Analytical Biochemistry 392 (2009) 70–76

Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

Obtaining highly purified intrinsically disordered protein by boiling lysis and single step ion exchange

Alexandra M. Livernois [1], Daniel J. Hnatchuk [1], Emma E. Findlater, Steffen P. Graether *

Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ont., Canada N1G 2W1

ARTICLE INFO

Article history:
Received 21 March 2009
Available online 21 May 2009

Keywords:
Intrinsically disordered protein
Purification
Boiling
NMR
Ion exchange
HPLC
Dehydrin

ABSTRACT

Intrinsically disordered proteins (IDPs) is a term used to describe proteins that do not have a well-defined tertiary structure. IDPs have many roles such as in cell cycle control (p53), neuronal signal transmission (myelin basic protein), and protein stability (dehydrins). Producing recombinant IDPs in bacteria for nuclear magnetic resonance (NMR) studies is problematic because the lack of stable tertiary structure makes them excellent substrates for bacterial proteases, which will cause loss in yield. We have developed a two-step method to produce the grape dehydrin K_2 and YSK_2 using *Escherichia coli*. Dehydrins are expressed by certain plants in response to dehydration, increased salinity, or low temperatures. Purification of 10 mg/L (K_2) and 15 mg/L (YSK_2) was performed by boiling bacterial pellets to lyse the cells, remove most of the contaminating proteins, and denature bacterial proteases. This resulted in protein purity comparable to that produced by sonication and nickel affinity chromatography. Boiling was followed by cation exchange chromatography to remove the remaining trace contaminants. The sample was shown to be more than 95% pure by reversed-phase high-performance liquid chromatography. The method presented here can easily be adapted to the purification of other IDPs and heat-stable proteins without requiring multiple chromatography steps or the use of protease inhibitors.

Structure determination techniques, such as X-ray crystallography and nuclear magnetic resonance (NMR)[2], have long focused on the analysis of proteins with well-defined folds. During the past decade, many proteins without a well-defined structure have been identified and studied. These intrinsically disordered proteins (IDPs, also known as intrinsically unstructured proteins) have major biological roles in cell signaling [1], transcription [2], and fatty acid synthesis [3]. IDPs are also important in disease pathogenesis such as Parkinson's [4,5]. The lack of structure has been attributed to the types of amino acids in an IDP protein sequence [6]; these proteins tend to be rich in polar and charged amino acids, whereas hydrophobic residues are rare. Because of this, IDPs do not form globular structures with hydrophobic residues in the core [7].

NMR is an excellent technique with which to study folded, unfolded, and disordered protein structures [8–10]. When applied to studying IDPs, NMR can determine whether there is residual structure and quantitate the amount of flexibility (dynamics) present. One challenge of using NMR is the requirement for relatively large amounts of highly purified protein (milligrams of >90% purity). With IDPs, this is complicated by the need for isotopically labeled protein. Currently, the most successful method to generate labeled protein is via heterologous expression in a bacterial host, especially because the cost of using eukaryotic expression systems is still quite high. One of the main problems is proteolytic cleavage. Although in vivo IDPs are not highly susceptible to degradation in their native hosts [11], they are excellent targets for bacterial proteases. Such complex protein mixtures during IDP purification can require multiple ion exchange chromatography and gel filtration steps [12,13] to obtain pure protein. In another case, an IDP was purified by metal affinity, hydrophobic interaction, and blue Sepharose chromatography [14]. Therefore, we set out to increase expression yields by reducing the amount of protein degradation that occurs during purification and reducing the number of chromatography steps involved. The IDPs we have chosen to test our purification are plant dehydrins, specifically the K_2 and YSK_2 proteins from *Vitis riparia* [15].

Plants exhibit a number of responses to dehydration stress, including the expression of dehydrins (for reviews, see Refs. [16–18]). One of the major roles of dehydrins is to act as a stress response protein that protects plants from damage due to drought,

* Corresponding author. Fax: +1 519 837 1802.
E-mail address: graether@uoguelph.ca (S.P. Graether).

[1] These authors contributed equally to this work.

[2] *Abbreviations used:* NMR, nuclear magnetic resonance, IDP, intrinsically disordered protein; AFP, antifreeze protein; HPLC, high-performance liquid chromatography; PCR, polymerase chain reaction; LB, Luria–Bertani; DMSO, dimethyl sulfoxide; IPTG, isopropyl β-d-thiogalactopyranoside; EDTA, ethylenediaminetetraacetic acid; PVDF, polyvinylidene difluoride; FPLC, fast protein liquid chromatography; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; RI, recrystallization inhibition.

0003-2697/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2009.05.023

high salinity, or extracellular freezing [16]. Dehydrin proteins can vary from 9 to 200 kDa in size [19]; this large range arises from the modular assembly of several sequence segments. The resulting dehydrin subfamilies are defined by the number of three different conserved segments and the presence of a fourth segment that is rich in polar residues. The K-segment, present in all dehydrins, is a lysine-rich 15-amino-acid sequence (EKKGIMDKIKEKLPG), although its conservation is not absolute [16]. This segment has been proposed to form an amphipathic α-helix that binds to proteins and membranes during periods of stress [19,20]. The number of K-segments present varies between 1 and 11 among dehydrin proteins. The S-segment is a serine-rich tract that can be phosphorylated [21]. It is located before or after the K-segment, and usually only 1 S-segment (or none) is found in dehydrins. The Y-segment has homology to the nucleotide binding site of plant and bacterial chaperones, although a clear functional role has not been described for this region. This segment contains the sequence repeat (V/T)DEYGNP. When present, 1–3 copies of this segment are located at the N terminus. Lastly, there are repeat domains known as ϕ-segments that are found in between the sequence-conserved segments. ϕ-Segments are often rich in Gly, Pro, or Ala and polar residues (especially Thr). Overall, the dehydrin sequence is highly hydrophilic and does not contain any Cys or Trp residues [17].

Dehydrins are thought to protect plants from dehydration effects by preventing or reversing protein denaturation and membrane disruption [19]; they may also lessen the dehydrative stress by binding large quantities of water [22]. Dehydrin proteins may also have various cryoprotective roles; several studies have shown that they are able to protect enzymes from activity loss due to cold denaturation [23]. In addition, two dehydrins have also been shown to have antifreeze protein (AFP) activity [24,25], suggesting that members of the dehydrin family may have the ability to bind to ice. AFPs are expressed in cold-tolerant organisms to prevent them from freezing at sub-zero temperatures. Although all AFPs appear to bind to ice, their structures are not well conserved and many different folds have been identified [26].

We set out to develop a system suitable for purifying milligram quantities of recombinant dehydrins in *Escherichia coli* for structure–function characterization by NMR and other biophysical methods. We constructed a pET22b plasmid containing the K_2 and YSK_2 gene sequence and successfully expressed the protein in *E. coli* in rich and minimal media. We exploited the heat stability property of IDPs such as dehydrins by using boiling to lyse the cells and remove many of the contaminating proteins, followed by one step of ion exchange chromatography. This allowed us to obtain highly purified proteins free of proteolytic cleavage in only two steps. This method can be readily applied to other IDPs and other heat-stable proteins that have been expressed in a recombinant bacterial system.

Materials and methods

Chemicals and reagents

All chemicals were purchased from Acros (Morris Plains, NJ, USA) or ThermoFisher Scientific (Montreal, Canada) unless otherwise indicated and were of reagent grade or higher. Cation exchange chromatography was performed on a GE Healthcare HiTrap SP FF column (Baie d'Urfe, Canada), and reversed-phase high-performance liquid chromatography (HPLC) was performed on a BioBasic C18 column (ThermoFisher Scientific). The grape dehydrin gene YSK_2 was a generous gift from Annette Nassuth (University of Guelph), and the K-peptide antibodies were a generous gift from Timothy J. Close (University of California, Riverside).

Cloning

The YSK_2 gene (GenBank AY706987) in the pGEM-T plasmid was amplified by polymerase chain reaction (PCR) using PfuTurbo (Stratagene, La Jolla, CA, USA) following the manufacturer's recommended protocol. The YSK_2 primers (5′-GCACGTCATATGGCATAT CAGCAAGATCCATG-3′ and 5′-CCTGTTCTCGAGGTGGGCCCCAGGCA GC-3′) and the K_2 primers (5′-CAGTACCTCGAGGTGGGCC CCAGGCA GCTTCTCCTTGA TCTTC-3′ and 5′-CAGTACCTCGAGGTGGGC CCCAG GCAGCTTCTCCTTGA TCTTC-3′) were used to amplify the gene of interest. The underlining represents the location of the restriction enzyme cut sites. Sense primers introduced an NdeI (CATATG) site at the start codon, whereas antisense primers introduced an XhoI (CTCGAG) site 5′ to the stop codon. The PCR products were digested with NdeI and XhoI restriction enzymes and ligated into the pET22b expression vector (Novagen, Gibbstown, NJ, USA) that had been digested with NdeI and XhoI. Note that the NdeI digestion of pET22b removes the pelB leader sequence, whereas digestion with XhoI retains the 3′ His6 tag encoded in the plasmid. Ligation products were transformed into *E. coli* NovaBlue (Novagen) for plasmid storage and propagation. Screening for positive clones containing the K_2 or YSK_2 gene was performed by sequencing transformed colonies from ampicillin Luria–Bertani (LB) agar plates. The resulting constructs were named pET22b–K_2–His6 and pET22b–YSK_2–His6.

To prevent the His6 tag sequence in the gene from being expressed, stop codons were added to the ends of the K_2 and YSK_2 gene sequences. The XhoI restriction site in both constructs was changed from CTCGAG to CTCCTA using the following pair of primers: 5′-GTGGTGGTGGTGCTCCTAGTGGGCCCCAGGCAGCTTCTCC-3′ and 5′-GCCTGGGGCCCACTAGGAGCACCACCACCACCACCACTGAGAT CC-3′. These mutations were made using the QuikChange kit (Stratagene) by following the manufacturer's instructions with the exception that 1% dimethyl sulfoxide (DMSO) was added to the mutagenesis reaction due to the GC-rich content of the primers. After transformation, screening for positive clones was performed by sequencing. Constructs without the His6 tags were named pET22b–K_2 and pET22b–YSK_2.

Protein expression: rich medium

The pET22b–K_2 and pET22b–YSK_2 plasmids were transformed into *E. coli* BL21(DE3) competent cells for expression in LB medium supplemented with 100 μg/ml ampicillin. A colony from a freshly transformed plate was used to start the 5-ml LB inoculum cultures. These were grown overnight at 37 °C. Cell cultures from the overnight growths were diluted 1:1000 into LB and grown at 37 °C with shaking at 250 rpm in an incubator shaker. Once an OD_{600} of 0.7–0.8 was reached (after 3–4 h), protein expression was induced with 0.4 mM isopropyl β-D-thiogalactopyranoside (IPTG) and the cells were grown for 3 h at 37 °C before being harvested.

Protein expression: minimal medium

To maximize M9 expression, the cells were first grown in rich LB medium before being transferred to our modified M9 minimal medium. Cells from the overnight LB growths were diluted 1:1000 into 1 L of LB medium and grown at 37 °C with shaking at 250 rpm in an incubator shaker. Once an OD_{600} of 0.7–0.8 was reached, the cells were centrifuged at 5000*g* for 30 min. Residual LB medium was removed by washing the pellets with 100 ml of modified M9 medium (for 1 L of M9 medium: 6.78 g Na_2HPO_4, 3.0 g KH_2PO_4, 0.5 g NaCl, 1.0 g NH_4Cl, 0.99 g $MgSO_4 \cdot 7H_2O$, 3.0 g d-glucose, 50 mg thiamine hydrochloride, trace elements (6.0 mg $CaCl_2 \cdot 2H_2O$, 60 mg $Fe_3SO_4 \cdot 7H_2O$, 11.5 mg $MnCl_2 \cdot 4H_2O$, 8.0 mg $CoCl_2 \cdot 6H_2O$, 7.0 mg $ZnSO_4 \cdot 7H_2O$, 3.0 mg $CuCl_2 \cdot 2H_2O$, 0.2 mg