Purification and analysis of colorful hypothetical open reading frames: An inexpensive gateway laboratory

ARTICLE *in* BIOCHEMISTRY AND MOLECULAR BIOLOGY EDUCATION · MARCH 2011

Impact Factor: 0.65 · DOI: 10.1002/bmb.20459 · Source: PubMed

2 AUTHORS, INCLUDING:

Jeffrey L Reinking

State University of New York at New Paltz

14 PUBLICATIONS **579** CITATIONS

BIOCHEMISTRY AND MOLECULAR BIOLOGY EDUCATION
Vol. 39, No. 2, pp. 141–144, 2011

Laboratory Exercise

Purification and Analysis of Colorful Hypothetical Open Reading Frames

AN INEXPENSIVE GATEWAY LABORATORY

Received for publication, June 16, 2010, and in revised form, August 6, 2010

Kara A. DeSantis and Jeffrey L. Reinking‡

From the Department of Biology, State University of New York at New Paltz, New Paltz, New York 12561

This laboratory exercise is an inquiry-based investigation developed around the core experiment where students, working alone or in groups, each purify and analyze their own prescreened colored proteins using immobilized metal affinity chromatography (IMAC). Here, we present reagents and protocols that allow 12 different proteins to be purified in parallel without specialized equipment and within a 2.5- to 3-hour undergraduate teaching laboratory. The visual feedback of purifying a colored biomolecule provides real-time emphasis of the power and simplicity of recombinant DNA technology and IMAC. As presented here in its simplest form, this laboratory occupies two laboratory periods: purification followed by SDS-PAGE analysis. As such, it can be easily inserted into the existing curriculum of a Biochemistry, Molecular Biology, Biotechnology, or even Genetics course to illustrate core concepts of central dogma and protein purification. Furthermore, the proteins in hand at the end of this 2-week module can also be used for follow-up experiments tailored to the needs, timeframe, and facilities available.

Keywords: Undergraduate, protein purification, affinity chromatography, colored proteins, low cost.

Protein purification is a skill and concept that is taught in a variety of different undergraduate classes. When taught as part of an undergraduate teaching laboratory, protein purifications give students an appreciation of the milieu of different protein species within the cell as well as a taste of the technology and chemistry required to isolate and characterize a single protein species. When the purification is performed on a recombinantly expressed transgene, students also gain an appreciation of biotechnology—how similar techniques might be applied—for example, to purify a medically relevant biological such as recombinant insulin. However, educating students in laboratory implementations of protein science presents many conceptual hurdles. Proteins are submicroscopic and, therefore, inherently intangible. Our students often have trouble relating the clear, colorless liquids in laboratory plastic ware to the abstract concepts taught in the companion lectures. There are ways to allow students to eventually "see" their specific protein of interest—for example, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) or enzymatic activity assays, if the protein happens to be catalytic—but most involve delay, often to a substantial degree, and are indirect, generating conceptual separation. One of the goals of this laboratory is to eliminate that separation by providing students with real-time feedback of being able to directly see the whereabouts of their protein of interest throughout the purification.

The plasmids used in this laboratory were all produced by Structural Proteomics in Toronto (SPiT, Canada). The targets are hypothetical open reading frames (ORFs) from a variety of microbial sources that were originally selected as part of the Protein Structure Initiative [1]. Tens of thousands of clones have been produced by SPiT. We mined this amazing resource for the 12 expression constructs used in this laboratory (Table I) by choosing only highly expressed, soluble proteins that are visibly colored when purified.

The laboratory presented here is essentially the consensus protein purification protocol that has been adopted by the structural genomics community [5], optimized to be taught to a laboratory section of 24 students (working in pairs), taught once per week for a time period of 2 hours and 40 min (the typical laboratory class taught in the Biology Department at the State University of New York at New Paltz). Versions of this laboratory have been taught as part of multisection upper-division undergraduate laboratory courses Basic Molecular Biology and Genetics. Of particular note, this laboratory is designed to be cost-effective and use only equipment that can easily be found in most biochemistry and molecular biology facilities.

‡To whom correspondence should be addressed. Tel.: +1 845 257 3771; Fax: +1 845 257 3791. E-mail: reinkinj@newpaltz.edu.

 DOI 10.1002/bmb.20459

TABLE I
Protein information

Arbitrary designation	NCBI designation	Genetic source	M_w[a] (kDa)	Color	References
Control	EC2477	*Escherichia coli*	54.9	Clear	–
C4	PH0175	*Pyrococcus horikoshii*	21.3	Green	–
B1	NE2056	*Nitrosomonas europaea*	22.9	Brown/yellow	–
B2	NE1107	*Nitrosomonas europaea*	18.8	Red	–
B6	TM0134	*Thermotoga maritima*	33.8	Light yellow	–
B3	NE0605	*Nitrosomonas europaea*	10.9	Lilac	–
B4	TM0144	*Thermotoga maritima*	15.2	Pink	–
B10	Atu1337	*Agrobacterium tumefaciens*	15.9	Red/brown	[2]
C5	TVG1377730	*Thermoplasma volcanium*	21.9	Lavender	[3]
C1	SMc00371	*Sinorhizobium meliloti*	20.3	Magenta	–
A10	AF2256	*Archaeoglobus fulgidus*	29.4	Light Yellow	[4]
A6	PA3697	*Pseudomonas aeruginosa*	16.9	Red	–
B9	Atu0013	*Agrobacterium tumefaciens*	24.5	Yellow	–

[a] Theoretical M_ws based on the amino acids expected to be expressed including N-terminal residues derived from vector.

EXPERIMENTAL OUTLINE

Prior to Laboratory Period 1

The expression vectors were constructed using a pET15 derivative with a tobacco etch virus protease cleavable N-terminal His Tag (construction originally reported in ref. 6). The plasmids are transformed into the expression strain BL21 (DE3) (Invitrogen). We have the instructor grow and harvest 375 mL cultures for each purification; alternatively cultures could also be grown and harvested by students as part of the curriculum. Although we have produced good results by inducing rich media, such as premixed Terrific Broth (Fisher), with 0.25 mM isopropyl-β-D-thiogalactopyranoside (IPTG) followed by 15 °C overnight growth, using the auto-inducing media ZYP-5052 with 1X trace metals [7] greatly simplifies the procedure of growing 12 different cultures and significantly improves yield. After overnight induction, cells are pelleted and transferred to 50 mL conical tubes and stored at −80 °C.

Laboratory Period 1

During the first laboratory period, students lyse bacterial cells and purify a mystery target protein, collecting samples during the procedure for analysis by SDS-PAGE at a later time. At the beginning of the laboratory period, students are given a frozen cell pellet in a 50-mL conical centrifuge tube to purify. Chilled Binding Buffer (10 mM Imidazole at pH 8.0, 500 mM NaCl) supplemented by Lysozyme (0.25 mg/mL) and the protease inhibitor phenylmethylsulfonyl fluoride (PMSF) (2 mM) are added to the cells to a total volume of 30 mL. To lyse 12 different cultures in parallel, we employed three freeze-thaw cycles. Students use a combination of vortexing, pipetting, and manual agitation to thaw and resuspend their cell pellet. The resuspended solution is flash frozen in liquid nitrogen—a very popular moment for the students. The cells are thawed under running cold water, and the freeze-thaw cycle is repeated two more times. To conserve time, we often have the instructor perform the first freeze-thaw cycle in the hour preceding the start of class. The resultant lysate is relatively viscous compared with lysis techniques that break down bacterial DNA (e.g. sonication). We choose not to reduce this viscosity with the addition of a DNAse such as Benzonase (EMD, Gibbstown, NJ) due to added cost and relatively poor efficiency of the enzyme in salt concentrations >150 mM. Rather, we address this viscosity by assisting gravity flow of the chromatographic column using easily constructed syringe-based negative pressure "pumps" (Fig. 1).

The crude extract is clarified by centrifugation at 30,000× g for 20 min in a refrigerated, high-speed centrifuge. During centrifugation, students pour and equilibrate 1 mL of nickel affinity resin (Ni-NTA Superflow, Qiagen, Valencia, CA) in an open-air plastic column fitted with a two-way stopcock with 10 mL of binding buffer. On completion of centrifugation, the high speed supernatant is applied to the affinity column. Once the supernatant flows through the column, either by gravity flow or syringe "pump" assistance, a total of 200 mL of chilled wash buffer (30 mM imidazole at pH 8.0, 500 mM NaCl) is applied. During the purification, a colored band can be observed accumulating in the nickel resin (Fig. 2*a*) and

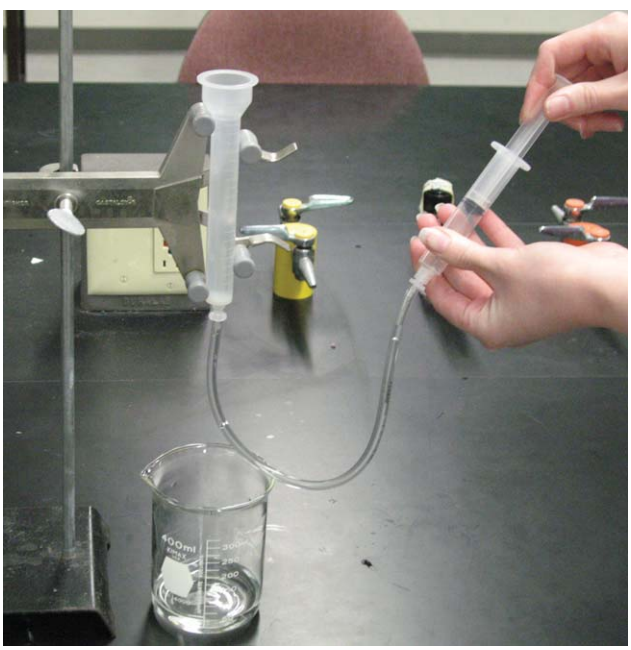

FIG. 1. **Chromatography apparatus in action**. The affinity resin is packed in an inexpensive plastic column (Bio-Rad 732-1010). To increase flow rate, a disposable Luer-Lok syringe is fitted using a short piece of 1/8″ ID tubing with two barbed male Luer adapters (Bio-Rad 731-8222).

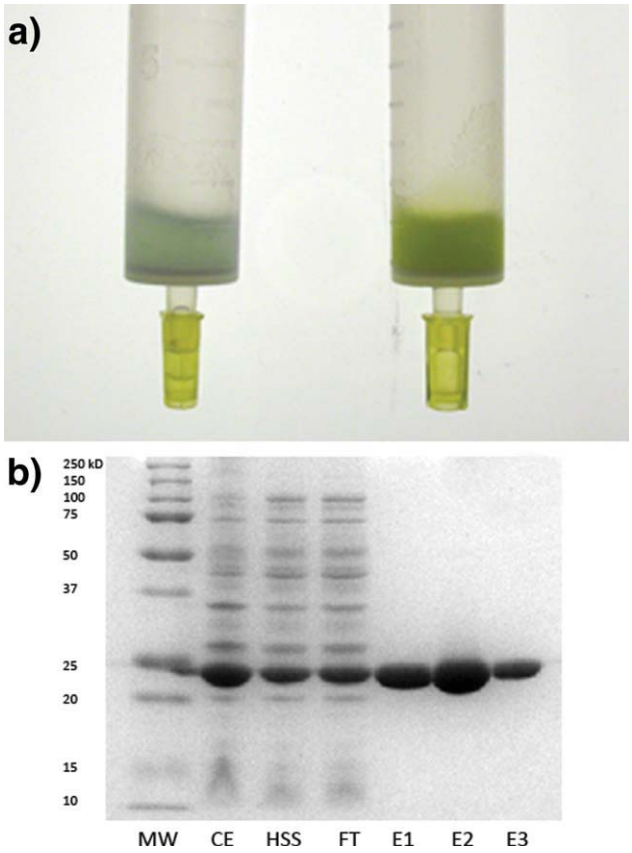

FIG. 2. **Sample purification of Atu0013**. (*a*) Protein bound to the column induces a visible color change on the affinity column. On the left are nickel beads washed only with buffer. On the right the beads are saturated with protein. (*b*) SDS-PAGE of representative purification. Lane labels correspond to crude extract (CE), high speed supernatant (HSS), flow-through (FT), and elutions (E1, E2, and E3). The M_w marker used was Precision Plus Protein Standard (Bio-Rad 161-0374). Note that the presence of target protein (24.5 kDa) in the FT fraction implies that the beads have been saturated, leading to extremely clean elution fractions.

eventually moving through the resin, when the protein is finally eluted with three separate 1 mL applications of chilled elution buffer (250 mM imidazole at pH 8.0, 500 mM NaCl). All of the samples reserved for gel analysis are mixed with 2X SDS Laemmli sample buffer, boiled for 10 min, and frozen for the next laboratory period.

Laboratory Period 2

Before the second laboratory period, we have the instructor pour a discontinuous 12% resolving/5% stacking denaturing acrylamide gel for each student group. Students load the gel and run at 100 V for ~1 hour in a tris-glycine buffer system. The gels are then stained using BioSafe Commassie (Bio-Rad), according to manufacturer's directions. A typical student purification profile is reported as Fig. 2*b*.

FIG. 3. **Protein gallery rainbow**. Small samples of eluted protein are shown in transparent tubes. Samples, left to right are in the same order as the samples in Table I.

RESULTS

At the end of the laboratory, students have produced a veritable rainbow of highly purified proteins within the class (Fig. 3). The proteins appear even more vividly colored, when concentrated within the nickel resin before elution. The degree of purification using this simple methodology is impressive (Fig. 4). Students are assigned to write a brief report that emphasizes the ability to interpret a gel and some rudimentary web-based bioinformatics. Given the gel image of their purification (such as Fig. 2*b*) and the hypothetical protein sequence of their target (including N-terminal vector-specific residues), the students are asked to calculate the theoretical molecular weight (M_w), identify and estimate actual M_w of the bands on the gel that correspond to their protein, identify the genetic source, and speculate as to why their protein might be colored. Students learn to use on-line tools such as M_w calculators and pBLAST to arrive at the relatively straightforward components of the analysis, but the final open-ended question requires higher order problem solving. Many of the proteins are only annotated as "hypothetical ORF" in Genbank, challenging students to use homology and pursue other creative avenues to hypothesize what the function of the protein might be and why it is colored.

SUGGESTED EXTENSIONS

Although we limit characterization of the protein simply to visual observation of color and SDS-PAGE to fit this

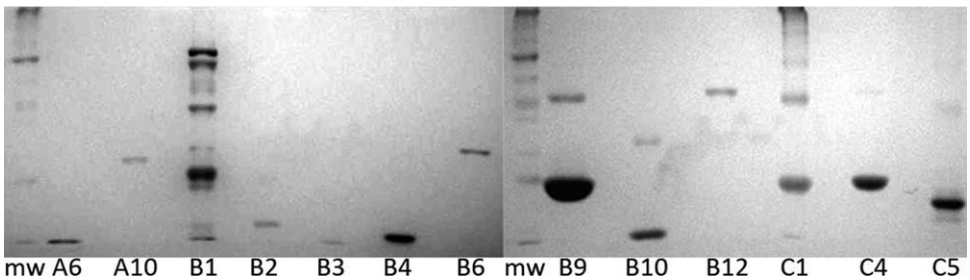

FIG. 4. **SDS-PAGE of eluted proteins**. Lane labels correspond to arbitrary designation from Table I. The M_w marker used was Prestained Protein Marker, Broad Range (New England Biolabs P7708S).

laboratory into two class periods, we can easily imagine extending characterization of the proteins produced. Techniques that could be applied to a generic protein could be employed. For example, the purified proteins can be quantitated using a Bradford assay [8]. Small-volume dialysis could be used to survey a variety of buffers to determine optimal solubility conditions of the protein [9]. Alternate forms of electrophoresis could be employed, such as native acrylamide or isoelectric focusing. The proteins could be run on additional columns such as gel filtration or ion exchange. In essence, any technique or specialized equipment available could be employed.

Students could also design and perform experiments to investigate the specific colored nature of the proteins. The UV-Visible spectrum could be measured; the peak absorbance is determined and used to calculate an extinction coefficient at that wavelength. The extinction coefficient calculated at various stages of the purification could be used to determine protein purity (as a substitute for specific activity) and compared with the gel results. Any transition metal ions contributing to the protein color could be detected using analytical instruments, if available. There are a myriad of potential follow-up experiments that could be performed with these intriguing proteins, extending the exercise well beyond the two periods described here.

STUDENT PITFALLS

During the purification, gravity flow alone is insufficient to allow completion of the laboratory within the prescribed time period, requiring the use of the syringe "pumps" depicted in Fig. 1. We find that, once given the pumps, some students attempt to pump through the liquid volumes as quickly as possible, leading to suboptimal purifications. We hand out the pumps only after the clarified lysate has run through the column, and ~30 min remains in the class. We also instruct the students not to use the pumps during the elution phase of the purification, to maximize yield and avoid cross-contamination with material in the syringe pump tubing. Use of the pump can lead to compaction of the affinity resin and very slow flow during the elution phase, which can be alleviated by resuspending the nickel beads with a pipette in the capped column before elution.

SUMMARY

Green fluorescent protein (GFP) is often purified in teaching laboratories, because it can be tracked visually with the aid of a near UV light source, addressing some of the concerns stated early about purification of colorless proteins. One advantage all of the proteins used herein have over GFP is that they are inherently colored rather than fluorescent and, therefore, do not require specialized light sources. The value of purifying colored proteins in the teaching laboratory has been explored previously using more complex chromatographic techniques on endogenous [10] and recombinant [11] protein mixtures. In contrast, this laboratory employs tagged proteins and a single-step affinity chromatography purification. Allowing the class to purify a large variety of proteins is also advantageous over a single or small set of proteins. Students are exposed to the dazzling variety of the protein world, develop a sense of ownership over their particular protein, and cannot rely on the analysis of their classmates when completing their assignment.

Many of the proteins in this laboratory are "unsuccessful" byproducts of a structural genomic pipeline in which they have not led to the solution of a three-dimensional protein structure or publication. We have found a new life for these abandoned clones as part of this laboratory exercise, creating a very successful, robust, affordable, and engaging laboratory to perform with our students. Equipment required is limited to a shaking incubator, high speed floor centrifuge, ultralow freezer, and six dual protein electrophoresis rigs. Our students have found this laboratory to be an invaluable educational experience. The laboratory handout, detailed instructions, hypothetical amino acid sequences, and information of how to obtain plasmid reagents are available on request to the corresponding author.

Acknowledgments— The authors gratefully acknowledge Alexei Savchenko and Alexander Yakunin (University of Toronto) for providing suggestions, plasmids, and access to the SPiT clone library.

REFERENCES

[1] Protein Structure Initiative–National Institute of General Medicine Sciences. available at http://www.nigms.nih.gov/Initiatives/PSI (revised 2010). Accessed August 5, 2010.

[2] S. Graslund, et al. (2008) Protein production and purification, *Nat. Methods* **5**, 135–146.

[3] Y. M. She, X. Xu, A. F. Yakunin, S. Dhe-Paganon, L. J. Donald, K. G. Standing, D. C. Lee, Z. Jia, T. D. Cyr (2010) Mass spectrometry following mild enzymatic digestion reveals phosphorylation of recombinant proteins in *Escherichia coli* through mechanisms involving direct nucleotide binding, *J. Proteome. Res.* **9**, 3300–3318.

[4] M. Proudfoot, S. A. Sanders, A. Singer, R. Zhang, G. Brown, A. Binkowski, L. Xu, J. A. Lukin, A. G. Murzin, A. Joachimiak, C. H. Arrowsmith, A. M. Edwards, A. V. Savchenko, A. F. Yakunin (2008) Biochemical and structural characterization of a novel family of cystathionine beta-synthase domain proteins fused to a Zn ribbon-like domain, *J. Mol. Biol.* **375**, 301–315.

[5] B. Nocek, E. Evdokimova, M. Proudfoot, M. Kudritska, L. L. Grochowski, R. H. White, A. Savchenko, A. F. Yakunin, A. Edwards, A. Joachimiak (2007) Structure of an amide bond forming F(420):gamma-glutamyl ligase from *Archaeoglobus fulgidus*—a member of a new family of non-ribosomal peptide synthases, **372**, 456–469.

[6] R. G. Zhang, T. Skarina, J. E. Katz, S. Beasley, A. Khachatryan, S. Vyas, C. H. Arrowsmith, S. Clarke, A. Edwards, A. Joachimiak, A. Savchenko (2001) Structure of *Thermotoga maritima* stationary phase survival protein SurE: a novel acid phosphatase, *Structure* **9**, 1095–1106.

[7] F. W. Studier (2005) Protein production by auto-induction in high density shaking cultures, *Protein Expr. Purif.* **41**, 207–234.

[8] M. M. Bradford (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* **7**, 248–254.

[9] S. Bagby, K. I. Tong, D. Liu, J. R. Alattia, M. Ikura (1997) The button test: a small scale method using microdialysis cells for assessing protein solubility at concentrations suitable for NMR, *J. Biomol. NMR* **10**, 279–282.

[10] M. T. Bes, J. Sancho, M. L. Peleato, M. Medina, C. Gómez-Moreno, M. F. Fillat (2006) Purification of colored photosynthetic proteins for understanding protein isolation principles, *Biochem. Mol. Biol. Educ.* **31**, 119–122.

[11] T. M. Cohen, A. E. Rohs, B. G. Lefebvre (2008) A laboratory demonstration illustrating bioseparations using colorful proteins, *Biochem. Mol. Biol. Educ.* **26**, 287–291.