

# Structural and Functional Analysis of ProQ: An Osmoregulatory Protein of *Escherichia coli*<sup>†</sup>

Michelle N. Smith,<sup>‡</sup> Stanley C. Kwok,<sup>§</sup> Robert S. Hodges,<sup>§</sup> and Janet M. Wood<sup>\*,‡</sup>

Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario N1G 2W1, Canada, and Department of Biochemistry and Molecular Genetics, University of Colorado at Denver and Health Sciences Center, P.O. Box 6511, Mail Stop 8101, Aurora, Colorado 80045

Received November 16, 2006; Revised Manuscript Received January 16, 2007

**ABSTRACT:** Transporter ProP of *Escherichia coli* senses extracellular osmolality and responds by mediating cytoplasmic accumulation of organic solutes such as proline. Lesions at the *proQ* locus reduce ProP activity *in vivo*. ProQ was previously purified and characterized. Homology modeling predicted that ProQ possesses an  $\alpha$ -helical N-terminal domain (residues 1–130) and a  $\beta$ -sheet C-terminal domain (residues 181–232) connected by an unstructured linker. In this work, we tested the structural model for ProQ, explored the solubility and folding of full length ProQ and its domains in diverse buffers, and tested the impacts of the putative ProQ domains on ProP activity *in vivo*. Limited tryptic proteolysis of ProQ revealed protease resistant fragments corresponding to the predicted N-terminal and C-terminal domains. Polypeptides corresponding to the predicted N- and C-terminal domains could be overexpressed and purified to near homogeneity using nickel affinity, size exclusion and reversed phase chromatographies. Circular dichroism spectroscopy of the purified proteins revealed that the N-terminal domain was predominantly  $\alpha$ -helical, whereas the C-terminal domain was predominantly  $\beta$ -sheet, as predicted. The domains were soluble and folded in neutral buffers containing 0.6 M NaCl. The N-terminal domain was soluble and folded in 0.1 M MES (2-[N-morpholino]-ethane sulfonic acid) at pH 5.6. Despite high solubilities, the proteins were not well folded in Na citrate (0.1 M, pH 2.3). The ProQ domains and the linker were expressed at physiological levels, singly and in combination, in bacteria lacking the chromosomal *proQ* locus. Among these proteins, the N-terminal domain could partially complement the *proQ* deletion. The full length protein and a variant lacking only the linker restored full activity of the ProP protein.

*Escherichia coli* cells regulate their hydration by accumulating and releasing K<sup>+</sup> or organic solutes (1). Increases in the external osmolality cause water to flow out of the cell, leading to dehydration of the cytoplasm, whereas decreases in the external osmolality cause water to flow into the cell. Cytoplasmic hydration is maintained through the collective efforts of osmotically activated solute transporters and channels. These proteins act to move solutes into or out of the cytoplasm in order to control the direction of net water flux across the membrane. Solute that can be concentrated in the cytoplasm at high levels without impairing growth are termed compatible solutes and include compounds such as proline, glycine betaine, and ectoine. Those that stimulate bacterial growth when available in high osmolality media are called osmoprotectants (1).

ProP is a H<sup>+</sup>/osmoprotectant symporter that can sense changes in the osmolality of the external environment and

respond by transporting a variety of structurally related osmoprotectants into the cell (2). ProP responds to changes in external osmolality in whole cells, membrane vesicles and proteoliposomes (3, 4), indicating that the membrane is the only other cellular component required for its activation and activity. Thus, ProP is both an osmosensor and an osmoregulator (4). ProP is a member of the major facilitator superfamily (MFS<sup>1</sup>) of transporters and a homology model for ProP, based on the crystal structure of GlpT from *E. coli*,

<sup>1</sup> Abbreviations: BCA, bicinchoninic acid; BLAST, basic local alignment search tool; C, ProQ C-terminal domain; CD, circular dichroism; DNase I, deoxyribonuclease I; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; ESMS, electrospray mass spectrometry; FPLC, fast protein liquid chromatography; H<sub>6</sub>C, histidine tagged ProQ C-terminal domain; HEPES, 4-[2-hydroxyethyl]-1-piperazine ethane sulfonic acid; H<sub>6</sub>N, histidine tagged ProQ N-terminal domain; H<sub>6</sub>NC, histidine tagged ProQ N-terminal–C-terminal domain fusion; HPLC, high performance liquid chromatography; IPTG, isopropyl- $\beta$ -D-thiogalactopyranoside; L, ProQ linker domain; LC, ProQ linker and C-terminal domain; N, ProQ N-terminal domain; NC, ProQ N-terminal–C-terminal domain fusion; NCBI, National Center for Biotechnology Information; NL, ProQ N-terminal and linker domains; LB, Luria broth; MALDI-TOF MS, matrix assisted laser desorption/ionization time of flight mass spectrometry; MES, 2-[N-morpholino]-ethane sulfonic acid; MOPS, 3-[N-morpholino]-propane sulfonic acid; MFS, major facilitator superfamily; MW, molecular weight; Ni(NTA), nickel nitrilotriacetate; NMR, nuclear magnetic resonance; PCR, polymerase chain reaction; PIPES, 1,4-piperazine diethane sulfonic acid; PVDF, polyvinylidene fluoride; Q, ProQ protein; QH<sub>6</sub>, histidine tagged ProQ; RNase A, ribonuclease A; RP HPLC, reversed phase HPLC;

<sup>†</sup> This work was supported by Research Grant OGP0000508 awarded to J.M.W. by the Natural Sciences and Engineering Research Council of Canada, by a postgraduate scholarship (PGSB) awarded to M.N.S. by the Natural Sciences and Engineering Research Council of Canada, and by grants awarded to R.S.H. (NIH P01 AI059576 and the John Stewart Chair in Peptide Chemistry).

\* Corresponding author. Tel: (519) 824-4120 ext. 53866. Fax: (519) 837-1802. E-mail: jwood@uoguelph.ca.

<sup>‡</sup> University of Guelph.

<sup>§</sup> University of Colorado at Denver and Health Sciences Center.

was created (5, 6). Like related MFS members, ProP contains 12 transmembrane domains and has cytoplasmic N- and C-termini. The cytoplasmic C-terminal domain of ProP is extended with respect to those of its closest paralogues, and a peptide corresponding to residues 468–497 of ProP forms an antiparallel, homodimeric,  $\alpha$ -helical coiled-coil (7–9). Cross-linking studies indicate that ProP dimers observed *in vivo* may include the coiled-coil structure (10). Mutations that disrupt or impair coiled-coil formation *in vitro* decrease ProP activity and increase the osmolality required to activate ProP *in vivo* (7, 11).

A point mutation, an insertion or an in-frame deletion at the *proQ* locus, dramatically decreased ProP activity but did not lead to its complete inactivation. The *proQ* lesions do not affect the transcription of *proP* or the translation of the ProP protein, as determined by *proP::lacZ* fusions and Western blot analysis, respectively (12, 13). Thus, ProQ may act post-translationally, through a direct protein–protein interaction, to stabilize the active conformation of ProP, or it may link ProP activation to other cellular processes. ProP is embedded in the membrane with its C-terminus in the cytoplasm, where it forms a homodimeric coiled-coil with a neighboring ProP monomer. The conformation of ProP is believed to change as osmolality increases, giving the transporter partial activity. ProQ, or a protein whose function depends on ProQ, may then stabilize the active conformation of ProP, amplifying transport activity (1). Bacteria containing ProQ orthologues do not always contain ProP orthologues, and bacteria that contain ProP orthologues often lack ProQ (14). Thus, the function of ProQ may not be confined to the amplification of ProP activity.

ProQ is a 26 kDa, basic, cytoplasmic protein with a calculated pI of 9.7 (12, 13). BLAST analysis reveals no ProQ homologues with known functions (13, 14). A homology model for residues 1–129 of ProQ was derived from the crystal structure of FinO, a basic mRNA binding protein implicated in the regulation of F-pilus biogenesis in *E. coli* (14, 15). This model was based on a very strong match of the ProQ N-terminal sequence to the FinO structure by fold recognition server 3D-PSSM and an almost gapless structure to sequence alignment for residues 7–121. A homology model for residues 178–232 of the protein was derived from the crystal structure of an SH3-like domain of a myosin motor protein in *Dictyostelium discoideum* (14, 16). This model was based on multiple, weak matches to the five stranded  $\beta$ -meander, classified as an SH3-like domain or an Sm motif, identified by 3D-PSSM. The 50 amino acid, hydrophilic peptide linking these two regions of the ProQ protein was predicted to be mostly unstructured.

Experiments involving *in vitro* systems, such as proteoliposomes co-reconstituted with purified ProP and ProQ, are needed to test the hypothesis that the ProQ protein interacts directly with the ProP protein. Such *in vitro* work requires that ProQ be purified and stable in solutions compatible with ProP function. Smith et al. reported that native ProQ (abbreviated below as Q) and histidine tagged ProQ (ProQ-His<sub>6</sub>, abbreviated below as QH<sub>6</sub>) could be kept in solution

and purified by increasing the NaCl concentration of the cell lysis buffer to at least 0.6 M and maintaining high salinity during subsequent chromatographic steps (14). However, such conditions may not be appropriate for studies designed to test ProQ–ProP interactions. It is possible that only the N- or C-terminal domain of ProQ is required to amplify ProP activity and that the functional domain is more soluble than the full length protein. If so, the N- or C-terminal domain of ProQ may serve as a proxy for the full length protein in future *in vitro* studies.

In this study, we confirmed the homology models for ProQ by overexpressing, purifying, and characterizing ProQ fragments, including the putative domains comprising residues 1–130 and 181–232. Assays performed *in vivo* were used to show that the N-terminal domain of ProQ alone can partially amplify ProP activity. The linker is not required, but the N- and C-terminal domains are both required for the full amplification of ProP activity in the absence of full length ProQ.

## EXPERIMENTAL PROCEDURES

**Bacterial Strains and Plasmids.** The genotypes of the *E. coli* strains used for this study are listed in Table 1S (Supporting Information). Plasmid isolation was performed using QIAprep Spin Miniprep Kits (Qiagen, Mississauga, ON). Routine DNA manipulation, plasmid construction, electrophoresis, and transformation were carried out as described previously (17, 18). The polymerase chain reaction (PCR) was performed as described previously (19) using *Pfu* turbo polymerase (Invitrogen, Burlington, ON) and oligonucleotides were purchased from Cortec DNA Services (Kingston, ON). The Molecular Biology Supercenter (University of Guelph, Guelph, ON) performed DNA sequencing to verify all plasmid constructs.












Plasmids encoding most ProQ fragments were constructed by PCR amplifying the required *proQ* sequence (plus desired flanking sequences) with plasmid pDC77 as the template or by excising the required *proQ* sequence from an existing plasmid, then inserting the amplicon or oligonucleotide in an appropriate vector, and recovering the resulting plasmid by transformation into *E. coli* DH5 $\alpha$  followed by subsequent transformation into the relevant *E. coli* genetic background (Table 1 and Table 1S (Supporting Information)). The PCR primers are listed in Table 2S (Supporting Information).

To construct plasmid pMS14, the oligonucleotide encoding residues M1–E130 of ProQ was PCR amplified, adding a 5' *Bam*HI site and a 3' *Hind*III site but no termination codon (Table 2S, Supporting Information). The amplified sequence was ligated into vector pQE82L to create an intermediate plasmid, pMS9. The oligonucleotide encoding residues V180–F232 of ProQ was amplified using primers ProQ2 and ProQ Fusion 1 (Table 2S, Supporting Information). Digestion of the resulting PCR product and plasmid pMS9 with *Eco*RI allowed for the insertion of the C-terminal coding sequence in-frame with the N-terminal coding sequence. The orientation of the insert was confirmed through restriction analysis and sequencing.

**Culture Media.** Bacteria were cultivated in solid or liquid LB (20) or in a MOPS based minimal medium (21). The MOPS medium was supplemented with 43.5 mM glycerol

SDS, sodium dodecyl sulfate; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; SH3, Src homology 3; TAPS, [(2-hydroxy-1,1-bis(hydroxymethyl) ethyl) amino]-1-propane sulfonic acid; TFA, trifluoroacetic acid.

Table 1: Plasmids and Encoded Proteins

Plasmid <sup>a</sup>	Vector <sup>b</sup>	Encoded Protein		
		Name	Structure	Schematic <sup>c</sup>
pDC77	pBAD24	Q	ProQ	
pMS11	pBAD24	N	ProQ(M1-E130)	
pMS15	pBAD24	C	M-ProQ(S181-F232)	
pMS18	pBAD24	NL	ProQ(M1-I183)	
pMS19	pBAD24	LC	M-ProQ(E131-F232)	
pMS21	pBAD24	NC	ProQ(M1-E130)-AW-ProQ(V180-F232)	
pMS22	pBAD24	L	M-ProQ(E131-V180)	
pMS1	pQE60	QH <sub>6</sub>	ProQ-H <sub>6</sub>	
pMS10	pQE82L	H <sub>6</sub> N	MRGSH <sub>6</sub> GS-ProQ(M1-E130)	
pMS13	pQE82L	H <sub>6</sub> C	MRGSH <sub>6</sub> GSM-ProQ(S181-F232)	
pMS14	pQE82L	H <sub>6</sub> NC	MRGSH <sub>6</sub> GS-ProQ(M1-E130)-AW-ProQ(V180-F232)	

<sup>a</sup> The details of plasmid construction are provided in Experimental Procedures and Table 1S (Supporting Information). The construction of pDC77 (13) and pMS1 (14) was reported previously. The other plasmids were constructed for this study. <sup>b</sup> The vectors are pBAD24 (33), pQE60 (Qiagen, Mississauga, ON), and pQE82L (Qiagen, Mississauga, ON). <sup>c</sup> Filled rectangles indicate the segments of ProQ encoded by the listed plasmids. N-terminal open squares indicate the N-terminal tag MRGSHHHHHHGS. The C-terminal gray square for the protein encoded by pMS1 denotes an RSHHHHHH tag.

and 9.5 mM ammonium chloride as carbon and nitrogen sources, respectively, and 1.6 mM Vitamin B<sub>1</sub> and 0.24 mM tryptophan to satisfy auxotrophic requirements of the strains. Ampicillin (Amp, 100 µg/mL), kanamycin (Km, 50 µg/mL), or both were added as required to maintain plasmids based on vectors pBAD24 (Amp), pREP4 (Km), pQE60 (Amp), or pQE82L (Amp).

**Transport Assays.** Bacteria were grown in MOPS based minimal medium (as above), and the initial rates of proline uptake were determined as described previously (22).

**SDS-Polyacrylamide Gel Electrophoresis and Western Blotting.** SDS-PAGE was performed as described by Laemmli (23) with gels containing 12% or 15% (w/v) polyacrylamide with 0.9% or 1.1% (w/v) bis-acrylamide, respectively. Tricine SDS-PAGE was performed as described by Schagger and von Jagow (24) with a 16.5% T, 6% C separating gel, a 10% T, 3% C spacer gel, and a 4% T, 3% C stacking gel. Gels were stained with Gel-Code Blue (Pierce, Rockford, IL) according to the manufacturer's instructions.

Western blotting was performed as previously described (14) using anti-ProQ antibodies (13). Quantitative Western blotting was used to adjust the expression levels of ProQ fragments as follows. Full length ProQ (QH<sub>6</sub>) and the ProQ fragments (H<sub>6</sub>N, H<sub>6</sub>C, and H<sub>6</sub>NC, Table 1) were purified by Ni (NTA) chromatography for use as standards. Cells cultivated in media supplemented with arabinose at various levels (0–13.3 mM) were solubilized and the extracts

analyzed by Western blotting in comparison with these standards, as follows. The quantity of full length ProQ expressed from plasmid pDC77 in *E. coli* WG914, without arabinose induction, was determined by comparing it with purified QH<sub>6</sub> on Western blots. Expression of the ProQ fragments at levels equimolar with full length ProQ, as indicated by standards representing the ProQ fragments, was attained by arabinose supplementation of the relevant cultures at the following concentrations (mM): N, 5.3; NL, 1.3; LC, 4.0; and NC, 0.1 (fragment names are defined in Table 1). Fragments C and L were not detected by Western blotting even when their expression was fully induced (arabinose at 13.3 mM; see Results). These induction conditions were used for both competition and complementation experiments.

**Protein Overexpression and Purification.** The ProQ fragments expressed and purified during this study are listed in Table 1. QH<sub>6</sub> was overexpressed as described (14). H<sub>6</sub>N, H<sub>6</sub>C, and H<sub>6</sub>NC were overexpressed by growing *E. coli* BL-21 Gold harboring the relevant plasmid (Table 1) overnight in ampicillin supplemented LB medium and subculturing the bacteria to an initial OD<sub>600</sub> of 0.4 in the same medium. Cultures were grown to an OD<sub>600</sub> of 0.6, and expression of the recombinant protein was induced by adding IPTG to 1 mM. Incubation continued for 4 h at 37 °C, and protein expression stopped when cells were harvested by centrifugation and washed once in ice cold saline (0.85% (w/v)). The resulting pellets were stored at –40 °C.



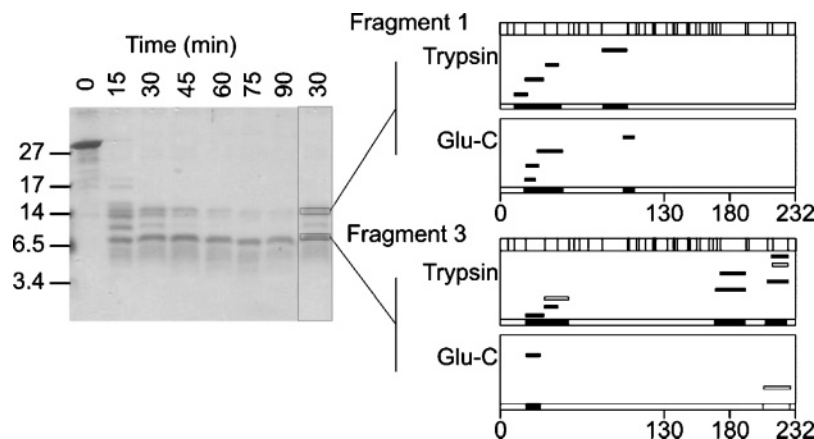


FIGURE 1: Limited trypsin proteolysis. QH<sub>6</sub> was purified by Ni-NTA and gel exclusion chromatography. (Left Panel) Limited proteolysis and Tricine SDS-PAGE were performed as described in Experimental Procedures. The final lane replicates that shown for 30 min of digestion with Fragments 1 and 3 outlined. The numbers to the left are the sizes of the molecular weight markers. (Right Panel) Fragments 1 and 3 were digested to completion with Trypsin or Glu-C, the resulting peptides were extracted from the gel as outlined in Experimental Procedures, and these peptides were analyzed by MALDI-TOF mass spectrometry (Table 2). The identities of the detected peptides were determined by searching the NCBI database using ProFound (26) and are shown as horizontal black bars. Partial oxidation of methionine residues during sample preparation led to the identification of the oxidized peptide species (open bars). The box above the coverage maps represents the distribution of the trypsin cleavage sites over ProQ, and the numbers below the coverage maps indicate the residues at the limits of the putative ProQ domains.

Proteins were purified by Ni(NTA) affinity chromatography as described for QH<sub>6</sub> (14) and by gel exclusion chromatography. Buffers used for the purification of QH<sub>6</sub>, H<sub>6</sub>N, H<sub>6</sub>C, and H<sub>6</sub>NC by Ni(NTA) affinity chromatography were based on buffer A (50 mM Na phosphate, 1 M NaCl, and 1 mM DTT at pH 8.0). The lysis buffer was buffer A supplemented with lysozyme (1 mg/mL), DNase I (30  $\mu$ g/mL), RNase A (30  $\mu$ g/mL), and Roche EDTA free Protease Inhibitor (5 mg/mL) (Roche, Laval, Quebec). The column wash buffer and elution buffer were buffer A supplemented with 20 mM and 250 mM imidazole, respectively. Gel exclusion chromatography was performed using a Superdex 75 HR 10/30 gel exclusion column with 50 mM Na phosphate, 0.6 M NaCl, 10 mM imidazole, 1 mM DTT, and 1 mM EDTA at pH 8 as the elution buffer at 4 °C.

Protein purity was analyzed, and the proteins were further purified by reversed phase HPLC (RP HPLC). Analytical and preparative scale RP HPLC were carried out on an Agilent 1100 series HPLC with a diode array and a Zorbax C8 column (15 cm  $\times$  4.6 mm, 5  $\mu$ m pore size). Varying gradients of eluents A (0.2% trifluoroacetic acid in water) and B (0.2% trifluoroacetic acid in acetonitrile) were used with a constant flow rate of 1 mL/min. Two milliliter fractions were collected, and HPLC purified proteins were further analyzed by analytical RP HPLC, SDS-PAGE (described above), and electrospray mass spectrometry (ESMS), performed using a Mariner Biospectrometer Workstation (Perceptive Biosystems, Forester City, CA). Further details of protein purification are provided in the legend to Figure 3.

**Limited Trypsin Digestion.** QH<sub>6</sub> was purified from *E. coli* SG13009 pMS1 using a combination of Ni (NTA) affinity and gel exclusion chromatographies (as above). Purified protein at a concentration of 0.5 mg/mL in 50 mM Na phosphate, 0.6 M NaCl, 10 mM imidazole, and 1 mM DTT at pH 8.0 was diluted to a final concentration of 0.3 mg/mL and combined with trypsin present in the same buffer (Promega, Madison, WI) at a ratio of 500 ProQ/1 trypsin

(w/w). The digestion was carried out at 37 °C. Samples were taken at 15 min intervals over the first 90 min and immediately boiled in SDS-PAGE sample buffer for 10 min. Protease resistant fragments were visualized via Tricine SDS-PAGE (as above). Gel slices containing protein fragments that were trypsin resistant after 30 min of digestion were excised. Control gel slices were obtained from empty lanes of the same gels. The protein fragments in these gel slices were digested to completion with trypsin or Glu-C (Promega, Madison, WI) and the resulting peptides extracted from the gel as described (25). Peptides were desalted using Ziptips (Millipore, Mississauga, ON) and resuspended in a saturated  $\alpha$ -cyano-4-hydroxy cinnamic acid matrix in 0.1% TFA in 50% acetonitrile. The masses of the resulting peptides were determined by MALDI-TOF mass spectrometry using Reflex III MALDI TOF (Bruker Daltonics, Billerica, MA). The identities of the resistant fragments and the resulting coverage of ProQ (accession number P45577) were determined by using ProFound (26) to search the NCBI database.

**Determination of Protein Solubilities.** Protein solubilities were explored as described by Collins et al., employing the sodium salts of buffer bases 0.1 M citrate at pH 2.3, phosphate at pH 4.2, MES at pH 5.6, PIPES at pH 6.5, HEPES at pH 7.5, and TAPS at pH 8.5 (27). These buffer bases were supplemented with the following salts at 0.2 and 0.5 M: MgCl<sub>2</sub>, CaCl<sub>2</sub>, NH<sub>4</sub>Cl, LiCl, KCl, KSCN, NaNO<sub>3</sub>, or NaCl. Briefly, purified proteins at concentrations of 1 to 2 mg/mL in the elution buffer were dialyzed into deionized water overnight. The resulting precipitates were resuspended, aliquots of the suspension were put into microcentrifuge tubes, the protein was recovered by centrifugation, and the supernatant was removed. The pellets were resuspended by repeated pipetting in 0.1 mL of the buffers specified above or in water, the suspensions incubated at room temperature for 30 min, and the insoluble protein again pelleted by centrifugation. The amount of protein in solution in the trial buffer was measured with the Bradford Assay (28). Significant solubilization of the protein occurred when the quantity

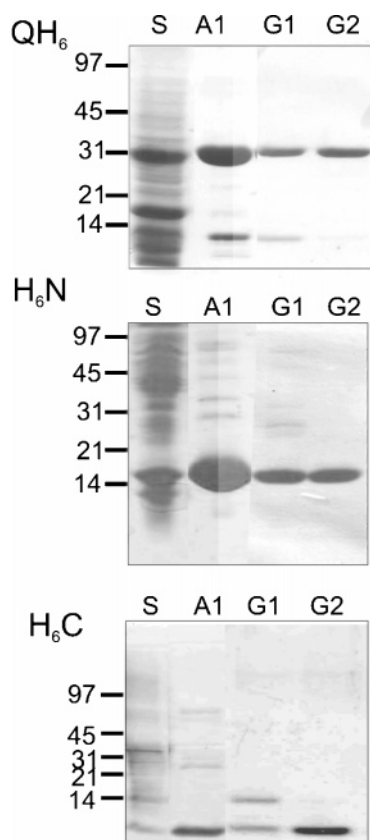


FIGURE 2: Analysis of proteins purified by Ni(NTA) and gel exclusion chromatographies. Analysis of fractions derived from the purification of QH<sub>6</sub>, H<sub>6</sub>N, and H<sub>6</sub>C. Proteins were purified, and SDS-PAGE (QH<sub>6</sub> and H<sub>6</sub>N) and Tricine SDS-PAGE (H<sub>6</sub>C) were performed as described in Experimental Procedures. The numbers to the left indicate the sizes of the molecular weight markers: S, soluble fractions; A1, proteins purified by affinity chromatography; and G1 and G2, fractions obtained by gel exclusion chromatography. Gel exclusion chromatography of QH<sub>6</sub> yielded two peaks, and both peak fractions are shown (G1 and G2). Gel exclusion chromatography of the ProQ fragments partially resolved them from contaminants, and representative fractions are shown (G1 and G2).

of protein present in the supernatant was at least 3-fold greater than the quantity of protein present in water in the same experiment.

**Circular Dichroism Spectroscopy.** Circular dichroism (CD) spectroscopy was performed on a Jasco-810 spectrometer with constant nitrogen flushing (Jasco, Inc., Easton, MD). Circular optical cells with a path length of 0.05 cm were used to determine the spectra of proteins in 0.1 M K phosphate at pH 7.4 or 0.1 M Na citrate at pH 2.3 in the presence and absence of 0.6 M NaCl over a wavelength range of 195–250 nm in 1 nm increments. The concentrations of the protein preparations were determined by amino acid analysis using a Beckman Model 6300 Amino Acid Analyzer (Beckman Coulter Inc, Fullerton, CA). Mean residue ellipticity was calculated using the following equation:

$$\theta = (\theta_{\text{obs}} \times \text{MRW}) / (10 \times \text{path length (cm)} \times \text{concentration (mg/mL)})$$

where  $\theta$  is the mean residue molar ellipticity,  $\theta_{\text{obs}}$  is the observed ellipticity, MRW is the mean residue weight, and concentration is the protein concentration determined by

amino acid analysis. Each spectrum is the average of 8 wavelength scans.

**Protein Assays.** Protein concentrations were determined using the BCA assay (29) with reagents from Pierce, (Rockford, IL) or the Bradford assay (28) with reagents from Sigma (Oakville, ON) according to manufacturers' instructions.

## RESULTS

**Limited Proteolysis Supports the Predicted Domain Structure of ProQ.** Limited proteolysis was used to compare protease resistant fragments of ProQ with the predicted domains. Tryptic digestion was selected to define protease resistant regions because ProQ contains 36 potential tryptic cleavage sites with 18 in the predicted N-terminal domain (residues 1–130), 13 in the predicted linker domain (residues 131–179), and 5 in the predicted C-terminal domain (residues 180–232).

These experiments identified four protease resistant fragments after trypsin digestion for 30 min, 2 of which remained after 90 min of digestion (referred to as Fragment 1 and Fragment 3) (Figure 1). The identities of these two fragments were further assessed by independent in-gel digestion with trypsin or Glu-C (trypsin and Glu-C cleave peptide bonds C-terminal to the basic amino acid residues and to glutamic acid, respectively). Fragment 1 had an apparent molecular weight of approximately 14 kDa (Figure 1, Tricine SDS-PAGE). In-gel digestion of this fragment with trypsin and analysis with MALDI-TOF MS identified four peptides unique to the digest (not also obtained from control gel slices, see Experimental Procedures) (Table 2). If the identified peptides were all derived from sequences near the N-terminus of ProQ, then they would span residues 12–100 (Figure 1). Their coverage of the putative N-terminal domain (residues 1–130) was calculated to be 42% by dividing the sum of the numbers of amino acids in the peptides (34 + 20 = 54, Table 2) by the number of amino acids in the putative N-terminal domain (130) and multiplying by 100%. However, the peptide with a mass of 1745 Da could correspond to either N-terminal residues 21–35 or C-terminal residues 211–226 (Table 2). This ambiguity was resolved by in-gel digestion with Glu-C, which produced four unique peptides (Table 2) that spanned residues 20–105 and covered 32% of the putative N-terminal domain (Figure 1). Thus, Fragment 1 corresponded to the putative N-terminal domain of ProQ.

Fragment 3 had an apparent molecular weight of approximately 9.3 kDa (Figure 1, Tricine SDS-PAGE). In-gel tryptic digestion of this fragment and analysis by MALDI-TOF MS revealed seven unique peptides (Table 2), only three of which could be assigned unambiguously to the predicted C-terminal domain of ProQ (Table 2 and Figure 1). In-gel digestion of Fragment 3 with Glu-C and MALDI-TOF MS revealed two unique peptides, one derived from a sequence near the N-terminus and another derived from a sequence near the C-terminus (Table 2 and Figure 1). Fragment 3 has an apparent molecular weight of 9 kDa, and thus, it is not large enough to span amino acids 21–228 of ProQ (Figure 1). Fragment 3 is more likely made up of two fragments with similar molecular weights, which were not fully resolved by Tricine SDS-PAGE. One of these fragments would be derived from the predicted C-terminal

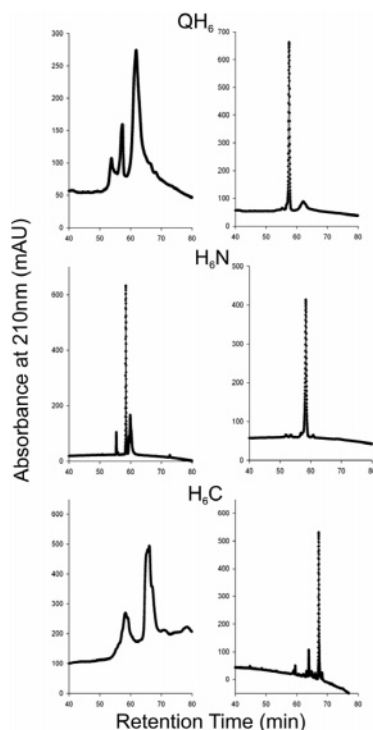


FIGURE 3: Analysis of proteins purified by reversed phase HPLC. Analytical RP HPLC was performed with a linear gradient of eluents A (0.2% trifluoroacetic acid in water) and B (0.2% trifluoroacetic acid in acetonitrile) at a rate of 2% per minute over 40 min. No peaks were seen with retention times below 40 min or above 80 min for any of the protein preparations. (QH<sub>6</sub>, Left) QH<sub>6</sub> (at 10.6 mg/mL) purified by Ni(NTA) affinity and gel exclusion chromatography was dialyzed into 0.1 M Na-citrate at pH 2.3 and diluted to 1 mg/mL in solvent A. A sample (20  $\mu$ L) was analyzed by analytical RP HPLC (as above). (QH<sub>6</sub>, Right) QH<sub>6</sub> (12.7 mg at 10 mg/mL) in 0.1 M Na citrate at pH 2.3 was diluted 4-fold into 0.2% TFA in water for further purification by RP HPLC using a discontinuous linear gradient of solvents A and B as follows. A 2% B per min linear gradient was used from 1 to 20% B, a 1% B per min linear gradient was used from 21 to 28% B, a 0.1% B per min linear gradient was used from 28 to 48% B, and a 4.2% B per min gradient was used from 48 to 90% B. The trace on the right was obtained by analytical RP HPLC (performed as above) on a fraction corresponding to the middle peak in the chromatogram of the partially purified preparation (QH<sub>6</sub>, left). (H<sub>6</sub>N, Left) Purified H<sub>6</sub>N (0.06 mg at 30 mg/mL) in 0.05 M sodium phosphate, 0.6 M NaCl, and 0.25 M imidazole at pH 8.0 was diluted 200-fold into buffer A. A sample (20  $\mu$ L) was analyzed by RP HPLC as outlined above. (H<sub>6</sub>N, Right) The same preparation (0.4 mL) was diluted 10-fold with solvent A and purified further using RP HPLC with a discontinuous linear gradient of solvents A and B as follows. A 2% B per min linear gradient was used from 1 to 20% B, a 1% B per min linear gradient was used from 21 to 28% B, a 0.1% B per min linear gradient was used from 28 to 48% B, and a 4.2% B per min gradient was used from 48 to 90% B. The trace on the right was obtained by analytical RP HPLC (outlined above) on a fraction corresponding to the first major peak in the chromatogram of the partially purified protein (H<sub>6</sub>N, left). (H<sub>6</sub>C, Left) H<sub>6</sub>C, purified only by affinity chromatography and present at 1.6 mg/mL in 0.05 M Na phosphate, 0.6 M NaCl, and 0.25 M imidazole at pH 8.0, was diluted 50-fold into buffer A. A sample (20  $\mu$ L) of this mixture was analyzed by RP HPLC as outlined above. (H<sub>6</sub>C, Right) H<sub>6</sub>C (7 mg in 5 mL of 0.05 M Na phosphate, 0.6 M NaCl, and 0.25 M imidazole at pH 8.0) was mixed with 10 mL of 8 M urea in 0.2% TFA (the final buffer was 0.017 M Na phosphate, 0.2 M NaCl, 0.083 M imidazole, 5.33 M urea, and 0.013% TFA) and purified using a discontinuous linear gradient as follows: 2% B per min gradient from 0 to 24% B, 1% B per min from 24 to 32% B, 0.1% B per min from 32 to 52% B, and 4.8% B per min from 52 to 100% B. The right trace represents the chromatogram obtained by analytical RP HPLC (outlined above) on a fraction corresponding to the second major peak in the chromatogram of the partially purified protein (H<sub>6</sub>C, left).

Table 2: Masses and Identities of Peptides Derived from Tryptic ProQ Fragments by In-Gel Trypsin and Glu-C Digestions

sample <sup>a</sup>	enzyme <sup>b</sup>	measured peptide mass (Da) <sup>c</sup>	peptide identity (ProQ residues) <sup>d</sup>
fragment 1	trypsin	1047.203	12–20
		1175.207	36–45
		1745.221 <sup>e</sup>	21–35
			211–226
		2289.169	81–100
fragment 1	Glu-C	1010.693	95–105
		1150.731	20–28
		1336.823	20–30
		2281.5	29–49
fragment 3	trypsin	1175.215	36–45
		1316.252	215–226
		1332.222	215–226 (oxidized)
		1745.238 <sup>f</sup>	211–226
			21–35 (oxidized)
		2122.356 <sup>g</sup>	174–193
			36–54 (oxidized)
fragment 3	Glu-C	2446.419	171–193
		1336.648	20–30
		2302.530	208–228

<sup>a</sup> ProQ protein fragments 1 and 3 were obtained as described in the text and Figure 1. <sup>b</sup> In-gel digestion of each fragment with the indicated enzyme was performed as described in Experimental Procedures. <sup>c</sup> The peptide masses were determined by MALDI-TOF MS as described in Experimental Procedures. <sup>d</sup> The identities of the detected peptides were determined by searching the NCBI database using ProFound (26). <sup>e</sup> This mass was found to correspond to ProQ residues 21–35 by GluC digestion (see text). <sup>f</sup> It was not possible to unambiguously assign these masses to either of the listed ProQ peptides (see text).

domain of ProQ. It would include amino acids 171–228. The tryptic peptides would then cover 74% of the putative C-terminal domain (residues 180–232), and the Glu-C peptide would cover 40% of that domain (Figure 1). The other fragment present in band 3 would include residues 20–54 in the predicted N-terminal domain of ProQ. The tryptic peptides would cover 26% of that domain, whereas the Glu-C peptide would cover 8% (Figure 1).

Thus, limited trypsin digestion of purified ProQ yielded protease resistant fragments corresponding to the predicted N- and C-terminal domains. In contrast, fragments corresponding to predicted tryptic fragments within the linker region (amino acids 130–173) were not found. These observations are consistent with the predicted domain structure of ProQ.

**Putative ProQ Domains Could Be Overexpressed and Purified.** A multiple sequence alignment of 12 ProQ homologues (14) was used to define the boundaries of the putative N- and C-terminal domains. Each domain showed sequence and secondary structure conservation, whereas the linker varied in length and amino acid composition. Predominantly  $\alpha$ -helical and  $\beta$ -sheet secondary structures were predicted for peptides corresponding to *E. coli* ProQ residues M1–E130 and V180–F232, respectively. This was true even though the sequence identity with *E. coli* ProQ varied from 99% for the *Shigella flexneri* homologue to only 39% for the *Shewanella oneidensis* homologue (14). Thus, for further studies, M1–E130 and V180–F232 were chosen as boundaries for the N- and C-terminal domains, respectively.

The putative N- and C-terminal domains of ProQ (H<sub>6</sub>N and H<sub>6</sub>C, Table 1) were overexpressed and purified to determine whether each could fold as a protease resistant



Table 3: Molecular Weights of Purified ProQ Fragments

protein		mass (Da)	
name	sequence	predicted	measured <sup>a</sup>
QH <sub>6</sub>	ProQ-H <sub>6</sub>	26958.6	26963.62
H <sub>6</sub> N	MRGSH <sub>6</sub> GS-ProQ(M1-E130)	15980.9	15999.2
H <sub>6</sub> C	MRGSH <sub>6</sub> GSM-ProQ(S181-F232)	7125.2	7127.4

<sup>a</sup> The masses of proteins purified by RP HPLC were determined by ESMS as described in Experimental Procedures and Results.

domain *in vivo*. Preliminary studies showed that H<sub>6</sub>N and H<sub>6</sub>C were expressed maximally when strain BL21-Gold containing the relevant plasmid (Table 1) was grown in LB medium to an OD<sub>600</sub> of 0.6, and expression was induced with 1 mM IPTG for 4 h at 37 °C (data not shown). QH<sub>6</sub> was overexpressed as previously described (14).

QH<sub>6</sub> and H<sub>6</sub>N were purified by Ni(NTA) affinity and gel exclusion chromatographies (Figure 2), and the purity of the resulting preparations was determined by analytical RP HPLC (Figure 3, left column of chromatograms). This material was further purified by preparative RP HPLC and the purity of the resulting preparations determined by analytical RP HPLC (Figure 3, right column of chromatograms) (see Experimental Procedures and the legend to Figure 3). In each of the resulting RP HPLC preparations, a single protein of the expected molecular mass was evident upon SDS-PAGE and dominated the mass spectrum (Table 3). Analytical RP HPLC of QH<sub>6</sub> purified by Ni(NTA) affinity and gel exclusion chromatographies (Figure 2) yielded a chromatogram that had three large peaks with retention times of 53.8, 57.3, and 61.9 min (Figure 3, left). Protein recovered from the second and third peaks was identified as QH<sub>6</sub> by ESMS and SDS-PAGE (data not shown). The basis for the resolution of these protein fractions was not determined. Protein from the peak with a retention time of 57.3 min was used during subsequent analyses. The composition of that material is illustrated in the right panel of Figure 3.

H<sub>6</sub>N was purified to near homogeneity by Ni(NTA) affinity chromatography (Figure 2, lane A1). Gel exclusion chromatography removed minor contaminants present in the preparation (Figure 2, fraction G2). Analytical RP HPLC detected some minor contaminants (Figure 3, left). H<sub>6</sub>N was resolved from the minor contaminants by preparative RP HPLC, as shown by subsequent analytical RP HPLC (Figure 3, right), ESMS (Table 3), and SDS-PAGE (data not shown).

H<sub>6</sub>C could be partially purified, but many contaminants were evident when it was viewed by Tricine SDS-PAGE (Figure 2, lane A1). Gel exclusion chromatography resolved part of the H<sub>6</sub>C from the contaminants (Figure 2, lane G2); however, the protein recovery was poor. For further experimentation, H<sub>6</sub>C was only purified by Ni(NTA) affinity chromatography to reduce protein losses. Analytical RP HPLC showed that many contaminants remained (Figure 3, left), but SDS-PAGE and ESMS detected only protein with the predicted molecular weight of H<sub>6</sub>C, and that protein was only found in the major peak. Instability in solution prevented further purification of H<sub>6</sub>C by RP HPLC using the buffers employed for the other ProQ fragments. It was, therefore, denatured in 0.05 M Na phosphate, 0.6 M NaCl, and 0.25 M imidazole at pH 8.0 containing 6 M urea prior to RP HPLC purification. The resulting preparation was predomi-

nantly H<sub>6</sub>C as determined by RP HPLC (Figure 3, right), Tricine SDS-PAGE (not shown), and ESMS (Table 3). These experiments showed that N- and C-terminal fragments of ProQ folded *in vivo* and could be purified, supporting the hypothesis that they represent domains present in full length ProQ.

**Solubilities of ProQ and Its Fragments.** The solubilities of H<sub>6</sub>N, H<sub>6</sub>C, and full length QH<sub>6</sub> were explored as described by Collins et al. (27). Among the buffers tested, only Na citrate (0.1 M, pH 2.3) could resolubilize aggregated H<sub>6</sub>N, H<sub>6</sub>C, and full length QH<sub>6</sub>. Purified QH<sub>6</sub> and H<sub>6</sub>N could be dialyzed into this citrate buffer and concentrated above 25 mg/mL using Millipore YM-10 (MWCO 10,000 Da) centricon devices (Millipore, Mississauga, ON). H<sub>6</sub>N was also soluble in MES (0.1 M, pH 5.6), but H<sub>6</sub>C and QH<sub>6</sub> were not.

**Secondary Structures of the ProQ Fragments.** Circular dichroism (CD) spectroscopy was used to determine whether H<sub>6</sub>N and H<sub>6</sub>C fold *in vivo* to form domains with the secondary structures predicted by Smith et al. (14) and to compare the secondary structures of these proteins in buffers that enhance their solubilities. To determine whether the spectra were influenced by contaminants present before RP HPLC or whether the denaturation of the protein during purification via RP HPLC was irreversible, the CD spectra of proteins purified by Ni (NTA) affinity and gel exclusion chromatographies were compared with those of proteins purified by RP HPLC, then lyophilized, and redissolved in the indicated buffers (see the legend to Figure 4). No major differences in spectra were observed, and the RP HPLC purified proteins were used to further analyze each protein's secondary structure in various buffers. Smith et al. reported that NaCl at concentrations above 0.6 M solubilized QH<sub>6</sub> in Na phosphate buffer at neutral pH and that the CD spectrum of untagged ProQ was pH- but not NaCl-sensitive (14). The CD spectrum of QH<sub>6</sub> in 0.1 M K phosphate buffer at pH 7.4 (Figure 4, top) was similar to that of ProQ in 10 mM K phosphate at pH 7.4 (14). These spectra were taken to represent the native secondary structure of ProQ. The addition of 0.6 M NaCl did not noticeably change the CD spectrum (Figure 4, top). Similar spectra were also obtained with the low pH Na citrate buffer (0.1 M, pH 2.3) in the presence of 0.6 M salt. However, a different spectrum, which showed considerably more random coil structure, was obtained in the absence of salt at pH 2.3.

In CD spectra,  $\alpha$ -helical secondary structure is characterized by strong ellipticity minima at 208 and 222 nm. In highly  $\alpha$ -helical proteins, the ellipticities at 222 and 208 nm are similar. In proteins with mixed  $\alpha$ -helical and random structure, the ellipticity at 222 nm is much less than that at 208 nm because random structure exhibits a CD spectral minimum below 200 nm (as shown in Figure 4 for QH<sub>6</sub>). Canonical  $\beta$ -sheet structure exhibits a weak minimum at 217 nm. Like QH<sub>6</sub>, H<sub>6</sub>N was expected to assume its native conformation in 0.1 M K phosphate buffer at pH 7.4. The CD spectrum of H<sub>6</sub>N indicated that it is mostly  $\alpha$ -helical (Figure 4, middle), in agreement with the model for N based on the crystal structure of FinO (14, 15). There was very little change in the CD spectrum when 0.6 M NaCl was added to this preparation (Figure 4, middle). As for QH<sub>6</sub>, the change of pH to 2.3 in the presence of 0.6 M NaCl had no effect on the spectrum. Also as for QH<sub>6</sub>, the spectrum

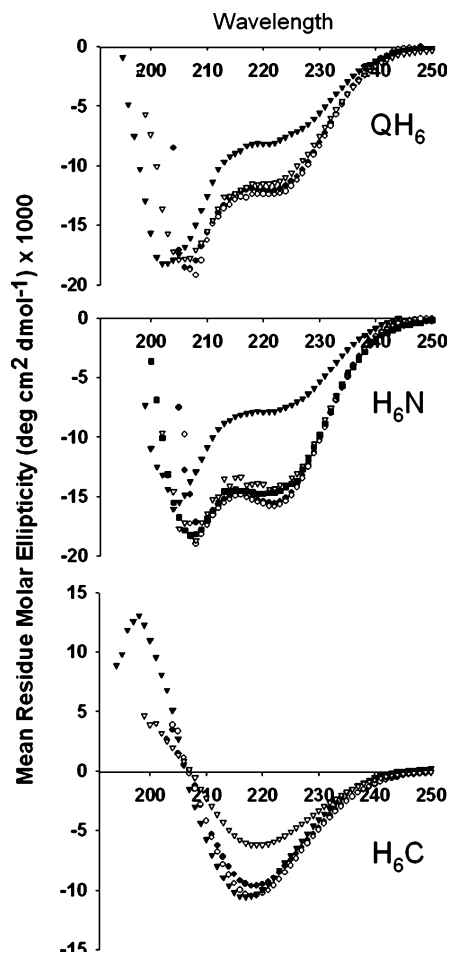


FIGURE 4: Circular dichroism spectra of QH<sub>6</sub>, H<sub>6</sub>N, and H<sub>6</sub>C. RP HPLC purified and lyophilized QH<sub>6</sub>, H<sub>6</sub>N, and H<sub>6</sub>C were dissolved at 1.6 mg/mL in 0.1 M K phosphate at pH 7.4. The solutions were dialyzed into 0.1 M K phosphate at pH 7.4 or 0.1 M Na citrate at pH 2.3. The dialysates were diluted with the same buffer or with the same buffer supplemented with NaCl to give 0.04 mg/mL protein in 0.1 M K phosphate buffer at pH 7.4 (circles) or 0.1 M Na-Citrate buffer at pH 2.3 (triangles), alone (closed symbols) or supplemented with 0.6 M NaCl (open symbols). H<sub>6</sub>N was also analyzed in 0.1 M MES at pH 5.6 (squares). CD spectra were recorded as described in Experimental Procedures.

showed a large decrease in  $\alpha$ -helical structure in the absence of salt at pH 2.3. CD analysis of H<sub>6</sub>N in 0.1 M MES showed that it folded similarly in that buffer and in K phosphate (Figure 4, middle). Thus, H<sub>6</sub>Q and H<sub>6</sub>N behaved similarly, being sensitive to low pH in the absence of salt. CD spectroscopy of H<sub>6</sub>C showed that its structure was mostly  $\beta$ -sheet. The spectra were similar at pH 7.4 in the presence and absence of salt and at pH 2.3 in the absence of salt. However, the addition of 0.6 M NaCl to the pH 2.3 buffer disrupted the structure (Figure 4, bottom).

These experiments showed that tagged, full length ProQ (QH<sub>6</sub>) and the expressed N- and C-terminal fragments of ProQ (H<sub>6</sub>N and H<sub>6</sub>C, Table 1) folded to form domains with the anticipated secondary structures in 0.1 M K phosphate buffer at pH 7.4. CD spectroscopy offered no evidence for the formation of an SH3-like domain by H<sub>6</sub>C. Although these proteins were soluble in Na citrate buffer at pH 2.3, the absence of salt increased the random coil structure of QH<sub>6</sub> and H<sub>6</sub>N. Interestingly, the presence of salt at pH 2.3 decreased the structure of H<sub>6</sub>C. H<sub>6</sub>N was also found to be soluble and structured in 0.1 M MES buffer at pH 5.6, where

it can be concentrated to more than 30 mg/mL. Thus, 0.1 M MES is a suitable buffer for further physiological studies of the N-terminal domain of ProQ.

**Complementation and Competition Analysis.** Untagged homologues of various ProQ fragments (Table 1) were expressed at physiological levels to test their *in vivo* functions. To determine if these proteins could complement a chromosomal *proQ* deletion or impair the function of full length ProQ *in vivo*, it was necessary to ensure that they were expressed at comparable levels. This study employed the vector pBAD24, which allows for the tight control of protein expression from the arabinose inducible *araBAD* promoter. ProQ was present at 2 nmol/mg cell protein when *proQ* was expressed from the pBAD24 derived plasmid pDC77 in *E. coli* WG914 cultivated in MOPS medium (see Experimental Procedures). The levels of arabinose supplementation yielding equivalent expression levels for the ProQ fragments were determined by quantitative Western blotting with purified full length ProQ (QH<sub>6</sub>), the ProQ domains (H<sub>6</sub>N and H<sub>6</sub>C), and a variant lacking the linker region (H<sub>6</sub>NC) as standards (see Experimental Procedures). The fragments that included a linker (NL, LC, and L) and that lacking the linker (NC) were degraded *in vivo* (Figure 5B). This degradation yielded subfragments of similar size to the N- or the C-terminal domain identified by limited proteolysis *in vitro* (Figure 1), suggesting that these two ProQ domains were also present *in vivo*.

Transposon insertion in or deletion of *proQ* decreases ProP activity. Expression of ProQ or QH<sub>6</sub> from a plasmid complements *proQ* lesions, indicating that *proQ* is responsible for the effects on ProP (13, 14). In this study, the impacts of ProQ fragments on ProP activity were determined by measuring the proline uptake activities of bacteria in which ProP was the only proline transporter expressed, and *proQ* was either expressed from the chromosome (*E. coli* WG210) or absent because of an in-frame chromosomal deletion (*E. coli* WG914) (Table 1S, Supporting Information) (14).

As previously observed, the *proQ* deletion decreased ProP activity, and the plasmid based expression of *proQ* from plasmid pDC77 restored it. Plasmid based *proQ* expression yielded higher ProQ levels than chromosome based *proQ* expression, yet the resulting ProP activities were similar (Figure 5AI). The expression of L, LC, or C did not amplify ProP activity, whereas the expression of N or NL partially restored ProP activity (ProQ fragments are defined in Table 1) (Figure 5AI). Further increasing the expression of N or NL did not increase ProP activity above this intermediate level (data not shown). The expression of a ProQ variant lacking the 50 amino acid linker region (NC) fully restored ProP activity (Figure 5AI). As anticipated, N, NL, LC, and NC were all expressed at levels at least equivalent to that of ProQ (Figure 5BI). Neither C nor L could be detected despite full arabinose induction (Figure 5BI). However, the failure of LC to complement the *proQ* deletion suggests that these subfragments would also be ineffective. Thus, the N-terminal domain of ProQ may contain residues that are important for ProQ function because this domain gives partial complementation. However, both the N- and C-domains of ProQ are important for its complete *in vivo* function.

Because ProQ is predicted to interact with ProP, the expression of an interacting ProQ domain in the presence



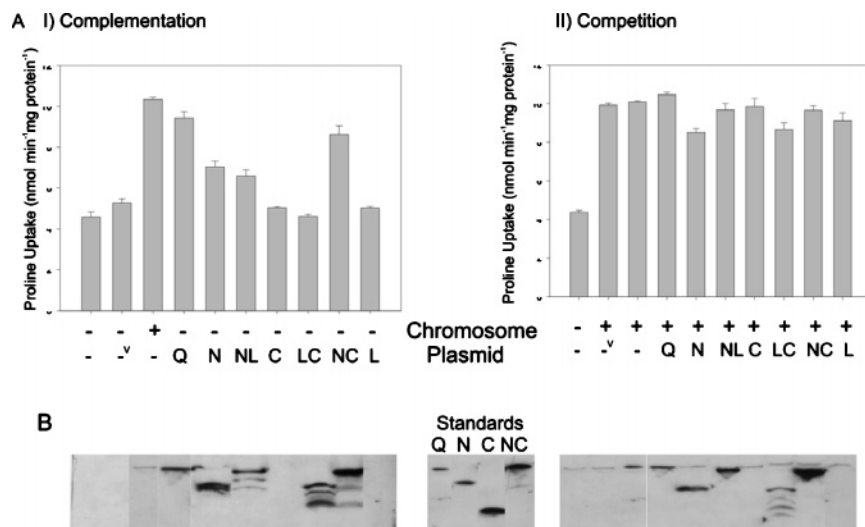


FIGURE 5: Functional analysis of ProQ fragments. (Panel A) The proline uptake activities of bacteria expressing ProQ fragments were determined as outlined in Experimental Procedures. Proline uptake rates were measured in the presence (+) or absence (–) of chromosomally encoded ProQ (Chromosome) and in the absence of plasmid (–), in the presence of the empty pBAD24 vector (–<sup>v</sup>), or in the presence of plasmid-encoded Q, N, C, L, NL, LC, or NC (Plasmid). (The ProQ fragments are defined in Table 1.) Cells harbored a *proQ* deletion (I) or wild type *proQ* (II) at the position of the chromosomal *proQ* locus. (Panel B) Protein expression levels were determined by Western blotting with anti-ProQ antibodies as outlined in Experimental Procedures. Samples were loaded at a concentration of 0.08 mg of total cell protein. The central panel in B shows histidine tagged protein standards, purified via Ni(NTA) affinity chromatography. Quantities of these proteins equimolar to the quantity of ProQ in cells with plasmid pDC77 were used. These quantities were 0.1  $\mu$ g (QH<sub>6</sub>), 0.082  $\mu$ g (H<sub>6</sub>QC), 0.06  $\mu$ g (H<sub>6</sub>QN), and 0.026  $\mu$ g (H<sub>6</sub>QC).

of the full length protein may be expected to decrease ProP activity. However, none of the tested ProQ fragments impaired the ability of full length ProQ to amplify ProP activity (Figure 5AII and BII).

## DISCUSSION

Multiple studies have shown that the activity of the osmoregulatory transporter ProP is attenuated when ProQ is absent (12–14, 30). ProQ does not amplify ProP activity by increasing *proP* transcription or ProP levels (12, 13). Thus, ProQ may act directly on ProP (1, 14).

The full length ProQ protein was previously purified and found to have limited solubility. Smith et al. proposed that ProQ comprises an N-terminal  $\alpha$ -helical domain connected to a C-terminal SH3-like domain by an unstructured linker (14). The mechanism by which ProQ acts on ProP cannot be predicted via sequence analysis because ProQ has no sequence homologues with known functions. To test the structural model and design functional tests, we first assessed the domain structure of ProQ by limited proteolysis. As predicted, N-terminal and C-terminal fragments of ProQ were protease resistant and hence apparently folded, whereas the central linker was protease sensitive and hence apparently unstructured (Figure 1). The protease resistant domains persisted even after 90 min treatments with trypsin, even though many trypsin cleavage sites are present inside each domain.

The identification of functional ProQ domains may allow us to engineer a proxy protein that is more stable and soluble in a wider variety of buffer systems than full length ProQ *in vitro*. The boundaries of the putative N- and C-terminal domains of *E. coli* ProQ were estimated to be M1–E130 and V180–F232, respectively (see Results). These fragments could be overexpressed and purified using a combination of Ni(NTA) affinity, gel exclusion, and reversed phase chromatographies. Each had the predicted secondary structure,

as determined by CD spectroscopy, because the N-terminal domain (H<sub>6</sub>N) was  $\alpha$ -helical, whereas the C-terminal domain (H<sub>6</sub>C) included  $\beta$ -sheet structure (Figure 4). By performing simple solubility screens, QH<sub>6</sub>, H<sub>6</sub>N, and H<sub>6</sub>C were found to be soluble at high concentrations in a second buffer, 0.1 M Na citrate at pH 2.3. High concentrations of H<sub>6</sub>N but not QH<sub>6</sub> or H<sub>6</sub>C could be attained in a third buffer, 0.1 M MES at pH 5.6. The impacts of these buffers on the secondary structures of the purified proteins were examined. Each protein was soluble in 0.1 M Na citrate at pH 2.3. QH<sub>6</sub> and H<sub>6</sub>N lost secondary structure in this buffer, but the structure could be restored in the presence of 0.6 M NaCl (Figure 4). However, H<sub>6</sub>C maintained its structure in this buffer but lost structure in the presence of salt (Figure 4). Previous studies showed that high salt cell lysis buffers improved the yields of ProQ and QH<sub>6</sub> by enhancing their solubilities (14). The secondary structures of QH<sub>6</sub> and H<sub>6</sub>N were maintained in high salt buffers, but that of H<sub>6</sub>C was perturbed (Figure 4). These data suggest that the native structure is maintained when the N-terminal domain of ProQ is expressed separately, but the secondary structure of the C-terminal domain is more labile. Future co-reconstitution studies to assess the interactions of ProQ with ProP in proteoliposomes may be possible using the N-terminal domain of ProQ, alone, in 0.1 M MES buffer at pH 5.6.

H<sub>6</sub>C was predicted to be SH3-like and hence to possibly mediate interactions of ProQ with other proteins (31). Although SH3 domains are predominantly  $\beta$ -sheet in structure, their CD spectra are atypical in having positive ellipticity with a maximum at 220 nm (32). Because the CD spectrum of H<sub>6</sub>C had a minimum at 220 nm in the tested buffers (Figure 4, bottom), this work does not indicate that H<sub>6</sub>C is SH3-like, but it does support the prediction that the C-terminal domain of ProQ is composed of mostly  $\beta$ -sheet structure.

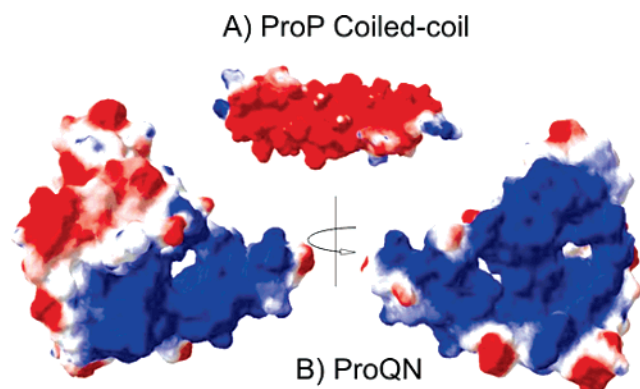


FIGURE 6: Electrostatic surfaces of ProP and the N-terminal domain of ProQ. The SPDV program Deep view was used to calculate the electrostatic surface potentials of the C-terminal coiled-coil of ProP and the N-terminal domain of ProQ. (Blue is positive, red negative, and white uncharged.) (A) NMR structure of ProP residues 468–497, which corresponds to the C-terminal homodimeric, antiparallel coiled-coil structure (pdb accession code 1R48) (9). (B) N-terminal domain of ProQ (amino acid residues 1–124) modeled on the crystal structure of FinO (pdb code 1DVO) as previously described (14). The modeled structure was rotated through 180° to show two surfaces.

To identify functionally important regions of ProQ, fragments N, NL, LC, C, L, and NC (Table 1) were expressed at physiologically relevant levels using vector pBAD24 (33). Some were degraded *in vivo* to subfragments corresponding in size to those identified as the N- and C-terminal domains after limited proteolysis *in vitro* (Figures 5 and 1). This further suggested the presence of two ProQ domains.

Functional analysis revealed that the N-terminal domain is important for the amplification of ProP activity. The expression of fragment N, alone, partially restored ProP activity to the *proQ* null mutant (Figure 5AI). Interestingly, the C-terminal fragment (C) did not complement but the ProQ fragment lacking the linker (NC) completely complemented the *proQ* null mutation (Figure 5AI). In contrast to the other fragments, neither C nor L could be detected within the cells used for these assays by Western blotting (Figure 5BI). However, fragments corresponding in size to L and C were produced by cells expressing fragment LC, (Figure 5BI), and the expression of LC and these fragments did not affect ProP activity (Figure 5AI). It is, therefore, unlikely that fragment L or C, alone, can amplify ProP activity. We predicted that fragments N and NL, those that partially complemented a ProQ deficiency, might also compete with ProQ *in vivo* and hence decrease ProP activity. None of the fragments, when expressed in the *proQ*<sup>+</sup> background, reduced ProP activity (Figure 5AII). This may indicate that the N-terminal domain contains residues that are important for functional interactions, but these interactions may be weak and require stabilization by another part of ProQ, possibly the C-terminal domain.

The sequence and predicted structure of the N-terminal domain of ProQ suggest two possible modes of interaction with ProP. The N-terminal domain of ProQ may interact electrostatically with the C-terminal coiled-coil domain of ProP, or the N-terminal domain of ProQ may form a coiled-coil with the C-terminal coiled-coil domain of ProP. The C-terminal domain of ProQ may stabilize either interaction,

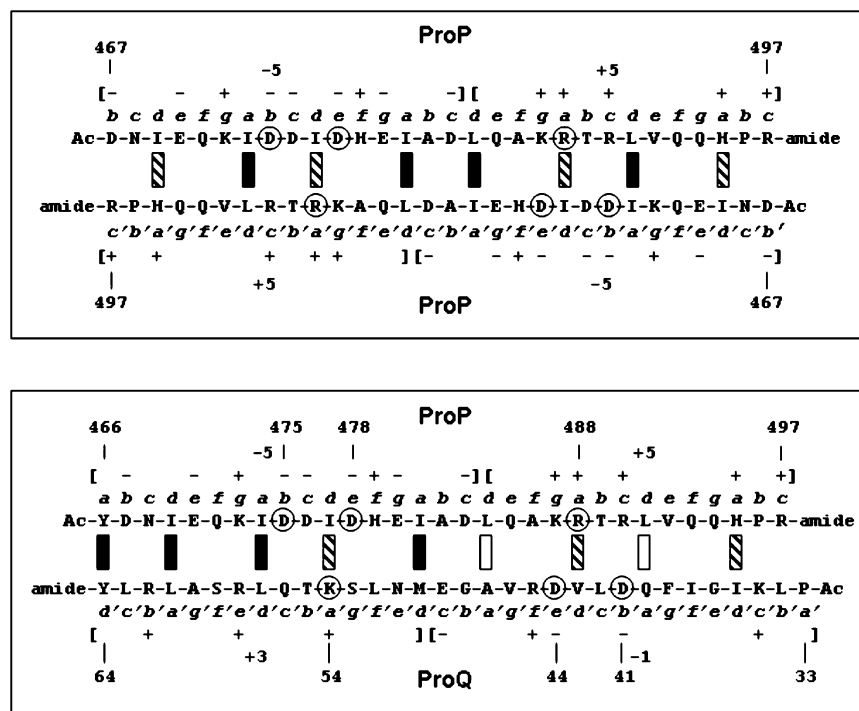


FIGURE 7: Comparison of the demonstrated ProP/ProP and the predicted ProP/ProQ coiled-coils. The top panel shows the antiparallel homodimeric coiled-coil formed by a peptide corresponding to the C-terminal sequence of ProP (9). The bottom panel shows the predicted antiparallel heterodimeric coiled-coil formed by the same region of ProP and residues 33–64 of ProQ. The hydrophobic core positions *a* and *d* are denoted by three types of boxes (solid, two hydrophobic side chains in the core; hatched, a hydrophobic side chain across from a destabilizing, positively charged residue in the core; open, a hydrophobic side chain across from a destabilizing residue, Ala or Gln, in the core). Each coiled-coil would be stabilized by electrostatic interactions arising from the overall peptide polarity (charges in square brackets flanked by the net charge of each subsequence). In addition, the positively charged core Arg or Lys could form salt bridges with negatively charged Asp. (These residues are circled.)

as noted above, or it may link the activity of the N-terminal domain to other cellular processes.

ProQ is basic (pI 9.7), and the majority of its basic residues are in the N-terminal domain that can be modeled on the mRNA binding protein FinO (13, 14) (Figure 6). A homology model of ProP predicts that its cytoplasmic face is basic (5). The NMR structure of the homodimeric coiled-coil formed by peptide replicas of ProP residues 468–497 (pdb code 1R48) shows an antiparallel structure with an acidic surface (9) (Figure 6). ProP peptides form this coiled-coil structure *in vitro*, and evidence suggests that it is present within ProP dimers *in vivo* (10). Electrostatic interactions with ProQ may stabilize the ProP coiled-coil and hence the active conformation of ProP.

The ProP coiled-coil contains six to seven of the heptad sequence repeats characteristic of  $\alpha$ -helical coiled-coils. Typically, amino acids with hydrophobic side chains are present at most heptad *a* and *d* positions. The antiparallel orientation of the coiled coil may be favored by the formation of salt bridges between residues R488 (*a* position) of one strand and residues D475 and D478 of the opposing strand and by the overall polarity of the peptide (9). It is possible that ProQ interacts with ProP by forming heterodimeric coiled-coil structures. ProQ residues 33–64 comprise heptad repeats that could form an antiparallel coiled-coil structure with the C-terminal coiled-coil domain of ProP. Salt bridges analogous to those that stabilize the ProP antiparallel homodimer could similarly stabilize the ProP/ProQ antiparallel heterodimer (Figure 7). For example, salt bridges could form between R488 of ProP and D41 and D44 of ProQ as well as between K54 of ProQ and D475 and D478 of ProP. Acidic and basic residues within this ProQ sequence would thereby complement charges present in the ProP sequence (Figure 7).

This study shows that ProQ of *E. coli* is composed of two domains connected by a protease accessible linker. Residues important for ProQ activity may occur within the N-terminal, 130 amino acid domain of the protein. Full complementation of a *proQ* deletion requires the presence of the N-terminal 130 and the C-terminal 52 residues but not the 50 amino acid linker. When expressed separately, the N-terminal domain appears to retain the secondary structure present within the full length protein. Thus, the N-terminal domain or the fused N- and C-terminal domains of ProQ may serve as valid proxies for further *in vitro* studies of ProP–ProQ interactions.

## ACKNOWLEDGMENT

We are grateful to Robert A. B. Keates (University of Guelph) for his help with the ProQ sequence and structure alignments, to Brian Tripet (University of Colorado at Denver and Health Sciences Center) for the amino acid analysis, and to Dyanne Brewer (Mass Spectrometry Facility, University of Guelph) for assistance with the in-gel digestions and mass spectrometry analysis.

## SUPPORTING INFORMATION AVAILABLE

*E. coli* strains and plasmids used for this study (Table 1S) and the oligonucleotide primers used in plasmid construction (Table 2S). This material is available free of charge via the Internet at <http://pubs.acs.org>.

## REFERENCES

- Wood, J. M. (1999) Osmosensing by Bacteria: signals and membrane-based sensors, *Microbiol. Mol. Biol. Rev.* 63, 230–262.
- 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.
- Milner, J. L., Grothe, S., and Wood, J. M. (1988) Proline porter II is activated by a hyperosmotic shift in both whole cells and membrane vesicles of *Escherichia coli* K12, *J. Biol. Chem.* 263, 14900–14905.
- Racher, K. I., Voegelé, 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.
- Wood, J. M., Culham, D. E., Hillar, A., Vernikovska, Ya. I., Liu, F., Boggs, J. M., and Keates, R. A. B. (2005) Structural model for the osmosensor, transporter, and osmoregulator ProP of *Escherichia coli*, *Biochemistry* 44, 5634–5646.
- Huang, Y., Lemieux, M. J., Song, J., Auer, M., and Wang, D.-N. (2003) Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*, *Science* 301, 616–620.
- Culham, D. E., Tripet, B., Racher, K. I., Voegelé, R. T., Hodges, R. S., and Wood, J. M. (2000) The role of the carboxyl terminal  $\alpha$ -helical coiled-coil domain in osmosensing by transporter ProP of *Escherichia coli*, *J. Mol. Recognit.* 13, 1–14.
- Hillar, A., Tripet, B., Zoetewey, D., Wood, J. M., Hodges, R. S., and Boggs, J. M. (2003) Detection of  $\alpha$ -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.
- 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.
- Hillar, A., Culham, D. E., Vernikovska, Ya. I., Wood, J. M., and Boggs, J. M. (2005) Formation of an antiparallel, intermolecular coiled-coil is associated with *in vivo* dimerization of osmosensor and osmoprotectant transporter ProP in *Escherichia coli*, *Biochemistry* 44, 10170–10180.
- Tsatskis, Y., Khambati, J., Dobson, M., Bogdanov, M., Dowhan, W., and Wood, J. M. (2005) The osmotic activation of transporter ProP is tuned by both its C-terminal coiled-coil and osmotically induced changes in phospholipid composition, *J. Biol. Chem.* 280, 41387–41394.
- Milner, J. L., and Wood, J. M. (1989) Insertion *proQ220::Tn5* alters regulation of proline porter II, a transporter of proline and glycine betaine in *Escherichia coli*, *J. Bacteriol.* 171, 947–951.
- 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.
- Smith, M. N., Crane, R. A., Keates, R. A. B., and Wood, J. M. (2004) Overexpression, purification and characterization of ProQ, a post-translational regulator for osmoregulatory transporter ProP of *Escherichia coli*, *Biochemistry* 43, 12979–12989.
- 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.
- Bauer, C. B., Kuhlman, C. A., Bagshaw, C. R., and Rayment, I. (1997) X-ray crystal structure and solution fluorescence characterization of Mg<sup>2+</sup>·(3′)-O-(N-methylanthraniloyl) nucleotides bound to the *Dictyostelium discoideum* myosin motor domain, *J. Mol. Biol.* 274, 394–407.
- Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids, *J. Mol. Biol.* 166, 557–569.
- 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.
- Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.



21. Neidhardt, F. C., Bloch, P. L., and Smith, D. F. (1974) Culture medium for enterobacteria, *J. Bacteriol.* 119, 736–747.
22. 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.
23. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature (London)* 227, 680–685.
24. Schagger, H., and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa, *Anal. Biochem.* 166, 368–379.
25. Zachertowska, A., Brewer, D., and Evans, D. H. (2006) Characterization of the major capsid proteins of myxoma virus particles using MALDI-TOF mass spectrometry, *J. Virol. Methods* 132, 1–12.
26. Zhang, W., and Chait, B. T. (2000) ProFound: an expert system for protein identification using mass spectrometric peptide mapping information, *Anal. Chem.* 72, 2482–2489.
27. Collins, B. K., Tomanicek, S. J., Lyamicheva, N., Kaiser, M. W., and Meuser, T. C. (2004) A preliminary solubility screen used to improve crystal trials: crystallization and preliminary X-ray structure determination of *Aeropyrum pernix* flap endonuclease-1, *Acta Crystallogr., Sect. D* 60, 1674–1678.
28. Bradford, M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* 72, 248–254.
29. 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.
30. Stalmach, M. E., Grothe, S., and Wood, J. M. (1983) Two proline porters in *Escherichia coli* K-12, *J. Bacteriol.* 156, 481–486.
31. Mayer, B. J. (2001) SH3 domains: complexity in moderation, *J. Cell Sci.* 114, 1253–1263.
32. 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.
33. 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.

BI6023786