Analytical Biochemistry 514 (2016) 8–11

Contents lists available at ScienceDirect

Analytical Biochemistry

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

Technical note

Separating full-length protein from aggregating proteolytic products using filter flow-through purification

Kelly A. Churion, Robert E. Rogers, Kayla J. Bayless, Sarah E. Bondos*

Department of Molecular and Cellular Medicine, Texas A&M Health Science Center, College Station, TX 77843-1114, USA

ARTICLE INFO

Article history:
Received 11 July 2016
Received in revised form
4 September 2016
Accepted 9 September 2016
Available online 10 September 2016

Keywords:
Protein purification
Aggregation
Proteolysis
Intrinsically disordered protein
Precipitation
Filtration

ABSTRACT

Separation of full-length protein from proteolytic products is challenging, since the properties used to isolate the protein can also be present in proteolytic products. Many separation techniques risk non-specific protein adhesion and/or require a lot of time, enabling continued proteolysis and aggregation after lysis. We demonstrate that proteolytic products aggregate for two different proteins. As a result, full-length protein can be rapidly separated from these fragments by filter flow-through purification, resulting in a substantial protein purity enhancement. This rapid approach is likely to be useful for intrinsically disordered proteins, whose repetitive sequence composition and flexible nature can facilitate aggregation.

1. Introduction

Proteolysis commonly occurs during the production and purification of recombinant proteins. Degradation or modification of proteins by proteases during expression and/or purification reduces yield and introduces contaminants similar to the full length protein. For example, proteolytic products that contain a tag or epitope will bind resin as well as the full length protein during affinity purification. For a protease to cut, the protein target site must be embedded in a protein region that is at least 10 amino acids long and unfolded [1]. Some proteins are more susceptible to proteolysis than others, including unstable proteins, proteins with long exposed loops, mutant proteins, molten globules, and intrinsically disordered proteins. A large proportion of gene sequences code for long stretches of amino acids that are likely to be unfolded in solution [2–4]. The occurrence of unstructured regions greater than 40 residues in length is especially common in proteins that regulate critical cell responses [5–10]. Since these regions often play regulatory roles, *in vitro* characterization of the impact of these intrinsically disordered sequences is essential to understand protein function [11–13]. However, intrinsically disordered regions are more readily cut by proteases, and the presence of proteolytic contaminants can alter the apparent function of the purified protein. Because the presence of these contaminating proteolytic products can confound interpretation of downstream characterization assays, removal of these proteolytic products in a rapid and facile manner is crucial.

General approaches to the proteolysis problem include prevention of proteolysis before or during purification and removal of proteolytic products after the initial purification steps. Slowing protein expression can often decrease aggregation and proteolysis. Lowering the incubation temperature after induction, shortening expression time, and/or changing to a less rich media can sometimes minimize or prevent proteolysis before protein purification [14].

A frequently used strategy to prevent proteolysis during purification is to add a variety of protease inhibitors to the extraction buffer. However, this technique is not completely effective for proteins extremely susceptible to proteolysis, such as intrinsically disordered proteins. Finally, if the proteolytic products bind the full-length protein, then they will co-purify with the full-length protein even if they lack affinity tags. An effective method to remove contamination by proteolytic products is to purify under denaturing conditions, yet this harsh method could render the protein of interest non-functional if it is not able to refold properly *in vitro*.

* Corresponding author.
E-mail address: sebondos@tamhsc.edu (S.E. Bondos).

http://dx.doi.org/10.1016/j.ab.2016.09.009
0003-2697/

Existing methods to remove proteolytic products can be slow, providing time for the protein of interest to further degrade and possibly aggregate. In dialysis methods, the full length protein is retained by the membrane, whereas proteolytic products smaller than the pores pass through. This process requires hours to days and thus is not ideal for a protein that continues to degrade or aggregates rapidly [15,16]. Gel filtration chromatography, which separates proteins by size, is also slow, generally taking hours to perform. Affinity chromatography using a tag on the protein or a specific antibody on the resin will capture full-length protein as well as any proteolytic products also containing the tag or epitope. Alternatively, iterative affinity purification using different tags on each of the protein termini can ensure that only full-length protein, which contains both tags, is collected. Although this is an effective method, these fusion tags could potentially alter protein function and/or activity [17,18].

A fast, reliable method is needed to retain protein activity as well as to safely and efficiently remove proteolytic products. In this paper, we demonstrate that two intrinsically disordered proteins generate aggregating proteolytic products that can be removed in minutes by filter flow-through purification. In this rapid technique, full-length protein passes through the filter and aggregated proteolytic products create particles larger than the pores and are thus retained by the filter. Our approach has been demonstrated with two distinct proteins: the putative human lipid-binding protein Sec14-like 1 (Sec14L1), and the *Drosophila* Hox transcription factor, Ultrabithorax (Ubx). Both of these proteins contain intrinsically disordered regions.

2. Results

The first protein tested was the 80 kDa human protein, Sec14L1, in which approximately 32% of the sequence is intrinsically disordered (Fig. 1A). Sec14L1 was expressed in *E. coli* for 48 h at 4 °C after induction with 300 μM IPTG. After centrifugation, cells were lysed using a French press at 1250 psig. After additional centrifugation to remove insoluble material, the supernatant was applied to a Ni^{2+}-binding agarose resin column (Novagen) pre-equilibrated with wash buffer (200 mM NaCl, 50 mM NaH_2PO_4, pH 8.1) and washed for 50 column volumes. Sec14L1 was eluted in wash buffer plus 200 mM imidazole. Gel electrophoresis showed Sec14L1 is on average 57% pure, and western blots identified proteolytic product as the major contaminants (Fig. 2B). Detailed protocols for all methods reported in this paper can be found in Supplemental Material.

To examine if the filter flow-through purification assay would improve purity, we filtered Sec14L1 protein through an Amicon concentrator with 100 nm pores (EMD-Millipore). The supernatant was added to the filter column prior to centrifuging at 10,000×*g* for 10 min at 4 °C. Soluble full length Sec14L1 and a soluble variant (either a long proteolysis product or phosphorylated protein) passed through the filter. However, aggregated proteolytic products were retained by the filter. Quantification of each fraction by western blot demonstrated that soluble Sec14L1 reached approximately 98% purity with no visible contamination from aggregating proteolytic products (Fig. 2B).

To test whether this assay was applicable to other types of proteins or specific to Sec14L1, we examined the 40 kDa *Drosophila* Hox transcription factor, Ultrabithorax (Ubx). Ubx is 53% intrinsically disordered, and contains a predominately structured DNA-binding homeodomain (HD) that binds to the sequence 5′-TAAT-3′ (Fig. 1A) [11,12]. Most of the Ubx sequence is intrinsically disordered, and these regions are extremely susceptible to proteolysis [11]. Since the homeodomain binds the target DNA with higher affinity than the full length Ubx protein, proteolytic contaminants containing this structured domain can out-compete full-length Ubx

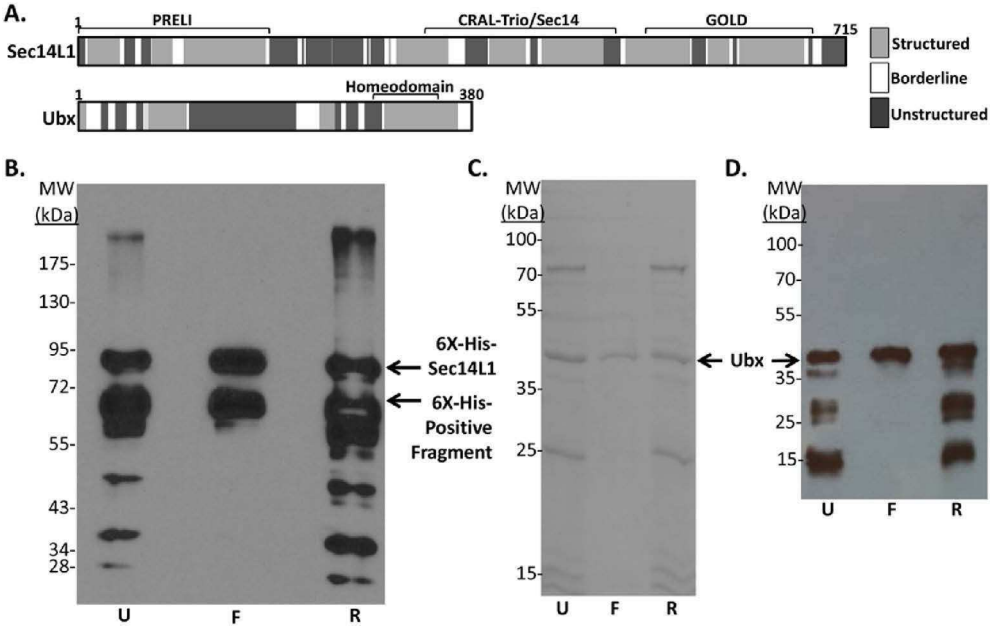

Fig. 1. Filter flow-through purification of the Sec14L1 and Ubx proteins. A) Protein schematics depicting the structured (medium grey, score ≤ 0.4) and disordered (dark grey, score ≥ 0.6) regions of Sec14L1 (top) and Ubx (bottom). White boxes indicate ambiguous regions, in which the score lies between 0.4 and 0.6. Ordered and disordered regions for Sec14L1 were determined from the disorder prediction algorithm PONDR, and scores for Ubx were the average of the IuPred, DisEMBL and PONDR programs as well as native state proteolysis experiments [11]. B) Western blot analysis with Mouse anti-6X His (GeneTex) of a filter flow-through purification of Sec14L1, comparing unfiltered protein (U), the filtrate (F), and retentate (R). Band quantification revealed that the dual Sec14L1 bands were 98% pure, and that 97% of the total Sec14L1 protein was recovered in the filtrate. C,D) SDS-PAGE analysis of filter flow-through purification of Ubx, visualized by coomasie stain (C) and western blot using the anti-Ubx primary antibody FP3.38 (D). Based on band quantification, Ubx purified by filtrate was >90% pure, and 96% of the total full-length Ubx protein was recovered in the filtrate. In both cases, the filter captures aggregated protein fragments, and filtration substantially increases the purity of the protein.

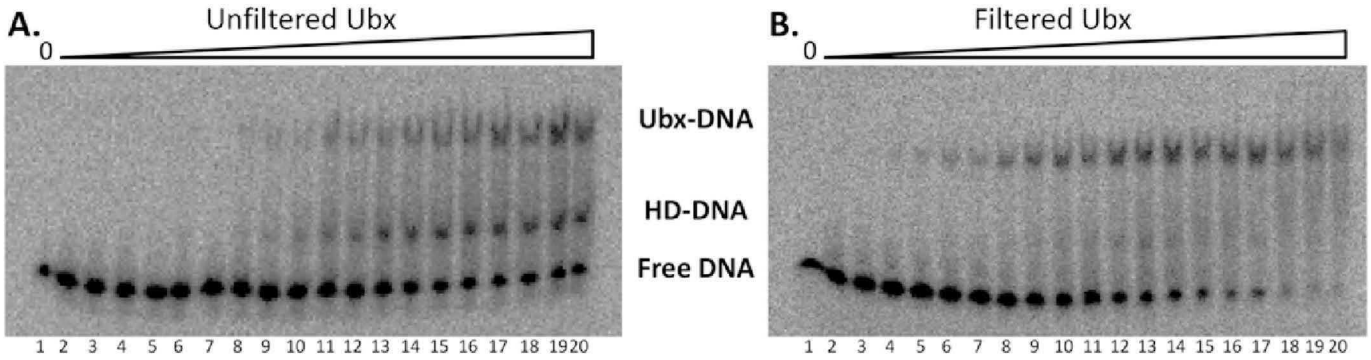

Fig. 2. Comparison of DNA binding by unfiltered and filtered UbxIa. Both panels show equilibrium DNA gel retardation (gel shift) experiments. All lanes contain radiolabeled 2 pM 40AB DNA [11], and lanes 2–20 also include Ubx, increasing from 2.5 pM to 10 nM. A) Unfiltered full-length UbxIa binding the 40AB DNA. B) Filtered Ubx binding 40AB. Filtration largely removes binding by contaminating proteolytic products, most notably the Ubx homeodomain (HD).

in DNA binding assays, preventing accurate affinity measurements [11].

Ultrabithorax splicing isoform Ia (Ubx) was expressed in *E. coli* for at 25 °C for 150 min after 1 mM IPTG induction at mid-log phase as previously described [11,12]. Cell pellets corresponding to 2 L of fermentation were frozen at −20 °C. To purify, a cell pellet was lysed in the presence of DNase. Cell debris was removed by centrifugation at 25,000×*g* for 30 min. Nucleic acids, precipitated with polyethyleneimine, were subsequently removed by centrifugation. After adjusting the supernatant pH to 6.8 and re-centrifuging, the supernatant was loaded on a Ni^{2+} NTA column (Thermo Scientific) equilibrated 10 column volumes with wash buffer containing 500 mM NaCl, 10 mM dithiothreitol, 5% glucose, 20 mM imidazole, and 50 mM NaH_2PO_4, pH 8.0. The column was iteratively washed in wash buffer containing 0 mM, 10 mM, 20 mM, 40 mM and 80 mM imidazole and Ubx was eluted in wash buffer containing 200 mM imidazole. All steps in the protein purification were performed at 4 °C.

Gel electrophoresis and Western blot analysis showed Ubx to be less than 34% pure, with the major contaminants being proteolytic products of Ubx (Fig. 1C,D). Because Ubx is prone to aggregation [19] and will adsorb non-specifically to many chromatography resins, removing aggregates by further chromatography was not feasible. To determine whether filter flow-through purification can remove these proteolytic products, we filtered 500 μL of Ubx protein through an Amicon concentrator with a 100 kDa cutoff filter (EMD-Millipore). The filter was centrifuged at 10,000×*g* for 10 min at 4 °C samples were collected for subsequent analysis. The SDS-PAGE results were quantified to determine purity (Fig. 1C,D). The full-length protein increased from <35% to >90% pure. The filter purified protein had little detectable contamination when visualized by coomasie stain or the more sensitive western blot analysis (Fig. 1C,D). However, any trace proteolytic products can be detrimental in DNA-binding assays. We used DNA binding electrophoretic mobility shift assays, performed as previously described [8,9] as a more sensitive method to detect Ubx proteolytic products. When assessing filter-purified Ubx, only major 2 bands were detected in the gel, corresponding to free DNA and DNA bound to full length protein Ubx (Fig. 2B). In contrast, the gel with unfiltered Ubx displayed additional bands, located in between these extremes, which correspond to the proteolytic products of Ubx bound to DNA (Fig. 2A). The dominant band in this region corresponds to the Ubx HD binding to DNA. Although the HD has a higher affinity for DNA than Ubx, it is present at lower concentrations, and thus the HD-DNA band first appears in the same sample as the Ubx-DNA band. We conclude that filter purification yields full length Ubx protein with no detectable proteolytic products.

3. Discussion

Purification is required to quantitatively assess the function of specific proteins. However, many proteins, especially those harboring intrinsically disordered regions, are prone to proteolysis during expression, purification, and storage. In this paper we demonstrate that filter flow-through purification can be a fast, efficient and inexpensive method for separating proteolytic product contaminants from full length purified protein. For filter purification to work, the proteolytic products must aggregate, causing them to be retained in a filter with pores sufficiently large for the full length protein to pass to the filtrate. We tested this approach on two proteins with significant intrinsically disordered regions, and found both proteins generated aggregating proteolytic products. Although Sec14L1 and Ubx have distinct properties and features, filter flow-through purification successfully removed proteolytic product contaminants from both preparations. Importantly, different procedures were used to express and initially purify these proteins and thus the success of filter purification is not dependent on the protocols used to produce either protein. This fast and efficient method can purify full-length proteins to near homogeneity, without compromising protein activity.

Acknowledgements

The authors would like to thank the members of the Bayless and Bondos labs for helpful discussions and comments on the manuscript. Funding was provided by NSF CAREER 1151394, the Ted Nash Long Life Foundation M1500779, and TAMHSC RDEAP to S.E.B. as well as the PHS grant HL095786 to K.J.B.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ab.2016.09.009.

References

[1] S. Hubbard, R.J. Benyon, in: R. Benyon, J.S. Bond (Eds.), In Proteolytic Enzymes, second ed., Oxford University Press, Oxford, UK, 2001, pp. 248–249.
[2] H.J. Dyson, P.E. Wright, Intrinsically unstructured proteins and their functions, Nat. Rev. Mol. Cell Biol. 6 (2005) 197–208.
[3] P.E. Wright, H. Dyson, Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm, J. Mol. Biol. 293 (1999) 321–331.
[4] P. Romero, Z. Obradovic, C.R. Kissinger, J.E. Villafranca, E. Garner, S. Guilliot, A.K. Dunker, Thousands of proteins likely to have long disordered regions, Pac Symp. Biocomput. 3 (1998) 437–448, 1998.
[5] A.K. Dunker, C.J. Brown, J.D. Lawson, L.M. Iakoucheva, Z. Obradović, Intrinsic disorder and protein function, Biochemistry 41 (2002) 6573–6582.

[6] V.N. Uversky, Natively unfolded proteins: a point where biology waits for physics, Protein Sci. 11 (2002) 739–756.
[7] A.K. Dunker, S.E. Bondos, F. Huang, C.J. Oldfield, Intrinsically disordered proteins and multicellular organisms, Semin. Cell Dev. Biol. 37 (2015, Jan) 44–55.
[8] P.E. Wright, H.J. Dyson, Intrinsically disordered proteins in cellular signalling and regulation, Nat. Rev. Mol. Cell Biol. 16 (2014) 18–29.
[9] A.K. Dunker, E. Garner, S. Guilliot, P. Romero, K. Albrecht, J. Hart, Z. Obradovic, C. Kissinger, J.E. Villafranca, Protein disorder and the evolution of molecular recognition: theory, predictions and observations, Pac Symp. Biocomput. 3 (1998) 473–484.
[10] E. Garner, P. Cannon, P. Romero, Z. Obradovic, A.K. Dunker, Predicting disordered regions from amino acid sequence: common themes despite differing structural characterization, in: Genome Inform Ser Workshop Genome Inform vol. 9, 1998, pp. 201–213.
[11] Y. Liu, K.S. Matthews, S.E. Bondos, Multiple intrinsically disordered sequences alter DNA binding by the homeodomain of the *Drosophila* Hox Protein Ultrabithorax, J. Biol. Chem. 283 (2008) 20874–20887.
[12] Y. Liu, K.S. Matthews, S.E. Bondos, Internal regulatory interactions determine DNA binding specificity by a Hox transcription factor, J. Biol. Chem. 390 (2009) 760–774.
[13] K.J. Niklas, S.E. Bondos, A.K. Dunker, S.A. Newman, Rethinking gene regulatory networks in light of alternative splicing, intrinsically disordered protein domains, and post-translational modifications, Front. Cell Dev. Biol. 3 (8) (2015).
[14] M. Lebendiker, T. Danieli, Production of prone-to-aggregate proteins, FEBS Lett. 588 (2014) 236–246.
[15] S. Harper, D.W. Speicher, Purification of proteins fused to glutathione S-transferase, Methods Mol. Biol. (2010) 259–280.
[16] A. Beloff, C.B. Anfinsen, The products of proteolysis of some purified proteins, J. Biol. Chem. (1948) 863–872.
[17] X. Zhao, G. Li, S. Liang, Several affinity tags commonly used in chromatographic purification, J. Anal. Methods Chem. (2013) 1–8.
[18] S.P. Tsai, D.W. Howell, Z. Huang, H.C. Hsiao, Y. Lu, K.S. Matthews, J. Lou, S.E. Bondos, The effect of protein fusions on the production and mechanical properties of protein-based materials, Adv. Funct. Mater 25 (2015) 1442–1450.
[19] S.E. Bondos, A. Bicknell, Detection and prevention of protein aggregation before, during, and after purification, Anal. Biochem. 316 (2003) 223–231.

Supplementary Material

Ubx Protein Expression

His-tagged was expressed in BL21 (DE3) pLysS and purified as described by Lui *et al.2008* **[11]** with the following modifications. Cells expressing UbxIa were harvested 150 min after induction. Cells were resuspended in 2 mL of collection buffer (100 mM NaCl, 20 mM NaH_2PO_4, 80 mM Na_2HPO_4, pH=7.5 one tablet of Complete Proteinase Inhibitor Mixture EDTA-free (Roche Applied Science) and frozen at -20 °C.

Ubx Protein Purification

Expressed from the pET19b plasmid to provide a His10 tag at the UbxIa N terminus. A cell pellet corresponding to 2 L of fermented *E.coli* (final OD_{600nm} ~1.1) was thawed at room temperature and lysed in 20 mL of buffer containing 800 mM NaCl, 10 mM DTT, two tablets of Complete Proteinase Inhibitor Mixture EDTA free (Roche Applied Science), 5% glucose, 5 mM imidazole, and 50 mM NaH_2PO_4, pH 8.0, 1 mM PMSF in ethanol. Released DNA in the cell lysate was digested with DNase I (20 mg/mL, 40 µl) prior to centrifugation at 17,000 rpm for 30 min. The supernatant was mixed with polyethyleneimine (200 µl of 50% (w/v)) and centrifuged at 17,000 rpm for 20 mins. The supernatant pH was then adjusted to column pH=8.0 and recentrifuged at 17,000 rpm for 20 mins. To prepare the affinity column, 4 mL of Ni^{+2} nitrilotriacetic acid-agarose resin (Thermo Scientific) was pre-equilibrated with wash buffer containing 500 mM NaCl, 10 mM-DTT, 5% glucose, 20 mM imidazole, and 50 mM NaH2PO4, pH 8.0. The clear lysate was incubated with gentle agitation at 4 °C for 20 min before being poured into the column. Packed resin was washed with 50 mL of wash buffer containing 0 mM imidazole followed by 100mL wash of buffer containing 10 mM imidazole, 50 mL wash buffer containing 20 mM imidazole, 50 mL wash containing 40 mM imidazole and finally 25 mL wash buffer containing 80 mM imidazole. Protein was eluted with 14 mL wash buffer containing 200 mM imidazole.

Ubx Aggregate Removal:

Ubx protein (500 µL) was filtered through an Amicon concentrator with a 100 kDa cutoff filter (EMD-Millipore). The filter was centrifuged at 10,000 x g for 10 minutes at 4 ^{o}C. Samples of filtrate and retentate were collected for subsequent analysis.

Sec14L1 Protein Expression:

Sec14L1 was amplified from Human Umbilical Vein Endothelial Cell cDNA using Accuprime Taq HIFI according to manufacturer's protocol (Invitrogen; Forward Primer: AGCGGATATCGATGGTGCAGAAATACCAG; Reverse Primer: AGCGCTCGAGAGAACCAGAACCCCTGGAGATCATGGAGCTGGAGTG). The amplicon was digested, purified, and ligated into the pQE-TriSystem vector using EcoRV and XhoI restriction enzymes. The resulting ligated product was transformed via heat shock into BL21(DE3)pLysS cells (Promega), and positive clones were selected by colony PCR, restriction digest, and sequencing.

A single colony of transformed BL21(DE3)pLysS cells were grown overnight at 37 ^{o}C in a 5 mL LB starter culture with 25 ug/mL carbenicillin and 25 ug/mL chloramphenicol. The starter culture was diluted 1:100 into 200 mL LB with 25 ug/mL carbenicillin and 25 ug/mL chloramphenicol and grown at 37 ^{o}C until the

OD_{600} measurement of the bacterial growth was 0.6. The bacterial culture was induced with 300 µM IPTG for 48 hours at 4 ^{o}C.

Sec14L1 Protein Purification:
The 200 mL bacterial culture was centrifuged at 2,800 xg for 15 minutes at 4 ^{o}C to pellet the induced bacteria. The pellet was resuspended in 10 mL lysis buffer (500 mM NaCl, 20 mM Na_2HPO_4, pH 8.0 with 1 mM PMSF and Roche protease inhibitor cocktail) and French-pressed at 1,250 psig before centrifuging at 20,000 xg for 15 minutes at 4 ^{o}C to pellet insoluble debris. The resulting supernatant was incubated with 2 mL Nickel charged His-binding resin (Novagen) pre-equilibrated in 10 column volumes lysis buffer at 4 ^{o}C for 2 hours. The resin was washed with 50 column volumes wash buffer (200 mM NaCl, 20 mM Na_2HPO_4, pH 8.0), and loosely-bound contaminants were removed with 5 column volumes of wash buffer with 20 mM Imidazole, then 5 column volumes of wash buffer with 40 mM Imidazole, then 5 column volumes of wash buffer with 80 mM Imidazole. Protein was then eluted with 4 column volumes of wash buffer with 200 mM Imidazole. Fractions containing protein were determined by coomassie and pooled.

Sec14L1 Aggregate Removal:
To remove aggregated proteins from the final protein product, 500 µL of purified protein solution was loaded into a 100 nm pore-sized Durapore PVDF spin column (Millipore) and centrifuged at 10,000 xg for 10 minutes to remove aggregated material.

Image Analysis
Band density for SDS PAGE gels and western blot were quantitated using Image Studio Lite version 5.2 software for both proteins.

Electrophoretic mobility shift analysis (gel shift) assays. Gel shift assays were performed as previously described **[11]**.