NIH Public Access
Author Manuscript
Anal Biochem. Author manuscript; available in PMC 2014 November 14.

Published in final edited form as:
Anal Biochem. 2010 May 15; 400(2): 203–206. doi:10.1016/j.ab.2010.01.011.

A unified method for purification of basic proteins

Sanjay Adhikari, **Praveen Varma Manthena**, **Kamal Sajwan**, **Krishna Kiran Kota**, and **Rabindra Roy**
Department of Oncology, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC 20057

Abstract

Protein purification is still very empirical, and a unified method to purify proteins without an affinity tag is not available yet. In the post-genomic era, functional genomics, however, strongly demands such a method. In this paper we have formulated a unique method that can be applied to purify any recombinant basic protein from *E. coli*. Here, we have found that if the pH of the buffer is merely one pH unit below the isoelectric point (pI) of the recombinant proteins, most of the latter bind to the column. This result supports the Henderson-Hasselbalch principle. Considering that *E. coli* proteins are mostly acidic, and based on the pI determined theoretically, apparently all recombinant basic proteins (at least pI-1 6.94) may be purified from *E.coli* in a single-step using a cation-exchanger resin, SP-sepharose, and a selected buffer pH depending on the pI of the recombinant protein. Approximately, two-fifths of human proteome, including many if not all nucleic acid interacting proteins, have a pI of 7.94 or higher; virtually all these 12,000 proteins may be purified using this method in a single-step.

Keywords

Protein purification; single step; ion-exchange chromatography; basic protein

INTRODUCTORY STATEMENT

Despite the large body of knowledge accumulated on recombinant protein expression, production remains a challenge. The biggest obstacle in obtaining large amounts of a given protein is multistep purification process. Each protein is expressed in different amounts and has different properties. Thus, to apply the same purification protocol across a broad range of proteins, researchers have engineered proteins with affinity tags that will bind to a specific ligand. Widely used tags include a small peptide of six histidine residues (6xHis), a calmodulin-binding peptide, the streptavidin / biotin, the cellulose-binding tag, the maltose-binding protein (NusA), Fc fusions and glutathione *S*-transferase (GST). Vectors designed for expressing tagged proteins can be purchased from several manufacturers. Some commercial organizations, however, are now trying to develop high-throughput methods for

Address Correspondence to: Rabindra Roy, Lombardi Comprehensive Cancer Center, LL level, S-122, 3800 Reservoir Road, NW, Georgetown University Medical Center, Washington, DC 20057, Ph: 202-687-7390, Fax: 202-687-1068 rr228@georgetown.edu or Sanjay Adhikari, Lombardi Comprehensive Cancer Center, LL level, S-120, 3800 Reservoir Road, NW, Georgetown University Medical Center, Washington, DC 20057, Ph: 202-687-5010, Fax: 202-687-1068 sa354@georgetown.edu.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

purifying proteins that are not tagged. "Typically most scientists work with tagged proteins through the discovery process. However, it is possible to lose some information because of tag interference in the biological assays. Removing the tag means removing some of the question marks [1]."

In this post-genomic era, the functional genomics researchers are attempting to make use of the vast wealth of data produced by genomic projects to describe gene (and protein) functions and interactions. Thus, the basic need now is to purify a large number of proteins, and eventually the whole human proteome, which is ~30,000 proteins, in a highthroughput (HT) manner [2]. Unlike genomics and proteomics, functional genomics focuses on various dynamic aspects, such as protein-protein interactions in addition to gene transcription and translation, and thus, requires a vast number of purified proteins.

In the “scaling up” of industrially important proteins, it is also not realistic for industries to use an affinity tag, such as 6X His- or GST-tag for cost-effectiveness. Several strategies are being used to increase the yields. Any given method may help to increase expression and purification for a given protein, but often more than one purification strategy is employed. To perform several different rescue strategies on multiple proteins, HT methodologies are applied.

Ion-exchange chromatography is one of the most common procedures for protein purification. It relies on the charge-charge interactions between the proteins in the lysate and the charges immobilized on the ion-exchange resins. The resins are of two types: cation and anion-exchange. If the net surface charge of the proteins is positive, they bind to the cation-exchange column. Moreover, the pH of the binding buffer should be below the isoelectric point (pI) of that protein for sucessful binding to the column. Commonly used cation-exchange resins contain sulfate derivatives (S-resins), whereas the anion-exchangers (CM resins) have carboxylate derived ions. However, it rarely provides single step purification due to lack of specificity [3].

In this study, we have purified four different basic proteins, using an SP-sepharose (cation-exchange) column, to a purity level of 80-99% in a single-step. We have also described the importance and the utility of the degree of alteration of the pH of cation-exchanger below the pI of the protein to be purified. This paper provides the first formulation for a group of tagless proteins that follow a general principle, pI-1 6.94; these proteins can be purified in a single-step using the ion-exchange chromatography. Approximately, two-fifths of human proteome, including many if not all nucleic acid interacting proteins, have a pI of 7.94 or higher; virtually all these 12,000 proteins may be purified using this method in a single-step.

MATERIALS AND METHODS

Construction of expression vectors and purification of mouse Apurinic/ Apyrimidinic endonuclease (mAPE1), human 8oxodG-DNA glycosylase (hOGG1), fragment of Breast cancer susceptibility gene 1 (BRCA1; aa. 502-802) fused with Glutathione S-Transferase (GST-BRCA1), human N-methylpurine-DNA glycosylase (hMPG), N 100 mouse N-methylpurine-DNA glycosylase (N 100 mMPG)

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

All the expression constructs, except mAPE1, were made as described previously [4–7]. mAPE1 and hOGG1 were expressed as His-tag, whereas the GST-BRCA1 was expressed as a GST fusion protein, following the published methods as described previously [4–7]. An expression construct encoding mAPE1 was prepared by ligating a PCR product containing the mAPE1 coding sequence at the Nde I and Bam HI sites of the pET15b vector as described in supplementary method.

For purification of mAPE1 and hOGG1 with His-tag and the GST-BRCA1 with GST-tag, we used Ni-NTA and glutathione columns following the protocols suggested by the manufacturer (Amersham Pharmacia Biotech, Piscataway, NJ) [4, 5]. For pI method, the cells were harvested from 2 liters of culture in 100 ml of Buffer A (120 mM NaCl, 10% glycerol, 1 mM DTT, 0.1 % Tween-20 and respective buffering compound with different pHs as described in Table 1). The cells were then lysed as described previously (7). The lysate was clarified by centrifugation at 30000 rpm for 30 min, and the supernatant was applied at a flow rate of 1 ml/min onto a 1 ml ion-exchange SP-sepharose column (Amersham Pharmacia Biotech, Piscataway, NJ), which was pre-equilibrated with Buffer A. The column was washed with 30 column volumes of Buffer A and then eluted with a gradient of 0–100% of Buffer B (Buffer A plus 500 mM NaCl) in Buffer A. The peak fractions containing 80-99% pure proteins after testing by SDS-PAGE, were pooled, dialysed against the corresponding storage buffer (if necessary) and stored at −20°C or −80°C for future use.

In-gel tryptic digestion and protein identification by mass spectrometry

The proteins were identified by mass spectrometry. The details are described in supplements.

N-terminal amino acid sequencing and western blot analysis

N-terminal amino acid sequencing and western blot analysis were performed as described previously [6]

Oligonucleotide substrate preparation and excision activity Aasay

The mAPE1, N 100mMPG, hMPG and hOGG1 proteins' specific activities were measured in crude extracts or with purified proteins as described previously [4, 8, 9, 10]

RESULTS

Purification and identification of proteins

In this study, we have purified different proteins from 2 L of *E. coli* cultures in different amounts and purity levels. Human *N*-Methylpuine DNA-glycosylase (hMPG; pI, 9.65), a Base Excision Repair (BER) enzyme, was purified in different pH conditions (1, 2 and 3 pH units below pI) and the results showed that the condition when pH equals to pI-1 provides best purity; lowering the pH more than 1 unit increases the yield insignificantly, but invited contaminating proteins and thus, reduces the purity (Table 1a and Supplementary Fig. S1A, B and C). Then, we compared this pI-based method with the popular affinity tag-based methods and found better or similar purification profiles: 25 mg mAPE1 (pI, 8.33), another

BER enzyme, was purified with 99% purity by the pI-based method, whereas only 13 mg by Ni-NTA affinity column. hOGG1 (pI, 9.01) was purified with similar purity and yield using both Ni-NTA affinity and pI-based methods (Table 1b and Supplementary Figs. S2A and B, S3A and B). Even a fragment of Breast cancer susceptibility gene 1 (BRCA1; aa. 502-802) fused with GST (GST-BRCA1; pI, 7.94) protein was purified with significantly better yield and purity (at least 80%) by pI-based method (Table 1c and Supplementary Fig. S4A and B) than by the glutathione affinity column. The latter purification quality was significantly poor due to concomitant expression of various truncated GST-BRCA1 proteins (Supplementary Fig. S4C). Notably, the pI-based method, which does not depend on the presence of GST tag, could provide better resolution and purity. However, N 100mMPG (pI, 6.94) could only be purified partially (Table 1d and Supplementary Fig. S5). Thus, the proteins having the pI 7.94 when combined with the appropriate pH of the purification buffer were purified at ~80-99% purity in single step by ion-exchange chromatography.

DISCUSSION

Functional genomics, the systematic characterization of the functions of an organism's genes, includes the study of the gene products, the proteins. Such studies require methods to express and purify a large number of proteins in a parallel, time- and cost-effective manner. Moreover, a method for the efficient over-expression and purification of recombinant proteins is of paramount importance for biotechnology applications. In particular, for the era of functional genomics that we have entered after the sequencing of complete genomes, this has become a routine matter. HT protein purification is, therefore, essential. To facilitate the procedure of protein purification, several tags to generate fusion proteins are available (e.g., polyHis, GST, MBP, CBP, etc.) for parallel purification using matrices coupled with affinity anchors, such as Ni^{2+}-nitrilotriacetic acid (Ni^{2+}-NTA). Among affinity column options, Ni-NTA and glutathione are the most powerful and popular for purification [11, 12]. Nonetheless, the mild elution conditions employed make affinity tags useful for purifying individual proteins in native form. Remarkably, affinity tags allow diverse proteins to be purified using a generalized protocol, which is in sharp contrast to highly customized procedures associated with conventional chromatography, and obviously provide a compelling consideration for proteomics or structural genomics ventures. Most of the available protein and peptide affinity tags were developed within the last 20 years, and they can be categorized according to the nature of the affinity tag and its target.

Ion-exchange resin is another class of matrix widely used in different separation, purification, and decontamination process. The trapping of ions takes place only with simultaneous releasing of other ions; thus the process is called ion-exchange. There are multiple types of ion-exchange resins, which are fabricated to selectively prefer one or several types of ions. They rarely provide single-step purification due to lack of specific affinity.

However, few reports showed for 1 or 2 protein(s) has/have been purified by single-step in ion-exchange chromatography. For example, reversing the flow direction (“back flash”) or “codon optimization” in combination with “strikingly high isoelectric point” helped purify 1 or 2 proteins in a single-step [3, 6]. However, none of those studies provided adequate

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

rationale or attempted to formulate to establish a unified method for protein purification. Overall, those studies were very empirical.

Even though many, if not most, recombinant proteins are now expressed with and purified using affinity tags, there are still serious issues. For example, how those tags are affecting the structure and functions is a growing concern. In fact, removing the tag actually can solve some of the problem. However, the tag removal process requires digestion of the fusion proteins with some restriction grade proteases. This digestion step is often highly challenging as it compels the proteins to be exposed to higher temperature and/or prolonged incubation time. Even then, another extra purification step is required to remove the cleaved tag. Thus, this process affects many times stability and function of the target proteins, and as a result defeats the purpose of single-step purification. Another issue in using affinity tags is the cost of the column. Ion-exchange columns are in general comparatively less expensive. Less cost is highly desirable, but this option only exists if it is possible to purify the proteins without any affinity tag. Recently, several new approaches are taken to help purify proteins without a tag. For example, it is possible to use aptamer to purify protein but they are extremely complex, uncertain and not yet suitable for HT system [13].

This study aimed to develop a method that will use ion-exchange resin but provide the specificity of affinity chromatography and thus, remove the tag-related concerns. Notably, in humans, the complete proteome consists of 30,000 proteins, and at least 12,000 proteins, including many if not all nucleic acid interacting proteins, have a pI of 7.94 or higher [14]. Thus, the method described here provides proof of the concept that over two-fifths of whole human proteome may be purified by this method. This method gives at least a comparable level, if not more, of purity and yield to the commonly used affinity tag method.

This study had two important considerations for protein purification: First, most of the *E. coli* proteins are acidic [15] and secondly, the pH of the buffer in use should not be very low. According to the Henderson-Hasselbalch equation, there should be enough binding (~90%) even at 1 pH unit below the pI (pI-1) of the protein. Now, further decreasing the pH, in spite of improving the specific binding, invites contaminating proteins to bind. In fact, selection of a narrow pH window and a precise salt gradient helped significantly improve our protein purification efficiency by providing high specificity in regard to binding and elution of proteins to and from the ion-exchange column. In conclusion, we have shown that at least four proteins having pI above or equal to 7.94 (pI 7.94) can be purified in a single-step using this method. Therefore, we predict that this method may work for proteins that come under the criteria of pI-1 6.94. This method, to our knowledge, is the first report of purification of a group of proteins, based on theoretically predicted and consequent selection of pH of the purification buffer and pI of the protein. This new method was developed by blending the classical concepts of chemistry, such as the Henderson-Hasselbalch principle with the information from the whole human and *E. coli* proteome, a great progress in modern biology, will definitely be useful and have an impact in the future progress of functional genomics.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Prof. Sankar Mitra and Dr. Tapas Hazra of UTMB, Galveston, Texas for providing the mAPE1 and hOGG1 cDNA. We also thank Prof. Eliot Rosen of Georgetown University for GST-BRCA1 cDNA. We thank Dr. Amrita Cheema for proteomics experiments performed at the Proteomics and Metabolomics Shared Resource of the Lombardi Comprehensive Cancer Center. We thank Protein Core facility, UTMB, TX for amino acid sequence analysis. We also thank Mr. Cliff Chung and Ms. Karen Howenstein for expert editorial help. The work was supported in part by NIH grants RO1 CA 92306 and RO1 CA 113447 awarded to RR.

ABBREVIATIONS

mAPE1	mouse Apurinic / apyrimidinic endonuclease
hOGG1	human 8-Oxoguanine DNA-glycosylase
hMPG	human N-Methyllpurine DNA-glycosylase
mMPG	mouse N-Methyllpurine DNA-glycosylase
THF	Tetrahydrofuran
εA	1,N^6-ethenoadenine

REFERENCES

1. Bonetta L. Protein purification: Fast forward. Nature. 2006; 439:1017–1021.
2. Hubbard T, Barker D, Birney E, Cameron G, Chen Y, Clark L, Cox T, Cuff J, Curwen V, Down. T, Durbin R, Eyras E, Gilbert J, Hammond M, Huminiecki L, Kasprzyk A, Lehvaslaiho H, Lijnzaad P, Melsopp C, Mongin E, Pettett R, Pocock M, Potter S, Rust A, Schmidt E, Searle S, Slater G, Smith J, Spooner W, Stabenau A, Stalker J, Stupka E, Ureta-Vidal A, Vastrik I, Clamp M. The Ensembl genome database project. Nucleic Acids Res. 2002; 30:38–41. [PubMed: 11752248]
3. Chern MK, Shiah WJ, Chen JJ, Tsai TY, Lin HY, Liu CW. Single-step protein purification by back flush in ion-exchange chromatography. Anal Biochem. 2009; 392:174–176. [PubMed: 19497288]
4. Hill JW, Hazra TK, Izumi T, Mitra, S S. Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair. Nucleic Acids Res. 2001; 29:430–438. [PubMed: 11139613]
5. Ma Y, Katiyar P, Jones LP, Fan S, Zhang Y, Furth PA, Rosen EM. The breast cancer susceptibility gene BRCA1 regulates progesterone receptor signaling in mammary epithelial cells. Mol. Endocrinol. 2006; 20:14–34. [PubMed: 16109739]
6. Adhikari S, Manthena PV, Uren A, Roy R. Expression, purification and characterization of codon-optimized human N-methylpurine-DNA glycosylase from *Escherichia coli*. Protein Expr. Purif. 2008; 58:257–262. [PubMed: 18191412]
7. Adhikari S, Uren A, Roy R. N-terminal extension of N-methylpurine DNA glycosylase is required for turnover in hypoxanthine excision reaction. J. Biol. Chem. 2007; 282:30078–30084. [PubMed: 17716976]
8. Adhikari S, Toretsky JA, Yuan L, Roy R. Magnesium, essential for base excision repair enzymes, inhibits substrate binding of N-methylpurine-DNA glycosylase. J. Biol. Chem. 2006; 281:29525–29532. [PubMed: 16901897]
9. Adhikari S, Kennel SJ, Roy G, Mitra PS, Mitra S, Roy R. Discrimination of lesion removal of N-methylpurine-DNA glycosylase revealed by a potent neutralizing monoclonal antibody. DNA Repair (Amst). 2008; 7:31–39. [PubMed: 17768096]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

10. Adhikari S, Uren A, Roy R. Excised damaged base determines the turnover of human N-methylpurine-DNA glycosylase. DNA Repair (Amst). 2009; 8:1201–1206. [PubMed: 19616486]
11. Scheich C, Sievert V, Büssow K. An automated method for high-throughput protein purification applied to a comparison of His-tag and GST-tag affinity chromatography. BMC Biotechnol. 2003; 3:12. [PubMed: 12885298]
12. Lichty JJ, Malecki JL, Agnew HD, Michelson-Horowitz DJ, Tan S. Comparison of affinity tags for protein purification. Protein Expr. Purif. 2005; 41:98–105. [PubMed: 15802226]
13. Javaherian S, Musheev MU, Kanoatov M, Berezovski MV, Krylov SN. Selection of aptamers for a protein target in cell lysate and their application to protein purification. Nucleic Acids Res. 2009; 37:e62. [PubMed: 19304751]
14. Chen EI, Hewel J, Felding-Habermann B, Yates JR 3rd. Large scale protein profiling by combination of protein fractionation and multidimensional protein identification technology (MudPIT). Mol. Cell. Proteomics. 2006; 5:53–56. [PubMed: 16272560]
15. Champion KM, Nishihara JC, Joly JC, Arnott D. Similarity of the Escherichia coli proteome upon completion of different biopharmaceutical fermentation processes. Proteomics. 2001; 1:1133–1148. [PubMed: 11990508]

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

a: Purification of hMPG (pI, 9.65) at different pHs

Protein	pH of the buffer used[a]	Total Protein (mg)	Total activity (IU[b])	Specific activity (IU/mg)	Recovered Activity (IU)	Specific activity (IU/mg)	% purity of different fractions in coomassie blue stained gel[c]	Yield (mg)
hMPG	8.65	400	1600	4	800	78	90–95	10.0
	7.65	400	1600	4	960	60	55–89	12.0
	6.65	400	1600	4	985	55	50–80	12.3

b: Comparison of protein purification methods- Ni-NTA affinity vs. ion-exchange[a] method based on pI of the proteins

Protein	Purification procedure	Total Protein (mg)	Total activity (IU[b])	Specific activity IU/mg)	Recovered Activity (IU)	Specific activity (1U/mg)	% purity of different fractions in coomassie blue stained gel[c]	Yield (mg)
mAPE1 (pI, 8.33)	Ni-NTA	500	1000	2	442	32	78–95	13
	Ion-exchange/pI	500	1000	2	900	36	98–99.9	25
hOGG1 (pI, 9.01)	Ni-NTA	465	465	1	325	12	76–92	21
	Ion-exchange/pI	465	465	1	325	12	77–93	20

c: Comparison of protein purification methods- Glutathione affinity vs. ion-exchange[a] method based on pI of the proteins

Protein	Purification procedure	Total Protein (mg)	% purity different fractions in coomassie blue stained gel[b]	Yield (mg)
GST-BRCA1 (pI, 7.94)	Glutathione affinity	800	<10	<1
	Ion-exchange/pI	800	60–80	4.5

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

d: Purification of a neutral protein N 100mMPG (pI, 6.94)

Protein	pH of the buffer used[a]	Total Protein (mg)	Total activity (IU[b])	Specific activity (IU/mg)	Recovered Activity	Specific activity (IU/mg)	% purity different fractions in coomassie blue stained gel[c]	Yield (mg)
N 100 mMPG	6.1	600	2000	3.3	135	24.5	37–60	2.9 mg

[a]The buffering compounds used to prepare buffers of different pHs were 40 mM Tris-HCl for pH 8.65 and 7.65, and 20 mM PIPES-NaOH for pH 6.65. The other details of buffer compositions are described in "Online Methods".

[b]One International Unit (IU) of hMPG-mediated cleavage reaction is defined as the excision of 20 nmol of double stranded DNA containing 1,N^6ethenoadenine in 10 mins at 37°C.

[c]Best five fractions. See supplementary figure S1 for details

[a]Forty millimolar Tris-HCl of pH 7.33 and 8.01 was used as a buffering compound to prepare buffers for purification of mAPE1 and hOGG1, respectively by ion-exchange/pI method. The other details of buffer compositions are described in "Online Methods".

[b]One International Unit (IU) of mAPE-mediated cleavage reaction is defined as the excision of double stranded DNA containing THF (125 nmol) in 7 min at 37°C. For hOGG1 1IU is defined as the excision of 16 pmol of double stranded DNA containing 8-oxodG in 1 hr at 37°C.

[c]Best fractions. See supplementary figures S2 and S3 for details

[a]Twenty millimolar PIPES-NaOH of pH 6.94 was used as a buffering compound to prepare buffer for purification of GST-BRCA1 by ion-exchange/pI method. The other details of buffer compositions are described in "Online Methods".

[b]Best fractions. See supplementary figure S4 for details

[a]Forty millimolar MES-NaOH of pH 6.1 was used as a buffering compound to prepare buffer for purification of N 100mMPG by ion-exchange/pI method. The other details of buffer compositions are described in "Online Methods".

[b]One International Unit (IU) of hMPG-mediated cleavage reaction is defined as the excision of 20 nmol of double stranded DNA containing 1,N^6ethenoadenine in 10 min at 37°C.

[c]Best five fractions. See supplementary figure S5 for details