Overexpression, Purification, and Characterization of ProQ, a Posttranslational Regulator for Osmoregulatory Transporter ProP of *Escherichia coli*†

Michelle N. Smith,‡,§ Rebecca A. Crane,‡ Robert A. B. Keates,‡,§,|| and Janet M. Wood*,‡,§

Department of Microbiology, Department of Chemistry and Biochemistry, and Guelph-Waterloo Centre for Graduate Work in Chemistry and Biochemistry, University of Guelph, Guelph, ON N1G 2W1, Canada

Received July 7, 2004; Revised Manuscript Received August 10, 2004

ABSTRACT: ProP is an osmosensor and osmoregulatory transporter in *Escherichia coli*. Osmotic activation of ProP is attenuated 5-fold in the absence of soluble protein ProQ, but *proQ* lesions do not influence *proP* transcription or ProP levels. The mechanism by which ProQ amplifies ProP activity is unknown. Putative *proQ* orthologues are found in Gram-negative bacteria (only), but none have known functions. ProQ was overexpressed to low and high levels with and without a C-terminal histidine tag (His_6). Plasmid-encoded ProQ or ProQ-His_6 complemented in-frame chromosomal deletion Δ*proQ676*, restoring ProP activity. After overexpression, both proteins were poorly soluble unless cells were lysed in media of high salinity. ProQ copurified with DNA binding proteins of similar size (HU and a histone-like protein) by ion exchange and exclusion chromatographies, whereas ProQ-His_6 could be purified to homogeneity by nickel chelate affinity chromatography. Sequence-based analysis and modeling suggest that ProQ includes distinct N- and C-terminal domains linked by an unstructured sequence. The N-terminal domain can be modeled on the crystal structure of α-helical RNA binding protein FinO, whereas the C-terminal domain can be modeled on an SH3-like domain (β-structure). Both ProQ and ProQ-His_6 appeared to be monomeric, though the higher Stokes radius of ProQ-His_6 may reflect altered domain interactions. The measured secondary structure content of ProQ (circular dichroism (CD) spectroscopy) contrasted with sequence-based prediction but was as expected if the spectrum of the C-terminal domain is analogous to those reported for SH3 domains. The CD spectrum of ProQ was pH- but not NaCl-sensitive.

As unicellular organisms, bacteria face particular challenges to survival in harsh environments. They modulate their cytoplasmic composition to survive rapid changes in physical parameters, including temperature, pressure, pH, and osmolality. An osmotic downshift (a decrease in external solute concentration) causes water influx and increases turgor pressure and membrane strain, diluting intracellular solutes and causing them to be released via mechanosensitive channels (*1*). An osmotic upshift (an increase in external solute concentration) causes water efflux, a decrease in cytoplasmic volume and turgor pressure, membrane strain, and macromolecular crowding and activates osmoregulatory transporters (*2*−*4*). Multiple transporters are available to mediate the uptake of either K^+ or organic osmoprotectants, the latter restoring cellular hydration without disrupting cellular functions (*2*).

ProP of *Escherichia coli* is a 500 amino acid integral membrane protein and a member of the major facilitator superfamily (MFS)[1] (*5*). It can be modeled on the structure of anion antiporter GlpT (*6*). As an osmoprotectant/H^+ symporter, ProP transports a wide variety of zwitterionic organic solutes such as glycine betaine and proline (*7*). It can be activated by an osmotic upshift in intact cells and after purification and reconstitution into proteoliposomes (*8*, *9*). Thus, no cellular component other than membrane lipid is essential for its osmotic activation in the proteoliposome system. *E. coli* ProP has an extended C-terminal domain that was predicted to form an α-helical coiled-coil (*5*). Dimerization of ProP in vivo is indeed associated with formation of an antiparallel, homodimeric α-helical coiled-coil by adjacent cytoplasmic ProP C-termini (*10*−*13*). Variants of ProP that lack the coiled-coil forming peptide (26 C-terminal amino acids) or harbor amino acid replacements that disturb

† This work was supported by Research Grant OGP0000508 awarded to J.M.W. and an Undergraduate Summer Research Award granted to M.S. by the Natural Sciences and Engineering Research Council of Canada, as well as an Ontario Graduate Scholarship in Science and Technology and an Ontario Graduate Scholarship awarded to R.A.C.

* Corresponding author. Mailing address: Department of Microbiology, University of Guelph, Guelph, ON N1G 2W1, Canada. Telephone: 519 824 4120 ext. 53866. Fax: 519 837 1802. E-mail: jwood@uoguelph.ca.

‡ Department of Microbiology.

§ Guelph-Waterloo Centre for Graduate Work in Chemistry and Biochemistry.

|| Department of Chemistry and Biochemistry.

[1] Abbreviations: BCA, bicinchoninic acid; BLAST, basic local alignment search tool; CD, circular dichroism; DTT, dithiothreitol; FPLC, fast protein liquid chromatography; IPTG, isopropyl-β-D-thiogalactopyranoside; NCBI, National Center for Biotechnology Information; LB, Luria broth; MALDI-TOF MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry; MOPS, 3-[*N*-morpholino]-propane sulfonic acid; MFS, major facilitator superfamily; MW, molecular weight; Ni(NTA), nickel nitrilotriacetate; PCR, polymerase chain reaction; PVDF, polyvinylidene fluoride; R_S, Stokes radius; SDS, sodium dodecyl suphate; SDS−PAGE, SDS−polyacrylamide gel electrophoresis; SH3, Src homology 3.

10.1021/bi048561g CCC: $27.50
Published on Web 09/25/2004

Table 1: Occurrence of *proP* and *proQ* in Sequenced Bacterial Genomes[a]

		putative orthologue			
		ProP		ProQ	
gram stain	organism	% identity	% overlap	% identity	% overlap
-	*Escherichia coli* K-12	100	100	100	100
-	*Salmonella enterica* serovar Typhimurium	97	100	92	100
-	*Shigella flexneri*	98	100	99	100
-	*Pseudomonas putida*	82	99		
-	*Agrobacterium tumefaciens*	65	98	40	31
-	*Xanthomonas campestris*	52	92		
-	*Yersinia pestis*	32	85	74	100
-	*Pasteurella multocida*	30	85	45	99
-	*Haemophilus influenzae*	29	83	43	97
-	*Shewanella oneidensis*	26	27	39	100
-	*Photorhabdus luminescens*	23	71	63	100
-	*Vibrio cholerae*	22	72	53	99
-	*Haemophilus ducreyi*	21	86	43	63
+	*Streptomyces coelicolor*	51	87		
+	*Bacillus cereus*	43	91		
+	*Corynebacterium glutamicum*	39	92		
+	*Mycobacterium avium*	39	89		
+	*Staphylococcus aureus*	37	89		

[a] Putative ProP or ProQ orthologues or both were found via BLAST searches (blastp) of microbial genomes available in the NCBI database. Percent overlap is the number of residues of the query sequence aligned continuously divided by the length of the query sequence. Putative ProP orthologues with sequence identity lower than 37% are listed only if a putative ProQ orthologue was also detected. Among the ProP sequences listed, ProP function has been demonstrated for the proteins from *E. coli* K-12 (*8*), *Salmonella enterica* serovar Typhimurium (*52*), *Agrobacterium tumefaciens* (Y. Tsatskis and J. M. Wood, unpublished data), and *Corynebacterium glutamicum* (*20*). The accession numbers for the sequence orthologous to ProP and ProQ (where it is present) are as follows. For *Escherichia coli K-12*, ProP is AAC44538 and ProQ is NP_416345. For *Salmonella enterica* serovar Typhimurium, ProP is AAL23114 and ProQ is NP_460802. For *Shigella flexneri*, ProP is NP_709830 and ProQ is NP_707288. For *Pseudomonas putida*, ProP is NP_745058. For *Agrobacterium tumefaciens*, ProB is AAL45120 and Ydh is NP_059699. For *Xanthomonas campestris*, PM1259 is NP_635478. For *Yersinia pestis*, YP00465 is NP_404107 and YP01704 is CAC90524. For *Pasteurella multocida*, PM1259 is NP_246196 and ProQ is AAK02352. For *Haemophilus influenzae*, HI0281 is NP_438449 and ProQ is P44286. For *Shewanella oneidensis*, SO4029 is NP_719559 and SO2602 is NP_718188. For *Photorhabdus luminescens*, PLU3134 is NP_930365 and ProQ is NP_929918. For *Vibrio cholerae*, VCA0669 is NP_233058 and VC1497 is AAF94652. For *Haemophilus ducreyi*, HD0835 is NP_873340 and ProQ is NP_873499. For *Streptomyces coelicolor*, SC02336 is NP_626583. For *Bacillus cereus*, BCE3056 is NP_979359. For *Corynebacterium glutamicum*, NCg12961 is NP_602258. For *Mycobacterium avium*, MAP01336 is NP_959067. For *Staphylococcus aureus*, ProP is NP_373784.

coiled-coil formation are expressed but respond weakly, transiently, or both to osmotic upshifts in vivo (*10*).

A Tn*5* insertion in locus *proQ* was identified as conferring resistance to toxic proline analogues on *E. coli* (*14*, *15*). Mutation *proQ220*::Tn*5* impaired the osmotic activation of ProP but did not impair transcription or translation of *proP*, the latter determined by studying the β-galactosidase activity associated with a *proP::lacZ* fusion and by Western blotting (*15*, *16*). Lesion *proQ220*::Tn*5* did not exert polar effects on downstream locus *prc* (*16*). Although *proQ* lesions impair ProP activity in vivo, ProP-His_6 purified and reconstituted into liposomes functions as an osmosensor and a transporter, mediating the accumulation of compatible solutes in high-osmolality media (*9*, *17*, *18*). The turnover number (k_{cat}) for ProP-His_6 in proteoliposomes corresponds better with that observed in ProQ-deficient bacteria than that in wild-type *E. coli* (*18*). Thus ProQ is believed to amplify the osmolality response of ProP via posttranslational modification or by linking the osmotic activation of ProP to other processes within the cell (*2*). Although other transporters are known to act as osmosensors and osmoregulators, none other than ProP is known to require an auxiliary protein for full function (*4*, *19*).

Although *proQ* lesions most dramatically impair the osmotic activation of ProP, they also slightly diminish the activities of some transporters that are not implicated in osmoregulation (*15*). Putative orthologues of ProP and ProQ were sought by BLAST searching of the microbial genomes included in the NCBI database (Table 1). Not all genomes encode putative orthologues of these proteins and where *proQ* and *proP* do cooccur, they are not adjacent. It is difficult to unambiguously identify ProP orthologues given the high background sequence similarity among MFS members, regardless of their function. For example, transporters ProP of *E. coli* and *Corynebacterium glutamicum* are both osmosensors and osmoregulatory betaine transporters although they share only 39% sequence identity (*20*). On the other hand, KgtP and ShiA of *E. coli* are also quite similar in sequence to ProP (30% and 27% identity, respectively) yet KgtP is an α-ketoglutarate transporter (*21*) and ShiA is a shikimate transporter (*22*). ProP is certainly present without ProQ in at least one organism (*C. glutamicum*) (*20*) (Table 1 and Supporting Information, Figure 1S). Putative ProQ orthologues exist in many Gram-negative bacteria, but none has been found in a Gram-positive organism (Table 1). ProQ may also be present without ProP, since the putative ProP orthologues identified in some Gram-negative organisms (e.g., those in *Pasteurella multocida*, *Haemophilus influenzae*, *Shewanella oneidensis*, *Photorhabdus luminescens*, *Vibrio cholerae* and *H. ducreyi*, Table 1 and Figure 1S) are no more similar in sequence to ProP than is ShiA. Thus the role of ProQ may not be limited to the regulation of ProP.

ProQ is predicted to be a 232 amino acid, 26 kDa protein, both basic and hydrophilic in nature, with an estimated pI of 9.7 (*16*). Neither *E. coli* ProQ nor any of its orthologues has been purified or analyzed on a structural or functional level. This paper reports the overexpression, purification, and characterization of native and histidine-tagged ProQ (ProQ

Table 2: *Escherichia coli* Strains and Plasmids

strain/plasmid	relevant genotype	source/derivation[a]
DH5α	F$^-$ *ϕ80dlacZΔM15* Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (r_K^- m_K^+) *sup*E44 λ^- *thi*-1 *gyr*A *rel*A1	BRL (Burlington, ON, Canada)
K-12	F$^-$ *rph*−1 λ^-	CGSC
RM2	F$^-$ *trp lacZ thi rpsL* Δ*(putPA)101*	*53*
SG13009	F$^-$ *his pyr*D Δ(*lon*−100) *rps*L	*54*
WG174	F$^-$ *trp lacZ rpsL thi* Δ*putPA101 proQ220*::Tn*5*	*14*
WG210	F$^-$ *trp rpsL lacZ thi* Δ*putPA101 proU205*	*8*
WG822	K-12 Δ*proQ*676	this work
WG914	WG210 Δ*proQ*676	this work
pBAD24	vector, AraC-controlled ara_{BAD} promoter	*55*
pDC77	*proQ* in vector pBAD24 (encodes ProQ)	*16*
pKO3	RepA (T_S) Cm^r *sac*B$^+$	*26*
pMS1	*proQ* in vector pQE-60 (encodes ProQ-His_6)	this work
pMS2	*proQ* in vector pQE-60 (encodes ProQ)	this work
pQE-60	vector, IPTG-controlled phage T5 promoter	Qiagen
pRAC2	Δ*proQ676* with flanking DNA in vector pKO3	this work
pSS1	*proQ* in vector pQE−60 (encodes ProQ truncate)	this work

[a] CGSC is *E. coli* Genetic Stock Centre.

and ProQ-His_6, respectively). This work was done to determine whether ProQ copurifies with other proteins and whether such preparations would support further structural and functional characterization of ProQ.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Plasmids. Strains and plasmids are listed in Table 2. Plasmid isolation was performed using QIAprep Spin Miniprep Kit (Qiagen, Mississauga, ON). Routine DNA manipulation, plasmid construction, electrophoresis, and transformation were carried out as described previously (*23*, *24*). The polymerase chain reaction (PCR) was performed as described previously (*25*) using *Taq* DNA polymerase or *Pfu* turbo (Invitrogen, Burlington, ON). The Molecular Biology Supercenter (University of Guelph, Guelph, ON) performed DNA sequencing to verify the sequences of all plasmid constructs and as otherwise specified.

The *proQ* deletion vector pRAC2 was created via two-step PCR (*26*) using primers listed in Table 1S (see Supporting Information) and *Taq* DNA polymerase (Invitrogen, Burlington, ON). The flanking 5′ (596 bp) and 3′ (604 bp) DNA fragments were amplified using primer proQNo with proQNi and primer proQCo with proQCi, respectively, and chromosomal DNA from *E. coli* RM2 as template. The resulting PCR products were mixed and amplified using primers proQNo and proQCo to generate a proQNo−proQCo fragment (1175 bp) encoding Δ*proQ676*. This fragment was restricted with *Bam*HI and ligated into *Bam*HI-cleaved plasmid pKO3 (*26*). Chromosomal *proQ* was deleted from *E. coli* strains K-12 and WG210 by allelic exchange using the resulting plasmid, pRAC2, as described (*26*). Putative *proQ* deletion mutants were screened by PCR analysis using primer proQNo with primer proQCo and by radial streak analysis performed with toxic proline analogue 3,4-dehydro-D,L-proline (*15*). Vector pRAC2 had been designed to replace *proQ* with an ORF encoding the peptide MENKCLSLVDGSLIVRAEHLVF. Repeated sequencing of pRAC2 and PCR products representing the chromosomal Δ*proQ676* allele revealed an unexpected single base deletion, resulting in an altered ORF encoding the peptide MENQPKCLCLGMGL. BLAST analysis has revealed no homologue of this peptide.

To construct plasmid pMS1 (encoding ProQ-His_6) the *proQ* ORF was amplified with primers ProQ1 and ProQ3 (see Supporting Information, Table 1S) using plasmid pDC77 as template. Primer ProQ3 created an *Nco*I site at the 5′ end of the ORF, whereas primer ProQ1 introduced a *Bgl*II site 3′ of the ORF and changed the stop codon (TGA) to AGA (arginine) in frame with the six histidine codons and the termination codon provided by the vector. The PCR product and pQE-60 were cleaved with *Nco*I and *Bgl*II (Invitrogen, Burlington, ON); the desired DNA fragments purified, mixed, ligated, and transformed into *E. coli* SG13009 pREP4. To construct pSS1, the *proQ* ORF was amplified using primers ProQ2 and ProQ3 (see Supporting Information, Table 1S) with plasmid pDC77 as template. Primer ProQ2 created a *Hin*dIII restriction site 3′ of the ORF and primer ProQ3 created a *Bgl*II site 5′ of the ORF. The PCR product and pQE-60 were digested with *Bgl*II and *Hin*dIII (ensuring that the vector-encoded His-tag would not be added to the *proQ* open reading frame), purified, mixed, ligated using T4 DNA ligase (Invitrogen, Burlington, ON), and transformed into *E. coli* DH5α. Sequence analysis revealed that the resulting plasmid, pSS1, encoded a ProQ variant that was truncated because stop codon TGA replaced that encoding R57 to create allele *proQ213*. To construct plasmid pMS2 (encoding ProQ), that mutation was corrected using the QuickChange site-directed mutagenesis kit (Stratagene, LaJolla, CA) with primers pSS1-1 and pSS1-2 (see Supporting Information, Table 1S), and the resulting plasmid was transformed in *E. coli* SG13009 pREP4.

Culture Media. Bacteria were cultivated in LB (*27*) or a 3-[*N*-morpholino]-propane sulfonic acid (MOPS)-based minimal medium (*28*). Ampicillin (Amp, 100 μg/mL), kanamycin (Km, 50 μg/mL), or both were added as required to maintain plasmids based on vector pBAD24 (Amp), pREP4 (Km), or pQE-60 (Amp).

Protein Expression Tests. Bacteria were inoculated into LB (10 mL) and incubated overnight at 37 °C. Bacteria containing plasmid pDC77 were subcultured into LB (100 mL) supplemented with arabinose (0.2%, w/v) to attain an optical density at 600 nm (OD_{600}) of 0.4 and incubated, with shaking, at 37 °C until the OD_{600} reached 1. Bacteria harboring plasmids pREP4 and pQE60, pMS1, or pMS2 were subcultured into LB to attain an OD_{600} of 0.4 and incubated, with shaking, at 37 °C until the OD_{600} reached 0.7. IPTG (isopropyl-β-D-thiogalactopyranoside, 1 mM) was added and incubation continued for 2 or 4 h. Cells were harvested by centrifugation (12 000 × *g* for 10 min at 4 °C), and the resulting cell pellets were washed in 0.85% NaCl. Cell pellets were resuspended in lysis buffer (75 mM potassium phosphate, pH 7.4, 1 mM dithiothreitol (DTT), 5 mM $MgSO_4$, 30 μg/mL DNase, 1 mg/mL lysozyme) (5 mL per gram of wet weight), incubated on ice for 30 min, lysed by sonication on ice (macrotip, Sonicator XL2020 (Heat Systems, NY), power set to 4, 4 min total alternating 10 s on, 10s off and centrifuged (12 000 × *g*, 20 min, 4 °C) to produce soluble and particulate fractions, the latter resuspended in an equal

volume of 0.85% NaCl. To test the solubilities of ProQ and ProQ-His_6 expressed from pQE60-based vectors in buffers with a variety of salt concentrations, the relevant strains were cultured, and protein expression was induced as described above. The OD_{600} was measured, and aliquots were taken that corresponded with 1 mL of culture at an OD_{600} of 1.0. Cells were harvested from these aliquots by centrifugation (13 000 rpm, 5 min in the microfuge). The pellets were resuspended in appropriate buffers (100 μL of 50 mM Tris-HCl, 5 mM EDTA, pH 7.4, supplemented with lysozyme (1 mg/mL) and NaCl as specified), lysed by four freeze−thaw cycles (quick freezing on dry ice for 3 min, thawing at 42 °C for 3 min with intermittent vortex mixing), and centrifuged (13 000 rpm, 5 min in the microfuge) to produce soluble and particulate fractions. The latter was resuspended in 100 μL of the same buffer.

Protein Electrophoresis and Western Blotting. SDS−polyacrylamide gel electrophoresis was performed as described (*29*) using a Mini-Protean slab cell (Bio-Rad, Mississauga, ON) with gels comprised of 12% or 10% (w/v) acrylamide with 1.1% or 0.9% (w/v) bis-acrylamide, respectively. Gels were stained with Gel-Code Blue (Pierce, Rockford, IL) according to the manufacturer's instructions. Western Blotting was performed as described by Kunte et al. (*16*) (transfer for 1 h at 80 V) using purified anti-ProQ antibodies (*16*) and alkaline phosphatase-conjugated goat anti-rabbit immunoglobulin G in PBS (Sigma, Mississauga, ON) (*16*). Blots were visualized with the ECL kit (Amersham Life Science, Baie d'Urfe, QC) according to the manufacturer's instructions.

Transport Assays. Bacteria were cultivated in MOPS minimal medium, and initial rates of proline uptake were determined as previously described (*18*). Serine uptake was measured in the same way using L-serine at a concentration and specific radioactivity of 20 μM and 25 mC/mmol, respectively.

Overexpression and Purification of ProQ and ProQ(His)$_6$. For ProQ purification based on its expression using plasmid vector pBAD24, LB (10 mL) was inoculated with *E. coli* WG692 (DH5α pDC77 (*16*)) and incubated for 7 h at 37 °C, shaking at 200 rpm. This culture (0.24 mL) was used to inoculate MOPS medium (24 mL) supplemented with NH_4Cl (9.5 mM) as a nitrogen source, glycerol (5 mg/mL) as a carbon source, L-tryptophan (245 μM), thiamine hydrochloride (1 μg/mL), and ampicillin (100 μg/mL) and incubated overnight under the same conditions. The bacteria were subcultured in the same medium plus L-arabinose (0.2% (w/v)) to an optical density at 600 nm (OD_{600}) of 0.4 and incubated until an OD_{600} of 0.8−1.0 was reached. Cells were harvested by centrifugation at 12 000 × *g* for 20 min at 4 °C and washed two times with 0.1 M potassium phosphate, pH 7.4, and the pellet was frozen at −40 °C.

For ProQ or ProQ-His_6 purification based on their expression using plasmid vector pQE-60, LB (100 mL) was inoculated with *E. coli* WG806 (SG13009 pREP4 pMS2) or WG805 (SG13009 pREP4 pMS1), respectively, and grown overnight as described above. LB (2 L) was inoculated with the overnight culture to reach an OD_{600} of 0.2. Cultures were incubated at 37 °C until an OD_{600} of 0.7 was reached, then induced with IPTG (1 mM), and incubated for 4 h at 37 °C. The cells were harvested by centrifugation (12 000 × *g* for 20 min at 4 °C) and washed twice with 0.1 M potassium phosphate buffer, pH 7.4, or 0.85% NaCl. The resulting pellet was weighed and stored at −40 °C.

To purify ProQ, cell pellets were resuspended in lysis buffer 1 (75 mM potassium phosphate, pH 7.4, 1 mM dithiothreitol (DTT), 5 mM $MgSO_4$, 30 μg/mL deoxyribonuclease, 1 mM phenylmethylsulfonyl fluoride (PMSF)) (5 mL per gram of wet weight). (This purification was therefore based on the fraction of ProQ that is soluble in this buffer (see Results).) Cells were disrupted by two passes through a French pressure cell at an internal cell pressure of 1.6×10^8 Pa. The lysate was centrifuged in the Sorvall SS34 rotor for 20 min at 12 000 × *g* at 4 °C. The supernatant was centrifuged in a Beckmann Ti70 rotor for 2 h at 100 000 × *g* at 4 °C. The resulting soluble fraction was used for ProQ purification as follows. Chromatography was performed using a Pharmacia FPLC component system including an LCC500 controller (Pharmacia LKB, Sweden) at room temperature. Samples were loaded using a super loop with a P-500 pump (Pharmacia LKB, Sweden) at a flow rate of 0.5 mL/min. The soluble fraction was applied to a MonoS HR10/10 column (Pharmacia LKB, Sweden) equilibrated with 0.1 M potassium phosphate buffer, pH 7.4, containing 1 mM DTT. The elution protocol included an initial wash with approximately 45 mL of 100 mM potassium phosphate buffer, pH 7.4, containing 1 mM DTT followed by a 10 mL linear gradient, in the same buffer, to 1 M KCl. Fractions containing ProQ, as assessed by SDS−PAGE, were pooled and frozen at −40 °C.

Solutions used during the purification of ProQ-His_6 were based on buffer A (50 mM sodium phosphate, 1 M NaCl, pH 8.0). Cells were thawed on ice for at least 15 min before being resuspended in lysis buffer 2 (buffer A supplemented with 0.5 mM PMSF, 5 mM ϵ-amino caproic acid, 5 mM imidazole, 0.1 mg/mL lysozyme) (5 mL per gram of wet weight). Cells were disrupted by two passes through a French pressure cell at an internal cell pressure of 1.6×10^8 Pa. The lysate was centrifuged in the Sorvall SS34 rotor at 12 000 × *g* for 20 min at 4 °C. Ni(NTA) resin (Qiagen, Mississauga, ON) (1 mL of resin per 5 mL of cell lysate) was loaded into a Bio-Rad Econo-column (Bio-Rad, Mississauga, ON, 1.5 cm × 30 cm) and washed with 50 mL of water, 50 mL of buffer A containing 0.25 M imidazole, and 70 mL of buffer A containing 10 mM imidazole. The primed resin was mixed with the soluble fraction; the mixture was incubated at 4 °C for 1 h before loading into the column and collecting the flow-through. The column was washed (50 mL of buffer A containing 20 mM imidazole) and eluted with buffer A containing 0.25 M imidazole. Fractions (2 mL) were sampled and stored at 4 °C.

*Stokes Radius Determination***.** Size-exclusion chromatography was performed using a Superdex 75 column equilibrated with 0.1 M potassium phosphate, 1 mM DTT, pH 7.4. Bovine serum albumin (67 000 Da, $R_s = 3.55$ nm), ovalbumin (43 000 Da, $R_s = 3.05$ nm), chymotrypsinogen A (25 000 Da, $R_s = 2.09$ nm), ribonuclease A (13 700 Da, $R_s = 1.64$ nm), and blue dextran (2 000 000 Da) were used as standards (*30*).

Protein Analyses. Protein concentrations were determined by the BCA (bicinchoninic acid) assay (*31*) using a kit obtained from Pierce Biotechnology Inc. (Rockford, IL) with bovine serum albumin as the standard. Protein samples for amino acid analysis were frozen in liquid nitrogen and then

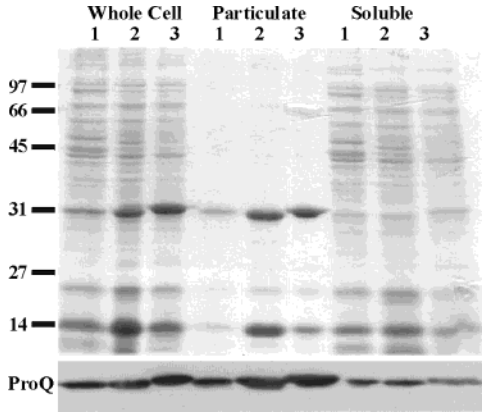

FIGURE 1: Overexpression of ProQ and ProQ-His$_6$. The top panel shows SDS−PAGE analysis of the whole cell, particulate, and soluble fractions of (1) ProQ expressed from plasmid pDC77 in *E. coli* DH5α after induction with 0.2% arabinose, (2) ProQ expressed from plasmid pMS2 in *E. coli* SG13009 pREP4 after induction with 1 mM IPTG, and (3) ProQ-His$_6$ expressed from plasmid pMS1 in *E. coli* SG13009 pREP4 after induction with 1 mM IPTG. Bacteria were resuspended, lysed by sonication, and fractionated by centrifugation as described in Experimental Procedures (Protein Expression Tests section). The bottom panel shows Western blot analysis of the fractions listed above (see Experimental Procedures).

freeze-dried for 48 h. Protein samples for N-terminal sequencing were fractionated by SDS−PAGE and then transferred from the gel to a 100% HPLC grade, methanol-equilibrated PVDF membrane (Bio-Rad, Mississauga, ON) for 2 h at 30 V using a Hoefer semi-phor TE-70 semidry transfer unit (Bio-Rad, Mississauga, ON) with the Towbin transfer buffer system (25 mM Tris base, pH 8.3, 192 mM glycine, 20% methanol) (*32*). After complete transfer, membranes were stained with 0.025% (w/v) Coomassie Brilliant Blue dissolved in 40% methanol for 15 min and then destained with 50% methanol. The protein band of interest was cut away from the remainder of the membrane for N-terminal analysis. The N-terminal and amino acid analyses were carried out by the Alberta Peptide Institute (University of Alberta, AB) using a HP G1005A sequencer and a Beckman 6300 analyzer, respectively.

Spectroscopy. The buffer composition of protein solutions was changed using Slide-A-Lyzer Mini Dialysis Units (Pierce, Rockford, IL). Protein solutions were concentrated by centrifugation at 3000 × *g* using Centricon 10 ultrafiltration membranes (Amicon, Danvers, MA). UV−visible absorption spectroscopy was performed with a Beckman DU-60 spectrophotometer (Mississauga, ON) in a 1 cm path length quartz cuvette. The Mass Spectroscopy facility, University of Guelph, performed matrix-assisted laser desorption ionization (MALDI) time-of-flight (TOF) mass spectrometry. The Bruker Reflex III (Bruker Daltonics Inc., Manning Park Billerica, MA) was operated in linear mode using horse heart cytochrome C (*m*/*z* 12 360) as the externally run mass calibration standard. Circular dichroism (CD) spectroscopy was performed with a model J-70 spectropolarimeter (Jasco, Tokyo, Japan) connected to a Jasco data processor using a quartz cell with a 1-mm path length. CD spectra were measured at 25 °C between 190 and 250 nm. In all cases, buffer baselines were recorded under identical conditions and subtracted from the spectra of the proteins. Each of the spectra presented was the average of five scans.

Secondary Structure Prediction and Comparative Modeling. The program CLUSTAL W (*33*) was used to derive multiple sequence alignments of prokaryotic ProQ and FinO homologues. ProQ secondary structure was predicted by combining results from PsiPred (*34*) and JPred (*35*). The 3D-PSSM web server (*36*) and GenThreader (*37*) were used to construct sequence-to-structure alignments and to derive templates for model construction. Models were then constructed in the program SPDBV (*38*). SPDBV was also used for side chain replacement and rotamer optimization, energy minimization, calculating electrostatic potential surfaces, and superimposition of structures. Resulting model structures were validated by PROCHECK (*39*) at a nominal resolution of 2.5 Å.

RESULTS

Deletion and Plasmid-Based Expression of proQ. Two-step PCR and allelic exchange (*26*) were used to create an in-frame chromosomal *proQ* deletion (Δ*proQ676*) and insert it into *E. coli* strains K-12 and WG210 as described in Experimental Procedures. PCR performed with primers proQNo and proQCo, using chromosomal DNAs from the resulting strains (WG822 and WG914) as templates, yielded the 1.2 kb DNA fragment expected for allele Δ*proQ676* and not the 1.8 kb fragment obtained with template DNA from *proQ*$^+$ bacteria. Western Blot analysis of whole cell extracts prepared from these strains revealed no ProQ protein. The ProP activities of these strains were also impaired (see below), as had been observed for bacteria harboring lesion *proQ220*::Tn*5* (*16*). Like insertion *proQ220*::Tn*5* (*16*), in-frame deletion Δ*proQ676* failed to produce the phenotype characteristic of mutations in downstream locus *prc*, substantiating the view that the *proQ* and *prc* phenotypes are distinct (data not shown). These experiments demonstrated that the *proQ* phenotype is independent of the Tn*5* insertion in *proQ* and pave the way for future work requiring a stable *proQ* null mutation.

When gene *proQ* was expressed from its native promoter in *E. coli* RM2, Western blot analysis showed ProQ to be present in the soluble fraction but not the membrane fraction obtained by cell disruption using a French pressure cell (data not shown). Kunte et al. demonstrated that ProQ could be expressed via the *araBAD* promoter under the control of arabinose and AraC in vector pBAD24 (*16*). We sought to further elevate the expression of ProQ and ProQ-His$_6$ by placing *proQ* downstream of the IPTG-inducible bacteriophage T5 promoter in expression vector pQE-60 (see Experimental Procedures). Control of gene expression from pQE-60 is achieved via incorporation of a second plasmid, pREP4, which encodes repressor LacI^q. Efforts to construct ProQ expression plasmids based on pQE-60 were unsuccessful unless ligation mixtures were transformed into bacteria harboring pREP4. IPTG induction of bacteria harboring the resulting plasmids (pMS2 and pMS1) resulted in high-level expression of proteins with the expected molecular weight (Figure 1). However, most of this protein was found in the particulate fraction after cell disruption. The solubilities of ProQ and ProQ-His$_6$ were examined after culturing bacteria at varying temperatures (15−37 °C) and with IPTG at various concentrations (0.05−1 mM). Under all conditions tested, up to one-half of the overexpressed protein was found in the insoluble fraction (data not shown). The maximum

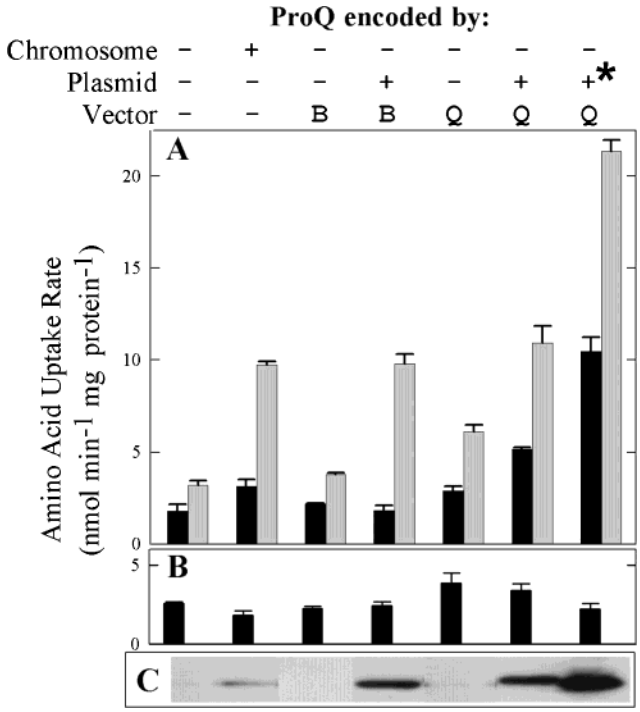

FIGURE 2: Complementation of lesion Δ*proQ676* by plasmid-based expression of ProQ or ProQ-His$_6$. Amino acid uptake activity and ProQ or ProQ-His$_6$ expression were measured in bacteria lacking or retaining the chromosomal *proQ* locus (chromosome − or chromosome +; *E. coli* strains WG914 and WG210, respectively). They did not or did contain a plasmid (plasmid − or plasmid +). The plasmids were based on vector pBAD24 (B) or pQE-60 (Q). Plasmid pMS1, denoted by an asterisk (∗), encoded ProQ-His$_6$ rather than ProQ. Proline uptake (panel A) or serine uptake (panel B) were measured as described in Experimental Procedures after cultivation in MOPS minimal medium supplemented with 0.05 M NaCl (black bars) or 0.17 M NaCl (gray bars). Expression of ProQ or ProQ-His$_6$ (panel C) was measured by Western blotting using anti-ProQ antibodies as described in Experimental Procedures.

protein yield was attained in these experiments when IPTG (1 mM) was added at a culture OD_{600} of 0.7 and growth was continued for 4 h at 37 °C.

ProQ and ProQ-His$_6$ Complement Δ*proQ676.* Expression of ProQ from pBAD24-based plasmid pDC77 complemented lesion *proQ220*::Tn*5*, restoring ProP activity to the level observed in *proQ*$^+$ bacteria (*16*). We further tested the ability of ProQ and ProQ-His$_6$ to complement lesion Δ*proQ676*. Expression of either ProQ or ProQ-His$_6$ from the pQE-60-derived plasmids restored the ProP activity of host strain WG914 (Δ*proQ676*) (Figure 2A, ProQ-His$_6$ is denoted by the asterisk (∗)). In this experiment, inducers (IPTG or D-arabinose) were omitted from the growth media to minimize protein expression. Nevertheless, ProQ-His$_6$ was expressed to a higher level than ProQ, and ProQ was expressed to higher levels than in cells with a single chromosomal copy of *proQ* (Figure 2C). ProQ-His$_6$ elevated ProP activity above that of bacteria expressing ProQ from the chromosome or a plasmid. Serine uptake activity was measured to determine whether this activation was ProP-specific. No such elevation of serine uptake activity was observed (Figure 2B). It is not clear whether the greater effect of ProQ-His$_6$ on ProP activity resulted from the His-tag or from its high expression level in this system. These results indicate that ProQ-His$_6$ can serve as a proxy for ProQ in subsequent studies and suggest that further investigation of the relationship between ProQ concentration and ProP activity would be warranted.

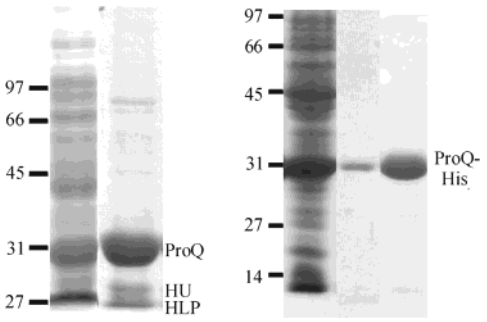

FIGURE 3: Purification of ProQ and ProQ-His$_6$. The left panel shows purification of ProQ from an extract of *E. coli* SG13009 pREP4 pMS2 by ion-exchange chromatography: left lane, soluble fraction obtained by ultracentrifugation of the cleared cell lysate; right lane, purified ProQ protein eluted from the MonoS column with the KCl gradient. The major band is ProQ; minor bands of lower molecular weight are HU and the histone-like protein (HLP) (see text). The right panel shows purification of ProQ-His$_6$ by Ni(NTA) chromatography: left lane, soluble faction obtained by ultracentrifugation of the cell lysate; right lanes, pure ProQ-His$_6$ eluted from the Ni-(NTA) resin with buffer containing 0.25 M imidazole. The numbers to the left of each panel indicate the positions of molecular weight markers (kDa).

Purification of ProQ and ProQ-His$_6$. Since ProQ was predicted to be basic (calculated pI of 9.7), the soluble fraction from bacteria overexpressing this protein was fractionated by cation-exchange chromatography using a MonoS column. Upon elution with a KCl gradient, a single peak was recovered at approximately 0.3 M KCl. SDS−PAGE of the eluted protein revealed a major component with an apparent molecular weight of 28 kDa (Figure 3, left panel) that reacted with anti-ProQ antibodies (data not shown). The N-terminal sequence of this protein was MENQPK, as predicted from the *proQ* sequence, and the amino acid composition was also consistent with prediction (data not shown). Two low molecular weight proteins copurified with ProQ. They were identified by N-terminal sequencing as HU (accession no. P02342) and a histone-like protein (accession no. P11457), both basic DNA-binding proteins with no sequence preference (*40*, *41*). Attempts to further separate ProQ from HU and the histone-like protein using DNA cellulose chromatography (*42*) or gel exclusion chromatography were not successful. The ability of ProQ to bind DNA cellulose may reflect its highly basic nature (pI 9.7) rather than its physiological role. The yield of ProQ was approximately 0.5 or 0.8 mg per gram of wet weight of cells when expressed from the pDC77- or the pMS2-based system, respectively, under these conditions.

The yield and purity of ProQ were further enhanced as follows. First, pMS1 was created so that ProQ-His$_6$ could be purified by nickel chelate affinity chromatography, eliminating other basic proteins. Second, the cell lysis conditions were adjusted in an effort to enhance the solubility of overexpressed ProQ and ProQ-His$_6$. Elevating the ionic strength of the lysis buffer with NaCl dramatically enhanced the solubility of both proteins (Figure 4, solubility was optimal at approximately 0.6 M NaCl). Ni(NTA) chromatography yielded ProQ-His$_6$ free of the low molecular weight DNA binding proteins and of other contaminants (Figure 3,

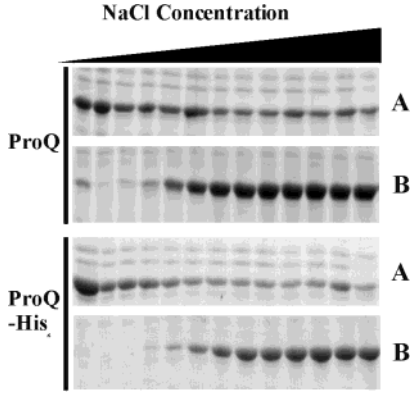

FIGURE 4: Impact of salinity on the solubility of ProQ and ProQ-His_6. *E. coli* strains WG806 (expressing ProQ, top panels) and WG805 (expressing ProQ-His_6, bottom panels) were grown, harvested, resuspended (buffer supplemented with NaCl to a final concentration of 0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, or 1.5 M), and lysed by freeze−thaw (see Experimental Procedures). For each protein, the particulate fraction (A) was then separated from the soluble fraction (B) by centrifugation, and the released proteins were analyzed by SDS−PAGE as described in Experimental Procedures.

right panel). The yield of ProQ-His_6 obtained in this way was 8 mg per gram of wet weight of cells.

ProQ and ProQ-His_6 Purify as Nonspherical Monomers or Dimers. The expected molecular weights (MW) of ProQ and ProQ-His_6, predicted from their sequences, are 25 876 and 26 872 Da, respectively. Each protein migrated as a single band. Their molecular weights were estimated to be 29 and 33 kDa, respectively, by SDS−PAGE (Figures 1 and 3). MALDI-TOF mass spectrometry of purified ProQ revealed a major peak at *m*/*z* 25 796, representing the molecular mass (see Supporting Information, Figure 2S). The second narrower peak at *m*/*z* 12 972 was assigned to M^{2+}, a doubly charged ion. Since the M^{2+} ion was most similar in mass to the standard used to calibrate the spectrometer, it provided a more reliable estimate of the molecular mass of the protein, 25 944 Da. This value is within experimental error of the calculated value. The first and third peak (*m*/*z* values of 9306 and 16 700) could represent either the two known minor protein contaminants (HU with a molecular mass of 9535 kDa and histone-like protein with a molecular mass of 17 688 kDa) or degraded ProQ fragments with a combined *m*/*z* value of 26 006. The Stokes radii of ProQ and ProQ-His_6 were determined by exclusion chromatography to be 3.1 and 3.6 nm, corresponding to apparent MW of 48 and 64 kDa, respectively (see Experimental Procedures). These values may indicate that these proteins are dimers, but they more likely reflect the domain structure of ProQ, as discussed below.

Spectroscopic Analysis of ProQ. The UV−visible spectrum of ProQ displayed a characteristic maximum at 280 nm with no absorbance in the visible wavelength range (data not shown). The extinction coefficient calculated from the absorbance at 280 nm and the protein concentration determined by BCA assay was 8700 M^{-1} cm^{-1}. This is reasonably consistent with the expected extinction coefficient of 9530 M^{-1} cm^{-1} based on its amino acid composition (*43*). The far-UV circular dichroism (CD) spectrum of ProQ is shown in Figure 5, top. The negative ellipticities at 208, 215, and 222 nm suggest that the protein is composed of a mixture

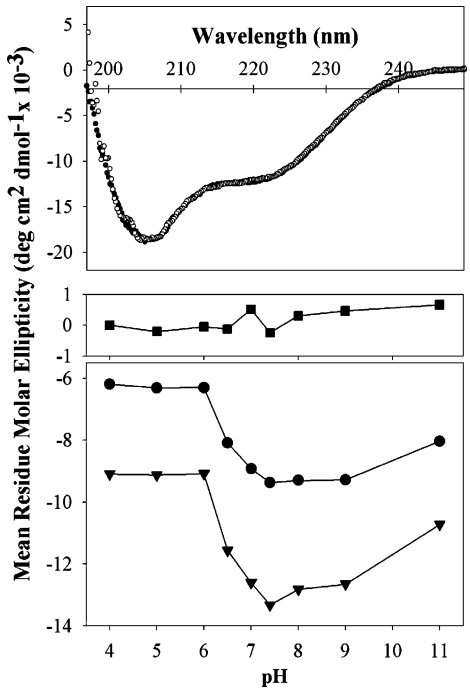

FIGURE 5: CD spectra of ProQ. The top panel shows the effect of sodium chloride on the secondary structure of ProQ. CD spectra were obtained using 0.1 mg/mL ProQ in 10 mM potassium phosphate buffer, pH 7.4, at room temperature. Mean residue molar ellipticity is plotted versus wavelength (●, no NaCl; ○, 1 M NaCl). The bottom panels show the effect of pH on the secondary structure of ProQ. CD spectra were obtained using 0.1 mg/mL ProQ in 10 mM potassium phosphate at the indicated pH. Mean residue molar ellipticity values at (■) 197, (▼) 206, and (●) 220 nm are plotted as a function of buffer pH.

of α-helices and *β*-sheets. Comparison with a database of CD spectra for proteins of known structure (elastase, flavodoxin, lactate dehydrogenase, myoglobin, lysozyme, ribonuclease A, papain, cytochrome C, hemoglobin, and chymotrypsin A) performed with JFIT (Jasco, Tokyo, Japan) suggested that ProQ is composed of 8% α-helix and 46% *β*-structure. A plot of the change in the molar ellipticity versus buffer pH for ProQ indicated that its secondary structure was pH-sensitive (Figure 5, bottom). In contrast, the CD spectrum of ProQ was insensitive to NaCl concentration in the range 0−1 M over the pH range 4−11 (see Figure 5 for representative spectra at pH 7.4). Finally, it should be noted that over the pH and ionic strength ranges tested, the secondary structure of ProQ was not wholly eliminated.

Comparative Modeling of ProQ. Secondary structure analysis for *E. coli* ProQ produced consistent results in different programs and delineated two domains: the N-terminal domain (1−124) is primarily helical (Figure 6) and is connected by a long unstructured segment (125−178) to a C-terminal domain (179−232), which is primarily *β*-sheet (see Supporting Information, Figure 3S). Similar structural domains are predicted for most of the putative ProQ orthologues listed in Table 1 and Figure 1S, though the genome of *Agrobacterium tumefaciens* appears to encode only the N-terminal domain.

The N-terminal and C-terminal sequences of *E. coli* ProQ were submitted separately to the fold-recognition server 3D-PSSM. The N-terminal sequence gave a single very strong match to the structure 1DVO (*44*) ($E = 1.6 \times 10^{-8}$, 25% sequence identity) and a structure to sequence alignment that

```
                1                                                        50
ProQ____PSS   CCCCCC.... .......... .......... .......... CCCHHHHHHH
ProQ____Seq   MENQPK.... .......... .......... .......... LNSSKEVIAF
-----------      -PK                                       L++++E++++
c1dvoa__Seq   ...PPKWKVK KQKLAEKAAR EAELTAKKAQ ARQALSIYLN LPTLDEAVNT
c1dvoa__SS    ...CCHHHHH HHHHHHHHHH HHHHHHHHHH HHHHHHHHCC CCCHHHHHHH

                50                                                       100
ProQ____PSS   HHHHCCCCCCC CCCCCCCHHH CCHHHH.HHH HHCCCCCCHH HHHHHHHHHH
ProQ____Seq   LAERFPHCFS AEGEARPLKI GIFQDL.VDR VAGEMNLSKT QLRSALRLYT
-----------   L++++P++F  +++++R+L++ GI+--L +D+ -+++++LS++ +LR+A++++T
c1dvoa__Seq   LKPWWPGLF. DGDTPRLLAC GIRDVLLEDV AQRNIPLSHK KLRRAMKAIT
c1dvoa__SS    HHCCCHHHE. ECCEECCCCC CHHHHHHHHH HHHCCCCCHH HHHHHHHHHH

                100                                                      150
ProQ____PSS   CCHHHHHHHH CCCEEECCCC CCCCCCCHHH HHHHHHHHHH HHHHHHHHHH
ProQ____Seq   SSWRYLYGVK PGATRVDLDG NPCGELDEQH VEHARKQLEE AKARVQAQRA
-----------   +S++YL+++K +GA+R+D++G ++++++++++ +-+A+++L-- --------+-
c1dvoa__Seq   RSESYLCAMK AGACRYDTEG YVTEHISQEE EVYAAERLDK IRRQNRIKAE
c1dvoa__SS    HCHHHHHHCC CCEEEECCCC CEEEEECHHH HHHHHHHHHH HHHHHHHHHH

                150                                                      200
ProQ____PSS   HHHHHHHHHC CCCCCCCCCC CCCCCCCCCC CCCCCCCCCC CCCCCCCCCC
ProQ____Seq   EQQAKKREAA ATAGEKEDAP RRERKPRPTT PRRKEGAERK PRAQKPVEKA
-----------   -Q+---
c1dvoa__Seq   LQAVLD.... .......... .......... .......... ..........
c1dvoa__SS    HHHHHC.... .......... .......... .......... ..........
```

FIGURE 6: Alignment of ProQ N-terminal sequence with that of FinO. ProQ residues 1−165 (row labeled ProQ___Seq) aligned with residues 33−184 of FinO (row labeled c1dvoa__Seq), terminating at the C-terminus of FinO. The alignment shows the close correspondence of the predicted secondary structure of ProQ (row labeled ProQ__PSS) and the actual secondary structure of FinO (row labeled c1dvoa__SS). The + and − symbols indicate positive and negative PSSM equivalence scores.

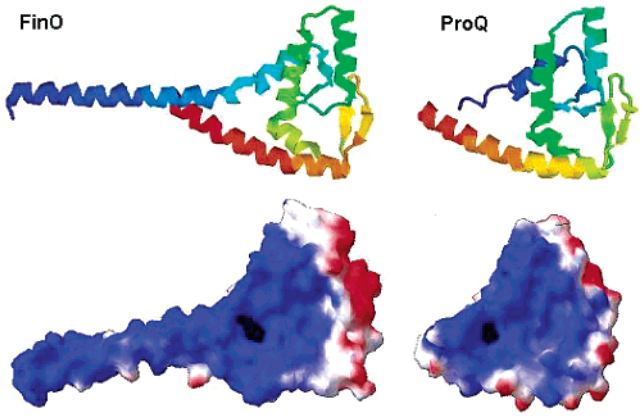

FIGURE 7: Stuctures of FinO and the derived model for ProQ N-terminal domain. The upper ribbon diagrams (left, FinO residues 33−184, PDB structure 1DVO; right ProQ residues 1−121) are colored blue at the N-terminal end through to red at the C-terminal end. The lower diagrams show the molecular surface computed by SPDBV colored by electrostatic potential (blue = positive, red = negative).

was almost gapless for residues 7−121. Structure 1DVO is the structure of FinO, which is an RNA binding protein involved in regulation of F-pilus biogenesis in *E. coli* (*45*). A weak relationship between FinO and ProQ sequences is also detected by PsiBlast and was noted previously (*16*). This structure-based alignment to FinO was reinforced by very similar results obtained from the program GenThreader. The alignment included the complete folded core of the FinO structure, omitting only the long N-terminal helix, which extends out from the core and which acts as part of the RNA binding surface of FinO (Figure 7). A model of the N-terminal domain was constructed from the 1DVO template and yielded a positive PROCHECK score of 0.14 (Figure 7). This model places the majority of polar and charged side chains on the protein surface. The buried side chains in the model are almost exclusively those of valine, leucine, isoleucine, phenylalanine, and alanine. Two serine and two tyrosine residues have low solvent accessibility in the model but appear to form H-bonds internally (S66 OH to L63 O; S67 OH to to Y70 N; Y64 to E30; Y70 to L34 N).

The C-terminal region of ProQ did not give strong sequence-to-sequence matches to any protein of known structure (see Supporting Information, Figure 1S). Sequence-to-structure matches returned by 3D-PSSM showed high *E*-values (best result 0.59), which would not normally be considered significant in a sequence-to-sequence alignment. However a single structural type dominated the top 20 hits, consisting of a five-stranded β-meander either classified as an SH3-like domain (9 out of 20 hits) or a closely related Sm motif (3 out of 20 hits). In the best cases, 1VIE, a bacterial dihydrofolate reductase, and 1LVK, the SH3-like domain of *Dictyostelium* myosin, the predicted secondary structure of ProQ closely matched the target structure, and alignments were almost gapless (see Supporting Information, Figure 3S). In sequence-to-structural alignments, these considerations may override an otherwise weak sequence match (*36*). The structure 1LVK (*46*) was chosen as template for modeling ProQ C-terminal domain (Figure 8) because the only gap in the alignment lay in a loop region and was easily accommodated by shortening the loop and adjacent strand. The resulting model must be considered speculative because of the low sequence similarity but shows remarkable adherence to the expectations of protein structure: polar and charged residues all face the exterior, and the protein core is entirely occupied by nonpolar amino acids.

DISCUSSION

ProQ is a soluble protein the absence of which from *E. coli* dramatically attenuates the response of osmoprotectant

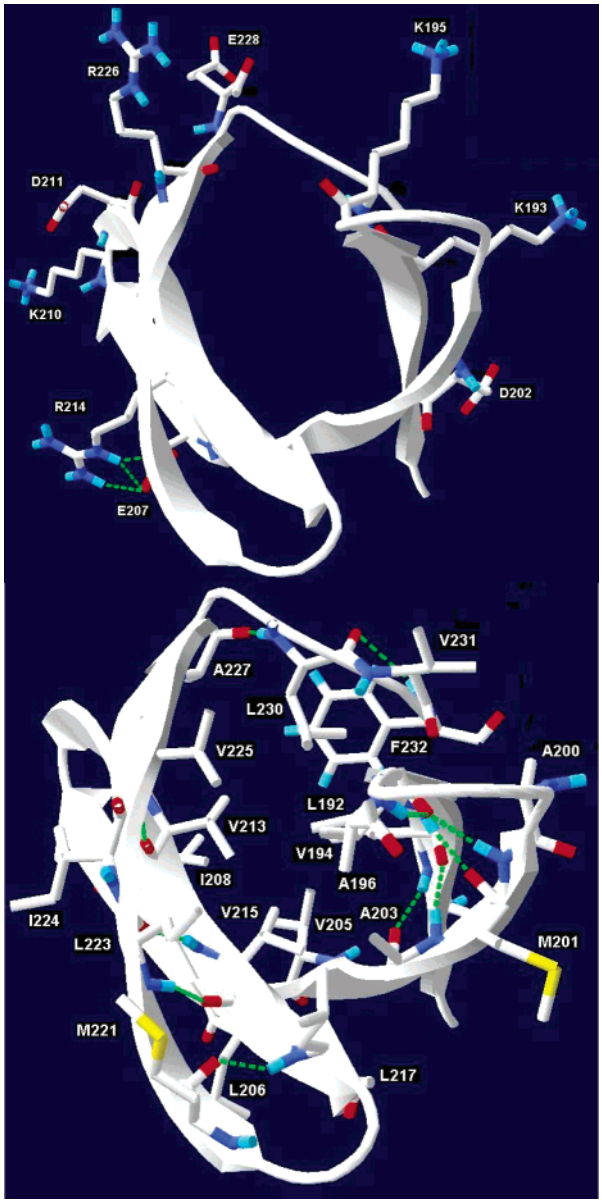

FIGURE 8: Proposed model of the ProQ C-terminal domain (residues 179−232). In the top panel, charged side chains face the exterior exclusively. In the bottom panel, the domain core consists exclusively of nonpolar side chains, although a few nonpolar side chains are surface-exposed.

transporter ProP to increasing extracellular osmolality (the "*proQ* phenotype") (*14*−*16*). Lesion *proQ220*::Tn*5* alters neither *proQ* transcription (*15*) nor cellular ProP protein levels (*16*) (at least under conditions eliciting the *proQ* phenotype). Thus ProQ may act directly on ProP, in turn an integral membrane protein and a member of the major facilitator superfamily (proteins catalyzing uniport, symport, or antiport) (*47*). Recent data suggest that ProP senses increasing extracellular osmolality by detecting the resulting changes in cytoplasmic composition, structure or both (*18*). Comparative analysis of the *E. coli* ProQ structure (Figures 6-8) suggests that it includes a cationic N-terminal domain (Figure 7) linked via an unstructured tether to an SH3-like domain (Figure 8). Eukaryotic SH3 domains bind with moderate affinity and selectivity to proline-rich ligands, thereby mediating protein−protein interactions that regulate enzymes, cluster proteins, or determine their subcellular localizations (*48*). For example, Sho1p acts as an osmosensor (or osmosensory signal transducer) within the high osmolarity glycerol (HOG) MAP kinase pathway of yeast (*49*). Integral to the cytoplasmic membrane, Sho1p includes a C-terminal SH3 domain, which recruits kinase Pbs2p to the membrane surface. SH3-like domains, structures with diverse ligand specificities, are also prevalent in prokaryotes and eukaryotes. For example, SH3-like domains are postulated to play an important role in transmembrane signaling during bacterial chemotaxis (*50*). Thus our structural predictions for ProQ are intriguing, and they must be tested as their functional implications are explored. To facilitate such studies, we set out to overexpress and purify ProQ.

Initial work suggested that ProQ became insoluble upon overexpression (Figure 1). However, most ProQ remained in solution if the overexpressing bacteria were lysed in buffers containing NaCl at a concentration in excess of 0.6 M (Figure 4). The native protein could be purified to near homogeneity in a single step (cation-exchange chromatography of cleared cell lysates) but basic DNA binding proteins remained as minor contaminants that were not readily removed (Figure 3). The identity of the major component was confirmed as ProQ by N-terminal sequencing, amino acid analysis, and MALDI-TOF mass spectroscopy (Supporting Information, Figure 2S) (the measured and calculated molecular weights (25 944 and 25 876 Da, respectively) were within experimental error). The UV−visible spectrum of ProQ provided no evidence for its association with nonprotein chromophores.

The Stokes radius of ProQ (3.1 nm) corresponds to an apparent MW of 48 kDa. ProQ may be a dimer, or this high value may reflect a nonspherical overall shape resulting from the folding of ProQ into two domains linked by an unstructured tether (Figures 6−8). The primary sequence and the structural models of ProQ illustrated in Figures 7 and 8 suggest that the N-terminal domain is predominantly α-helical and the C-terminal domain predominantly β structure with a net preponderance of the former (JPRED predicts 32% α-helix, 13% β-strand, and 55% random coil). In contrast, comparison of the CD spectrum of ProQ (Figure 5) with those of other proteins suggested a very different distribution of secondary structure elements (8% α-helix, 46% β-sheet). The structural models of ProQ (Figures 7 and 8) provide a potential explanation for this behavior. Some SH3 domains display positive ellipticity with a maximum at approximately 220 nm rather than the negative ellipticity with a minimum at approximately 217 nm characteristic of β-sheets (*51*). Linkage of an SH3-like domain with such a spectrum to the FinO-like N-terminal domain of ProQ would yield a spectrum like that shown in Figure 5 (top). In the 220 nm region, the second of two minima characteristic of α-helical proteins would coincide with the maximum characteristic of the SH3-like domain rather than with the minimum characteristic of other β structures, resulting in an underestimation of α-helix content. The CD spectrum of ProQ was NaCl-insensitive (up to 1 M) (Figure 5, top) and pH-sensitive, particularly in the range 7−9 (Figure 5, bottom). NaCl may alter the juxtaposition of the N- and C-terminal domains of ProQ without significantly altering the secondary structure of either domain, or it may induce secondary structure changes that are complementary in their effects on the CD spectrum. The pH sensitivity of the CD spectrum was not surprising given the

number and distribution of ionizable amino acids in ProQ (Figure 6 and Supporting Information, Figure 3S).

ProQ was also expressed with a histidine-tag to facilitate its purification by affinity chromatography. Like native ProQ, ProQ-His_6 reversed the effects of lesions *proQ220*::Tn*5* and Δ*proQ676* on ProP activity (Figure 2). As expected, ProQ-His_6 could be purified to homogeneity by nickel affinity chromatography (Figure 3). Its Stokes radius (3.6 nm) was larger than that of ProQ, suggesting a higher apparent molecular weight (64 kD). This could have arisen because the C-terminal six-histidine tag interferes with the association between distinct N- and C-terminal domains of the protein or ProQ-His_6 could be a homodimer. In fact, overexpression of ProQ-His_6 amplified ProP activity beyond the levels attained when ProQ was either expressed at physiological levels or overexpressed (Figure 2). This may have resulted from the particularly high expression level of ProQ-His_6 or from structural rearrangement due to attachment of the C-terminal tag. Separate expression of the N- and C-terminal domains of ProQ will allow these structural and functional predictions to be further explored.

Since they did not copurify with ProQ-His_6, protein HU and the histone-like protein likely copurified with ProQ because they share its highly basic nature, not via direct association. ProQ is unlikely to modulate translation (like its putative structural homologue, FinO) by binding *proP* mRNA, since *proQ* lesions do not affect *proP* expression or ProP levels. However, ProQ could act in this way to alter ProP activity indirectly by modulating the expression of another gene. Discovery of a putative SH3-like domain at the C-terminus of ProQ suggests that it is involved in protein−protein interactions that do not survive cellular disruption in high salt (but could account for lack of solubility at low ionic strength). In this study, ProQ and ProQ-His_6 were produced at purities and yields sufficient to support further structural and functional analyses. ProQ-His_6 will serve as a valid proxy for ProQ during those studies.

ACKNOWLEDGMENT

We are grateful to Sarah Sanowar for the construction of plasmid pSS1 and to Dr. R. Yada and the members of his laboratory (Department of Food Science, University of Guelph) for use of and assistance with the circular dichroism spectropolarimeter.

SUPPORTING INFORMATION AVAILABLE

The sequences of oligonucleotide primers used during this research (Table 1S), an alignment of the ProQ sequence with those of putative ProQ orthologues (Figure 1S), the MALDI-TOF mass spectrum of ProQ (Figure 2S), an alignment of the ProQ C-terminal sequence with that of 1LVK (Figure 3S), and a comparison of SH3-like domains structurally matched to the ProQ C-terminus by 3D-PSSM (Figure 4S). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Booth, I. R., and Louis, P. (1999) Managing hypoosmotic stress: aquaporins and mechanosensitive channels in *Escherichia coli*, *Curr. Opin. Microbiol. 2*, 166−169.
2. Wood, J. M. (1999) Osmosensing by Bacteria: Signals and Membrane-Based Sensors, *Microbiol. Mol. Biol. Rev. 63*, 230−262.
3. Wood, J. M., Bremer, E., Csonka, L. N., Krämer, R., Poolman, B., van der Heide, T., and Smith, L. T. (2001) Osmosensing and osmoregulatory compatible solute accumulation by bacteria, *Comp. Biochem. Physiol. 130*, 437−460.
4. Morbach, S., and Krämer, R. (2002) Body shaping under water stress: osmosensing and osmoregulation of solute transport in bacteria, *ChemBioChem 3*, 384−397.
5. Culham, D. E., Lasby, B., Marangoni, A. G., Milner, J. L., Steer, B. A., van Nues, R. W., and Wood, J. M. (1993) Isolation and sequencing of *Escherichia coli* gene *proP* reveals unusual structural features of the osmoregulatory proline/betaine transporter, ProP, *J. Mol. Biol. 229*, 268−276.
6. Keates, R. A. B., Culham, D. E., Hillar, A., Vernikovska, Ya. I., Boggs, J. M., and Wood, J. M. (2004) A structural model for transporter ProP of *Escherichia coli* yields new insights regarding its roles as an osmosensor, a transporter, and an osmoregulator, manuscript in preparation.
7. MacMillan, S. V., Alexander, D. A., Culham, D. E., Kunte, H. J., Marshall, E. V., Rochon, D., and Wood, J. M. (1999) The ion coupling and organic substrate specificities of osmoregulatory transporter ProP in *Escherichia coli*, *Biochim. Biophys. Acta 1420*, 30−44.
8. Grothe, S., Krogsrud, R. L., McClellan, D. J., Milner, J. L., and Wood, J. M. (1986) Proline transport and osmotic stress response in *Escherichia coli* K-12, *J. Bacteriol. 166*, 253−259.
9. Racher, K. I., Voegele, R. T., Marshall, E. V., Culham, D. E., Wood, J. M., Jung, H., Bacon, M., Cairns, M. T., Ferguson, S. M., Liang, W.-J., Henderson, P. J. F., White, G., and Hallett, F. R. (1999) Purification and reconstitution of an osmosensor: transporter ProP of *Escherichia coli* senses and responds to osmotic shifts, *Biochemistry 38*, 1676−1684.
10. Culham, D. E., Tripet, B., Racher, K. I., Voegele, R. T., Hodges, R. S., and Wood, J. M. (2000) The role of the carboxyl terminal α-helical coiled-coil domain in osmosensing by transporter ProP of *Escherichia coli*, *J. Mol. Recog. 13*, 1−14.
11. Zoetewey, D. L., Tripet, B. P., Kutateladze, T. G., Overduin, M. J., Wood, J. M., and Hodges, R. S. (2003) Solution structure of the C-terminal antiparallel coiled-coil domain from *Escherichia coli* osmosensor ProP, *J. Mol. Biol. 334*, 1063−1076.
12. Hillar, A., Tripet, B., Zoetewey, D., Wood, J. M., Hodges, R. S., and Boggs, J. M. (2003) Detection of α-helical coiled-coil dimer formation by spin-labeled synthetic peptides: a model parallel coiled-coil peptide and the antiparallel coiled-coil formed by a replica of the ProP C-terminus, *Biochemistry 42*, 15170−15178.
13. Hillar, A. P., Culham, D. E., Vernikovska, Ya. I., and Boggs, J. M. (2004) Formation of an antiparallel, intermolecular coiled-coil is associated with *in vivo* dimerization of osmosensor and osmoprotectant transporter ProP in *Escherichia coli*, manuscript in preparation.
14. Stalmach, M. E., Grothe, S., and Wood, J. M. (1983) Two proline porters in *Escherichia coli* K-12, *J. Bacteriol. 156*, 481−486.
15. Milner, J. L., and Wood, J. M. (1989) Insertion *proQ220*::Tn*5* alters regulation of proline porter II, a transporter of proline and glycine betaine in *Escherichia coli*, *J. Bacteriol. 171*, 947−951.
16. Kunte, H. J., Crane, R. A., Culham, D. E., Richmond, D., and Wood, J. M. (1999) Protein ProQ influences osmotic activation of compatible solute transporter ProP in *Escherichia coli* K-12, *J. Bacteriol. 181*, 1537−1543.
17. Racher, K. I., Culham, D. E., and Wood, J. M. (2001) Requirements for osmosensing and osmotic activation of transporter ProP from *Escherichia coli*, *Biochemistry 40*, 7324−7333.
18. Culham, D. E., Henderson, J., Crane, R. A., and Wood, J. M. (2003) Osmosensor ProP of *Escherichia coli* responds to the concentration, chemistry and molecular size of osmolytes in the proteoliposome lumen, *Biochemistry 42*, 410−420.
19. Poolman, B., Blount, P., Folgering, J. H. A., Friesen, R. H. E., Moe, P. C., and van der Heide, T. (2002) How do membrane proteins sense water stress? *Mol. Microbiol. 44*, 889−902.
20. Peter, H., Weil, B., Burkovski, A., Krämer, R., and Morbach, S. (1998) *Corynebacterium glutamicum* is equipped with four secondary carriers for compatible solutes: identification, sequencing, and characterisation of the proline/ectoine uptake system ProP and the ectoine/proline/glycine betaine carrier EctP, *J. Bacteriol. 180*, 6005−6012.

21. Seol, W., and Shatkin, A. J. (1993) Membrane topology model of *Escherichia coli* α-ketoglutarate permease by PhoA fusion analysis, *J. Bacteriol. 175*, 565−567.
22. Whipp, M. J., Camakaris, H., and Pittard, A. J. (1998) Cloning and analysis of the *shiA* gene, which encodes the shikimate transport system of *Escherichia coli* K-12, *Gene 209*, 185−192.
23. Sambrook, J., Frisch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
24. Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids, *J. Mol. Biol. 166*, 557−569.
25. Brown, E. D., and Wood, J. M. (1992) Redesigned purification yields a fully functional PutA protein dimer from *Escherichia coli*, *J. Biol. Chem. 267*, 13086−13092.
26. Link, A. J., Phillips, D., and Church, G. M. (1997) Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization, *J. Bacteriol. 179*, 6228−6237.
27. Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
28. Neidhardt, F. C., Bloch, P. L., and Smith, D. F. (1974) Culture medium for enterobacteria, *J. Bacteriol. 119*, 736−747.
29. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature (London) 227*, 680−685.
30. Amersham Pharmacia Biotech. (1998) *Gel Filtration Principles and Methods*, 8th ed., Amersham Pharmacia Biotech AB, Uppsala, Sweden.
31. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid, *Anal. Biochem. 150*, 76−85.
32. Towbin, H., Staehelin, T., and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications, *Proc. Natl. Acad. Sci. U.S.A. 76*, 4350−4354.
33. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
34. Jones, D. (1999) Protein secondary structure prediction based on position-specific scoring matrices, *J. Mol. Biol. 292*, 195−202.
35. Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., and Barton, G. J. (1998) JPred: a consensus secondary structure prediction server, *Bioinformatics 14*, 892−893.
36. Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. E. (2000) Enhanced Genome Annotation using Structural Profiles in the Program 3D-PSSM, *J. Mol. Biol. 299*, 499−520.
37. Jones, D. (1999) GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences, *J. Mol. Biol. 287*, 797−815.
38. Guex, N. and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling, *Electrophoresis 18*, 2714−2723.
39. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
40. Rouviere-Yaniv, J., and Gros, F. (1975) Characterization of a novel, low-molecular-weight DNA-binding protein from *Escherichia coli*, *Proc. Natl. Acad. Sci. U.S.A. 72*, 3428−3432.
41. Azam, T. A., and Ishihama, A. (1999) Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity, *J. Biol. Chem. 274*, 33105−33113.
42. Garke, G., Deckwer, W. D., and Anspach, F. B. (2000) Preparative two-step purification of recombinant human basic fibroblast growth factor from high-cell-density cultivation of *Escherichia coli*, *J. Chromatogr., B: Biomed. Sci. Appl. 737*, 25−38.
43. Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins, *Biochemistry 6*, 1948−1954.
44. Ghetu, A. F., Gubbins, M. J., Frost, L. S., and Glover, J. N. (2000) Crystal structure of the bacterial conjugation repressor FinO, *Nat. Struct. Biol. 7*, 565−569.
45. van Biesen, T., and Frost, L. S. (1994) The FinO protein of IncF plasmids binds FinP antisense RNA and its target, *traJ* mRNA, and promotes duplex formation, *Mol. Microbiol. 14*, 427−436.
46. Bauer, C. B., Kuhlman, C. A., Bagshaw, C. R., and Rayment, I. (1997) X-ray crystal structure and solution fluorescence characterization of Mg.2′(3′)-O-(*N*-methylanthraniloyl) nucleotides bound to the *Dictyostelium discoideum* myosin motor domain, *J. Mol. Biol. 274*, 394−407.
47. Saier, M. H., Jr. (2000) Families of transmembrane transporters selective for amino acids and their derivatives, *Microbiology 146*, 1775−1795.
48. Mayer, B. J. (2001) SH3 domains: complexity in moderation, *J. Cell Sci. 114*, 1253−1263.
49. Hohmann, S. (2002) Osmotic stress signaling and osmoadaptation in yeasts, *Microbiol. Mol. Biol. Rev. 66*, 300−372.
50. Stock, J., and Da Re, S. (1999) A receptor scaffolding mediates stimulus-response coupling in bacterial chemotaxis, *Cell Calcium 26*, 157−164.
51. Maxwell, K., and Davidson, A. R. (1998) Mutagenesis of a buried polar interaction in an SH3 domain: sequence conservation provides the best prediction of stability effects, *Biochemistry 37*, 16172−16182.
52. Cairney, J., Booth, I. R., and Higgins, C. F. (1985) *Salmonella typhimurium proP* gene encodes a transport system for the osmoprotectant betaine, *J. Bacteriol. 164*, 1218−1223.
53. Wood, J. M. (1981) Genetics of L-proline utilization in *Escherichia coli*, *J. Bacteriol. 146*, 895−901.
54. Gottesman, S., Halpern, E., and Trisler, P. (1981) Role of *sulA* and *sulB* in filamentation by Lon mutants of *Escherichia coli* K-12., *J. Bacteriol. 148*, 265−273.
55. Guzman, L.-M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter, *J. Bacteriol. 177*, 4121−4130.

BI048561G