology, and mass spectrometry our group has attempted to chromosomally affinity tag *E. coli* genes and purify protein complexes from the cell at their native levels. The components of these protein complexes are then characterized using mass spectrometry. We have recently published a report in the journal *Nature* where we demonstrate that these protein-protein interactions form a large protein interaction network that has scale-free characteristics that would make it robust against random perturbations. This high throughput study has also spawned several small biochemistry projects looking at the functional basis of selected protein-protein interactions. Some of these projects are being led by me and others by collaborators and many of these projects have been influenced by my previous interests in metalloproteins and the complex cofactors that they often contain. Publications describing many of these projects and initial findings are currently being submitted for peer review.

Proposed research

Functional Proteomics: Identification and Characterization of Protein Interaction Networks

Introduction

Our increased understanding of molecular processes is often closely coupled to the development of novel methodology. My group will utilize recent advances in molecular biology, protein chemistry and mass spectrometry to uncover novel protein-protein interactions and, by extension, protein complexes. My research focus will be multifaceted and I would like to extend these technologies to the area of microbial pathogenesis and endeavor to uncover protein interaction networks formed between bacterial proteins and mammalian host proteins during infection and how these protein-protein interactions affect the function of the naturally occurring complexes. A more detailed description of this project is given below. I would also like to continue to explore areas of research related to *E. coli* systems biology in collaboration with my current research group and follow up on several interesting observations resulting from the high-throughput *E. coli* protein-protein interaction project.

Principle scientific direction

Citrobacter rodentium (CR) is a pathogenic member of the Enterobacteriaceae and a close relative of *Escherichia coli*. CR is a mouse specific pathogen which is capable of colonizing the host

gastrointestinal tract using a mechanism of attaching and effacing (A/E) lesion formation also employed by the clinically significant diarrhoeagenic *E. coli* strains enteropathogenic *E. coli* (EPEC) and enterohemorragic *E. coli* (EHEC). EPEC and EHEC both lack a suitable small animal model for the in vivo study of host-pathogen interactions due to their narrow host range specificity, and studies of EPEC and EHEC pathogenesis with cultured epithelial cells is now thought not to be representative of the disease process.

CR, and its infection, colonization and pathogenesis in mouse have therefore become a more common model for the disease processes of A/E pathogens. Indeed, many key virulence factors are shared by A/E pathogens and are located on a pathogenicity island (PAI) known as the locus of enterocyte effacement (LEE). The LEE encodes a type III secretion system (TTSS), a macromolecular complex known to act as a 'syringe' and inject virulence effector proteins into host cells. In addition, the LEE encodes, transcriptional regulators, proteins involved in intimate attachment to the host, and several secreted proteins. These secreted proteins can be split into two groups, translocator proteins that are necessary for the translocation of proteins into the host, and effector proteins which carry out a function in the host which plays a role in pathogenesis.

Initial functional characterization of the CR LEE using a proteomics approach identified potential LEE encoded effector proteins and a handful of non-LEE encoded effector proteins, postulated to be encoded in other CR PAIs. This work was however conducted prior to the availability of the CR genome sequence, now complete to 99.99 %. The availability of the CR genome sequence will now greatly enhance the ability of mass spectrometry to characterize the extracellular CR proteome both inside and outside of the host, as peptides will no longer have to be sequenced de novo, a slow process which requires extremely high quality well fragmented peptide mass spectra.

My research group will first utilize modern MUltiDimensional Protein Identification Technology (MUDPIT) to comprehensively characterize the CR extracellular proteome and identify potential effector proteins. In order to obtain complete coverage of the disease process, samples will be taken from hosts at various stages post infection as well as from CR not introduced into the host. This work will hopefully help characterize effectors into temporal groups based on their presence or absence in different stages of disease. These proteins will then be further characterized using further proteomics and biochemical approaches.

Proteins often function as part of multi-protein complexes, and it has been shown that bacterial effector proteins modify and subvert host cell functions, quite possibly by binding to or modifying the composition or function of host protein complexes. Recent work has led to large scale characterization of many native eukaryotic and prokaryotic protein complexes using affinity purification to isolate intact assemblies and mass spectrometry to identify the constituent protein components [1-3]. This project will utilize this approach to look at the complexes formed primarily between CR effector proteins (both known effectors and those potential effectors uncovered by MUDPIT), introduced into the host cell, and the host cell machinery with which these effectors interact.

Recombineering (targeted *in vivo* recombination) methods developed for use in *E. coli* will allow the rapid chromosomal modification of CR genes suspected of encoding effector proteins [4,5] (Figure 1A). The modification of chromosomal alleles using this technology will be of paramount

importance in order to maintain native levels of effector proteins in the host cell at all times. If overproduced, or produced out of context these proteins may form unnatural complexes with host proteins, be potentially mis-localized and not represent an accurate model of the A/E mechanism. These genes will initially be C-terminally epitope tagged with TAP or SPA tandem affinity purification tags to allow for rapid two-step affinity purification (Figure 1B) [6,7]. Examples of purified complexes from *E. coli* and from Human cells demonstrate that this affinity purification procedure is applicable over a huge dynamic range (Figure 1C-F), making it ideal for detecting interactions between CR effectors and mouse proteins. Follow-up work could potentially utilize the same basic technology to create CR effector-Green Fluorescent Protein fusions which could then be used to localize effector proteins in the host cell. It is my hope that this data along with the work of other colleagues will provide insight into the molecular targets of CR effector proteins and ultimately answer questions as to how these effectors modify the function, location or composition of their targets to elicit the effacing mechanism.

Collaborative research projects

My research group will also continue a collaborative project with Professors Andrew Emili and Jack Greenblatt at the University of Toronto to uncover protein-protein interactions in *E. coli* on a genome-wide scale. During my latest Post-Doctoral Fellowship I directed a small team in the Emili and Greenblatt labs which performed the tagging and affinity purification experiments for over 1000 *E. coli* open reading frames [3]. This work is now continuing with the goal of completing the tagging and purification of all *E. coli* proteins within the next few years. Many interesting observations have emerged from this large scale analysis of *E. coli* protein complexes. It would be my intention to follow up on a few select observations, specifically where novel protein-protein interactions may give insight into the roles of proteins essential for cell viability for which little experimental information is available. The eventual goal of these investigations would be to uncover processes which may be the basis for development of novel antimicrobial compounds.

Teaching Philosophy

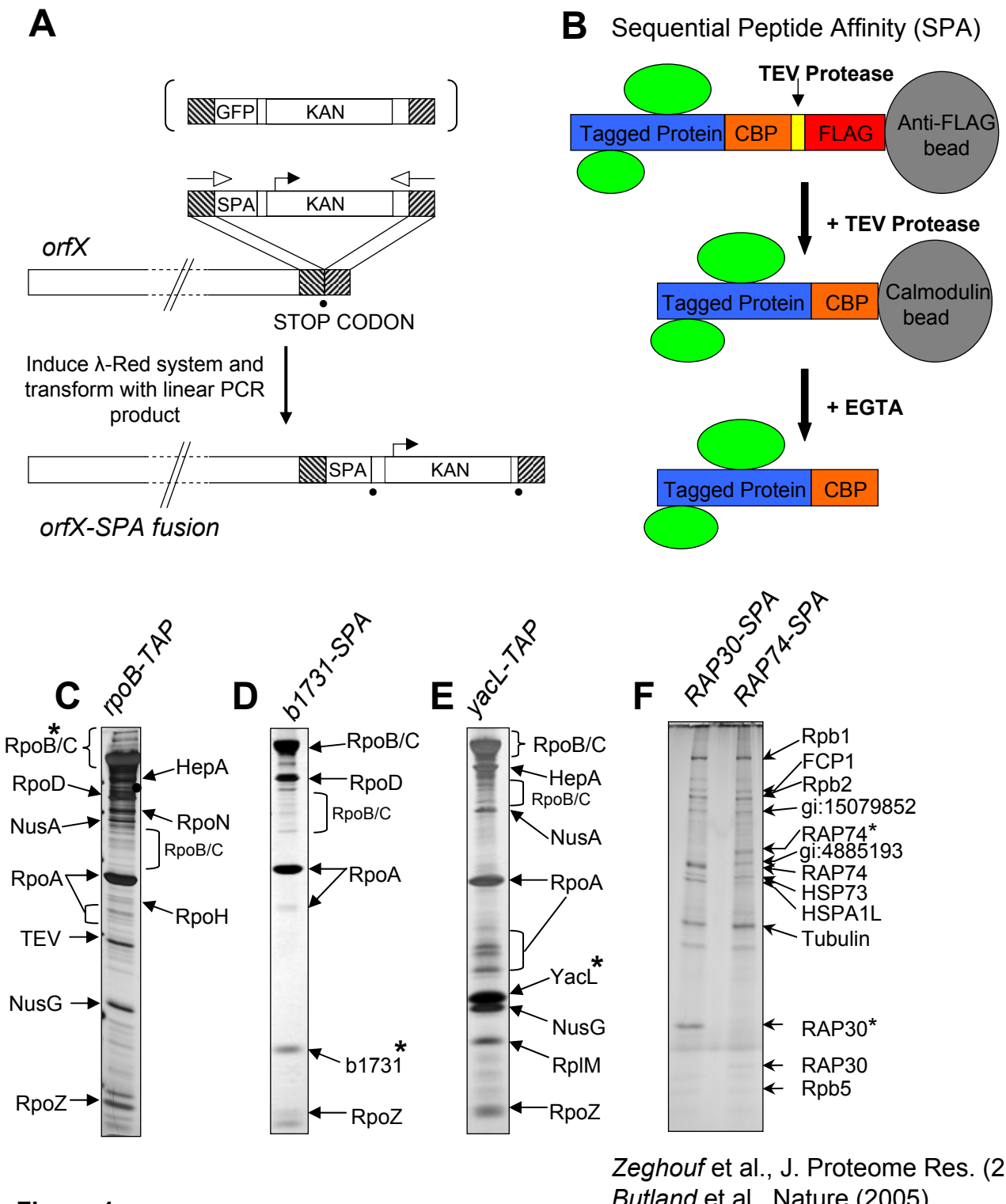
As both a senior graduate student and as a Post Doctoral Fellow I have had the opportunity to mentor both undergraduate students and graduate students on both my own and more diverse projects. I have always enjoyed this teaching aspect of my research, and have recently become more involved by giving a guest lecture to undergraduate biochemistry students. I would feel comfortable teaching many aspects of biochemistry, microbiology and proteomics / functional genomics.

Conclusion

My research focus will be on the identification and characterization of biologically important protein-protein interaction networks. A key goal will be to uncover novel protein-protein interactions between *Citrobacter rodentium* effectors of the attaching and effacing disease process and their molecular targets in the murine host. This hypothesis driven combination of proteomics and biochemistry will help advance our understanding of processes targeted by pathogens and may also be amenable for use in other bacterial species and host systems with only minor need for technological and methodological optimization. I feel that I could contribute both as a researcher and a teacher to the diversity and continued success of the department.

References

1. Gavin, A. C. et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141-7 (2002).
2. Ho, Y. et al. Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* **415**, 180-3 (2002).
3. Butland, G. et al. Interaction network containing conserved and essential protein complexes in Escherichia coli. *Nature* **433**, 531-7 (2005).
4. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-5 (2000).
5. Yu, D. et al. An efficient recombination system for chromosome engineering in Escherichia coli. *Proc Natl Acad Sci U S A* **97**, 5978-83 (2000).
6. Rigaut, G. et al. A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**, 1030-2 (1999).
7. Zeghouf, M. et al. Sequential Peptide Affinity (SPA) system for the identification of mammalian and bacterial protein complexes. *J Proteome Res* **3**, 463-8 (2004).



Zeghouf et al., J. Proteome Res. (2004),
Butland et al., Nature (2005)

Figure 1

(A) Schematic representation of how the λ-Red homologous recombination system can be used to modify chromosomal genes in *E. coli* and *Citrobacter rodentium.* (B) Two step Sequential Peptide Affinity (SPA) system, comprised of Calmodulin Binding Peptide (CBP) and FLAG peptide sequences, used to purify protein complexes from their native levels in the cell. (C-F) Examples of purified *E. coli* RNA polymerase complexes purified using a known tagged subunit (C; rpoB) or proteins of unknown function recently shown to associate with RNA Polymerase (D,E). (F) SPA purifications of tagged TFIIF sununits (RAP30/RAP74) from a stable human cell line. • indicates stop codon, * indicates tagged protein.

GeneAmp 5700 (Applied Biosystems) with a Quantitect SYBR Green RT–PCR kit (Qiagen). Expression levels were normalized on the basis of the amount of *polyubiquitin* transcripts. All values are means and standard deviations for three triplicates of three biological repeats.

Construction of fluorescence-tagged fusions

cDNA fragments of *CASTOR* and *POLLUX* were amplified by PCR with the primers *CAS*-F/*CAS*-R (5′-ACGCGTCGACATGTCCTTGGATTCGGAG-3′, 5′-CATGCCATGGATT CCTTTTCAGTAATTAC-3′) and *POL*-F/*POL*-R (5′-ACGCGTCGACATGATACCACT ACCAGTA-3′, 5′-CATGCCATGGAATCGCCTGAAGCAATCAC-3′), respectively. Each fragment was digested with *Sal*I and *Nco*I and ligated into the same restriction sites of pUC18-CaMV35S-sGFP (S65T)-nos[30]. For the construction of *AtrecA*-DsRed2 fusion, a fragment encoding the transit peptide of *AtrecA*[19] was amplified from *Arabidopsis thaliana* genomic DNA with the use of the primers *AtrecA*-*Sal*I-F/*AtrecA*-LFH-R (5′-ACGCGT CGACATGGATTCACAGCTAGTC-3′, 5′-ATCGAATTCAGAACTGATTTTGTG-3′). DsRed2 fragment was amplified from pDsRed2-1 (Clontech) with the use of DsRed2-LFH-F/DsRed2-*Not*I-R (5′-CTGAATTCGATCGCGACATGGCCTCCTCCGAGAA-3′, 5′-ATTTGCGGCCGCCTACAGGAACAGGTGGTG-3′) primers. *AtrecA*-LFH-R and DsRed2-LFH-F primers were designed to introduce overlapping nucleotides (underlined) at the 3′ and 5′ ends of the *AtrecA* and DsRed2 fragments, respectively. Using both fragments as templates, joint PCR was performed with *AtrecA*-*Sal*I-F and DsRed2-*Not*I-R primers. The resulting fusion fragment was digested with *Sal*I and *Not*I and cloned into the same restriction site of pUC18-CaMV35S-sGFP (S65T)-nos vector.

Microprojectile bombardment and confocal laser scanning microscopy

Microprojectile bombardment was performed with a Biolistic PDS-1000/He Particle Delivery System (Bio-Rad). Epidermis of *Allium cepa* scaly bulb and roots of *Pisum sativum* were bombarded with a rupture-disk pressure of 1,100 p.s.i. (~7.6 MPa) at a target distance of 6 cm. At 24–40 h after bombardment, they were analysed with a Bio-Rad Radiance2000 confocal laser scanning microscope. Green fluorescence of GFP and red fluorescence of DsRed2 were excited at 488 nm with an argon laser and collected sequentially with a filter set (HQ530/60 and E570LP). Images of both fluorescences were processed and merged with the Lasersharp2000 program system (Bio-Rad).

Computer analysis

Sequences were analysed by BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) and GENSCAN version 1.0 (http://genes.mit.edu/GENSCAN.html). Clustal W (http://www.ebi.ac.uk/clustalw/) was used for multiple alignment and evolutionary relationships. The target peptide and transmembrane regions were predicted by TargetP version 1.01 (http://www.cbs.dtu.dk/services/TargetP/) and TMHMM version 2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0/). For domain and structure analyses, both Pfam (http://www.sanger.ac.uk/Software/Pfam/) and FUGUE v.2.0 (http://www-cryst.bioc.cam.ac.uk/~fugue/) were applied.

Received 30 September; accepted 29 November 2004; doi:10.1038/nature03237.
Published online 22 December 2004.

1. Smith, S. E. & Read, D. J. *Mycorrhizal Symbiosis* (Academic, London, 1997).
2. Oldroyd, G. E. D. Dissecting symbiosis: developments in Nod factor signal transduction. *Ann. Bot.* **87,** 709–718 (2001).
3. Kistner, C. & Parniske, M. Evolution of signal transduction in intracellular symbiosis. *Trends Plant Sci.* **7,** 511–518 (2002).
4. Ehrhardt, D., Wais, R. & Long, S. Calcium spiking in plant root hairs responding to Rhizobium nodulation signals. *Cell* **85,** 673–681 (1996).
5. Truchet, G. *et al.* Sulphated lipooligosaccharide signals from *Rhizobium meliloti* elicit root nodule organogenesis in alfalfa. *Nature* **351,** 670–673 (1991).
6. Radutoiu, S. *et al.* Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* **425,** 585–592 (2003).
7. Cárdenas, L. *et al.* Ion changes in legume root hairs responding to Nod factors. *Plant Physiol.* **123,** 443–452 (2000).
8. Senoo, K. *et al.* Isolation of two different phenotypes of mycorrhizal mutants in the model legume plant *Lotus japonicus* after EMS-treatment. *Plant Cell Physiol.* **41,** 726–732 (2000).
9. Szczyglowski, K. *et al.* Nodule organogenesis and symbiotic mutants of the model legume *Lotus japonicus*. *Mol. Plant Microbe Interact.* **11,** 684–697 (1998).
10. Bonfante, P. *et al.* The *Lotus japonicus LjSym4* gene is required for the successful symbiotic infection of root epidermal cells. *Mol. Plant Microbe Interact.* **13,** 1109–1120 (2000).
11. Novero, M. *et al.* Dual requirement of the *LjSym4* gene for mycorrhizal development in epidermal and cortical cells of *Lotus japonicus* roots. *New Phytol.* **154,** 741–749 (2002).
12. Harris, J. M., Wais, R. & Long, S. R. *Rhizobium*-induced calcium spiking in *Lotus japonicus*. *Mol. Plant Microbe Interact.* **16,** 335–341 (2003).
13. Hayashi, M. *et al.* Construction of a genetic linkage map of the model legume *Lotus japonicus* using an intraspecific F2 population. *DNA Res.* **8,** 301–310 (2001).
14. Nakamura, Y. *et al.* Structural analysis of a *Lotus japonicus* genome. II. Sequence features and mapping of sixty-five TAC clones which cover the 6.5-mb regions of the genome. *DNA Res.* **9,** 63–70 (2002).
15. Kawasaki, S. & Murakami, Y. Genome analysis of *Lotus japonicus*. *J. Plant Res.* **113,** 497–506 (2000).
16. Kawaguchi, M. *et al.* Providing the basis for genomics in *Lotus japonicus*: the accessions Miyakojima and Gifu are appropriate crossing partners for genetic analyses. *Mol. Gen. Genomics* **266,** 157–166 (2001).
17. Stracke, S. *et al.* A plant receptor-like kinase required for both fungal and bacterial symbiosis. *Nature* **417,** 959–962 (2002).
18. Ane, J. M. *et al. Medicago truncatula DMI1* required for bacterial and fungal symbioses in legumes. *Science* **303,** 1364–1367 (2004).
19. Köhler, R. H. *et al.* Exchange of protein molecules through connections between higher plant plastids. *Science* **276,** 2039–2042 (1997).
20. Shi, J., Blundell, T. L. & Mizuguchi, K. FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J. Mol. Biol.* **310,** 243–257 (2001).
21. Jiang, Y. *et al.* Crystal structure and mechanism of a calcium-gated potassium channel. *Nature* **417,** 515–522 (2002).
22. Jiang, Y. *et al.* Structure of the RCK domain from the *E. coli* K^+ channel and demonstration of its presence in the human BK channel. *Neuron* **29,** 593–601 (2001).
23. Kwok, E. Y. & Hanson, M. R. Plastids and stromules interact with the nucleus and cell membrane in vascular plants. *Plant Cell Rep.* **23,** 188–195 (2004).
24. Kawaguchi, M. *et al.* Root, root hair, and symbiotic mutants of the model legume *Lotus japonicus*. *Mol. Plant Microbe Interact.* **15,** 17–26 (2002).
25. Perry, J. A. *et al.* A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. *Plant Physiol.* **131,** 866–871 (2003).
26. Schauser, L. *et al.* Symbiotic mutants deficient in nodule establishment identified after T-DNA transformation of *Lotus japonicus*. *Mol. Gen. Genet.* **259,** 414–423 (1998).
27. Niwa, S. *et al.* Responses of a model legume *Lotus japonicus* to lipochitin oligosaccharide nodulation factors purified from *Mesorhizobium loti* JRL501. *Mol. Plant Microbe Interact.* **14,** 848–856 (2001).
28. Broughton, W. J. & Dilworth, M. Y. Control of leghemoglobin synthesis in snake beans. *Biochem. J.* **125,** 1075–1080 (1971).
29. Firmin, J. L. *et al.* Resistance to nodulation of cv. Afghanistan peas is overcome by nodX, which mediates an O-acetylation of the *Rhizobium leguminosarum* lipo-oligosaccharide nodulation factor. *Mol. Microbiol.* **10,** 351–360 (1993).
30. Isono, K. *et al.* Leaf-specifically expressed genes for polypeptides destined for chloroplasts with domains of σ^{70} factors of bacterial RNA polymerases in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **94,** 14948–14953 (1997).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank K. Szczyglowski, J. Webb and J. Stougaard for providing mutant seeds; M. Hayashi for help with mapping; T. Kojima and R. Ohtomo for mycorrhiza analysis; Y. Niwa for providing pUC18-CaMV35S-sGFP (S65T)-nos vector; G. Oldroyd and J. Sun for help with Ca-spiking assays; J. Krüger and B. B. H. Wulff for critical reading of the manuscript; J. Soll for providing the pea root transformation protocol before publication; and M. Durrant for help with modelling the CASTOR pore structure. Part of this work was supported by the fund of Promotion of Basic Research Activities for Innovative Biosciences (BRAIN), and Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency. Research at the Sainsbury Laboratory is funded by the Gatsby Charitable Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.K. (kawasa@nias.affrc.go.jp).
The sequences have been deposited at the DNA Data Bank of Japan with the following accession numbers: LjT02K14a (AP006732), LjT02K14b (AP006733), LjT02K14c (AP006734) LjT45I15 (AP006736), LjT20F11 (AP006737), LjT46G19 (AP006735), LjT45B09a (AP006729), LjT45B09b (AP006730) and LjT45B09c (AP006731); genomic sequences (B-129 Gifu) of *CASTOR* (AB162016), *POLLUX* (AB162017), and mRNA sequences (B-129 Gifu) of *CASTOR* (AB162157) and *POLLUX* (AB162158).

Interaction network containing conserved and essential protein complexes in *Escherichia coli*

Gareth Butland[1], José Manuel Peregrín-Alvarez[2], Joyce Li[1], Wehong Yang[1], Xiaochun Yang[1], Veronica Canadien[3], Andrei Starostine[1], Dawn Richards[3], Bryan Beattie[3], Nevan Krogan[1], Michael Davey[1], John Parkinson[2,4,5], Jack Greenblatt[1,3,5] & Andrew Emili[1,5]

[1]*Banting and Best Department of Medical Research, University of Toronto, 112 College Street, Toronto, Ontario M5G 1L6, Canada*
[2]*Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M4K 1X8, Canada*
[3]*Affinium Pharmaceuticals, 100 University Avenue, Toronto, Ontario M5J 1V6, Canada*
[4]*Department of Biochemistry and* [5]*Department of Medical Genetics and Microbiology, University of Toronto, Medical Sciences Building, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada*

Proteins often function as components of multi-subunit complexes. Despite its long history as a model organism[1], no large-scale analysis of protein complexes in *Escherichia coli* has yet been reported. To this end, we have targeted DNA cassettes into the *E. coli* chromosome to create carboxy-terminal, affinity-tagged alleles of 1,000 open reading frames (~23% of the

genome). A total of 857 proteins, including 198 of the most highly conserved, soluble non-ribosomal proteins essential in at least one bacterial species, were tagged successfully, whereas 648 could be purified to homogeneity and their interacting protein partners identified by mass spectrometry. An interaction network of protein complexes involved in diverse biological processes was uncovered and validated by sequential rounds of tagging and purification. This network includes many new interactions as well as interactions predicted based solely on genomic inference or limited phenotypic data[2]. This study provides insight into the function of previously uncharacterized bacterial proteins and the overall topology of a microbial interaction network, the core components of which are broadly conserved across Prokaryota.

The yeast-based tandem affinity purification (TAP) procedure for isolating protein complexes makes use of site-specific recombination to introduce a dual tagging cassette into chromosomal loci[3]. *Escherichia coli* does not readily recombine exogenous linear DNA fragments into its chromosome, but expression of the lambda general recombination system (λ-Red) markedly enhances integration[4,5]. We adapted one such system[5] to introduce a DNA cassette, bearing a selectable marker and either the TAP or sequential peptide affinity (SPA) tags[3,6], into the C termini of open reading frames (ORFs) in the lysogenic *E. coli* strain DY330 (ref. 6), which harbours λ-Red under control of a temperature-sensitive repressor[5] (Fig. 1a). The tagged bait proteins, expressed at endogenous levels, were purified ~10^6-fold to homogeneity from log-phase cultures using two rounds of affinity chromatography[6]. To minimize nucleic-acid-mediated interactions, extracts were pre-treated with nuclease. Polypeptide components of isolated complexes were then identified using two forms of mass spectrometry: peptide mass fingerprinting was performed on all silver-stained polypeptide bands visible by SDS–polyacrylamide gel electrophoresis (PAGE) not seen in parallel control purifications, whereas gel-free shotgun sequencing was used to identify small and lower-abundance proteins. To minimize the false discovery rate (false positives), protein–protein interactions were deemed authentic provided a reciprocal interaction could be confirmed or if the data were reproducible. Figure 1b shows a schematic overview of the methodology.

The effectiveness of the procedure was confirmed in pilot purifications of DNA-dependent RNA polymerase (RNAP). Tagged core subunit β (RpoB) co-purified specifically with essential elongation factors (NusA and NusG), specialized sigma factors involved in promoter recognition (σ^{32} (RpoH), σ^{38} (RpoS), σ^{54} (RpoN), σ^{70} (RpoD)), and with accessory factors ω (RpoZ), HepA (RapA) and YacL (15 kDa acidic protein of unknown function not previously known to interact with RNAP) (Fig. 2a; see also Supplementary Table 1). Similarly, NusG co-purified with YacL, HepA, core enzyme and termination factor Rho, whereas ω bound σ^{70}, NusA and b1731, another small protein of unknown function. In reciprocal experiments, tagged b1731 co-purified with β, β′ (RpoC), α (RpoA), σ^{70} and ω, but not with Nus factors, HepA or YacL (Fig. 2b), implying an exclusive association with initiating holoenzyme. In contrast, tagged YacL bound ω, NusG and HepA together with core enzyme (Fig. 2c; see also Supplementary Table 1), suggesting a role in elongation. On the basis of their specific and reproducible association with RNAP, we suggest that b1731 and YacL be renamed Rap (RNAP-associated protein) B and C, respectively. YacL and b1731 are not required for viability, and homologues are restricted to γ-Proteobacteria (Supplementary Fig. 1a), suggesting a specialized function. Notably, tagged σ^{54} and σ^{38} bound to the sugar transporter ManX, whereas σ^{32} co-purified with quinone oxidoreductase Qor, indicating that protein–protein interactions may regulate alternative sigma factor activities.

Purification of DNA-dependent DNA polymerase indicated that low-abundance complexes are amenable to analysis. A nine-subunit holo-complex, DNA polymerase III*, which contains a core complex (α (DnaE), ε (DnaQ) and θ (HolE)) requiring only processivity factor β (DnaN) for full replicative activity[7], was readily isolated (Supplementary Table 1). Purification of tagged α, ε and θ yielded core complex as well as the clamp loader (γ and τ (dnaX), δ (HolA), δ′ (HolB), χ (HolC) and ψ (HolD)), which recruits β onto DNA (Fig. 2d; θ detected by gel-free mass spectrometry). In reciprocal experiments (Supplementary Table 1), tagged χ and ψ co-purified with core enzyme and the clamp loader. χ also bound PriA (primosomal DNA helicase), b1808 (putative ATP-dependent helicase) and TopA (DNA topoisomerase I), along with three conserved replication factors that act coordinately *in vitro*[8]: Ssb, single-strand DNA-binding protein previously reported to bind χ[9]; RecQ, a replicative DNA helicase; and TopB, DNA topoisomerase III (Fig. 2d). In turn, Ssb co-purified with TopB, exonucleases RecJ and SbcB, and helicases PriA, RecG and RecQ (Supplementary Table 1), consistent with multiple roles in chromosome dynamics.

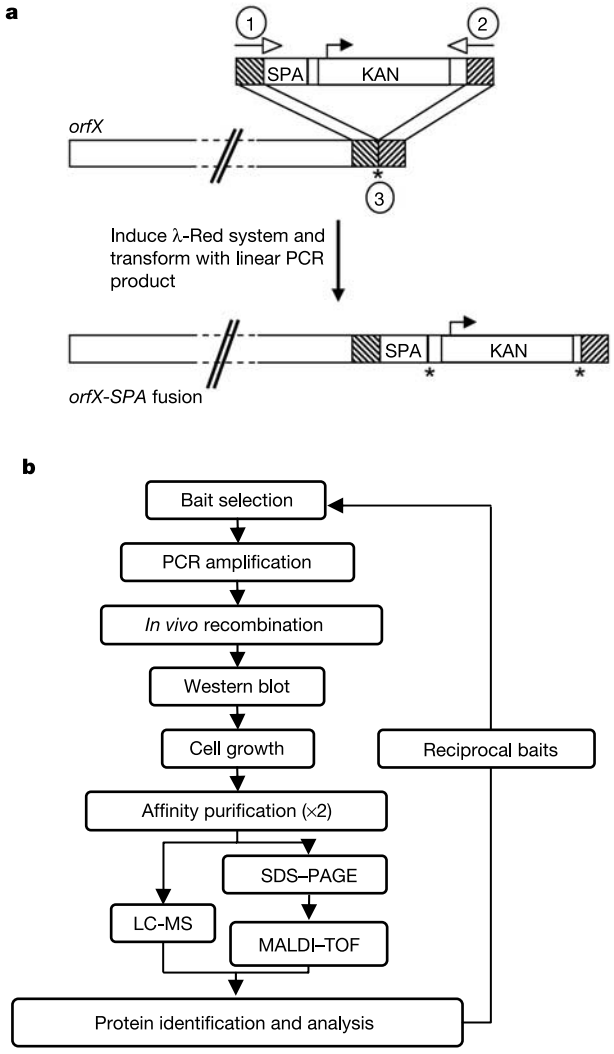


Figure 1 Systematic identification and validation of protein complexes in *E. coli*. **a**, Gene-specific affinity-tagging cassettes produced by polymerase chain reaction (PCR)[6] using primers (1, 2) homologous to a target translational termination codon (3) were integrated into the *E. coli* chromosome using the λ-Red recombination system[5]. KAN, kanamycin-resistance cassette. Asterisks indicate stop codons. **b**, Flow chart of steps in purifying and validating protein complexes. Interactions were confirmed either by reciprocal tagging and purification, or by repeat analysis. LC-MS, gel-free liquid chromatography-tandem mass spectrometry; MALDI-TOF, gel-based peptide mass fingerprinting using matrix-assisted laser desorption/ionization–time-of-flight mass spectrometry.

TopB co-purified with Ssb and RecQ, suggesting that these factors serve a similar role in coordinating DNA replication in prokaryotes as suggested for their eukaryotic homologues[10].

We expanded our analysis to cover about one-quarter of the *E. coli* genome, targeting 1,000 ORFs (Supplementary Fig. 2 and Supplementary Table 2), including 248 uncharacterized 'y' genes, 168 putative 'b' genes and 209 of the most broadly conserved genes encoding soluble (non-membrane) non-ribosomal proteins essential in *E. coli* or another bacterium (see Methods). We successfully tagged 857 proteins (86%; confirmed by western blot analysis), including 198 essential and conserved (essential-conserved) proteins, and were able to purify 648 (65%; detected by mass spectrometry), which compares favourably to analogous studies of protein complexes in yeast[3]. A total of 118 of these proteins had no detectable partners, whereas 5,254 putative protein–protein interactions were detected for the other 530 baits (Supplementary Table 1). To eliminate false-positives, we reciprocally tagged and purified a large subset of candidate partners. A validation rate of ~53% was achieved (see Supplementary Information), which compares favourably to studies in yeast[11], confirming the stringency of the methodology. A total of 716 non-redundant binary interactions, involving 83 essential (excluding the ribosome) and 152 non-essential proteins, have been validated so far (highlighted in Supplementary Table 1). The entire validated data set is shown graphically in Fig. 3a.

Eighty-five per cent of the validated interactions are new, as they are not described in the Database of Interacting Proteins (DIP)[12], Biomolecular Interaction Network Database (BIND)[13], STRING[14], or Prolinks databases[15], whereas only ten orthologous interactions (interologs) were reported in a two-hybrid interaction screen in *Helicobacter Pylori*[16] (Supplementary Table 3). The significance of these novel interactions is reinforced by functional annotation (Supplementary Table 2). For instance, acyl carrier protein (ACP), a key carrier of growing fatty acid chains, bound specifically and reproducibly to enzymes linked to biogenesis of fatty acids, phospholipids and lipid A (essential outer-membrane constituent) (Fig. 2e; see also Supplementary Table 1), including two 3-ketoacyl-ACP synthases (FabB, FabF), 3-ketoacyl-ACP reductase (FabG), 3-hydroxyacyl-ACP dehydrase (FabZ), LpxD (essential protein required for lipid A biogenesis), YbgC (*tol-pal* cluster hydrolase of short-chain acyl-CoA thioesters), AcpS (involved in transfer of 4′- phosphopantethein to ACP), Aas and PlsB (membrane proteins involved in phospholipid acylation), and YiiD (putative acetyltransferase). ACP also co-purified with GlmU (an essential bi-functional enzyme that converts glucosamine-1-phosphate to UDP-GlcNAc (lipid A precursor)), AidB (isovaleryl-CoA dehydrogenase), SecA (pre-protein translocase), as well as MukB and SpoT, as previously reported[17]. We did not detect IscS, which was predicted to interact with ACP in an overexpression study[17], possibly because transient or low-affinity interactions are missed by our method.

Many other informative complexes were detected (see Supplementary Information). These included notable interactions mediated by the cysteine desulfurase IscS (IscS–FdhD and IscS–YhhP), between two uncharacterized essential proteins (YgjD–YeaZ), and by a sizeable group of uncharacterized proteins with factors involved in ribosome function, RNA processing, and/or RNA binding.

Graph network analysis of the validated data set provided evidence of 'scale-free' behaviour[18]. Most of the proteins had few interacting partners, whereas a subset of 'hubs' formed a far greater number of connections (Fig. 3b). Comparable connectivity was observed for the essential-conserved proteins alone (Supplementary Fig. 3). Scale-free networks are predicted to be robust against random node removal but vulnerable to hub removal, a property that might be expected to be preserved across evolution[18]. Indeed, removal of the 20 most highly connected nodes (>15 interactions; Supplementary Table 4) markedly reduced the network connectivity (see Supplementary Figs 4a and b). Notably, these same hubs were all highly conserved (detected in ≥125 genomes; Supplementary Fig. 4c). Moreover, protein connectivity was proportional (positively correlated) to the number of genomes a homologue was detected in (Supplementary Fig. 4d). Although a previous analysis of bacterial two-hybrid data[19] failed to detect such a dependency, possibly due to a high false-positive rate, our results are in agreement with a more in-depth analysis of the relationship of protein evolutionary rates to the number of interactions in eukaryotes[20].

To investigate further the conserved nature of the bacterial network, we analysed co-occurrence of BLAST homologues of

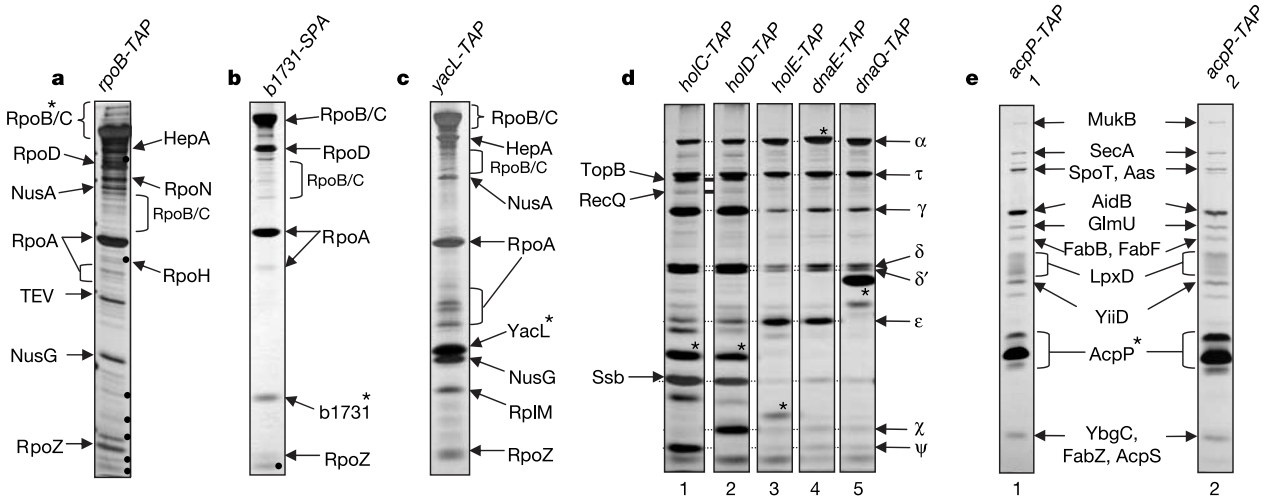


Figure 2 Analysis of affinity-purified protein complexes. SDS–PAGE silver-stain analysis of the components of affinity-purified complexes from *E. coli*. **a**–**c**, Purification of TAP-tagged *E. coli* RNAP subunit β (**a**) and two associated proteins: SPA-tagged b1731 (**b**) and TAP-tagged YacL (**c**). **d**, Purification of TAP-tagged HolD (ψ) and HolC (χ) subunits of the processivity clamp loader (lanes 1 and 2), and DNA polymerase III core subunits HolE (θ), DnaE (α) and DnaQ (ε) (lanes 3–5). **e**, Independent purifications of TAP-tagged AcpP (lanes 1 and 2). Only validated interacting proteins are labelled. Black circles indicate bands that failed to yield spectra or validated results; asterisks indicate tagged bait; and brackets indicate degradation products.

each pair of interacting proteins across all three domains of life (Archaea, Prokaryota and Eukaryota)[21] (see Methods). As seen in Supplementary Fig. 5, the interacting proteins were more likely to be co-conserved than control, randomly selected protein pairs, indicating that the interactions are similarly conserved. Notably, the most highly conserved proteins were highly connected, forming a single interconnected component (Fig. 3c, d). This core set of 154 interactions involving 71 proteins (including the ribosome;

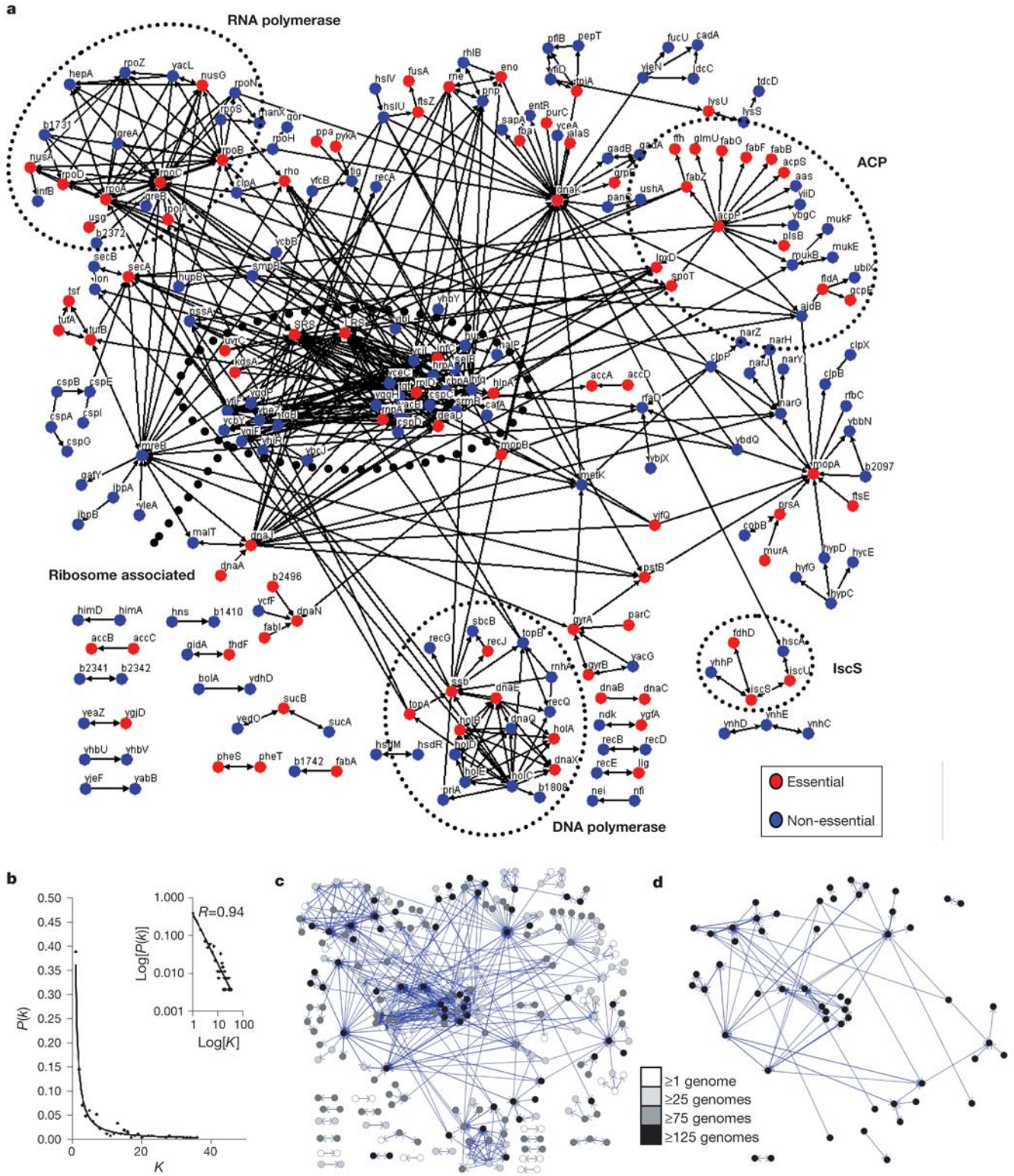


Figure 3 Network properties of bacterial protein–protein interactions. **a**, Network of validated protein complexes. Interactions are represented as directional edges extending from the tagged protein. Baits without partners are removed for clarity. Red nodes, essential proteins; blue nodes, non-essential proteins; black ovals, complexes discussed in text. **b**, Connectivity distribution of validated interactions (K) per protein plotted as a function of frequency, P(k). Inset: log-plot power law distribution, $P(k) \approx k^{-\gamma}$, where γ is the degree exponent. R, Pearson's correlation coefficient (see Methods). **c**, Network sequence conservation. Node shading (white-to-black) is scaled according to the number of genomes that pairs of interacting proteins co-occur in. **d**, Network of highly conserved proteins co-occurring in ≥125 genomes (homologue raw BLAST bit score ≥50).

Supplementary Table 5) potentially fulfils critical roles across all bacteria. Similar results were obtained using clusters of orthologous groups of proteins[22] (COGs; Supplementary Fig. 6a, b).

Co-occurrence of homologues across different genomes (phylogenetic profiles) has previously been used to explore functional links between genes[23,24]. Different approaches for constructing these profiles include the use of orthology assignments (for example, COGs) and sequence homology[25]. Recently, a study[26] introduced a mapping approach based on sequence homology methods to assess the degree of conservation of interologs between species. We adopted a similar approach to examine patterns of interolog conservation within the complexes detected in this study. Intriguingly, AcpP and several of its interacting partners displayed significant divergence in Bacilli, Actinobacteridae, *Mycoplasma* spp. as well as Archaea and eukaryotes (Fig. 4a). The lack of obvious homologues of AcpP in actinobacteria is consistent with the highly diverged nature of the predicted orthologues from this phylum in the COGs classification scheme. Evolutionary divergence was also evident with the DNA and RNA polymerase complexes. Although core RNAP subunits and cofactors *nusA*, *rpoA*, *rpoB*, *rpoC*, *rpoD*, *rpoH* and *rpoS* are found in virtually all eubacteria, other subunits (such as *hepA*, *rpoZ*, *b1731* and *yacL*) are restricted to γ-Proteobacteria (Supplementary Fig. 1a). Likewise, the DNA polymerase clamp loader subcomplex (*holA-D*) is similarly restricted (Supplementary Fig. 1b). These data suggest that sequence divergence may lead to functional diversification and the formation of novel modules. By clustering interacting proteins based on their phylogenetic profiles, it may be possible to identify new modules. For instance, the interacting PflB–PepT gene products cluster together (see Supplementary Fig. 2).

Statistical methods such as mutual information, Pearson correlation coefficient, Hamming distance (D) and the chance co-occurrence probability distribution (P) have been developed to predict functional relationships among genes based on phylogenetic profiles[23]. We applied the latter two metrics to quantify the extent of correlation between different phylogenetic profiles. Using relatively modest cutoff values ($D < 25$, $P < 10^{-11}$), a small but significant subset (~14%) of the interacting proteins showed closely correlated phylogenetic profiles relative to a control set of randomly selected pairs of bacterial proteins (Fig. 4b). Chromosomal proximity has also been used to infer functional linkages between evolutionarily conserved proteins[27]. Investigations into the distribution of relative

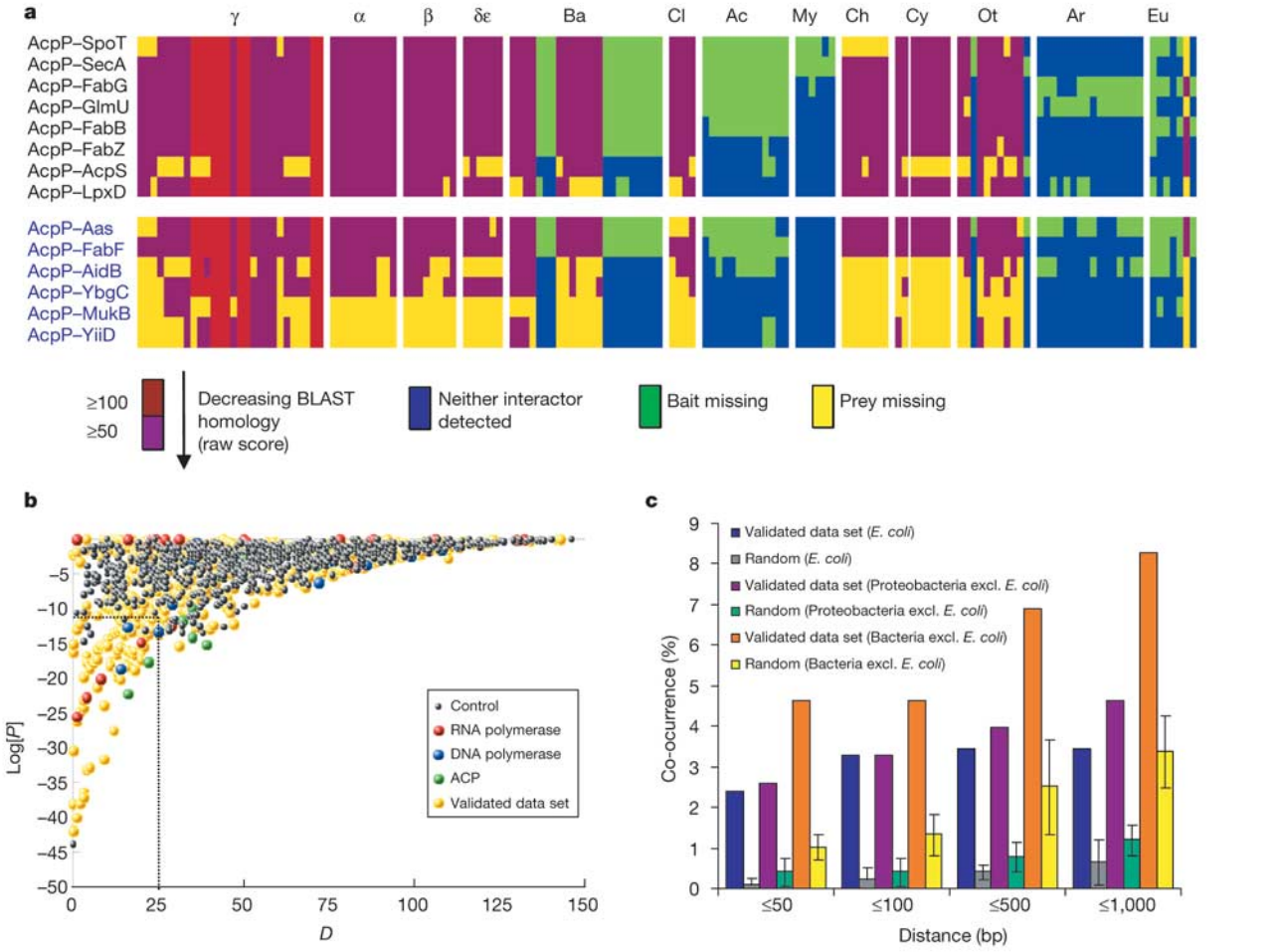


Figure 4 Bioinformatic analyses of interacting protein modules. **a**, Phylogenetic profile of ACP interactions (bait–partner) within 148 genomes. Coloured boxes indicate degree of BLAST homology. Black font, bait–partner both essential-conserved proteins; blue font, only bait is an essential-conserved protein. Phylogenies from the NCBI taxonomy database (see Supplementary Information). γ, α, β and δε indicate respective proteobaceria. Ba, Bacilli; Cl, Clostridia; Ac, Actinobacteridae; My, *Mycoplasma*; Ch, Chlamydiaceae; Cy, Cyanobacteria; Ot, unclassifed bacteria; Ar, Archaea; Eu, Eukaryota. **b**, Scatter log-plot of Hamming distance (D) versus chance co-occurrence probability distribution (P) for interactions associated with RNAP, DNA polymerase, ACP, the complete validated data set or random protein pairs (control). The dotted box indicates interactions (~14%) with significantly correlated profiles. **c**, Bar graph of relative genomic distances (base pairs) between genes encoding interacting proteins. Orange/yellow, orthologues in bacteria (excluding *E. coli*; see Supplementary Fig. 1); purple/green, orthologues in proteobacteria (excluding *E. coli*); blue/grey, *E. coli*. Error bars indicate standard deviation for 20 replicate controls.

chromosomal distances between loci encoding pairs of interacting proteins revealed that only a modest proportion (3.4%) of the interacting proteins were encoded by genes located within 500 base pairs (bp) of each other within the *E. coli* genome (Fig. 4c). A slightly larger fraction (6.9%) of putative orthologues was similarly separated by less than 500 bp in at least one other bacterial genome (Fig. 4c; see also Supplementary Table 6). Importantly, of the 42 protein pairs satisfying this criterion, 18 are encoded within the same operon in *E. coli* (see Supplementary Information), including 5 of the 15 interactions involving AcpP, further validating these data. These results indicate that only a fraction of our experimentally detected bacterial protein interactions could be readily predicted by genome-context methods (see above and Supplementary Table 3).

In summary, a reliable network of functionally diverse protein complexes was elucidated in *E. coli* using an experimental approach that can be readily adapted to other prokaryotes[6]. These data offer an insight into the function of uncharacterized proteins and outline the topological organization of a bacterial interactome. Because only about 30 bacterial proteins are currently targeted by prescription drugs[28], knowledge of physical interactions mediated by conserved, essential bacterial proteins should facilitate the design of broad-range antimicrobials. These data should also prove valuable for calibrating computational approaches designed to predict functional associations between proteins. Moreover, the tagged strain collection should facilitate biochemical studies using traditional or microarray-based assays. □

Methods

Construction of TAP/SPA-tagged *E. coli* strains

Escherichia coli strains bearing either TAP- or SPA-tagged alleles were constructed by targeted homologous recombination of DNA cassettes into the *E. coli* strain DY330 as previously reported[6]. Primer sequences are available upon request. The SPA tag was selected for use after initial trials due to the reduced protein degradation observed when compared with TAP-purified protein complexes.

Large-scale SPA/TAP purification of *E. coli* protein complexes

TAP or SPA purification was performed using 2–4 l log-phase cultures as previously described[6] except that Benzonase nuclease (Novagen; 3 U) was incubated with the cleared cell extract on ice for 30 min before purification. Purified complexes were split into aliquots for SDS–PAGE and liquid chromatography-tandem mass spectrometry analysis as detailed in the text.

Identification of proteins by mass spectrometry

Complex subunits were separated by SDS–PAGE on 12% acrylamide gels with a Whatman V16 vertical gel apparatus run at low current (9 mA) for 16 h. Gels were silver-stained using a standard protocol, except that formaldehyde crosslinking was not performed (details available upon request). Protein bands were excised and analysed as described previously[29]. Gel-free protein sequencing was performed by microcapillary-scale liquid chromatography–electrospray–ion trap tandem mass spectrometry as described[29]. Spectra were searched against an in-house database of predicted *E. coli* protein-coding sequences.

Selection of target gene products

ORFs were selected for study to obtain broad biological coverage, including highly conserved essential and non-essential proteins, proteins with putative functional assignments, and hypothetical uncharacterized ORFs. Proteins known or predicted to contain trans-membrane helices were avoided owing to technical difficulties associated with purifying membrane proteins. The essential-conserved set of genes was selected using a basic rule set (see Supplementary Information).

Connectivity distribution

To assess the correlation between connectivity (k) and frequency ($P(k)$), and between connectivity and the number of genomes a homologue was detected in, we calculated the Pearson's correlation coefficient R.

Phylogenetic analysis of protein complexes

For each *E. coli* sequence, a TBLASTN[30] search was performed against each of the different organism genome data sets. In addition, to avoid complications caused by intronic regions, a protein data set was obtained for each eukaryotic organism considered here and a BLASTP[30] performed. The raw score for the highest sequence match to each data set was extracted and stored in a local database. Phylogenetic interaction profiles were visualized using a java applet developed in house.

Phylogenetic distribution of proteins and interactions

The phylogenetic distribution of the proteins was obtained as previously described[21] using as reference the non-redundant protein sequence database SWALL (SwissProt plus TrEMBL; see Supplementary Methods). $P(h)$ represents the frequency of proteins with a homologue in the corresponding taxonomic groups. We used BLAST[30] with a threshold raw score of 50 and default parameters. We considered a protein interaction to be conserved if both interacting proteins have detectable homologues in any of the 148 complete genomes analysed (see legend to Supplementary Fig. 1). To analyse statistical significance, 30 control set samples of equal size were taken from the *E. coli* genome as previously described[21]. We used a two-tailed t-test, at a 95% confidence level.

Phylogenetic profiles

Each gene is represented by a vector representing the pattern of co-occurrence across 148 genomes (with a value of 1 assigned when a homologue is present, and 0 when one is not). The extent of correlation between phylogenetic profiles of pairs of interacting proteins was assessed by computing Hamming distance (D) and the chance co-occurrence probability distribution (P)[24]. If N is the total number of genomes over which we construct a phylogenetic profile for R genes, X and Y are the number of genomes in which homologues of two genes occur, and z is the number of genomes in which the genes co-occur, then $P = w_z \bar{w}_{\bar{z}} / W$, where w_z is the number of ways to distribute z co-occurrences over N genomes, $\bar{w}_{\bar{z}}$ is the number of ways to distribute the rest of the $X - z$ and $Y - z$ genes over the rest of the $N - z$ lineages, and W is the number of ways to distribute X and Y over N genomes without restriction, and $D = x + y - 2z$.

Genomic distance

Relative distances between genes encoding each pair of interacting proteins were calculated using the chromosome location coordinates of genes in COGs[22].

Received 6 October; accepted 3 December 2004; doi:10.1038/nature03239.

1. Neidhardt, F. (ed.) Escherichia coli *and* Salmonella*: Cellular and Molecular Biology* (ASM Press, Washington DC, 1996).
2. Serres, M. H. *et al.* A functional update of the *Escherichia coli* K-12 genome. *Genome Biol.* **2,** research0035.1–0035.7 (2001).
3. Gavin, A. C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415,** 141–147 (2002).
4. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* **97,** 6640–6645 (2000).
5. Yu, D. *et al.* An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **97,** 5978–5983 (2000).
6. Zeghouf, M. *et al.* Sequential Peptide Affinity (SPA) system for the identification of mammalian and bacterial protein complexes. *J. Proteome Res.* **3,** 463–468 (2004).
7. Stukenberg, P. T. & O'Donnell, M. Assembly of a chromosomal replication machine: two DNA polymerases, a clamp loader, and sliding clamps in one holoenzyme particle. V. Four different polymerase-clamp complexes on DNA. *J. Biol. Chem.* **270,** 13384–13391 (1995).
8. Harmon, F. G., Brockman, J. P. & Kowalczykowski, S. C. RecQ helicase stimulates both DNA catenation and changes in DNA topology by topoisomerase III. *J. Biol. Chem.* **278,** 42668–42678 (2003).
9. Witte, G., Urbanke, C. & Curth, U. DNA polymerase III chi subunit ties single-stranded DNA binding protein to the bacterial replication machinery. *Nucleic Acids Res.* **31,** 4434–4440 (2003).
10. Nakayama, H. RecQ family helicases: roles as tumor suppressor proteins. *Oncogene* **21,** 9008–9021 (2002).
11. von Mering, C. *et al.* Comparative assessment of large-scale data sets of protein–protein interactions. *Nature* **417,** 399–403 (2002).
12. Salwinski, L. *et al.* The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res.* **32** (Database issue), D449–D451 (2004).
13. Bader, G. D., Betel, D. & Hogue, C. W. BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res.* **31,** 248–250 (2003).
14. von Mering, C. *et al.* STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res.* **31,** 258–261 (2003).
15. Bowers, P. *et al.* Prolinks: a database of protein functional linkages derived from coevolution. *Genome Biol.* **5,** R35 (2004).
16. Rain, J. C. *et al.* The protein–protein interaction map of *Helicobacter pylori*. *Nature* **409,** 211–215 (2001).
17. Gully, D., Moinier, D., Loiseau, L. & Bouveret, E. New partners of acyl carrier protein detected in *Escherichia coli* by tandem affinity purification. *FEBS Lett.* **548,** 90–96 (2003).
18. Wuchty, S., Oltvai, Z. N. & Barabasi, A. L. Evolutionary conservation of motif constituents in the yeast protein interaction network. *Nature Genet.* **35,** 176–179 (2003).
19. Jordan, I. K., Wolf, Y. I. & Koonin, E. V. No simple dependence between protein evolution rate and the number of protein-protein interactions: only the most prolific interactors tend to evolve slowly. *BMC Evol. Biol.* **3,** 1 (2003).
20. Fraser, H. B., Wall, D. P. & Hirsh, A. E. A simple dependence between protein evolution rate and the number of protein-protein interactions. *BMC Evol. Biol.* **3,** 11 (2003).
21. Peregrin-Alvarez, J. M., Tsoka, S. & Ouzounis, C. A. The phylogenetic extent of metabolic enzymes and pathways. *Genome Res.* **13,** 422–427 (2003).
22. Tatusov, R. L., Koonin, E. V. & Lipman, D. J. A genomic perspective on protein families. *Science* **278,** 631–637 (1997).
23. Wu, J., Kasif, S. & DeLisi, C. Identification of functional links between genes using phylogenetic profiles. *Bioinformatics* **19,** 1524–1530 (2003).
24. Pellegrini, M., Marcotte, E. M., Thompson, M. J., Eisenberg, D. & Yeates, T. O. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc. Natl Acad. Sci. USA* **96,** 4285–4288 (1999).
25. Date, S. V. & Marcotte, E. M. Discovery of uncharacterized cellular systems by genome-wide analysis of functional linkages. *Nature Biotechnol.* **21,** 1055–1062 (2003).
26. Yu, H. *et al.* Annotation transfer between genomes: protein-protein interologs and protein-DNA regulogs. *Genome Res.* **14,** 1107–1118 (2004).
27. Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G. D. & Maltsev, N. The use of gene clusters to infer

functional coupling. *Proc. Natl Acad. Sci. USA* **96,** 2896–2901 (1999).
28. Haselbeck, R. *et al.* Comprehensive essential gene identification as a platform for novel anti-infective drug discovery. *Curr. Pharm. Des.* **8,** 1155–1172 (2002).
29. Krogan, N. J. *et al.* High-definition macromolecular composition of yeast RNA-processing complexes. *Mol. Cell* **13,** 225–239 (2004).
30. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402 (1997).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements The authors wish to thank C. J. Ingles and M. Shales for comments on the manuscript. This work was supported by funds from the Ontario Research and Development Challenge Fund and Genome Canada to A.E. and J.G. G.B. was a recipient of a Charles H. Best Post-Doctoral Fellowship. J.M.P.-A. acknowledges support from the Hospital for Sick Children (Toronto, Ontario, Canada) Research Training Centre. Computer analyses were undertaken at the Centre for Computational Biology, Hospital for Sick Children.

Authors' contributions Informatics studies were performed and analysed by J.M.P.-A. and J.P. Experimental design and data analysis were coordinated by G.B. Tagging and purification experiments were performed by W.Y., X.Y., J.L. and G.B. V.C., A.S., D.R., B.B., N.J.K. and M.D. performed and assisted with mass spectrometry analysis. The manuscript was jointly drafted by G.B., A.E., J.G., J.M.P.-A. and J.P. The project was conceived and designed by J.G. and was directed by A.E.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.E. (andrew.emili@utoronto.ca) or J.G. (jack.greenblatte@utoronto.ca).

Nanoarchaeum equitans creates functional tRNAs from separate genes for their 5′- and 3′-halves

Lennart Randau[1,2], Richard Münch[2], Michael J. Hohn[1,3], Dieter Jahn[2] & Dieter Söll[1,4]

[1]*Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, New Haven, Connecticut 06520-8114, USA*
[2]*Institut für Mikrobiologie, Technical University Braunschweig, P.O. Box 3329, D-38023 Braunschweig, Germany*
[3]*Lehrstuhl für Mikrobiologie und Archaeenzentrum, Universität Regensburg, Universitätsstrasse 31, D-93053 Regensburg, Germany*
[4]*Department of Chemistry, Yale University, 266 Whitney Avenue, New Haven, Connecticut 06520-8114, USA*

Analysis of the genome sequence of the small hyperthermophilic archaeal parasite *Nanoarchaeum equitans*[1,2] has not revealed genes encoding the glutamate, histidine, tryptophan and initiator methionine transfer RNA species. Here we develop a computational approach to genome analysis that searches for widely separated genes encoding tRNA halves that, on the basis of structural prediction, could form intact tRNA molecules. A search of the *N. equitans* genome reveals nine genes that encode tRNA halves; together they account for the missing tRNA genes. The tRNA sequences are split after the anticodon-adjacent position 37, the normal location of tRNA introns. The terminal sequences can be accommodated in an intervening sequence that includes a 12–14-nucleotide GC-rich RNA duplex between the end of the 5′ tRNA half and the beginning of the 3′ tRNA half. Reverse transcriptase polymerase chain reaction and aminoacylation experiments of *N. equitans* tRNA demonstrated maturation to full-size tRNA and acceptor activity of the $tRNA^{His}$ and $tRNA^{Glu}$ species predicted *in silico*. As the joining mechanism possibly involves tRNA *trans*-splicing, the presence of an intron might have been required for early tRNA synthesis.

The origin of the tRNA molecule is the subject of continuing discussions and has led to different models postulating that tRNA evolved by duplication or ligation of an RNA hairpin[3,4]. To examine these models further, the investigation of ancient tRNA genes was central. An interesting organism for this task was *Nanoarchaeum equitans*, currently the only characterized member of the kingdom Nanoarchaeota, which roots early in the archaeal lineage, before the emergence of Euryarchaeota and Crenarchaeota[5]. A significant fraction of the small number of *N. equitans* open reading frames consists of 'split genes' that are encoded as fused versions in other archaeal genomes. Our attention was caught by the 'absence' of four tRNA genes encoding the glutamate, histidine, tryptophan and initiator methionine acceptors[5].

We therefore developed a computational approach to search for tRNA signature sequences in the *N. equitans* genome. Our program, trained by an alignment of 4,000 tRNA gene sequences (taken from ref. 6), identifies sequences comprising the highly conserved T-loop region and defines the adjacent 3′-acceptor stem sequence. The reverse complementary sequence (defining the 5′-acceptor stem sequence) plus a D-stem position weight matrix identifies the corresponding 5′ half. The length of the position weight matrices can be adjusted and mismatches in the acceptor stem can be included. Finally, putative tRNA-halves are ligated *in silico* and analysed by COVE[7]. In addition to identifying the set of tRNAs predicted by the tRNAScan-SE program[8], our algorithm found nine tRNA halves spread throughout the chromosome. Surprisingly, these tRNA halves could be joined *in silico* to form the missing $tRNA^{His}$, $tRNA_i^{Met}$, $tRNA^{Trp}$ and two $tRNA^{Glu}$ species (Fig. 1). Further analysis of the tRNA half genes revealed several striking features. First, the location of the sequence separation that generated all nine tRNA half genes is after position 37, one nucleotide downstream of the anticodon and the common location of tRNA introns[9]. Second, a consensus sequence matching the highly conserved archaeal Box A promoter element[10] was found upstream of all 5′ tRNA halves. Third, this same consensus sequence (5′-TTTT/ATAAA-3′) was located 17–25 base pairs (bp) further upstream of the 3′ tRNA halves, resulting in a transcript with a 12–14-bp-long GC-rich leader sequence. Last, it is remarkable that this leading sequence is in all cases the exact reverse complement to a sequence following the corresponding 5′ tRNA half.

The existence of three $tRNA^{Glu}$ half genes was most exciting. Two 5′ tRNA halves were identified that differed solely by one anticodon base (isoacceptors with UUC and CUC anticodon), whereas only one 3′ $tRNA^{Glu}$ half gene was found. Both 5′ $tRNA^{Glu}$ half genes were followed by the identical 14-bp sequence that was the exact reverse complement of the single 3′ $tRNA^{Glu}$ half upstream sequence. All identified split tRNA genes contained the consensus bases of all archaeal elongator tRNAs[6], namely U8, A14, G15, G18, G19, C32, U33 and the T-loop GTTCA/GAATC (53–61), with the exception of the putative $tRNA^{Trp}$ harbouring an unusual GG sequence preceding the anticodon. The identified $tRNA_i^{Met}$ displays the consensus sequences of archaeal initiator tRNAs such as the anticodon stem/loop nucleotides (nt) 29–41 (GGGCU-CAUAACCC) and the R11:Y24 base pair (G11:C24), which is the reverse of the Y11:R24 base pair found in elongator tRNAs including the annotated *N. equitans* $tRNA^{Met}$. Therefore we define the split $tRNA_i^{Met}$ as the missing initiator tRNA. The sequences also reveal characteristic nucleotides in the respective tRNA species needed for recognition by the cognate aminoacyl-tRNA synthetase. For example, the $tRNA^{His}$ half genes encode the unique G-1:C73 base pair required for aminoacylation of $tRNA^{His}$ by histidyl-tRNA synthetase[11], and the $tRNA^{Glu}$ isoacceptors contain the characteristic D-loop nucleotides 20a and 20b and the deletion of base 47 essential for making the 'augmented D-helix'[12].

We performed reverse transcriptase polymerase chain reaction (RT–PCR) analysis of *N. equitans* total tRNA to verify the computationally predicted sequence of the newly discovered joined tRNAs. Our sequencing results confirmed the sequences for $tRNA^{Glu}$

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 6, Issue of February 11, pp. 4360–4366, 2005
Printed in U.S.A.

A Role for SlyD in the *Escherichia coli* Hydrogenase Biosynthetic Pathway*

Received for publication, October 18, 2004, and in revised form, November 23, 2004
Published, JBC Papers in Press, November 29, 2004, DOI 10.1074/jbc.M411799200

Jie Wei Zhang‡, Gareth Butland§, Jack F. Greenblatt§, Andrew Emili§, and Deborah B. Zamble‡¶

From the ‡Department of Chemistry, University of Toronto, Toronto, Ontario M5S 3H6, Canada and the §Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario M5G 1L6, Canada

The [NiFe] centers at the active sites of the *Escherichia coli* hydrogenase enzymes are assembled by a team of accessory proteins that includes the products of the *hyp* genes. To determine whether any other proteins are involved in this process, the sequential peptide affinity system was used. The analysis of the proteins in a complex with HypB revealed the peptidyl-prolyl cis/trans-isomerase SlyD, a metal-binding protein that has not been previously linked to the hydrogenase biosynthetic pathway. The association between HypB and SlyD was confirmed by chemical cross-linking of purified proteins. Deletion of the *slyD* gene resulted in a marked reduction of the hydrogenase activity in cell extracts prepared from anaerobic cultures, and an in-gel assay was used to demonstrate diminished activities of both hydrogenase 1 and 2. Western analysis revealed a decrease in the final proteolytic processing of the hydrogenase 3 HycE protein, indicating that the metal center was not assembled properly. These deficiencies were all rescued by growth in medium containing excess nickel, but zinc did not have any phenotypic effect. Experiments with radioactive nickel demonstrated that less nickel accumulated in Δ*slyD* cells compared with wild type, and overexpression of SlyD from an inducible promoter doubled the level of cellular nickel. These experiments demonstrate that SlyD has a role in the nickel insertion step of the hydrogenase maturation pathway, and the possible functions of SlyD are discussed.

The production of metalloenzymes frequently requires dedicated auxiliary proteins to assemble the functional metallocenters (1, 2). In the case of an enzyme with a single ion bound to unmodified protein ligands, maturation usually involves just one partner protein (1, 3). These factors, referred to as metallochaperones (3), deliver the correct metal ion to the target protein via protein-protein interactions (1). For the biosynthesis of more complex metallocenters, multiple accessory proteins are often required (2). These factors control a cascade of events that can include gathering and insertion of all of the inorganic and organic components, partial construction of the metal center, posttranslational modifications, electron transfer, protein folding, and/or hydrolysis of nucleotide triphosphates to drive the whole process forward (2). These molecular factories generate enzymes that are essential for a variety of fundamental cellular processes, but many of the protein components have not yet been identified or fully characterized.

The hydrogenase enzymes, which catalyze the reversible formation of dihydrogen (H_2) from two protons and two electrons, contain several different types of active sites (4, 5). In *Escherichia coli* the hydrogenases are all members of the [NiFe] class of enzymes that have nickel, iron, and three non-protein diatomic ligands in a deeply buried active site (6, 7). The outline of the general sequence of events during hydrogenase metallocenter assembly in *E. coli* has been largely derived from studies of the hydrogenase 3 large subunit (HycE), although the overall process is similar for the other isoenzymes (for recent reviews see Refs. 2 and 7–9). The main auxiliary proteins are encoded by the *hypA–F* genes (10–12), and the general purpose folding chaperones GroEL and GroES may also be required for optimal maturation (13). During the first steps of the pathway the iron and its diatomic ligands are prepared and inserted into the hydrogenase 3 precursor protein HycE by HypCDEF. Next, HypC remains associated with HycE while HypA and the GTPase HypB facilitate insertion of the nickel ion. The nickel serves as part of the recognition motif for the isoenzyme-specific protease HycI, and the processing of the [NiFe]-containing protein is completed following cleavage of a C-terminal fragment.

Many of the details of the hydrogenase biosynthetic pathway are not yet understood, and it is possible that not all of the individual components are known. For example, a nickel metallochaperone for *E. coli* hydrogenase 3 has not been identified. To search for other factors involved in this process we are using a proteomics approach to characterize native multiprotein complexes (14). These experiments led to the isolation of SlyD, an *E. coli* protein of unknown function that has not been previously associated with this pathway (see Fig. 1). SlyD is a member of the FK506-binding protein (FKBP)[1] family of peptidyl-prolyl isomerases (PPIases) that was originally cloned because genetic mutations provide resistance to lysis induced by the phage φX174 (*s*ensitivity to *ly*sis) (15). In addition to the PPIase domain at the N terminus, SlyD also has a 50-residue C-terminal domain that is rich in potential metal-binding residues (Fig. 1), containing 15 histidines, 6 cysteines, and 7 glutamates/aspartates (15, 16). Previous studies demonstrated that SlyD can bind a variety of different metals including up to three nickel ions (16, 17). The metal-binding domain of SlyD is not required for PPIase activity (17), although metal-ion binding inhibits the PPIase activity of the full-length protein (17)

* This work was supported in part by grants from the Natural Sciences and Engineering Research Council of Canada and the Canada Research Chairs Program (to D. B. Z.) and the Protein Engineering Network of Centers of Excellence and Genome Canada (to J. F. G. and A. E.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed. Tel./Fax: 416-978-3568; E-mail: dzamble@chem.utoronto.ca.

[1] The abbreviations used are: FKBP, FK506-binding protein; EDC, 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride); MS, mass spectrometry; PPIase, peptidyl-prolyl cis/trans-isomerase; SPA, sequential peptide affinity; TCEP, Tris(2-carboxyethyl)phosphine hydrochloride.

TABLE I
Bacterial strains

Strain	Genotype	Reference
DY330	W3110 ΔlacU169 gal490 lamda cI857 Δ(cro-bioA)	20
Δ*slyD*	*DY330* Δ*slyD*	This work
ST395	*DY330 slyD-SPA*	This work
ST713	*DY330 hypB-SPA*	This work
MC4100	F^- *araD139* Δ(*argF-lac*)U169 *ptsF25 deoC1 relA1 flbB5301 rpsL150*λ^-	21
HD709	MC4100 Δ*hycI*	22

and in practice it allows for purification by immobilized metal affinity chromatography (16).

In this report we show that SlyD interacts with the hydrogenase accessory protein HypB and that it has a role in hydrogenase metallocenter assembly. Knocking out the gene for SlyD results in reduced hydrogenase activity and the detection of unprocessed HycE protein, indicating that the maturation pathway is blocked. This deficiency can be complemented by the addition of nickel to the growth medium. Furthermore, the Δ*slyD* strain accumulates less nickel then the wild-type strain when grown under anaerobic conditions. How SlyD might participate in hydrogenase metallocenter assembly, including a possible role as a nickel source, is discussed.

MATERIALS AND METHODS

Materials—Restriction endonucleases were purchased from New England Biolabs. DNA oligonucleotides were purchased from Sigma. DNA ligase was purchased from MBI Fermentas. *Pfu* DNA polymerase was from Stratagene. All water was eluted from a Milli-Q water system (Millipore). The anti-HypB (18) and anti-HycE (19) rabbit polyclonal antibodies were a generous gift from Prof. A. Böck (University of Munich, Germany). Expression and purification of HypB will be described elsewhere.[2]

Bacterial Strains—The *slyD-SPA* (ST395) and *hypB-SPA* (ST713) strains (see Table I) were constructed from *E. coli DY330* (wild type) as described previously (14). The Δ*slyD* strain was prepared as described previously (20). The Δ*hycI* (HD709) and wild-type *MC4100* strains of *E. coli* (21, 22) were from the laboratory of Prof. A. Böck (University of Munich, Germany).

Sequential Peptide Affinity (SPA) Purification of Protein Complexes and Protein Identification—Cells were grown anaerobically for 16 h as described below. Extracts were prepared and SPA purification was performed as described (14). Tandem MS analysis of the samples was performed by microcapillary liquid chromatography electrospray tandem MS (liquid chromatography-electrospray ionization-MS/MS). Briefly, the protein mixture was resuspended in 100 mM NH_4HCO_3, 1 mM $CaCl_2$, pH 8.5, and digested by trypsin overnight at 37 °C with 2 μl of immobilized trypsin Poros beads (PerSeptive). The digested peptides were fractionated on a 7.5-cm (100 μm ID) reverse phase C18 capillary column attached inline to a ThermoFinnigan LCQ-Deca ion trap mass spectrometer. The entire digested sample was loaded and analyzed as described (23). All tandem mass spectra were searched using the SEQUEST computer algorithm against an *E. coli* protein sequence data base compiled in-house. Each spectral match was validated by using the STATQUEST probability algorithm and was significant at a 99% confidence level (24).

Growth Conditions and Preparation of Crude Cell Extracts for Biochemical Assays—Cells were grown anaerobically following inoculation with 1% (v/v) overnight culture at 37 °C in sealed flasks of buffered TGYEP medium containing 10 g of tryptone, 5 g of yeast extract, 69 mM K_2HPO_4, and 22 mM KH_2PO_4/liter (25) for the indicated times. The medium was supplemented with 1 μM sodium molybdate, 1 μM sodium selenite, 30 mM sodium formate, and 0.5% glucose, except for in the analysis of hydrogenase 1 and 2 or induction of the pBAD vectors, in which case the glucose and formate were replaced with 0.8% glycerol and 15 mM sodium fumarate (26). $NiSO_4$ or $ZnCl_2$ salts were added to the medium at the indicated concentrations. When required, the antibiotics kanamycin (50 μg/ml) or chloramphenicol (34 μg/ml) were included in the growth medium. Cells were harvested by centrifugation and washed with 50 mM potassium phosphate buffer, pH 7.0, and resuspended in the same buffer containing 1 mM dithiothreitol, 0.5 mM benzamidine, 0.2 mM phenylmethylsulfonyl fluoride, and trace amount of DNase I. Crude cell extracts were prepared by sonication and subsequent centrifugation at 15,000 × *g* for 20 min at 4 °C. The supernatant was quickly frozen in $N_{2(l)}$ and stored at −80 °C. The protein concentrations of crude cell extracts were determined by the BCA protein assay system (Pierce). The addition of extra nickel to the extracts did not affect any of the activities monitored (data not shown).

Plasmids—To clone into the pET24b plasmid (Novagen), the *slyD* gene was amplified by PCR from DH5α *E. coli* by using the forward 5′-CAGGAGATCATATGAAAGTAGCAAAAGACCTGGTG-3′ and reverse 5′-GTCACTTCTCGAGTATTAGTGGCAACCGCAAC-3′ primers. The PCR product was purified by using the QIAquick PCR purification kit (Qiagen) and digested with NdeI and XhoI. Following elution from an agarose gel with the QIAquick kit, the fragment was ligated with pET24b digested with the same enzymes and then transformed into XL-2 Blue *E. coli* (Stratagene). Plasmid DNA was isolated from cultures grown in kanamycin-supplemented LB by using a Qiagen HiSpeed Plasmid Miniprep kit. The pET24-SlyD construct was verified by dideoxynucleotide sequencing (ACGT, Toronto). The same procedure was used to clone *slyD* into the pBAD24 vector (American Type Culture Collection) to make pBAD24-SlyD except that the forward 5′-CTATTCTTCGCTAGCTCAGGAGATATCATGAAAG-3′ and reverse 5′-GTCACTTTCTAGATATTAGTGGCAACCGCAAC-3′ primers were used, the PCR product and plasmid DNA were cut with NheI and XbaI, and ampicillin was used in the medium.

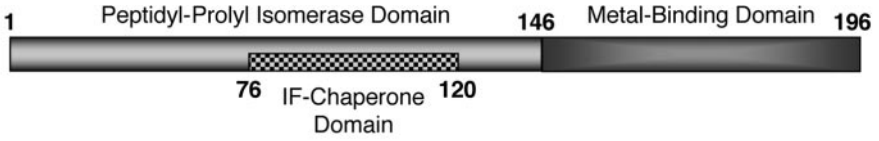
Purification of SlyD—BL21(DE3) *E. coli* transformed with pET24-SlyD were grown in kanamycin-containing LB to an A_{600} of 0.7, induced by the addition of 0.35 mM isopropyl 1-thio-β-D-galactopyranoside, and incubated for an additional 3 h at 37 °C. The bacteria were harvested by centrifugation (30 min at 4600 rpm), and the cell pellet was resuspended in 25 ml of lysis buffer/liter of culture (50 mM Tris, pH 8, 0.1 mM phenylmethylsulfonyl fluoride) and sonicated on ice. All subsequent steps were performed at 4 °C. The lysate was centrifuged for 30 min at 18,000 rpm in an SS34 rotor, and the supernatant was loaded onto a Ni^{2+}-nitrilotriacetic acid-agarose column (Qiagen) pre-equilibrated in lysis buffer containing 100 mM NaCl. The column was washed with 10 column volumes of lysis buffer containing 500 mM NaCl, the proteins were eluted in lysis buffer containing 100 mM imidazole, pH 7.4, and the eluate was dialyzed against 20 mM Tris, pH 8, 1 mM EDTA. The solution was loaded onto a Mono-Q HR 5/5 column (Amersham Biosciences) equilibrated with 20 mM Tris, pH 7.5, and eluted with linear gradient of 200–400 mM NaCl over 75 ml. The fractions containing SlyD were identified by SDS-PAGE, pooled, concentrated in a Amicon Ultra centrifugal filter, and loaded onto a Superdex 75 column (Amersham Biosciences) equilibrated in 10 mM HEPES, 100 mM KCl, pH 7.6. The fractions containing pure SlyD were identified by SDS-PAGE, pooled, treated with 100 mM TCEP and stored in the anaerobic glove box at 4 °C.

Chemical Cross-linking—SlyD and/or HypB were incubated in 100 mM potassium phosphate, pH 6.5, for 30 min at room temperature. A final concentration of 12.5 mM 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride) (EDC, Pierce) was added to the reactions, followed by a 1-h incubation at room temperature and analysis on 12.5% SDS-polyacrylamide gels that were stained with Coomassie dye.

Western Analysis—Cell extracts were resolved on SDS-polyacrylamide gels and transferred onto nitrocellulose or polyvinylidene difluoride (HycE Western) membranes. The blots were probed with the appropriate primary antibody at a 1:3000 dilution (mouse monoclonal anti-FLAG antibody, Sigma) or a 1:1000 dilution (anti-HycE, anti-HypB). The 2° antibody used, diluted 1:30,000, was either the goat anti-rabbit or goat anti-mouse horseradish peroxidase conjugate (Bio-Rad). The enhanced chemiluminescence technique (Pierce) was used for detection.

Hydrogenase Activity Assays—Total hydrogenase activity of crude cell extracts was measured by hydrogen-dependent reduction of benzyl viologen according to the procedure of Ballantine and Boxer (27). Re-

[2] M. R. Leach, S. Sandal, H. Sun, and D. B. Zamble, manuscript in preparation.



MKVAKDLVVS LAYQVRTEDG VLVDESPVSA PLDYLHGHGS LISGLETALE GHEVGDKFDV
AVGANDAYGQ YDENLVQRVP KDVFMGVDEL QVGMRFLAET DQGPVPVEIT AVEDDHVVVD
GNHMLAGQNL KFNVEVVAIR EATEEELAHG HVHGAHDHHH DHDHDGCCGG HGHDHGHEHG
GEGCCGGKGN GGCGCH

FIG. 1. **Primary sequence and putative domain organization of *E. coli* SlyD (15–17, 61).**

actions were prepared in a septum-sealed cuvette in an anaerobic glove box (95% N_2 and 5% H_2). One unit of activity is defined as 1 μmol of benzyl viologen reduced/min and an extinction coefficient of 7400 M^{-1} cm^{-1} was used (27). To measure the activities of hydrogenases 1 and 2 an in-gel assay was performed (28). Crude cell extracts were resolved on 10 or 12.5% polyacrylamide gels run at 4 °C for 4 h at 200 V. All buffers and gels contained 0.1% SDS. The gels were incubated in 100 mM potassium phosphate buffer, pH 7, containing 0.5 mM benzyl viologen (Sigma) and 1 mM triphenyltetrazolium chloride (Sigma) in the anaerobic glove box at room temperature for 16 h. The data were analyzed on a Fluorochem 8800 gel documentation system (Alpha Innotech).

Cellular Nickel Accumulation—The indicated concentrations of ^{63}Ni (specific activity 420 mCi/mmol; PerkinElmer life Sciences) were added to cells grown anaerobically. BL21(DE3) transformed with pET24-SlyD or pET24b were grown in LB, and *DY330* and Δ*slyD* cells were grown in TGYEP supplemented with glucose and formate. The cells were harvested and extensively washed with 50 mM potassium phosphate, pH 7.0, then crude extracts were prepared by sonication in the same buffer and subsequent centrifugation at 15,000 × *g* for 20 min. The radioactivity was measured by scintillation counting in 5 ml of UltimaGold scintillation fluid with a Liquid Scintillation Analyzer Tri-Carb 2100TR (Packard Instruments).

RESULTS

SlyD Directly Interacts with HypB—The SPA system is a method to identify the components of multiprotein complexes in bacterial cells under native conditions (14). The protein of interest is genomically tailored with a tag that encodes three modified FLAG sequences as well as a calmodulin-binding peptide. The target protein and any associated proteins are then isolated by two affinity chromatography steps. The identity of each protein is determined either by tryptic digestion of the whole eluent followed by liquid chromatography-electrospray ionization-MS/MS or resolution on SDS-PAGE followed by matrix-assisted laser desorption ionization MS/MS of individual gel bands. We are using this procedure to map out protein-protein interactions between the Hyp factors and other *E. coli* proteins[3] in bacteria grown under conditions that favor hydrogenase expression (11, 25). Upon SPA purification of tagged HypB from ST713, both the HypB protein and the histidine-rich protein called SlyD (Fig. 1) were identified (data not shown). Furthermore, inspection of data collected for another study revealed that HypB was identified in the purified SlyD-SPA complex isolated from strain ST395 (63). SlyD was not detected in a complex with any of the other accessory proteins investigated to date, including HypA, HypC, HypD, or HybG.[3]

To confirm the HypB and SlyD protein-protein interaction, chemical cross-linking of purified recombinant proteins was performed (Fig. 2). Upon incubation of SlyD with EDC, a slightly faster mobility band was observed, possibly because of an internal cross-link. In the cross-linking reaction containing both HypB and SlyD, a new band was observed at the mobility of a 1:1 HypB-SlyD complex. The presence of HypB in this covalent complex was confirmed by Western analysis with an anti-HypB antibody (18) (data not shown). A faint band at a larger molecular weight than the heterodimer was also observed in the HypB reactions both with or without SlyD (Fig. 2, *lanes 5* and *6*), and it is most likely a HypB homodimer (18).

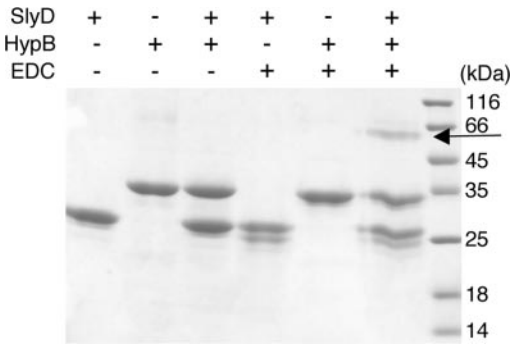


FIG. 2. **SlyD associates with HypB.** Purified HypB (5 μM) and/or SlyD (10 μM) were incubated in the presence or absence of EDC. The reactions were resolved on a 12.5% SDS-polyacrylamide gel, and the proteins were stained with Coomassie dye. The HypB-SlyD cross-link is indicated by the *arrow*.

Expression of slyD—It is not known what factors control the expression of *slyD*, although complementation experiments indicated uninduced constitutive expression driven by a cryptic promoter (15), and homology to the promoters that utilize σ^{70} was noted (16). The hydrogenases are regulated by various metabolites, but expression of all of the hydrogenase isoenzymes is repressed by oxygen (29, 30), and the *hyp* operon is regulated by the anaerobic transcription factor FNR (31). Furthermore, nickel regulates the expression of the nickel uptake transporter in *E. coli* and various nickel-containing enzymes in other organisms (Ref. 7 and references therein). To determine whether these growth conditions also affect SlyD production, Western analysis was performed on extracts from strain ST395 with an anti-FLAG antibody. The amount of SlyD-SPA detected was not affected by excess nickel (up to 300 μM) in the growth medium whether cells were grown open to the atmosphere instead of in sealed anaerobic flasks or whether LB medium was used instead of TGYEP containing formate (data not shown).

SlyD Contributes to Hydrogenase Activity—Many auxiliary proteins for metallocenter assembly in enzymes were originally identified, because genetic mutants are deficient in the enzyme activity (9). To test whether SlyD plays a role in hydrogenase maturation, the *slyD* gene was disrupted (20). This deletion resulted in a noticeably slower growth rate under our anaerobic conditions (data not shown) to a similar degree as that previously reported for Δ*slyD* strains grown aerobically (32), and this difference was not investigated further. Wild-type and Δ*slyD* cells were grown under anaerobic conditions for 6 or 17 h, and hydrogenase activity was measured in cell extracts by monitoring anaerobic benzyl viologen reduction in a solution containing hydrogen gas (27). The activity detected in the wild-type extracts prepared after 6 h of growth was similar to that previously reported (29, 33), and in either strain relatively less hydrogenase activity was routinely detected after the longer growth period (Fig. 3 and data not shown). At both time points, however, hydrogenase activity was significantly abrogated by

[3] G. Butland, J. W. Zhang, J. F. Greenblatt, A. Emili, and D. B. Zamble, unpublished data.

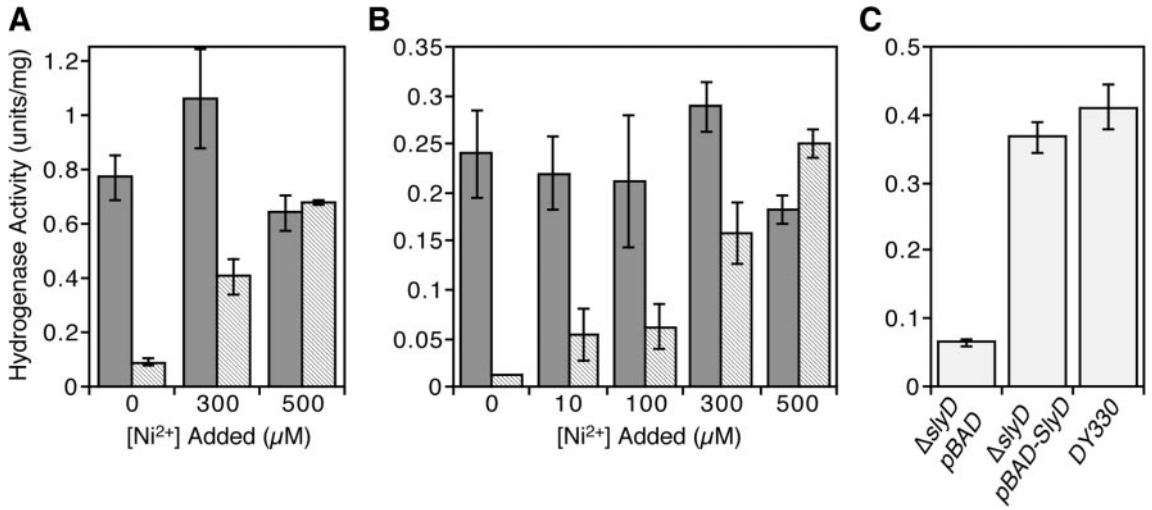


FIG. 3. **Δ*slyD* cells are deficient in soluble hydrogenase activity and complemented with nickel.** *A*, cultures were anaerobically grown in TGYEP medium containing glucose, formate, and the indicated concentrations of added nickel for 6 h. Cell extracts were prepared from wild-type strain *DY330* (*solid bars*) and the Δ*slyD* deletion strain (*striped bars*) and tested for hydrogenase activity with the anaerobic benzyl viologen solution assay. *B*, same as *A* except the cells were grown for 17 h. *C*, cell extracts were prepared from wild-type *DY330*, Δ*slyD* transformed with the pBAD vector, or Δ*slyD* transformed with pBAD-SlyD vector, and tested for hydrogenase activity with the anaerobic benzyl viologen solution assay. Cultures were anaerobically grown in TGYEP media supplemented with glycerol and fumarate as well as 100 μM arabinose. The results represent the average of two independent measurements, and *error bars* indicate ± 1 S.D.

the *slyD* deletion (Fig. 3, *A* and *B*). To confirm that this phenotype was a result of the *slyD* deletion and not a downstream polar effect, *slyD* was expressed in-trans from a pBAD24 vector (34) transformed into the Δ*slyD* strain. The P_{BAD} promoter is induced with arabinose and inhibited by glucose (34), so for these experiments cells were grown in TGYEP supplemented with glycerol and fumarate instead of glucose and formate. Arabinose titration of the Δ*slyD* strain transformed with pBAD-SlyD produced increasing amounts of hydrogenase activity (data not shown), with wild-type levels restored upon exposure to 100 μM arabinose (Fig. 3*C*). An increase in hydrogenase activity was not observed if the cells were transformed with the empty pBAD24 plasmid (Fig. 3*C*).

The decrease of hydrogenase activity in the Δ*slyD* strain is a similar phenotype, although not as extreme, as that observed following deletion of the individual *hyp* genes (11). However, the spectrophotometric hydrogenase assay of crude cell extracts is not specific for any particular hydrogenase isoenzyme (35), so it does not necessarily reflect whether a protein is required for the maturation of all three expressed hydrogenases. To investigate whether SlyD affects the activity of hydrogenases 1 and 2, fumarate and glycerol were added to the growth media (26), cell extracts were resolved on a polyacrylamide gel, and the activity of these two enzymes were analyzed with an in-gel benzyl viologen assay (28). In this strain of *E. coli* (*DY330*) the activity of hydrogenase 2 was much stronger than hydrogenase 1, so 50 μg of cell extracts were loaded to quantitate the activity for hydrogenase 1, and only 2 μg of cell extracts were used for the analysis of hydrogenase 2 (Fig. 4). The disruption of *slyD* caused a reduction in the activity of both hydrogenase 1 and 2 (Fig. 4), and the activity was restored by induction of *slyD* from the pBAD-SlyD vector (data not shown).

Complementation of Δ*slyD with Nickel*—HypB is implicated in the nickel insertion step of hydrogenase biosynthesis, because the hydrogenase deficiency of *hypB* mutants is complemented by excess nickel in the growth medium (10, 11, 36). Of the other Hyp factors, only *hypA* lesions are partially complemented with nickel (26). The fact that SlyD can bind nickel ions (17)[4] and that it interacts directly with HypB (Fig. 2) suggested

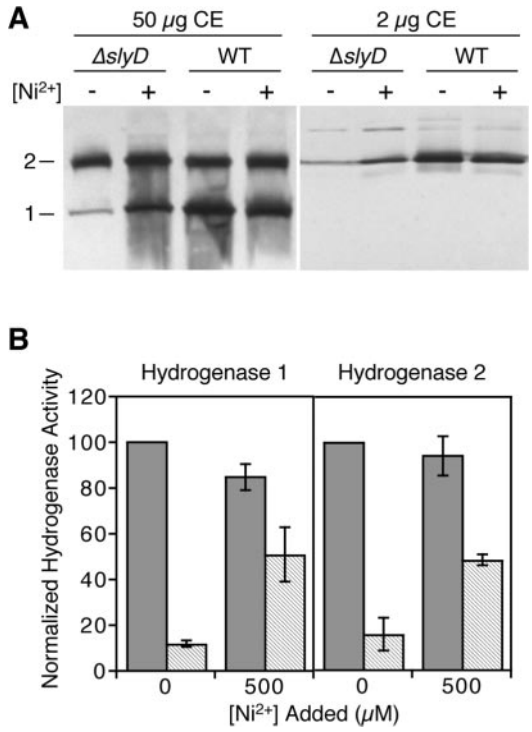


FIG. 4. **Hydrogenase 1 and 2 deficiency in the Δ*slyD* mutant is complemented with nickel.** *A*, activity staining of hydrogenase 1 and 2 in extracts prepared from wild-type (*WT*) or Δ*slyD* cells. Either 50 or 2 μg of crude cell extracts (*CE*) were analyzed from cells grown anaerobically with or without 500 μM nickel added to the TGYEP media containing glycerol and fumarate. *B*, quantitation of two experiments such as that shown in *A*. The activities detected for each enzyme in cell extracts from wild-type (*solid bars*) and Δ*slyD* (*striped bars*) cells were normalized in each experiment to that of the wild-type strain grown in the absence of added nickel. *Error bars* indicate ± 1 S.D.

that SlyD may also be involved at this step of the pathway. To test this hypothesis, $NiSO_4$ was added to the media of Δ*slyD* and wild-type cells, and hydrogenase activity was analyzed in cell extracts by using the solution benzyl viologen assay (Fig. 3). The concentrations of nickel used did not have a significant affect on the activity detected in the wild-type extracts. The

[4] J. W. Zhang and D. B. Zamble, unpublished data.

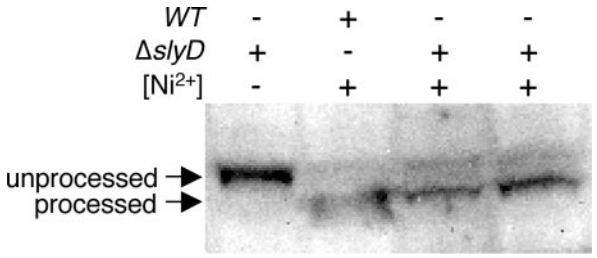


FIG. 5. **C-terminal processing of HycE.** The Δ*slyD* mutant strain was grown anaerobically in TGYEP media containing glucose and formate either in the absence or presence of 500 μM added $NiSO_4$. Crude cell extracts were resolved on a 10% SDS-polyacrylamide gel followed by Western analysis with the anti-HycE antibody. The wild-type (*WT*) *DY330* extract used in the experiment shown was prepared from cells grown with 500 μM nickel in the media, but nickel did not affect the mobility of HycE in the wild-type strain.

addition of 1 μM nickel did not have a noticeable affect on the hydrogenase activity of the Δ*slyD* extracts (data not shown), but an increase in activity was observed at higher concentrations of metal. Wild-type levels of activity were restored to the Δ*slyD* cells grown for either 6 or 17 h in the presence of 500 μM nickel (Fig. 3). Similarly, excess nickel partially complemented the hydrogenase 1 and 2 deficiency of the Δ*slyD* strain (Fig. 4). This effect is specific for nickel, because the addition of the same amounts of zinc to the growth medium did not complement the Δ*slyD* deficiency (data not shown).

In one of the final steps of hydrogenase 3 maturation, the protease HycI cleaves 32 residues from the C terminus of the HycE precursor protein (37), and the processed and unprocessed forms of HycE can be resolved by Western analysis (37). Nickel insertion is a prerequisite for this proteolysis step so processing the C-terminal tail is diagnostic of whether the metal center has been correctly incorporated (37). To determine whether SlyD affects the processing of hydrogenase 3, the HycE protein was examined in both the wild-type and Δ*slyD* cells. Cell extracts from a Δ*hycI* strain of *E. coli* and the corresponding wild-type strain were also examined to verify the mobility of the processed and unprocessed HycE proteins (data not shown) (21, 22). In the absence of SlyD, the HycE protein migrates at a slightly larger molecular weight (Fig. 5). This observation that the C-terminal fragment of the protein has not been processed in the Δ*slyD* strain indicates incomplete metallocenter assembly. Growth in the presence of excess nickel complements this deficiency and restores processing (Fig. 5) in agreement with the results from the hydrogenase activity assays. The mobility of the HycE protein in the wild-type strain was not affected by nickel in the medium (data not shown). No significant difference in the total amount of HycE protein was detected in the SlyD mutant extracts when compared with wild-type extracts.

SlyD Increases Nickel Accumulation—Finally, experiments with the radioactive isotope ^{63}Ni were performed to determine whether SlyD influences the nickel content of *E. coli* cells grown under anaerobic conditions. Cells were exposed to radioactive nickel at the start of anaerobic growth and at each time point aliquots were removed, washed, and the radioactivity was measured in either whole cells (data not shown) or in crude cell extracts (Fig. 6*A*). Most of the observed uptake occurred in the first 2 h. The time course of nickel uptake in the Δ*slyD* cells was similar to wild type, but the final level of nickel accumulation was only about half that detected in the wild-type cells. Similarly, the nickel content was monitored in anaerobically grown BL21(DE3) cells transformed with pET24-SlyD. Induction of SlyD with isopropyl 1-thio-β-D-galactopyranoside resulted in overexpression of the protein, visibly detectable in cell extracts on a Coomassie-stained SDS-polyacrylamide gel (data not shown), and an increase in nickel accumulation compared with cells transformed with the control pET24 vector (Fig. 6*B*).

DISCUSSION

Three [NiFe] hydrogenase enzymes are expressed in *E. coli* as components of various types of anaerobic metabolism (35). There is a fourth encoded isoenzyme (38), but this operon is silent in wild-type cells (30). SlyD is a protein that was not previously linked with these systems, but the experiments described in this report demonstrate that it is required for optimal activity of all three hydrogenases. The only aspect of production that is common to all isoenzymes is metallocenter biosynthesis by the accessory proteins, a subset of which act pleiotropically (7–9). Thus the influence of SlyD on the three isoenzymes suggests that the protein is involved in metallocenter assembly. Furthermore, complementation of Δ*slyD* by growth in excess nickel indicates that this protein has a role in nickel insertion and does not influence the preceding steps of diatomic ligand preparation, iron incorporation, holding the protein in a competent state for Ni^{2+} insertion, or the subsequent proteolysis of the C-terminal tail and internalization of the complete metal center. Exactly how SlyD participates in nickel insertion is not clear. It is possible that SlyD affects the nickel uptake or export mechanisms of *E. coli* and thus indirectly influences hydrogenase production, but the specific interaction with HypB suggests that SlyD is more directly involved.

As with many members of the PPIase superfamily (39, 40), a cellular function has not been identified for SlyD. It has been suggested that PPIases are important for protein folding and stability, for the formation of multiprotein complexes, or as switch mechanisms that regulate the activity of the substrate proteins and contribute to cell signaling (39–41). Furthermore, the actual PPIase activity of these factors is not always required (39, 41). None of the other known members of the FKBP family has a metal-binding domain similar to the C-terminal sequence of SlyD, but several have been implicated in metal homeostasis pathways. For example, a recent study demonstrated that FKBP52 interacts with the copper metallochaperone Atox1 and has a role in copper efflux (42). Mouse FKBP23 contains two Ca(II)-binding EF-hand motifs, and the ER-localized protein may act as a Ca(II)-dependent chaperone (43). Furthermore, mutant mice deficient in another homologue, FKBP12, had symptoms that mimic human congenital heart disorder, and it was suggested that the phenotype was because of an effect on the calcium release activity of muscle receptors (44).

The only *E. coli* hydrogenase accessory protein that has been shown to bind nickel is HybF (28), a protein that functions in the maturation of hydrogenases 1 and 2 and is replaced by HypA in the hydrogenase 3 pathway (26). The *Helicobacter pylori* HypA protein binds stoichiometric nickel and forms a heterodimer with HypB (45), but this activity has not yet been demonstrated for the *E. coli* protein. The *E. coli* HypA and HypB are implicated in the nickel insertion step because, like SlyD, mutations produce a hydrogenase-deficient phenotype that is complemented by growth in high nickel concentrations (10, 11, 26), and mutants that decrease the GTPase activity of HypB inhibit nickel incorporation into HycE (36). HypB homologues from many organisms have histidine-rich regions and *in vitro* studies of HypB proteins from *Bradyrhizobium japonicum* (46) and *Rhizobium leguminosarum* (47) established that they bind multiple nickel ions. Further examination of the *B. japonicum* protein established that the polyhistidine sequence is involved in nickel storage (48), but a mutant lacking the His-rich region could still bind a single nickel ion and support hydrogenase production (48, 49). Similar polyhistidine stretches are observed in accessory proteins for other nickel enzymes, such

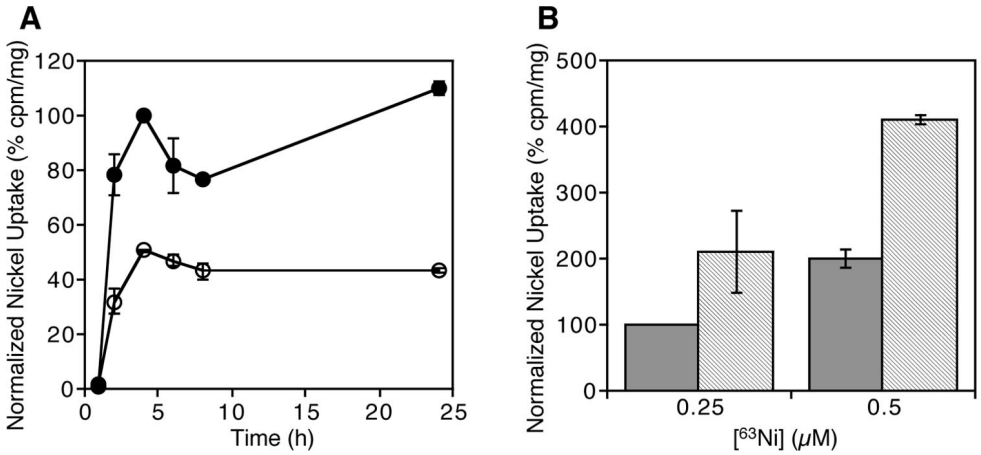


FIG. 6. **SlyD influences cellular nickel accumulation.** *A*, *DY330* (*filled circles*) and Δ*slyD* (*open circles*) cells were grown under anaerobic conditions in TGYEP media in the presence of 0.25 μM ^{63}Ni. At the indicated time points crude cell extracts were prepared, and nickel accumulation was measured by scintillation counting. Data from individual experiments were normalized to the wild-type sample taken at 4 h. *B*, BL21(DE3) transformed with the pET24b (*solid bars*) or pET24b-SlyD plasmid (*striped bars*) were grown anaerobically in LB supplemented with 0.1 mM isopropyl 1-thio-β-D-galactopyranoside and the indicated concentrations of ^{63}Ni. After 20 h, crude cell extracts were prepared, and nickel accumulation was measured by scintillation counting. Data were normalized to the pET24b, 0.25 μM ^{63}Ni control. Values represent the average of two separate experiments, and *error bars* represent ± 1 S.D.

as UreE and CooJ from the urease and carbon monoxide dehydrogenase biosynthetic pathways, respectively (7). Again the His-rich regions appear to be for nickel storage and can be separated from the nickel insertion activity (7). The fact that neither HypB nor any of the other *E. coli* hydrogenase proteins contain a similar region led to the suggestion that an alternative protein could fulfill the role of a nickel supplier for hydrogenase biosynthesis, and such a function was proposed for SlyD (17).

Thus one potential role for SlyD is that of a nickel source for the hydrogenase metallocenter assembly pathway. Previous studies have clearly shown that SlyD can bind multiple nickel ions (16, 17), and the inherent metal-binding activity is sufficiently strong to cause SlyD to be a common contaminant on immobilized metal affinity chromatography of *E. coli* extracts (16, 50, 51). Furthermore, an estimated 10^4 molecules of SlyD are expressed in normally growing cells (32), the same order of magnitude as the 25,000 molecules/cell of HypB synthesized under anaerobic conditions (18). A nickel source for the hydrogenase enzymes would mean that production does not rely on the availability of freely diffusible metal ions and that cellular exposure to potentially toxic ions is minimized. A recent study of the *E. coli* transcription factor NikR, which regulates expression of the nickel membrane transporter (52, 53), demonstrated that it responds to nickel concentrations corresponding to only a few free ions (54), suggesting that it would not allow uncomplexed metal to accumulate in the cell. It will be interesting to examine whether SlyD plays a role in production of Glyoxylase I, the only other known nickel enzyme in *E. coli* (55). It is also possible that SlyD is not specific for nickel but plays a more general role in cellular metal homeostasis, in analogy with the eukaryotic and cyanobacterial metallothioneins (56–58).

The lack of SlyD results in diminished but not negligible total cellular nickel, which could indicate a reduction in the level of available metal. However, this result could also be due to less nickel incorporation into the hydrogenase enzymes. If nickel insertion is deficient when SlyD is not expressed, the cellular demand for nickel will be met and the uptake will be turned off at much lower levels of accumulation. Thus it is possible that another property of SlyD participates in hydrogenase biosynthesis to promote optimal activity of HypB and facilitate nickel insertion, working either in cooperation with or instead of the metal binding activity. This hypothesis is supported by the observation that the levels of hydrogenase activity in the Δ*slyD* strain can be restored to wild-type levels when grown in the presence of excess nickel, whereas the Δ*hypB* mutation is only partially complemented (36). The fact that metal binding inhibits the PPIase activity of SlyD (17) makes it tempting to speculate that the two functions are coupled together in some type of switch mechanism, but additional experiments are required to test this hypothesis. SlyD sensitizes *E. coli* to the bacteriophage φX174 lysis protein E by stabilizing the protein (59), possibly through the isomerization of an essential proline residue (60), so perhaps SlyD stabilizes HypB or one of the other proteins in the hydrogenase biosynthetic pathway. It is interesting to note that SlyD belongs to a subfamily of FKBPs with an extra domain (Fig. 1), named an IF domain (61). This domain is required for a chaperone-like folding activity in a homologous FKBP from *Methanococcus thermolithotrophicus* (62), an activity that is independent of the PPIase activity (62). Finally, some FKBPs modulate signal transduction pathways as architectural factors that bind and correctly orient interacting components of multiprotein complexes (40). In a similar fashion, SlyD may interact with the hydrogenase accessory proteins and be responsible for transiently anchoring and orienting them during metal transfer.

In summary, this study reveals that SlyD is a new component of the hydrogenase metallocenter assembly pathway in *E. coli*. This protein interacts specifically with HypB and influences the nickel insertion step to promote optimal hydrogenase production. At this time, the function of this metal-binding protein is not clear; however there are four putative non-exclusive properties of SlyD that could act in hydrogenase metallocenter assembly, metal binding, PPIase, folding chaperone, and architectural assembly. Experiments are now underway to characterize the role of SlyD in the maturation pathway. SlyD homologues are encoded in a variety of prokaryotic genomes including *H. pylori*, *Haemophilus influenzae*, and *Pseudomonas aeruginosa*, although with variable amounts of the C-terminal metal-binding domain. It will be interesting to address whether or not SlyD is also involved in metallocenter biosynthesis in these organisms.

Acknowledgments—We thank S. C. Wang for the pET24-SlyD vector, Dr. M. R. Leach for HypB and a critical reading of the manuscript, and X. Yang, W. Yang, and A. Starostine for technical assistance. We also

thank Prof. A. Böck for the generous donation of the anti-HycE and anti-HypB antibodies as well as *E. coli* strains MC4100 and HD709, and we thank Dr. M. Blokesch for advice.

REFERENCES

1. Rosenzweig, A. C. (2002) *Chem. Biol.* **9,** 673–677
2. Kuchar, J., and Hausinger, R. P. (2004) *Chem. Rev.* **104,** 509–526
3. O'Halloran, T. V., and Culotta, V. C. (2000) *J. Biol. Chem.* **275,** 25057–25060
4. Ragsdale, S. W. (2000) in *Enzyme-catalyzed Electron and Radical Transfer* (Holzenburg, A., and Scrutton, N., eds) Vol. 35, pp. 487–518, Kluwer Academic/Plenum Publishers, New York
5. Vignais, P. M., Billoud, B., and Meyer, J. (2001) *FEMS Microbiol. Rev.* **25,** 455–501
6. Fontecilla-Camps, J.-C., Frey, M., Garcin, E., Higuchi, Y., Montet, Y., Nicolet, Y., and Volbeda, A. (2001) in *Hydrogen As a Fuel: Learning from Nature* (Cammack, R., Frey, M., and Robson, R., eds) pp. 93–109, Taylor & Francis, New York
7. Mulrooney, S. B., and Hausinger, R. P. (2003) *FEMS Microbiol. Rev.* **27,** 239–261
8. Casolot, L., and Rousset, M. (2001) *Trends Microbiol.* **9,** 228–237
9. Blokesch, M., Paschos, A., Theodoratou, E., Bauer, A., Hube, M., Huth, S., and Böck, A. (2002) *Biochem. Soc. Trans.* **30,** 674–680
10. Lutz, S., Jacobi, A., Schlensog, V., Bohm, R., Sawers, G., and Böck, A. (1991) *Mol. Microbiol.* **5,** 123–135
11. Jacobi, A., Rossmann, R., and Böck, A. (1992) *Arch. Microbiol.* **158,** 444–451
12. Maier, T., Binder, U., and Böck, A. (1996) *Arch. Microbiol.* **165,** 333–341
13. Rodrigue, A., Batia, N., Muller, M., Fayet, O., Bohm, R., Mandrand-Berthelot, M. A., and Wu, L. F. (1996) *J. Bacteriol.* **178,** 4453–4460
14. Zeghouf, M., Li, J., Butland, G., Borkowska, A., Canadien, V., Richards, D., Beattie, B., Emili, A., and Greenblatt, J. F. (2004) *J. Proteome Res.* **3,** 463–468
15. Roof, W. D., Horne, S. M., Young, K. D., and Young, R. (1994) *J. Biol. Chem.* **269,** 2902–2910
16. Wülfing, C., Lombardero, J., and Plückthun, A. (1994) *J. Biol. Chem.* **269,** 2895–2901
17. Hottenrott, S., Schumann, T., Plückthun, A., Fischer, G., and Rahfeld, J.-U. (1997) *J. Biol. Chem.* **272,** 15697–15701
18. Maier, T., Jacobi, A., Sauter, M., and Bock, A. (1993) *J. Bacteriol.* **175,** 630–635
19. Sauter, M., Böhm, R., and Böck, A. (1992) *Mol. Microbiol.* **6,** 1523–1532
20. Yu, D., Ellis, H. M., Lee, E.-C., Jenkins, N. A., Copeland, N. G., and Court, D. L. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 5978–5983
21. Casadaban, M. J., and Cohen, S. N. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76,** 4530–4533
22. Binder, U., Maier, T., and Böck, A. (1996) *Arch. Microbiol.* **165,** 69–72
23. Krogan, N. J., Peng, W.-T., Cagney, G., Robinson, M. D., Haw, R., Zhong, G., Guo, X., Zhang, X., Canadien, V., Richards, D. P., Beattie, B. K., Lalev, A., Zhang, W., Davierwala, A. P., Mnaimneh, S., Starostine, A., Tikuisis, A. P., Grigull, J., Datta, N., Bray, J. E., Hughes, T. R., Emili, A., and Greenblatt, J. F. (2004) *Mol. Cell* **13,** 225–239
24. Kislinger, T., Rahman, K., Radulovic, D., Cox, B., Rossant, J., and Emili, A. (2003) *Mol. Cell. Proteomics* **2,** 96–106
25. Begg, Y. A., Whyte, J. N., and Haddock, B. A. (1977) *FEMS Microbiol. Lett.* **2,** 47–50
26. Hube, M., Blokesch, M., and Böck, A. (2002) *J. Bacteriol.* **184,** 3879–3885
27. Ballantine, S. P., and Boxer, D. H. (1985) *J. Bacteriol.* **163,** 454–459
28. Blokesch, M., Rohrmoser, M., Rode, S., and Böck, A. (2004) *J. Bacteriol.* **186,** 2603–2611
29. Sawers, R. G., Ballantine, S. P., and Boxer, D. H. (1985) *J. Bacteriol.* **164,** 1324–1331
30. Self, W. T., Hasona, A., and Shanmugam, K. T. (2004) *J. Bacteriol.* **186,** 580–587
31. Messenger, S. L., and Green, J. (2003) *FEMS Microbiol. Lett.* **228,** 81–86
32. Roof, W. D., Fang, H. Q., Young, K. D., Sun, J., and Young, R. (1997) *Mol. Microbiol.* **25,** 1031–1046
33. Theodoratou, E., Paschos, A., Magalon, A., Fritsche, E., Huber, R., and Böck, A. (2000) *Eur. J. Biochem.* **267,** 1995–1999
34. Guzman, L.-M., Belin, D., Carson, M. J., and Beckwith, J. (1995) *J. Bacteriol.* **177,** 4121–4130
35. Böck, A., and Sawers, G. (1996) in *Escherichia coli and Salmonella, Cellular and Molecular Biology* (Neidhardt, F. C., ed) pp. 262–282, ASM Press, Washington, D. C.
36. Maier, T., Lottspeich, F., and Böck, A. (1995) *Eur. J. Biochem.* **230,** 133–138
37. Rossman, R., Sauter, M., Lottspeich, F., and Böck, A. (1994) *Eur. J. Biochem.* **220,** 337–384
38. Andrews, S. C., Berks, B. C., McClay, J., Ambler, A., Quail, M. A., Golby, P., and Guest, J. R. (1997) *Microbiology* **143,** 3633–3647
39. Göthel, S. F., and Marahiel, M. A. (1999) *Cell. Mol. Life Sci.* **55,** 423–436
40. Harrar, Y., Bellini, C., and Faure, J.-D. (2001) *Trends Plant Sci.* **6,** 426–431
41. Fischer, G., and Aumuller, T. (2003) *Rev. Physiol. Biochem. Pharmacol.* **148,** 105–150
42. Sanokawa-Akakura, R., Dai, H., Akakura, S., Weinstien, D., Fajardo, J. E., Lang, S. E., Wadsworth, S., Siekierka, J., and Birge, R. B. (2004) *J. Biol. Chem.* **279,** 27845–27848
43. Nakamura, T., Yabe, D., Kanazawa, N., Tashiro, K., Sasayama, S., and Honjo, T. (1998) *Genomics* **54,** 89–98
44. Shou, W., Aghdasi, B., Armstrong, D. L., Guo, Q., Bao, S., Charng, M.-J., Mathews, L. M., Schneider, M. D., Hamilton, S. L., and Matzuk, M. M. (1998) *Nature* **391,** 489–492
45. Mehta, N., Olson, J. W., and Maier, R. J. (2003) *J. Bacteriol.* **185,** 726–734
46. Fu, C., Olson, J. W., and Maier, R. J. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 2333–2337
47. Rey, L., Imperial, J., Palacios, J. M., and Ruiz-Argueso, T. (1994) *J. Bacteriol.* **176,** 6066–6073
48. Olson, J. W., and Maier, R. J. (2000) *J. Bacteriol.* **182,** 1702–1705
49. Olson, J. W., Fu, C., and Maier, R. J. (1997) *Mol. Microbiol.* **24,** 119–128
50. Scholz, C., Maier, P., Dolinski, K., Heitman, J., and Schmid, F. X. (1999) *FEBS Lett.* **443,** 367–369
51. Finzi, A., Cloutier, J., and Cohen, É. A. (2003) *J. Virol. Methods* **111,** 69–73
52. De Pina, K., Desjardin, V., Mandrand-Berthelot, M.-A., Giordano, G., and Wu, L.-F. (1999) *J. Bacteriol.* **181,** 670–674
53. Chivers, P. T., and Sauer, R. T. (2000) *J. Biol. Chem.* **275,** 19735–19741
54. Bloom, S. B., and Zamble, D. B. (2004) *Biochemistry* **43,** 10029–10038
55. Clugston, S. L., Barnard, J. F. J., Kinach, R., Miedema, D., Ruman, R., Daub, E., and Honek, J. F. (1998) *Biochemistry* **37,** 8754–8763
56. Chan, J., Huang, Z., Merrifield, M. E., Salgado, M. T., and Stillman, M. (2002) *Coord. Chem. Rev.* **233–234,** 319–339
57. Coyle, P., Philcox, J. C., Carey, L. C., and Rofe, A. M. (2002) *Cell. Mol. Life Sci.* **59,** 627–647
58. Cavet, J. S., Borrelly, G. P. M., and Robinson, N. J. (2003) *FEMS Microbiol. Rev.* **27,** 165–181
59. Bernhardt, T. G., Roof, W. D., and Young, R. (2002) *Mol. Microbiol.* **45,** 99–108
60. Witte, A., Schrot, G., Schön, P., and Lubitz, W. (1997) *Mol. Microbiol.* **26,** 337–346
61. Suzuki, R., Nagata, K., Yumoto, F., Kawakami, M., Nemoto, N., Furutani, M., Adachi, K., Maruyama, T., and Tanokura, M. (2003) *J. Mol. Biol.* **328,** 1149–1160
62. Furutani, M., Ideno, A., Iida, T., and Maruyama, T. (2000) *Biochemistry* **39,** 453–462
63. Butland, G., Peregrin-Alvarez, J. M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogari, N., Davey, M., Parkinson, J., Greenblatt, J. F., and Emili, A. (2005) *Nature*, in press

JOURNAL OF BACTERIOLOGY, Jan. 2001, p. 189–199 Vol. 183, No. 1
0021-9193/01/$04.00+0 DOI: 10.1128/JB.183.1.189–199.2001

Two Conserved Glutamates in the Bacterial Nitric Oxide Reductase Are Essential for Activity but Not Assembly of the Enzyme

GARETH BUTLAND,† STEPHEN SPIRO, NICHOLAS J. WATMOUGH, AND DAVID J. RICHARDSON*

Centre for Metalloprotein Spectroscopy and Biology, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom

Received 12 April 2000/Accepted 3 October 2000

The bacterial nitric oxide reductase (NOR) is a divergent member of the family of respiratory heme-copper oxidases. It differs from other family members in that it contains an Fe_B–heme-Fe dinuclear catalytic center rather than a Cu_B–heme-Fe center and in that it does not pump protons. Several glutamate residues are conserved in NORs but are absent in other heme-copper oxidases. To facilitate mutagenesis-based studies of these residues in *Paracoccus denitrificans* NOR, we developed two expression systems that enable inactive or poorly active NOR to be expressed, characterized in vivo, and purified. These are (i) a homologous system utilizing the *cycA* promoter to drive aerobic expression of NOR in *P. denitrificans* and (ii) a heterologous system which provides the first example of the expression of an integral-membrane cytochrome *bc* complex in *Escherichia coli*. Alanine substitutions for three of the conserved glutamate residues (E125, E198, and E202) were introduced into NOR, and the proteins were expressed in *P. denitrificans* and *E. coli*. Characterization in intact cells and membranes has demonstrated that two of the glutamates are essential for normal levels of NOR activity: E125, which is predicted to be on the periplasmic surface close to helix IV, and E198, which is predicted to lie in the middle of transmembrane helix VI. The subsequent purification and spectroscopic characterization of these enzymes established that they are stable and have a wild-type cofactor composition. Possible roles for these glutamates in proton uptake and the chemistry of NO reduction at the active site are discussed.

Many species of bacteria contain a nitric oxide reductase (NOR) which catalyzes the reaction $2NO + 2e^- + 2H^+ \rightarrow N_2O + H_2O$ (21, 26). The reduction of NO serves as a key step in denitrification (in which N-oxyanions and N-oxides are used as respiratory electron acceptors) and as a way of removing cytotoxic NO. The NOR of *Paracoccus denitrificans* is an integral-membrane protein that normally purifies as two-subunit complex NorCB (10–12, 15). NorC is a monoheme membrane-anchored *c*-type cytochrome. NorB is a divergent member of the family of catalytic subunits from respiratory heme-copper oxidases (HCOs) (21, 26). Typical features of catalytic subunits of the HCOs are a core functional unit of 12 transmembrane helices, which bind a magnetically isolated electron-transferring heme, and a dinuclear active site, formed by a second heme magnetically coupled to a copper ion (Cu_B). Seven conserved histidine residues, responsible for ligating the three redox-active metal centers, can be identified in helices II, VI, VII, and X. Each of these histidine residues is conserved in the NorB subunit of NOR (21, 26, 29).

The key difference between the catalytic subunit of NOR and those of other HCOs is the composition of the dinuclear center. In NorB there is a nonheme iron (Fe_B) at the active site rather than copper (Cu_B) (13), possibly because, under the highly reducing conditions of the primordial biosphere, ferrous ions were more readily available than insoluble cuprous ions to the ancestral enzyme from which both NOR and HCOs evolved. Since it is likely that denitrification preceded aerobic respiration in the biosphere, the primary function of the ancestral oxidase was probably the reduction of NO. Consequently, a key step in the evolution of aerobic life on earth may have been the replacement of iron by copper in the ancestral oxidase, allowing it to reduce oxygen more efficiently (4). Recent biochemical studies have begun to reveal additional differences in the catalytic pockets of NorB and HCOs. For example, resonance Raman spectroscopy of the CO adduct of reduced NOR has suggested that the catalytic pocket of NorB is more negatively charged than those of HCOs (18). In addition, redox potentiometry has indicated a midpoint potential of high-spin heme b_3 that is around 200 mV lower than that of high-spin heme a_3 of cytochrome oxidase (11). This may serve to prevent formation of a dead-end Fe(II)-NO complex during the catalytic cycle.

High-resolution X-ray analysis of the crystal structures of cytochrome *c* oxidase, together with site-directed mutagenesis, have led to the identification of amino acid residues that are important in the delivery of chemical and "pumped" protons from the cytoplasm to the dinuclear center during the catalytic cycle and that define the so-called K and D channels (1, 16, 32). The absence of these residues from NOR suggests that the enzyme is not a proton pump and that it takes the protons required for reduction of NO from the periplasm, and there is experimental evidence consistent with this idea (2, 3, 22). Hence, evolution of the ancestral NO-reducing enzyme into an HCO involved the acquisition of not only a Cu-containing dinuclear center but also a proton pumping mechanism. The

* Corresponding author. Mailing address: School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom. Phone: 44 1603 593250. Fax: 44 1603 592250. E-mail: d.richardson@uea.ac.uk.

† Present address: Department of Biochemistry, University of Oxford, Oxford OX1 3QU, United Kingdom.

primary structures of NorB subunits reveal a number of conserved glutamic acid residues in putative transmembrane helices and periplasmic loops, which are absent in other HCOs (29). These are (*P. denitrificans* numbering) E122 in the helix III/IV loop, E125 at the periplasmic surface of helix IV, E198 and E202, located one and two helical turns, respectively, below the putative Fe_B ligand (His194) in helix VI, and E267, located in the middle of helix VIII. This sequence conservation, together with the energetic cost of placing a charged residue in the lipid bilayer, suggests a functional importance for these glutamates. Possible roles include Fe_B binding, modulation of the charge of the catalytic pocket or of the catalytic heme redox potential, and mediation of proton movements.

In order to investigate the role of conserved residues in bacterial NORs, there has been a need to develop suitable expression systems for catalytically inactive enzymes. Heterologous and homologous expression systems would allow assessment of the physiological consequences of mutations in NorB, as well as the production of pure enzyme for structure-function studies. We have developed two such systems that meet these criteria, and the characterization of enzymes with E125A, E198A, and E202A substitutions in intact cells, membrane fractions, and purified preparations is reported. The results demonstrate that E125 and E198 are not required for the assembly of a stable holo-NorCB enzyme complex but have a critical role in NO reduction.

MATERIALS AND METHODS

Construction of a system for homologous expression of *P. denitrificans norCB*. Plasmid pKPD1 is a clone of the entire *cycA* (*P. denitrificans* cytochrome c_{550}) gene and promoter region in expression vector pKK223-3, which has been modified to contain a unique *Sal*I site (25). A 325-bp *Sal*I-*Eco*RI fragment containing the *cycA* promoter region was excised from pKPD1 and cloned into pUC18 to yield pGB1. A 7.8-kb *Hin*dIII fragment containing the *norCBQDEF* operon was excised from pEG8HI (a gift from R. J. M. van Spanning, Vrije Universiteit, Amsterdam, The Netherlands) and cloned into pUC18 to yield pNORHC. There are two *Bsa*BI restriction sites in the *cycA* promoter region in pGB1. The first site cuts 1 bp downstream of the ATG start codon of the *cycA* gene, and the second cuts 6 bp downstream of the first. The *norCBQDEF* coding region, minus the *norC* promoter, was excised from pNORHC on a blunt-ended 5.7-kb *San*DI-*Hin*dIII fragment and ligated into pGB1 that had been digested with *Bsa*BI. Clones containing the 5.7-kb insert in the correct orientation were selected and designated pCYCNOR1. The *cycA-nor* fusion was excised from pCYCNOR1 on a 6.0-kb *Eco*RI-*Pst*I fragment and cloned into pBluescript KS+. Recombinant clones were designated pCYCNOR2. The 6.0-kb *cycA-nor* fusion was excised from pCYCNOR2 by digestion with *Xba*I and *Hin*dIII and cloned into broad-host-range vector pEG400 to yield pCYCNOR3, generating a *cycA-nor* fusion that could be propagated in *P. denitrificans*.

Construction of a *norB*::Ω-*Km* mutant. The 4-kb *Bgl*II fragment containing the *nor* operon from pNORHC was cloned into the *Bam*HI site of pBluescript KS+ to yield pBg14. A 2.2-kb *Bam*HI Ω-Km fragment was excised from pHP45Ω-Km and cloned into the unique *Bam*HI site in pBg14, which is in *norB*, to generate pBglkm. The insert from pBglkm was cloned as a 6.1-kb *Xba*I-*Eco*RI fragment into pLITMUS28 to yield pLitBglkm, which was then digested with *Xba*I and *Spe*I, and the 6.1-kb fragment was cloned into the *Xba*I site of pUC18. A correctly oriented clone was selected, and the plasmid was named pUCBglkm. These cloning steps resulted in the location of the entire 6.1-kb DNA fragment containing the Ω-Km cassette and flanking DNA from the *nor* coding sequence on a single *Eco*RI fragment. This fragment was introduced into the suicide vector pRVS1 to yield pRVSBglkm. *Escherichia coli* S17.1 (pRVSBglkm) and *P. denitrificans* 1222 were then used in biparental filter matings on L agar (16 h at 30°C). Cells were removed from the filter by resuspension in L broth and plated onto L agar (36 h at 37°C) containing rifampin, spectinomycin, kanamycin, and X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside). Streptomycin-sensitive, kanamycin-resistant white colonies were designated *P. denitrificans* GB1 (genotype *norB*::Ω-*Km*). The presence of the 2.1-kb Ω cartridge in *norB* was confirmed by direct genomic PCR analysis.

Directed mutagenesis of *norB*. pCYCNOR3 was digested with *Xba*I, which cuts upstream of the *cycA* promoter, and *Xho*I, which cuts in the middle of *norB*. The 1.6-kb fragment (which contains the E125, E198, and E202 codons) was ligated into *Xba*I/*Xho*I-digested pBluescript KS+ to generate pNORXX16, which was transformed into *E. coli* DH5α. PCRs were set up with complementary primer pairs suitable for introduction of the E125A (GAA → GCG), E198A (GAG → GCC), E202A (GAG → GCC), and E198A plus E202A mutations. The following silent restriction sites were incorporated into the primers to allow for easy screening for mutations by restriction digests: E125A, *Fsp*I; E198A, *Sac*I; E202A, *Apa*I; E198A plus E202A, *Nae*I and *Apa*I. The template for PCR was pNORXX16, and reactions were performed using the Quickchange (E198) or ExSite mutagenesis kit (Stratagene). In the final stage of the protocol, *E. coli* XL1-Blue transformed in each of the four ligation reactions was spread onto L agar plates supplemented with 100 μg of ampicillin/ml. Potential mutants were selected and sequenced to establish their authenticity. The plasmids carrying codons leading to the E125A, E198A, E202A, and E198A plus E202A mutations were designated pNOR125A, pNOR198A, pNOR202A, and pNOR198202A, respectively. The 1.6-kb *Xba*I-*Xho*I fragments from the four mutant plasmids were cloned into *Xba*I/*Xho*I-digested pCYCNOR. The resulting plasmids were then digested with *Fsp*I (E125A), *Sac*I (E198A), or *Apa*I (E202A and E198A plus E202A), as appropriate. All clones were found to have the expected restriction pattern. These plasmids were designated p125CNOR, p198CNOR, p202CNOR, and p198202CNOR and were introduced into *P. denitrificans* GB1 by triparental matings with the corresponding *E. coli* DH5α transformants and *E. coli* JM83 harboring helper plasmid pRK2013.

Construction of pNOREX used for the expression of the *norCBQDEF* operon in *E. coli*. The 6.0-kb *Xba*I-*Hin*dIII fragment, containing the *cycA-nor* fusion from pCYCNOR3, was cloned into pUC18 to yield pNOREX, which has the *nor* operon in the correct orientation to allow expression from the *lac* promoter. pNOREX was found to be unstable in *E. coli* DH5α, so strain JM109 (*lacI*^{q}), in which pNOREX was more stable, was used as the host. Similar procedures were used to construct p125EX, p198EX, p202EX, and p198202EX, using the 6.0-kb *Xba*I-*Hin*dIII fragments from p125CNOR, p198CNOR, p202CNOR, and p198202CNOR.

Anaerobic growth of *P. denitrificans*. *P. denitrificans* strains were grown aerobically at 37°C in 50 ml of L broth supplemented with the appropriate antibiotics. For each strain, a 500-ml bottle of succinate-nitrate minimal medium, supplemented with the appropriate antibiotics, was inoculated with a 1% volume from the L broth cultures. The bottle was then sealed, and the cells were mixed thoroughly by inversion. The starting optical density at 610 nm (OD_{610}) was determined, and, to avoid further introduction of oxygen during anaerobic growth, the contents of the inoculated bottles were aliquoted into 25-ml bottles. Each 25-ml bottle was then tightly sealed and incubated at 30°C. A single bottle was opened for each time point in the growth curve and the OD_{610} was recorded. Also, at each time point, 1.5 ml of cells was centrifuged for 5 min at 13,000 rpm in a bench top microcentrifuge. The culture supernatant was then assayed colorimetrically for nitrite.

Analytical methodologies. NO reductase activity was measured amperometrically using a Clark-type electrode, essentially as previously described (10, 13), but using ascorbate, phenazine methosulfate, and horse heart cytochrome *c* as the electron donor/mediator system. Cytochrome oxidase activity was measured spectrophotometrically by monitoring the NOR-dependent oxidation kinetics of reduced horse heart cytochrome *c* in aerated cuvettes. Protein levels were estimated using the bicinchoninic acid method with bovine serum albumin as a standard. The stain for heme-linked peroxidase activity, which is specific for *c*-type cytochromes, was as previously described (20). The rates given in Tables 2 and 3 are representative data taken from samples prepared from at least two independent cultures.

EPR, UV-Vis, and mediated redox potentiometry. Electron paramagnetic resonance (EPR) spectra were recorded using an ER-200D X-band spectrometer (Bruker Spectrospin) interfaced to an ESP1600 computer and fitted with a liquid-helium flow cryostat (ESR-9; Oxford Instruments). UV-visible (UV-Vis) spectra were collected using an Aminco SLM DW2000 spectrophotometer. Samples for UV-Vis spectra and redox titrations were at 25°C in 50 mM Tris-HCl (pH 7.5). Mediated redox potentiometry was performed as previously described (11). Dithionite and ferricyanide were used as the reductant and oxidant, respectively. Redox mediators were phenazine methosulfate, phenazine ethosulfate, diaminodurene, 4-hydroxynaphthoquinone, 5-anthraquinone 2-sulfonate, 6-anthraquinone 2,6-disulfonate, and benzyl viologen (at a final concentration of 20 μM). Quinhydrone was used as a redox standard ($E_{m,7.0}$ = +295 mV). All potentials quoted are with respect to that of the normal hydrogen electrode.

TABLE 1. Strains and plasmids

Strain or plasmid	Description[a]	Reference or source
Strains		
E. coli DH5α	*supE44* Δ(*lacU169*) (ø80 *lacZΔM15*) *recA endA1 gyrA96 thi-1 hsdR17 relA1 deoR*	Gibco-BRL
E. coli JM109	*recA endA gyrA thi hsdR supE relA* Δ(*lac-proAB*) [F′ *traD proAB*$^+$ *lacI*^{q}*Z*ΔM15]	Gibco-BRL
E. coli S17-1	*thi pro hsdR hsdM*$^+$ *recA*; chromosomal insertion of RP4-2 (Tc::Mu Km::Tn*7*); Str^r	24
P. denitrificans 1222	Restriction modification deficient	6
P. denitrificans GB1	1222 *norB*::Ω-*Km*	This work
Plasmids		
pBluescript II KS+	Cloning vector; Amp^r	Stratagene
pRVS1	pBR322-derived suicide vector	28
pHP45Ω-Km	Plasmid carrying Ω-Km; Kan^r	7
pEG400	Broad-host-range cloning vector	9
pEC86	Plasmid carrying *ccm* gene cluster	19
pUC18	Cloning vector; Amp^r	Roche
pLITMUS28	Cloning vector; Amp^r	New England Biolabs
pKK223-3	Derivative of pKK223 with unique *Sal*I site	
pKPD1	*P. denitrificans cycA* gene and promoter region cloned into pKK223-3	25
pGB1	325-bp *Sal*I-*Eco*RI fragment containing the *cycA* promoter from pKPD1 in pUC18	This work
pEG8H1	*P. denitrificans nor* operon in pEG400	R. van Spanning
pNORHC	7.8-kb *Hin*dIII fragment from pEG9HI containing *norCBQDEF* cloned in pUC18	This work
pCYCNOR1	5.7-kb *San*DI-*Hin*dIII fragment from pNORHC cloned into pGB1	This work
pCYCNOR2	6-kb *Eco*R1-*Pst*I fragment from pCYCNOR1 containing the *cycA-nor* fusion cloned in pBluescript KS+	This work
pCYCNOR3	6-kb *Xba*I-*Hin*dIII fragment containing *cycA-nor* fusion cloned in pEG400	This work
pBg14	4-kb *Bgl*II fragment from pNORHC cloned in pBluescript KS+	This work
pBglkm	2.2-kb *Bam*HI Ω-Km fragment from pHP45Ω-Km cloned into *norB* gene of Bg14	This work
pLITBglkm	6.1-kb *Xba*I-*Eco*RI *norCB*::Ω*QEDF* insert from pBglkm cloned into pLITMUS28	This work
pUCBglkm	6.1-kb *Xba*I-*Spe*I *norCB*::Ω*QEDF* insert from pLITBglkm cloned in pUC18	This work
pRVSBglkm	6.1-kb *Eco*RI *norCB*::Ω*QEDF* insert from pUCBglkm cloned in pRVS1	This work
pNOREX	6.1-kb *Xba*I-*Hin*dIII *cycA-nor* fusion from pCYCNOR cloned in pUC18	This work
pNORXX16	1.6-kb *Xba*I-*Xho*I *nor* fragment from pCYCNOR clone into pBluescript KS+	This work
pNOR125A	E125A NorB mutation in pNORXX16	This work
pNOR198A	E198A NorB mutation in pNORXX16	This work
pNOR202A	E202A NorB mutation in pNORXX16	This work
pNOR198202A	E198A, E202A NorB mutations in pNORXX16	This work
p125CNOR	E125A NorB mutation in pCYCNOR	This work
P198CNOR	E198A NorB mutation in pCYCNOR	This work
p202CNOR	E202A NorB mutation in pCYCNOR	This work
p198202CNOR	E198A, E202A NorB mutations in pCYCNOR	This work
p125EX	E125A NorB mutation in pNOREX	This work
p198EX	E198A NorB mutation in pNOREX	This work
p202EX	E202A NorB mutation in pNOREX	This work
p198202EX	E198A, E202A NorB mutations in pNOREX	This work

[a] Str^r, streptomycin resistant; Kan^r, kanamycin resistant; Amp^r, ampicillin resistant.

Redox titrations were fitted using a customized program in table-curve 2D (Jandel Scientific) allowing estimates of E_m and multiple independent $n = 1$ components to float as appropriate (11). The error for each E_m was estimated from multiple titrations to be ±20 mV.

Purification of NOR. Recombinant NOR was purified from 15-liter cultures of *E. coli* grown in L broth in an aerated bioreactor. When the culture density reached an OD_{650} of 0.4, the culture was induced for 4 h with 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside). Cells were then harvested by cross-flow filtration and broken in a French press. Membranes were recovered by centrifugation and suspended in 100 mM Tris-HCl (pH 7.6)–50 mM NaCl–1 mM EDTA, sonicated, and solubilized in 1% (wt/vol) *n*-dodecyl-β-D-maltoside (4°C for 1 h). The sample was then centrifuged at 45,000 rpm in a Beckman 70Ti rotor for 1 h at 4°C. The supernatant was immediately diluted 10-fold with buffer to avoid precipitation of the protein. The protein sample was then purified using Q-Sepharose (0 to 500 mM NaCl gradient) and Cu-IMAC (2.5 to 50 mM imidazole gradient) chromatographies, as previously described (13). The elution buffer was 50 mM Tris-HCl (pH 7.5)–0.1% dodecyl maltoside. Fractions containing NOR were identified spectroscopically.

RESULTS AND DISCUSSION

Characterization of engineered NOR in intact cells and membrane fractions of *P. denitrificans*. The utility of pCYCNOR3 (Table 1) for *nor* expression in *P. denitrificans* depends on its ability to express *norCB* under growth conditions for which NOR is not essential. Thus, the expression of an engineered *norCB* that encodes an inactive or poorly active NOR is possible. The *nor* promoter is only active under anaerobic conditions and is dependent on the presence of NO (14, 17, 27). However, in pCYCNOR3 the *nor* genes are transcribed from the *P. denitrificans cycA* promoter (from the cytochrome c_{550} gene), which is known to be active under some aerobic growth conditions (20, 25). *P. denitrificans* strains GB1 (*norB*::Ω) and GB1(pCYCNOR3) were grown under a range of conditions, and the NOR activities of the membrane fractions were deter-

TABLE 2. NOR activities of membranes prepared from *P. denitrificans* GB1 grown under different conditions with different plasmids

Plasmid present	NOR expressed[a]	Growth condition	NOR activity (nmol · mg of protein^{-1} · min^{-1})
None	None	O_2, succinate	<5
None	None	O_2, methylamine	<5
None	None	O_2, L broth	<5
pCYCNOR3	WT	O_2, succinate	23
pCYCNOR3	WT	O_2, methylamine	230
pCYCNOR3	WT	O_2, L broth	300
pCYCNOR3	WT	Anaerobic, succinate + nitrate	440
p125CNOR	E125A	O_2, methylamine	<5
p198CNOR	E198A	O_2, methylamine	<5
P202CNOR	E202A	O_2, methylamine	200
p198202CNOR	E198A/E202A	O_2, methylamine	<5
p125CNOR	E125A	O_2, L broth	<5
p198CNOR	E198A	O_2, L broth	<5
P202CNOR	E202A	O_2, L broth	210
p198202CNOR	E198A/E202A	O_2, L broth	<5

[a] WT, wild type; E125A, NOR^{E125A} (similar for E198A and E202A).

mined. GB1 displayed no detectable NOR activity under any of the growth conditions tested (Table 2). By contrast, NOR activity was detectable in GB1(pCYCNOR3) under all growth conditions tested. Activity was lowest following aerobic growth on succinate medium and was an order of magnitude higher in membranes prepared from cells grown aerobically on methylamine, anaerobically on succinate-nitrate medium, or aerobically to late stationary phase on L broth (Table 2).

These expression studies demonstrated that the *cycA* promoter is capable of driving expression of the *nor* operon under growth conditions for which NOR is nonessential. Having established this, mutations leading to E125A, E198A, E202A, and E198A plus E202A substitutions were introduced into the *norB* gene (Table 1) and the engineered enzymes were expressed in aerobic methylamine-grown cultures of GB1 (*norB*::Ω). There was no detectable NOR activity in membranes from the strains expressing enzymes with the E125A, E198A, and E198A plus E202A substitutions and intermediate levels of activity in membranes from the strain expressing NOR with the E202A substitution (Table 2). Immunochemistry was employed to assess the expression of the catalytic NorB subunit. Membrane fractions were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) alongside a sample of purified NOR. The gel was Western blotted and then probed with an anti-NorB antibody. A strongly reactive band that corresponded to NorB could be observed in the lane containing purified NOR. This protein was absent from *norB*::Ω mutant GB1 but present in GB1(pCYCNOR3) (Fig. 1A). The same band was also present in GB1(p125CNOR), GB1(p198CNOR), GB1(p202CNOR), and GB1(p198202CNOR). Although there was some sample-to-sample heterogeneity, analysis of three independent membrane preparations suggested that levels of expression of the NorB polypeptide in all of the strains carrying either the wild-type or mutant forms of the *norB* gene were similar. This confirmed that NorB biosynthesis and stability were similar for all of the engineered enzymes and so could not account for the pronounced differences in NOR activity observed in membranes expressing these enzymes.

To assess the physiological competence of the engineered

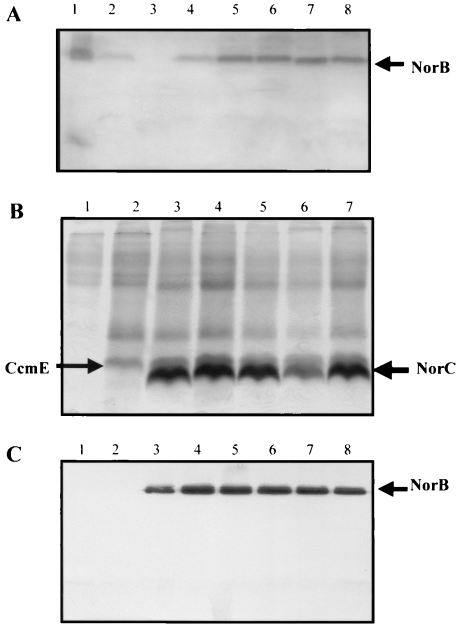


FIG. 1. Heme-stained SDS-PAGE gel and anti-NorB-probed Western blot of membrane fractions from *P. denitrificans* and *E. coli*. (A) Anti-NorB-probed Western blot of *P. denitrificans* membranes. Membranes were prepared from cells grown aerobically on L broth and solubilized in 1% dodecyl maltoside. Fifteen microliters (5 to 10 μg of protein) of each sample was loaded onto the SDS-PAGE gel, which was subsequently used for the Western blotting. Lane 1, purified NorCB; lane 2, strain 1222; lane 3, GB1; lane 4, GB1(pCYCNOR3); lane 5, GB1(p125CNOR); lane 6, GB1(p198CNOR); lane 7, GB1(p202CNOR); lane 8, GB1(p198202CNOR). (B) Heme-stained gel of *E. coli* membranes. Lane 1, JM109; lane 2, JM109(pEC86); lane 3, JM109(pNOREX, pEC86); lane 4, JM109(p125EX, pEC86); lane 5, JM109(p198EX, pEC86); lane 6, JM109(p202EX, pEC86); lane 7, JM109(p198202EX, pEC86). (C) Anti-NorB-probed Western blot of *E. coli* membranes. Lane 1, JM109; lane 2, JM109(pEC86); lane 3, JM109(pNOREX, pEC86); lane 4, JM109(p125EX, pEC86); lane 5, JM109(p198EX, pEC86); lane 6, JM109(p202EX, pEC86); lane 7, JM109(p198202EX, pEC86); lane 8, purified NorCB. Membranes were solubilized in 1% dodecyl maltoside, and 15 μl (5 to 10 μg of protein) was loaded into each well of the SDS-PAGE gels.

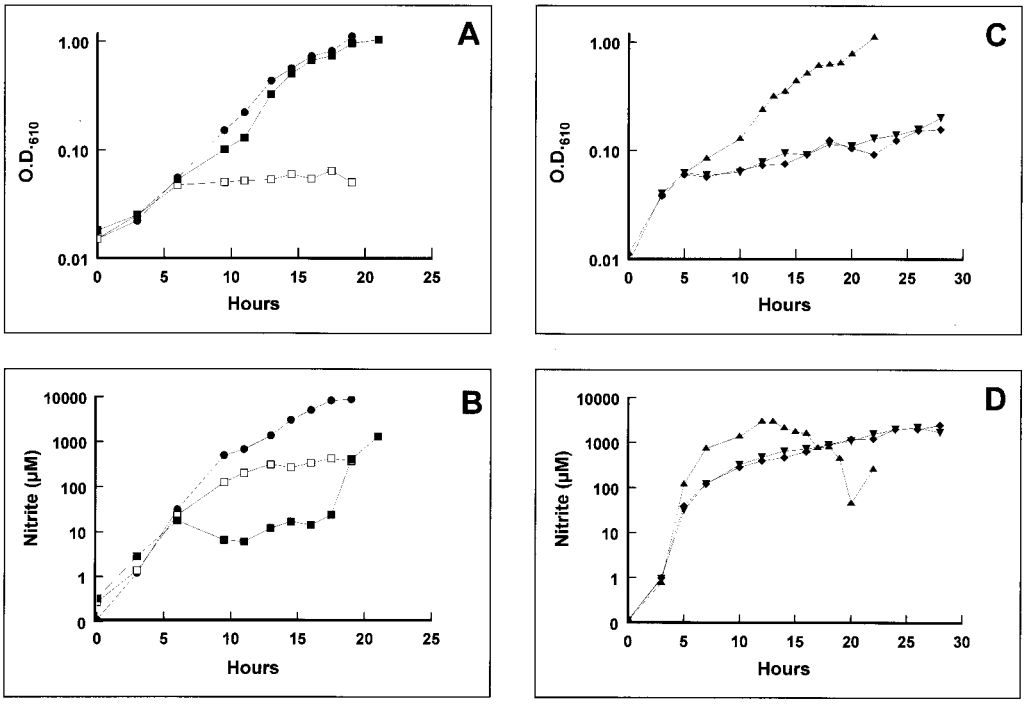


FIG. 2. Growth curves of *P. denitrificans* 1222 (●), GB1 (□), GB1(pCYCNOR3) (■), GB1(p125CNOR) (▼), GB1(p198CNOR) (◆), and GB1(p202CNOR) (▲). Cultures were grown under anaerobic denitrifying conditions. (A) Growth kinetics, monitored via OD_{610}, of 1222, GB1, and GB1(pCYCNOR3). (B) Nitrite accumulation kinetics during growth of 1222, GB1, and GB1(pCYCNOR3). (C) Growth kinetics, monitored via OD_{610}, of GB1(p125CNOR), GB1(p198CNOR), and GB1(p202CNOR). (D) Nitrite accumulation kinetics during growth of GB1(p125CNOR), GB1(p198CNOR), and GB1(p202CNOR).

NORs, *P. denitrificans* 1222, GB1, GB1(pCYCNOR3), GB1(p125CNOR), GB1(p198CNOR), and GB1(p202CNOR) were cultured under anaerobic denitrifying conditions with succinate as the carbon source, ammonium as the nitrogen source, and nitrate as a respiratory electron acceptor (Fig. 2A). During the first 6 h of anaerobic incubation, strains 1222 (wild type) and GB1 (*norB*::Ω) exhibited similar growth kinetics. This growth period was accompanied by a rapid accumulation of nitrite in the culture supernatant, which could be attributed to the respiratory reduction of nitrate to nitrite (Fig. 2B). After this period, growth of GB1 was almost completely attenuated (Fig. 2A), although a net accumulation of nitrite continued throughout the 20-h duration of the growth experiment (Fig. 1B). It is likely that after 6 h some of the nitrite initially produced by the culture was reduced to NO by the cytochrome cd_1 nitrite reductase. The NO cannot be further reduced in the absence of NOR and so inhibits growth. Introduction of the *nor*-expressing clone pCYCNOR3 into GB1 restored the wild-type capacity for anaerobic denitrifying growth (Fig. 2A). Nitrogen gas bubbles could be observed during growth of both 1222 and GB1(pCYCNOR3), indicative of complete denitrification. However, there were significant differences in the nitrite extrusion profiles during growth of strains 1222 and GB1(pCYCNOR3). The wild-type strain accumulated nitrite in the growth medium throughout growth, but in GB1(pCYCNOR3) the nitrite reached a steady concentration (7 to 12 mM) between 6 and 18 h, increasing rapidly again thereafter (Fig. 2B). These differences may be a consequence of expressing *nor* from the *cycA* promoter. The *nor* promoter is coregulated with the nitrite reductase genes by the NO-responsive activator NNR (14, 17, 27, 29); this coordinate regulation is lost in the recombinant expression system.

Strains GB1(p125CNOR) and GB1(p198CNOR) resembled GB1 in that they were able to grow during the first 6 h after inoculation (Fig. 2C) by virtue of the energy-conserving reduction of nitrate to nitrite, which accumulated in the culture supernatant (Fig. 2D). Thereafter, no further growth was apparent, presumably as a result of the failure to reduce the NO derived from nitrite reduction at sufficiently rapid rates to prevent toxicity. GB1 expressing NOR^{E202A} showed almost complete complementation (Fig. 2C). In all three cases the growth phenotypes reflect the relative levels of NOR activity observed in membrane fractions prepared from the methylamine-grown cells (Table 2).

Characterization of engineered *P. denitrificans* NOR in intact cells and subcellular fractions of *E. coli*. Homologous expression of *norCB* proved essential for assessing the physiological competence of engineered NORs and for establishing that they were synthesized and stable. To provide large quantities of NOR for purification and spectroscopic analysis, a heterologous *nor* expression system utilizing the IPTG-inducible *lacZ* promoter of pUC18 in *E. coli* JM109 was developed. To facilitate expression, *E. coli* was cotransformed with pNOREX (Table 1) and pEC86, which contains the cytochrome *c* assembly (*ccm*) genes (19, 22). Dithionite-reduced UV-Vis spectra revealed differences between JM109, JM109(pEC86), and JM109(pEC86/pNOREX) membrane extracts. The reduced spectrum of detergent-solubilized membranes from JM109

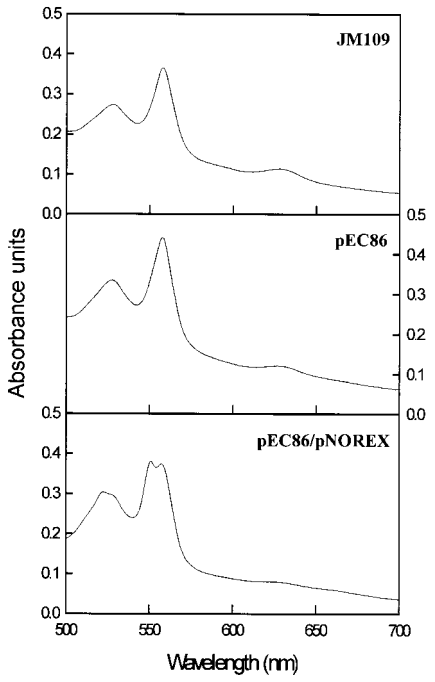


FIG. 3. UV-Vis absorption spectra of membranes prepared from *E. coli* JM109, JM109(pEC86), and JM109(pEC86/pNOREX). The strains were grown aerobically in 500-ml Luria-Bertani medium (in 2.5-liter baffled flasks) and induced at an OD_{610} of 0.4 with 1 mM IPTG. Cells were harvested 4 h after induction, and membranes were prepared as described in Materials and Methods. Spectra were acquired from suspensions of 100 μg of membranes per ml.

shows a peak at approximately 560 nm typical of *E. coli* respiratory complexes containing *b*-type heme, such as the cytochrome bo_3 oxidase and the formate dehydrogenase (Fig. 3). There is no indication of the presence of any *c*-type cytochromes, since there is no absorption peak at around 550 nm. The JM109(pEC86) extract shows an overall increase in the intensity of the spectrum and a slight shift of the peak towards 558 nm. This is probably due to the expression of the CcmE protein (from pEC86), which is known to absorb maximally at 558 nm when in the reduced form (22). The clearest spectral changes occurred in JM109(pEC86/pNOREX), where two peaks with almost equal intensities at approximately 550 and 558 nm can be resolved (Fig. 3). The peak at 550 nm is indicative of the presence of a *c*-type cytochrome and is likely to arise from NorC. The intensity at 558 nm is likely to arise from a combination of CcmE, various respiratory complexes, and also the *b* hemes of NorB.

Detergent-solubilized membrane extracts of JM109, JM109(pEC86), and JM109(pEC86/pNOREX) were subjected to SDS-PAGE, and the gel was stained for heme-dependent peroxidase activity to enable detection of *c*-type cytochromes. The JM109 membrane extract shows very little covalently attached heme (Fig. 1B). The faint high-molecular-weight bands are thought to arise from noncovalently bound *b* heme, which had not fully dissociated from some of the respiratory complexes. The JM109(pEC86) membrane extract contained a band at 19 kDa, which stained strongly for heme and which is likely to be CcmE (19, 22). In addition this extract contained a ≈30-kDa heme-staining polypeptide which is likely to arise from an endogenous *E. coli* protein, since it is also present in the JM109 membrane extract, albeit at extremely low levels. The membrane extract from cells containing both pEC86 and pNOREX has an additional 18-kDa heme-staining polypeptide that migrated slightly faster than the CcmE polypeptide. The molecular mass of this heme-staining polypeptide is consistent with it being NorC. To confirm the presence of NorB, all three membrane extracts resolved by SDS-PAGE were Western-blotted and probed with an anti-*P. denitrificans* NorB antibody (Fig. 1C). There was no cross-reacting band in the lanes loaded with JM109 and JM109(pEC86) extracts. However, a single strongly cross-reacting band that migrated to the same position as the NorB polypeptide of purified NOR could be clearly identified in the JM109(pEC86/pNOREX) extract. In agreement with the apparent expression pattern of NorC and NorB from the heme-staining and immunochemical analysis, membranes prepared from JM109 and JM109(pEC86) both displayed no detectable NOR activity, whereas JM109(pEC86/pNOREX) membranes displayed significant activity (Table 3). These data confirmed that *P. denitrificans* NOR was expressed and active in *E. coli* JM109 containing pNOREX and the *ccm* coexpression plasmid pEC86. This represents the first example of the expression of a large integral-membrane respiratory cytochrome *bc* complex in *E. coli*. The coexpression of the *ccm* genes from pEC86 was critical to the success of this strategy since the NorC subunit did not assemble efficiently in its absence (data not shown). It should also be noted that pNOREX contained the whole *norCBQDEF* operon. Attempts to express NOR in the absence of the *norQDEF* genes were unsuccessful, but a systematic study of the role of each of these genes in the assembly and/or stability of NOR was not undertaken at this stage.

TABLE 3. NOR and oxidase activities in membranes prepared from *E. coli* JM109 expressing wild-type or mutant forms of NOR

Plasmids present	NOR expressed[a]	NOR activity (nmol of NO mg of protein^{-1} min^{-1})	Oxidase activity (nmol of O mg of protein^{-1} min^{-1})
None	None	<5	
pEC86	None	<5	
pEC86, pNOREX	Wild type	470	40
pEC86, p125EX	E125A	20	2
pEC86, p198EX	E198A	<5	<1
pEC86, p202EX	E202A	140	30
pEC86, p198202EX	E198A/E202A	<5	<1

[a] E125A, NOR^{E125A} (similar for E198A and E202A).

To assess the activity of the engineered NORs in *E. coli*, p125EX, p198EX, p202EX, and p198202EX were all introduced into JM109 with pEC86. Solubilized membrane extracts were subjected to SDS-PAGE, and the gel was stained for covalently bound heme (Fig. 1B). The 18-kDa polypeptide identified as NorC was present in all of the extracts. The presence of NorB was confirmed using the NorB antibodies (Fig. 1C). Membrane extracts of the *E. coli* strains expressing engineered NorB were assayed for NOR activity (Table 3). The activities followed a pattern similar to those obtained for the mutant enzymes expressed in *P. denitrificans*; no activity was detected for the E198A mutant, and the E202A mutant had the highest activity. The only major discrepancy was that the E125A mutant had no detectable NOR activity when expressed in *P. denitrificans* but did show a very low (5% of wild-type) activity when expressed in *E. coli*. This residual activity may have been too low to detect in *P. denitrificans* as a consequence of the background NOR activity of cytochrome oxidases.

It was also possible to determine whether the *P. denitrificans* NOR expressed in *E. coli* possessed an oxidase activity. This assay had previously only been possible with purified enzyme from *P. denitrificans*, because of the high levels of cytochrome *c* oxidase activity that are present in *P. denitrificans* membranes. *E. coli*, however, contains only quinol oxidases. Low levels of cytochrome *c* oxidase activity were detected in JM109(pNOREX, pEC86) (Table 3). There was no activity in cells that did not harbor pNOREX, confirming that the activities detected were due to NOR. Significantly, no cytochrome oxidase activity could be detected in NORs with the E125A or E198A substitution.

UV-Vis and EPR characterization of purified preparations of NOR^{REC}, NOR^{E125A}, and NOR^{E198A}. Recombinant NOR (NOR^{REC}), NOR^{E125A}, and NOR^{E198A} were purified from 15-liter L broth cultures of *E. coli* JM109 carrying the appropriate plasmid. The key step in the purification (described in Materials and Methods) involved the Q-Sepharose column from which NorCB eluted in two cytochrome-containing peaks at around 320 and 450 mM NaCl (Fig. 4). The first of these also contained large amounts of CcmE. In the second peak, NorCB was separated from other contaminating cytochromes and proteins. The NOR from this second peak was collected and separated on the IMAC column prior to characterization. The ratio of the two elution peaks from the Q-Sepharose column varied from preparation to preparation and influenced the final yield of purified protein, which was 5 to 10 mg per 15 liters of culture.

The patterns of NOR and oxidase activities in purified NOR^{REC} and NOR^{E125A} and NOR^{E198A} reflected those observed in membrane fractions. The turnover number for NOR^{REC} was comparable to that of native NOR (purified from *P. denitrificans*; NOR^{NAT}) determined in side-by-side experiments and was in the range of 40 to 70 electrons s^{-1} for NO reduction and 2 to 5 electrons s^{-1} for oxygen reduction. NOR^{E198A} had no detectable activity, and NOR^{E125A} had a low turnover number in the range of 3 to 5 electrons s^{-1} for NO reduction and around 1 electron s^{-1} for oxygen reduction. Comparison of the UV-Vis spectra of the oxidized forms of NOR^{NAT}, NOR^{REC}, NOR^{E125A}, and NOR^{E198A} revealed absorption features typical of the Soret band (411 nm) and αβ

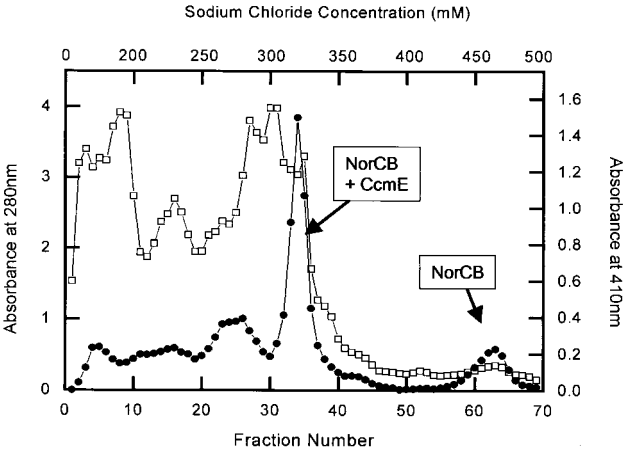


FIG. 4. The elution profile of NorCB from a Q-Sepharose column. Fraction volumes were 12 ml. A 0 to 200 mM salt gradient was run for the first 120 ml. This was then switched to a 200 to 500 mM gradient for the following 720 ml. ●, cytochrome absorbance at 410 nm; □, protein absorbance at 280 nm.

absorption bands (520 to 570 nm) of low-spin ferric hemes. On reduction with dithionite, the Soret band shifted to 420 nm and an increase in absorption at 550 and 560 nm characteristic of the α bands of low-spin ferrous *c* and *b* hemes, respectively, was observed. These features were essentially identical for all four enzymes, and representative spectra for NOR^{REC} are shown in Fig. 5A.

A major difference in the spectra of the four enzymes was observed in a charge transfer (CT) band at around 600 nm, which arises from the high-spin heme b_3 of the dinuclear center (11, 13). This CT band disappears on reduction of the enzyme, enabling it to be resolved most clearly in "oxidized minus reduced" spectra (Fig. 6). In NOR^{NAT} the wavelength for maximum absorbance (λ_{max}) is 595 nm (Fig. 6A and E), in NOR^{REC} (Fig. 6B and F) and NOR^{E125A} (Fig. 6D and H) it is red shifted to 606 nm, and in NOR^{E198A} (Fig. 6C and G) it is a mixture of the 595- and 606-nm forms. This CT band has been seen in the visible absorption spectrum of NOR in a number of published preparations, but variations in its position and intensity have been noted (8, 10, 11, 13, 15), and it is barely visible at all in enzymes from *Pseudomonas stutzeri* (12, 15). The position of this CT band cannot be correlated with differences in enzyme activity, and its variable λ_{max} probably arises from differences in the coordination environment of the high-spin heme *b* of the resting enzymes. The precise nature of these differences cannot be resolved at present, but is likely to involve the sixth coordination position of the ferric heme iron in the resting-state enzymes. Significantly, this difference in the resting state of the enzymes cannot account for the low activity of NOR^{E125A} and NOR^{E198A}, since NOR^{NAT} ("595" species) and NOR^{REC} ("606" species) both exhibit high enzymatic activity. The differences in resting states of the NOR enzymes are reminiscent of those of the *E. coli* cytochrome *bo* quinol oxidase, in which the active-site dinuclear center can exhibit considerable heterogeneity (30, 31).

NOR^{NAT}, NOR^{REC}, NOR^{E125A}, and NOR^{E198A} were also examined by EPR spectroscopy. The EPR spectrum of NOR^{NAT}

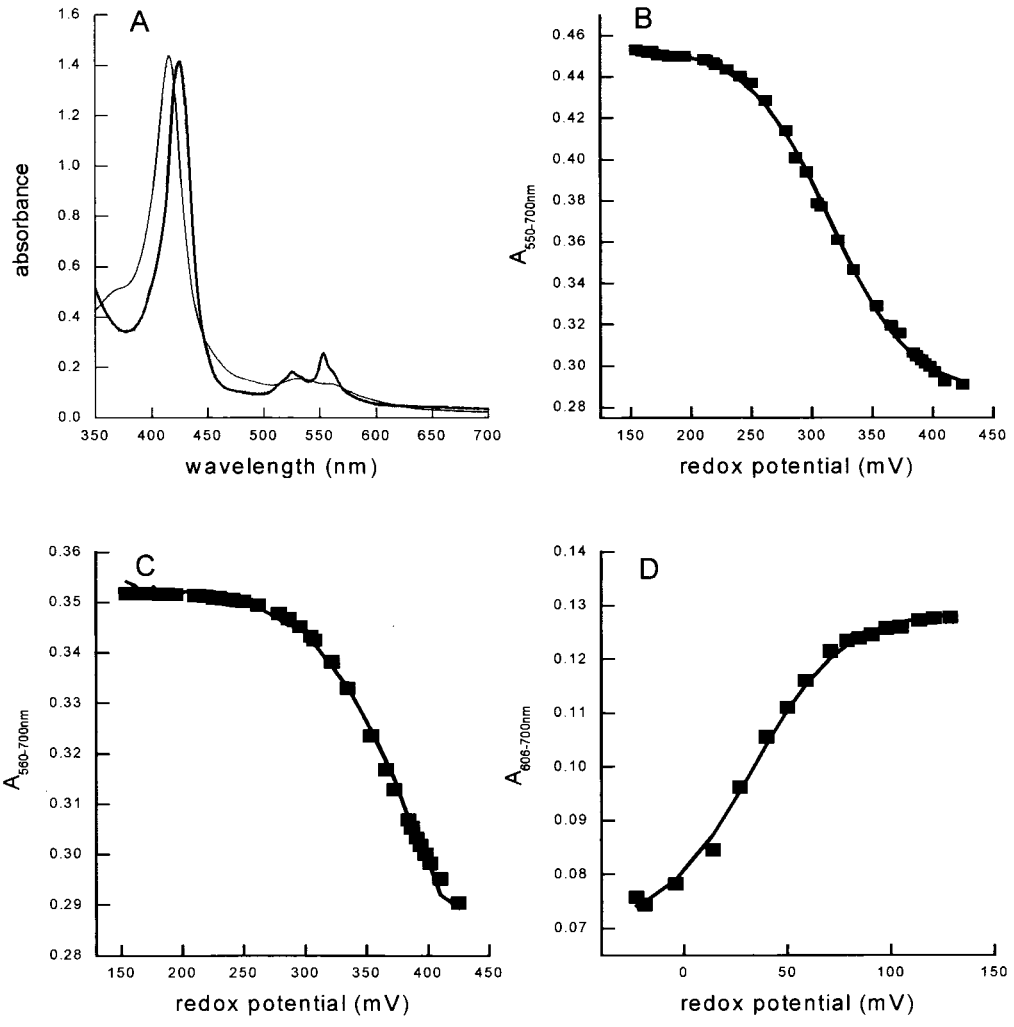


FIG. 5. Visible absorption spectra and redox potentiometry of NOR^{E198}. (A) "Air-oxidized" (thin line) and "dithionite-reduced" (thick line) absorption spectra. (B) Redox titration of NOR^{REC} monitored at 550 to 700 nm. (C) Redox titration of NOR^{REC} monitored at 560 to 700 nm. (D) Redox titration of NOR^{REC} monitored at 606 to 700 nm. The solid curves (B to D) show fits with $n = 1$ Nernstian curves using midpoint potentials of +316, +366, and +32 mV, respectively. All spectra and titrations were performed on samples incubated at 20°C in 20 mM Tris-HCl (pH 7.5)–0.02% dodecyl maltoside–340 mM NaCl–0.5 mM EDTA.

shows the presence of two low-spin ($s = ½$) ferric hemes (Fig. 7A), one with a typical rhombic spectrum ($g_z = 3.00, g_y = 2.25, g_x = 1.46$) and the other with a high g_{max} signal ($g_z = 3.55$). These signals have previously been ascribed to the low-spin *bis*-His-coordinated heme *b* and the low-spin His-Met-coordinated heme *c* of NOR and yield spin quantitations of 1:1 (5). Both signals could also be resolved in NOR^{REC} and were present at relative intensities similar to those of NOR^{NAT} (Fig. 7B). The major difference between NOR^{NAT} and NOR^{REC} was in the signals at $g \approx 6$ and $g \approx 4.2$. The signal at $g \approx 6$ arises from $s = 5/2$ high-spin ferric heme, most likely a small proportion of the heme *b* from the dinuclear center that is not magnetically coupled to the nonheme iron. The small increase of this uncoupled population of the dinuclear center in NOR^{REC} is also reflected by an increase in the structured $g \approx 4.2$ resonance that arises from the uncoupled Fe_B nonheme iron. The EPR spectra of NOR^{E125A} and NOR^{E198A} were essentially identical to that of the NOR^{REC} enzyme (not shown). Given that NOR^{REC} is fully active, the increased population of the non-magnetically coupled dinuclear center compared to that for NOR^{NAT} cannot account for the low activity of the NOR^{E125A} and NOR^{E198A} enzymes.

Spectropotentiometric characterization of NOR^{REC}, NOR^{E125A}, and NOR^{E198A}. Visible absorption spectra of NOR^{NAT}, NOR^{REC}, NOR^{E125A}, and NOR^{E198A} were collected at a number of defined redox potentials. In all cases, increases in the intensities of the α bands of the low-spin *c* heme (550 nm) and low-spin *b* heme (560 nm) were observed between ca. +400 and +200 mV. The absorption differences at 550 to 700 nm and 560 to 700 nm over this potential range were plotted as a function of redox potential and the midpoint potentials were derived by fitting single component $n = 1$ Nernstian curves to the data (a representative data set for NOR^{REC} is shown in Fig. 5B and C). For all four enzymes, the midpoint potentials of the low-spin *c* heme lay in the range of +310 to +322 mV (Table 4). Those of the low-spin *b* heme

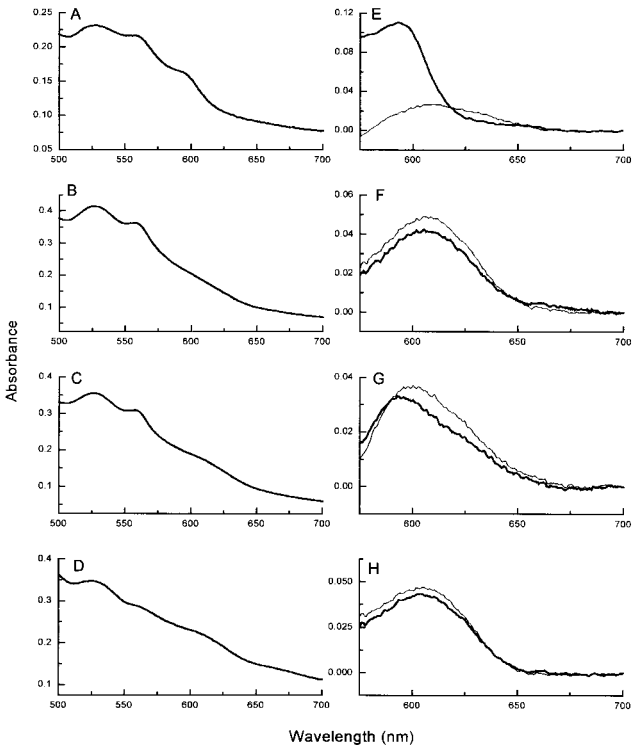


FIG. 6. Absorption spectra of NOR^{NAT}, NOR^{REC}, NOR^{E198A}, and NOR^{E125A} showing the CT band arising from high-spin heme b_3. (A to D) Air oxidized spectra of NOR^{NAT} (A), NOR^{REC} (B), NOR^{E198A} (C), and NOR^{E125A} (D). (E to H) Oxidized-minus-reduced difference spectra (thick lines) of NOR^{NAT} (E), NOR^{REC} (F), NOR^{E198A} (G), and NOR^{E125A} (H) and three-electron-reduced minus four-electron-reduced difference spectra (thin lines) of NOR^{NAT} (E), NOR^{REC} (F), NOR^{E198A} (G), and NOR^{E125A} (H). These were obtained by subtracting spectra collected at around −50 mV from spectra collected at around +140 mV.

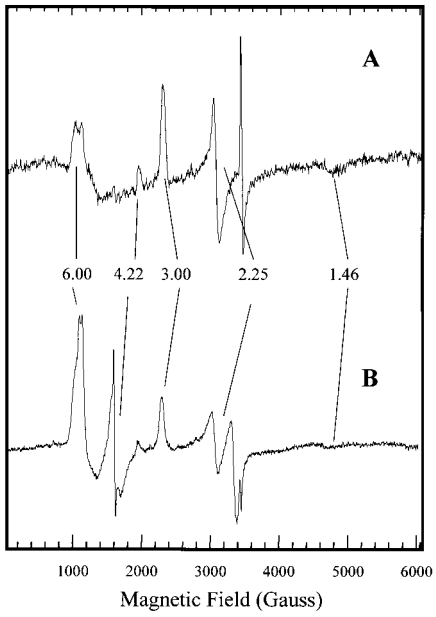


FIG. 7. X-band EPR spectra of air-oxidized native (A) and recombinant wild-type (B) NOR. The spectra were recorded at 10 K, 9.44 GHz, and 2 mW microwave power. The feature at ca. 2.01 in spectrum A may arise from a small (<1%) contaminant of a $[3Fe4S]^{1+}$ center from the *P. denitrificans* membrane-bound nitrate reductase.

were slightly higher, in the range +345 to +401 mV (Table 4). These values are all consistent with these low-spin heme centers mediating electron transfer from the physiological electron donors, periplasmic cytochrome c_{550} (E_m = +265 mV) and pseudoazurin (E_m = +230 mV), to the dinuclear center.

Analysis of the 595/606-nm CT band arising from high-spin heme *b* is more complex. We have previously reported that the λ_{max} of this CT band in NOR^{NAT} is shifted from 595 to 606 nm following reduction of the low-spin *c* and *b* hemes and Fe_B to yield the "three-electron-reduced" enzyme (11). We have argued that this shift in λ_{max} reflects a change in the coordination environment of high-spin heme *b* that accompanies the reduction of the spectroscopically silent Fe_B (E_m = +320 mV) (11). This idea is supported by the present study, and an illustrative absorption spectrum of the three-electron-reduced form collected at +140 mV is presented in Fig. 6E. This reveals that the extinction coefficient of the 595-nm CT band is around fivefold greater (approximately 6 mM^{-1} cm^{-1}) than that of the 606-nm CT band (approximately 1.2 mM^{-1} cm^{-1}).

In the light of these data, consideration of the oxidized spectrum of NOR^{E198A} (Fig. 6G) suggests that around 20% of the enzyme is in the 595-nm form and around 80% is in the 606-nm form. The spectrum is largely unchanged when the E_h is lowered to +250 mV (not shown). However, lowering the E_h from +250 to +140 mV results in the loss of the 595-nm feature of the NOR^{E198A} spectrum and a 20% increase in the intensity of the 605-nm band (Fig. 6G). This increase can be ascribed to the reduction of Fe_B in the 20% of the enzyme population that is in the 595-nm form. The absorption changes are too small to allow the plotting of a Nernstian curve, but the data place the E_m of NOR^{E198A} Fe_B at around +200 mV, considerably lower than that of the Fe_B in NOR^{NAT} (+320 mV) (11). Both NOR^{REC} and NOR^{E125A} are in an almost homogenous 606-nm form in the fully oxidized enzyme, and reduction to +140 mV does not significantly change the form of this CT band (Fig. 6F and H). Thus, consideration of NOR^{NAT}, NOR^{REC}, NOR^{E125A}, and NOR^{E198A}, poised at around +140 mV, reveals that all four enzymes are in a similar spectroscopic state in which it is likely that low-spin hemes *c* and *b* and Fe_B are reduced and high-spin heme *b* remains oxidized in a 606-nm state. The 606-nm bands of NOR^{NAT}, NOR^{REC}, NOR^{E125A}, and NOR^{E198A} all disappear as the E_h is lowered from +140 mV to −80 mV. This reflects the reduction of high-spin heme b_3 to yield the fully (four-electron-reduced) enzyme. The reduction of this center can be fitted to single *n* =

TABLE 4. Midpoint redox potentials of the NOR heme centers

Enzyme	E_m (mV) of:		
	c heme	Low-spin *b* heme	High-spin *b* heme
NOR^{NAT}	310	345	40
NOR^{REC}	316	366	32
NOR^{E125A}	320	380	20
NOR^{E198A}	322	401	2

1 Nernstian curves (Fig. 5D; shown for NOR^{REC} only) which yield midpoint redox potentials that lie in the range of +2 to +40 mV for all four forms of NOR studied (Table 4). We have previously argued that the low potential of high-spin heme b_3 provides a thermodynamic barrier that prevents reduction of this site during the catalytic cycle (11). If heme b_3 becomes reduced, then, potentially, a dead-end ferrous nitrosyl complex could form. This thermodynamic barrier would exist in each of the four enzyme types discussed in this paper, since in each case the difference in reduction potential between low-spin heme *b* and the active-site heme is at least 300 mV.

In conclusion, the accumulated spectroscopic data for purified NOR^{NAT}, NOR^{REC}, NOR^{E125A}, and NOR^{E198A} demonstrate that the inactivity of NOR^{E125A} and NOR^{E198A} is unlikely to be accounted for by either enzyme instability, failure to insert cofactors, or perturbation of the redox potentials of the heme cofactors. The location of E198 one helical turn below a likely Fe_B ligand (H194) in helix VI strongly implicates it as contributing to the immediate environment of Fe_B, which is also consistent with the preliminary suggestion that the E_m of the Fe_B is perturbed in NOR^{E198A}. Certainly Fe_B, which unlike Cu_B prefers an octahedral coordination environment, is likely to have at least one extra protein-derived ligand. However, E198 is also well placed to serve as a base for NO radical chemistry or for delivery of catalytic protons. More-detailed spectroscopic and electrochemical studies on NOR^{E198A} will now be undertaken to explore these possibilities. The importance of E125 is perhaps most surprising given that it is located towards the periplasmic face of helix IV and is not predicted to be close to the Fe_B. Previous studies have indicated that NOR is not proton translocating (2, 3, 23) and that the two chemical protons required for NO reduction are taken up from the periplasm. Given the conservation of E125, a role in proton uptake should be considered.

ACKNOWLEDGMENTS

We are grateful to James Moir for valuable ideas at the outset of this work; Lola Roldán, Andrew Thomson, and Myles Cheeseman for discussion and help in strategy development; Karen Grönberg and Jeremy Thornton for provision of native NOR; Matti Saraste for provision of antibody to NorB; Werner Klipp, Rob van Spanning, and Lynda Thöney-Meyer for provision of strains and plasmids; and Adam Baker and Louise Prior for collecting the EPR spectra.

The work was funded by BBSRC grant 83/C10160, the award of a BBSRC/EPSRC special studentship to G.B., the UEA innovation fund, and CEC grant EC BIO-CT98-0507. N.J.W. is a Wellcome Trust University Award Lecturer (054798/Z/98Z).

REFERENCES

1. **Adelroth, P., R. B. Gennis, and P. Brzezinski.** 1998. Role of the pathway through K(I-362) in proton transfer in cytochrome c oxidase from *R. sphaeroides.* Biochemistry **37:**2470–2476.
2. **Bell, L. C., D. J. Richardson, and S. J. Ferguson.** 1992. Identification of nitric oxide reductase activity in *Rhodobacter capsulatus*: the electron transport pathway can either use or bypass both cytochrome c_2 and the cytochrome bc_1 complex. J. Gen. Microbiol. **138:**437–443.
3. **Carr, G. J., M. D. Page, and S. J. Ferguson.** 1989. The energy-conserving nitric-oxide-reductase system in *Paracoccus denitrificans*. Distinction from the nitrite reductase that catalyses synthesis of nitric oxide and evidence from trapping experiments for nitric oxide as a free intermediate during denitrification. Eur. J. Biochem. **179:**683–692.
4. **Castresana, J., and M. Saraste.** 1995. Evolution of energetic metabolism: the respiration-early hypothesis. Trends Biochem. Sci. **20:**443–448.
5. **Cheesman, M. R., W. G. Zumft, and A. J. Thomson.** 1998. The MCD and EPR of the heme centers of nitric oxide reductase from *Pseudomonas stutzeri*: evidence that the enzyme is structurally related to the heme-copper oxidases. Biochemistry **37:**3994–4000.
6. **de Vries, G. E., N. Harms, J. Hoogendijk, and A. H. Stouthamer.** 1989. Isolation and characterisation of *Paracoccus denitrificans* mutants with increased conjugation frequencies and pleiotropic loss of a (nGATCn) DNA modifying property. Arch. Microbiol. **152:**52–57.
7. **Fellay, R., J. Frey, and H. Krisch.** 1987. Interposon mutagenesis of soil and water bacteria: a family of DNA fragments designed for in vitro insertional mutagenesis of gram-negative bacteria. Gene **52:**147–154.
8. **Fujiwara, T., and Y. Fukumori.** 1996. Cytochrome *cb*-type nitric oxide reductase with cytochrome *c* oxidase activity from *Paracoccus denitrificans* ATCC 35512. J. Bacteriol. **178:**1866–1871.
9. **Gerhus, E., P. Steinrucke, and B. Ludwig.** 1990. *Paracoccus denitrificans* cytochrome c_1 gene replacement mutants. J. Bacteriol. **172:**2392–2400.
10. **Girsch, P., and S. de Vries.** 1997. Purification and initial kinetic and spectroscopic characterization of NO reductase from *Paracoccus denitrificans*. Biochim. Biophys. Acta **1318:**202–216.
11. **Grönberg, K. L., M. D. Roldán, L. Prior, G. Butland, M. R. Cheesman, D. J. Richardson, S. Spiro, A. J. Thomson, and N. J. Watmough.** 1999. A low-redox potential heme in the dinuclear center of bacterial nitric oxide reductase: implications for the evolution of energy-conserving heme-copper oxidases. Biochemistry **38:**13780–13786.
12. **Heiss, B., K. Frunzke, and W. G. Zumft.** 1989. Formation of the N-N bond from nitric oxide by a membrane-bound cytochrome *bc* complex of nitrate-respiring (denitrifying) *Pseudomonas stutzeri*. J. Bacteriol. **171:**3288–3297.
13. **Hendriks, J., A. Warne, U. Gohlke, T. Haltia, C. Ludovici, M. Lubben, and M. Saraste.** 1998. The active site of the bacterial nitric oxide reductase is a dinuclear iron center. Biochemistry **37:**13102–13109.
14. **Hutchings, M. I., and S. Spiro.** 2000. The nitric oxide regulated *nor* promoter of *Paracoccus denitrificans*. Microbiology **146:**2635–2641.
15. **Kastrau, D. H., B. Heiss, P. M. Kroneck, and W. G. Zumft.** 1994. Nitric oxide reductase from *Pseudomonas stutzeri*, a novel cytochrome *bc* complex. Phospholipid requirement, electron paramagnetic resonance and redox properties. Eur. J. Biochem. **222:**293–303.
16. **Konstantinov, A. A., S. Siletsky, D. Mitchell, A. Kaulen, and R. B. Gennis.** 1997. The roles of the two proton input channels in cytochrome *c* oxidase from *Rhodobacter sphaeroides* probed by the effects of site-directed mutations on time-resolved electrogenic intraprotein proton transfer. Proc. Natl. Acad. Sci. USA **94:**9085–9090.
17. **Kwiatkowski, A. V., W. P. Laratta, A. Toffanin, and J. P. Shapleigh.** 1997. Analysis of the role of the *nnrR* gene product in the response of *Rhodobacter sphaeroides* 2.4.1 to exogenous nitric oxide. J. Bacteriol. **179:**5618–5620.
18. **Moenne Loccoz, P., and S. deVries.** 1998. Structural characterisation of the catalytic high spin heme *b* of nitric oxide reductase. J. Am. Chem. Soc. **120:**5147–5152.
19. **Reincke, B., L. Thony-Meyer, C. Dannehl, A. Odenwald, M. Aidim, H. Witt, H. Ruterjans, and B. Ludwig.** 1999. Heterologous expression of soluble fragments of cytochrome *c*552 acting as electron donor to the *Paracoccus denitrificans* cytochrome *c* oxidase. Biochim. Biophys. Acta **1411:**114–120.
20. **Roldan, M. D., H. J. Sears, M. R. Cheesman, S. J. Ferguson, A. J. Thomson, B. C. Berks, and D. J. Richardson.** 1998. Spectroscopic characterization of a novel multiheme *c*-type cytochrome widely implicated in bacterial electron transport. J. Biol. Chem. **273:**28785–28790.
21. **Saraste, M., and J. Castresana.** 1994. Cytochrome oxidase evolved by tinkering with denitrification enzymes. FEBS Lett. **341:**1–4.
22. **Schulz, H., R. A. Fabianek, E. C. Pellicioli, H. Hennecke, and L. Thöny-Meyer.** 1999. Heme transfer to the heme chaperone CcmE during cytochrome *c* maturation requires the CcmC protein, which may function independently of the ABC-transporter CcmAB. Proc. Natl. Acad. Sci. USA **96:**6462–6467.
23. **Shapleigh, J. P., and W. J. Payne.** 1985. Nitric oxide-dependent proton translocation in various denitrifiers. J. Bacteriol. **163:**837–840.
24. **Simon, R., V. Priefer, and A. Pühler.** 1983. Vector plasmids for in vivo and in vitro manipulations of gram-negative bacteria, p. 99–108. *In* A. Pühler (ed.), Molecular biology of bacteria and plant interaction. Springer-Verlag KG, Berlin, Germany.
25. **Stoll, R., M. D. Page, Y. Sambongi, and S. J. Ferguson.** 1996. Cytochrome *c*550 expression in *Paracoccus denitrificans* strongly depends on growth condition: identification of promoter region for *cycA* by transcription start analysis. Microbiology **142:**2577–2585.
26. **van der Oost, J., A. P. de Boer, J. W. de Gier, W. G. Zumft, A. H. Stouthamer, and R. J. van Spanning.** 1994. The heme-copper oxidase family consists of three distinct types of terminal oxidases and is related to nitric oxide reductase. FEMS Microbiol. Lett. **121:**1–9.
27. **van Spanning, R. J., E. Houben, W. N. Reijnders, S. Spiro, H. V. Westerhoff, and N. Saunders.** 1999. Nitric oxide is a signal for NNR-mediated transcription activation in *Paracoccus denitrificans*. J. Bacteriol. **181:**4129–4132.
28. **van Spanning, R. J., C. W. Wansell, W. N. Reijnders, N. Harms, J. Ras, L. F. Oltmann, and A. H. Stouthamer.** 1991. A method for introduction of unmarked mutations in the genome of *Paracoccus denitrificans*: construction of strains with multiple mutations in the genes encoding periplasmic cytochromes c_{550}, c_{551i}, and c_{553i}. J. Bacteriol. **173:**6962–6970.

29. **Watmough, N. J., G. Butland, M. R. Cheesman, J. W. Moir, D. J. Richardson, and S. Spiro.** 1999. Nitric oxide in bacteria: synthesis and consumption. Biochim. Biophys. Acta **1411:**456–474.
30. **Watmough, N. J., M. R. Cheesman, C. S. Butler, R. H. Little, C. Greenwood, and A. J. Thomson.** 1998. The dinuclear center of cytochrome bo_3 from *Escherichia coli*. J. Bioenerg. Biomembr. **30:**55–62.
31. **Watmough, N. J., M. R. Cheesman, R. B. Gennis, C. Greenwood, and A. J. Thomson.** 1993. Distinct forms of the haem *o*-Cu binuclear site of oxidised cytochrome *bo* from *Escherichia coli*. Evidence from optical and EPR spectroscopy. FEBS Lett. **319:**151–154.
32. **Zaslavsky, D., and R. B. Gennis.** 1998. Substitution of lysine-362 in a putative proton-conducting channel in the cytochrome *c* oxidase from *Rhodobacter sphaeroides* blocks turnover with O_2 but not with H_2O_2. Biochemistry **37:** 3062–3067.