Snags with tags: Some observations made with $(His)_6$-tagged proteins

P. Ramage[1], R. Hemmig[1], B. Mathis[1], S.W. Cowan-Jacob[1], J.M. Rondeau[1], J. Kallen[1], M.J.J. Blommers[1], M. Zurini[2], and S. Rüdisser[1]
Novartis Pharma Ltd., Basel, Switzerland, [1]Protein Structure Unit, and [2]Biomolecules Production Unit

PROTEIN PURIFICATION

The unusual behavior during purification of four affinity-tagged proteins—a MAP kinase, a transcription factor, intracellular binding protein, and a receptor tyrosine kinase—led the authors to devise alternate methods of obtaining their end-product. They found it necessary to either remove the affinity tag (by recloning or by cleavage of the tag) or omit the traditional immobilized metal ion adsorption chromatography (IMAC) step altogether from the purification scheme.

Introduction

Over the last decade, the cloning of an affinity tag into a target protein's gene has revolutionized the purification of recombinant proteins. Such tags have greatly facilitated protein purification and sometimes improved expression levels. Recent advances in genomics and proteomics have meant that huge numbers of proteins are now being expressed, usually in a tagged format that allows standard purification schemes to be used in multiple serial or parallel purifications. The most widely used affinity tag is the poly-histidine tag, usually comprising 6 or 10 histidines (plus several additional amino acids) and positioned at the N- and/or C-terminus of the protein.

The tagged protein is purified by passing the supernatant (or even solubilized inclusion bodies) obtained after cell lysis over a chelating chromatography medium to which metals ions such as zinc, cobalt, or nickel have been attached. The histidine residues in the tag will adsorb to the charged metal ions, and thus the proteins containing the His-tag will be selectively retained on the column. After the bound proteins have been washed with high-salt or other buffers, they are eluted either by competition with imidazole or by reducing the pH of the elution buffer. This method of purification is referred to as immobilized metal ion adsorption chromatography (IMAC). The purity of the eluted protein is dependent on the level of expression and elution scheme used, but is usually in the range of 80–90%.

Despite its popularity, there are drawbacks to this kind of purification approach, including the possibility of co-elution of metalloproteinases and oxidation of the eluted protein, catalyzed by leached metal ions. One additional drawback that is seldom described, is the influence of the His-tag itself on the behavior of the recombinant protein.

In the following article, we describe the unusual behavior of four recombinant proteins due either to the presence of His-tags or to the IMAC method used to purify them. All the proteins described were being prepared for structural studies and therefore were required to meet the highest specifications for purity, homogeneity, and monodispersity at high concentrations, while still maintaining function. In these specific instances, in order to meet these requirements it became necessary to either remove the His-tag (by recloning or by cleavage of the tag) or to omit IMAC altogether from the purification scheme.

Protein 1: A MAP kinase

Protein 1 expressed well in *E. coli* (~30 mg/liter) and was purified over a combination of nickel-charged nitrilotriacetic acid (Ni-NTA) and ion exchange chromatography columns. The ion exchange profile and subsequent analysis by denaturing polyacrylamide gel electrophoresis showed unexpected charge heterogeneity. Attempts to phosphorylate the protein produced incomplete and heterogeneous results. LC-MS was unable to provide conclusive evidence for oxidation. After extensive efforts, crystals diffracting to 2.5 Å were obtained with non-phosphorylated protein.

When the initial purification step was changed to chromatography using Blue Sepharose™ 6 Fast Flow, highly purified kinase was obtained. In subsequent anion exchange chromatography using a Mono Q™ column, a single peak (Fig 1) was produced. This protein could be fully phosphorylated, and crystals diffracting to 1.8 Å were obtained almost immediately.

From these results we concluded that modification to the protein occurs either during or after purification on Ni-NTA. The kinase purified using Blue Sepharose 6 Fast Flow was purer, more stable, and crystallized more

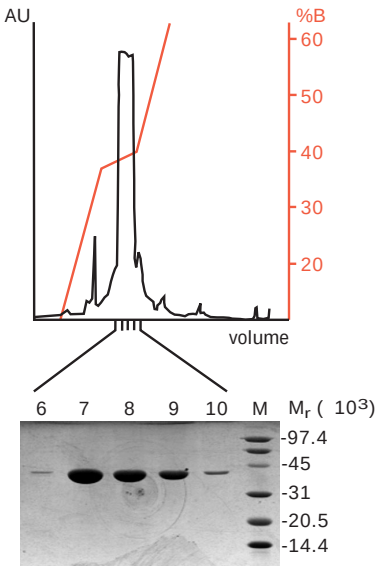

Fig 1. Mono Q anion exchange chromatography of MAP kinase pre-purified using Blue Sepharose 6 FF affinity chromatography. Mono Q HR 5/5 column was run with Buffer A (20 mM TrisHCl, pH 8, 1 mM DTT). A gradient of 0–100% 1 M NaCl in Buffer A was formed over 15 column volumes, with a reduced gradient between 35% and 40% (Buffer B). Pictured below the chromatogram is the reducing 12% SDS-polyacrylamide gel used to analyze eluted fractions. The first five lanes consist of fractions 6–10 from the Mono Q column, as indicated on the chromatogram. M = molecular weight marker.

easily, giving crystals with higher resolution.

Protein 2: A transcription factor

Protein 2 was well expressed as a soluble, N-terminally His-tagged protein in *E. coli*. Purification was carried out using Ni-NTA followed by gel filtration on Superdex™ 75. The resulting protein solution was turbid and could not be concentrated to above 3 mg/ml. Some initial NMR data suggested oxidation.

Based on previous experience, we decided that because

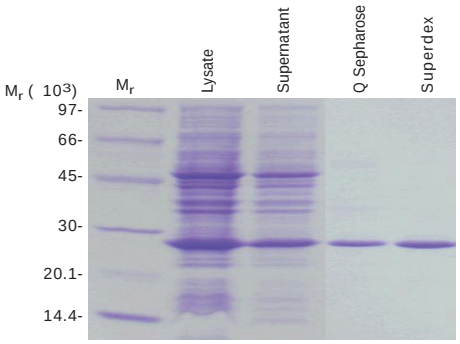

Fig 2. Purification of Protein 2 over Q Sepharose HP. Gel: 4–20% reducing SDS-polyacrylamide. M_r = molecular weight marker; lysate = cell lysate; Supernatant = supernatant after lysate; Q Sepharose = pooled peak from Q Sepharose 26/10 HP column; Superdex = final pooled material following Superdex 75 HR 10/30 column.

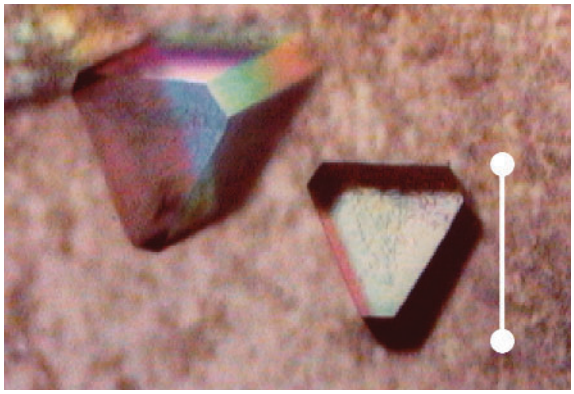

Fig 3. Protein 2/inhibitor-complex crystals. Crystals were grown in the presence of an inhibitor by vapor diffusion. Protein concentration: 16 mg/ml. Diffraction: 3.7 Å.

the expression was reasonable (~20% total protein), anion exchange chromatography using Q Sepharose High Performance ought to provide adequate purification. At pH 8.5 the protein eluted in the late unbound fraction, giving a highly purified product (Fig 2). The resulting pool was further purified using Superdex 75 and could be easily concentrated to 16 mg/ml for crystallization. Crystals diffracting to 2.5 Å were rapidly obtained, but not in the presence of ligand. NMR investigations were unable to show significant binding of ligand to the protein. Thereafter, the tag was removed by recloning, and the protein was repurified. At this point, NMR was able to show ligand binding, and crystals were obtained of protein/inhibitor complex (Fig 3).

Our results indicate that when purified using Ni-NTA -agarose, this $(His)_6$-tagged protein was unstable and oxidized, presumably due to leaching nickel ions. When purified using conventional chromatography, the protein was well behaved, stable, and crystallizable, but showed little or no ligand binding. Recloning to remove the tag resulted in an equally stable protein, which was shown to bind ligand and could be co-crystallized with ligand.

Protein 3: An intracellular binding protein

Protein 3 had previously been expressed in *E. coli* as a non-tagged soluble protein. The protein was extremely soluble and could only be crystallized at extremely high concentrations, and not with the desired ligands. Based on experience in other groups, the protein was point-mutated in several positions to make it less soluble, and an N-terminal His-tag that could be cleaved with PreScission™ Protease was added. Perhaps not surprisingly, the resulting protein was largely insoluble and could be expressed only in a soluble form by fermentation at lower temperature.

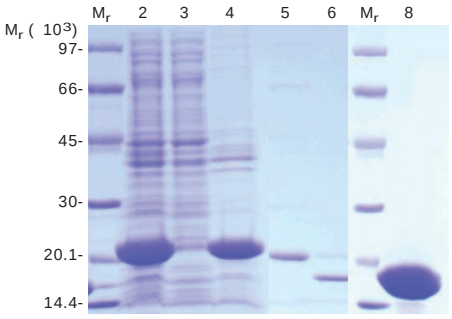

Fig 4. Expression of Protein 3 as inclusion bodies. Lane 1, M_r = molecular weight marker; lane 2, cell lysate; lane 3, supernatant after lysate; lane 4, final inclusion bodies; lane 5, refolded protein before treatment with PreScission Protease; lane 6, protein after digestion with PreScission Protease; lane 8, final concentration protein.

Even then, the protein was unstable following Ni-NTA agarose chromatography, precipitating readily.

We decided to try a refolding approach without affinity chromatography. Cells containing the recombinant protein were grown at 37 °C, and the protein expressed well as inclusion bodies (Fig 4). The inclusion bodies were harvested and washed, and parallel dilution- and dialysis-based refoldings were set up. The best recoveries of soluble protein were with a dialysis approach, allowing the protein to refold slowly. Because the protein was initially destined for NMR, and had a high pI, further purification over Mono S™ into phosphate buffers was envisaged. However, when the protein was fast-desalted into phosphate pH 6.5 (from Tris pH 8.5), it precipitated within minutes. A series of different buffers below pH 8 were tried, but none remedied the problem.

We decided to remove the His-tag using PreScission Protease while the protein remained in the refolding buffer (50 mM Tris, 5 mM DTT, pH 8.0) and then to repeat the desalting experiment. To our surprise, the protein remained in solution in 50 mM Hepes, pH 7.5 and could be purified over Mono S and Superdex 75 media. Furthermore, the resulting NMR spectrum (Fig 5) indicated that the protein was folded and that it bound ligand in a specific manner.

In this case, an N-terminal His-tag in some way interfered with the refolding of the protein, and destabilized it. Proteolytic removal of the His-tag resulted in a protein exhibiting the desired characteristics.

Protein 4: A receptor tyrosine kinase

Protein 4, in common with many of its family members, could not be expressed in an active, soluble form in *E. coli*. Instead, intracellular baculovirus expression

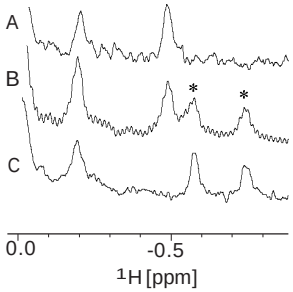

Fig 5. 1-D ^{1}H NMR spectrum of the upfield shifted methyl resonances of Protein 3. All spectra were acquired at 296 K in 50 mM sodium phosphate buffer at pH 7.00. A. spectrum of Protein 3 obtained from soluble expression in *E. coli*. B. Spectrum from the same batch of protein as for curve A, but with less than an equimolar amount of ligand added. Formation of the Protein 3/ligand complex can be observed by additional resonances indicated by stars. C. Spectrum of refolded Protein 3 with an excess of ligand. Binding of ligand to refolded Protein 3 is confirmed by signals characteristic for the Protein 3/ligand complex in curve C.

was used, employing an N-terminal $(His)_6$-tag. Good expression levels (~ 15 mg/liter) were obtained, and a single IMAC step using either cobalt or nickel resulted in ~ 90% purity. Following gel filtration chromatography, the protein was easily concentrated to > 10 mg/ml without significant losses and appeared highly suitable for crystallization screening. However, light-scattering studies showed a disturbing degree of polydispersity (~ 20%), which decreased over several days of storage at 4 °C. N-terminal sequence analysis over the same timeframe showed a progressive clipping of the N-terminus until the tag and linker had been fully removed.

Attempts to autophosphorylate the activation loop of the protein were hampered by low kinase activity, and mass spectrometry showed that the protein was partially oxidized. The progressive clipping of the His-tag was attributed to the activity of a contaminating metalloproteinase. EDTA added after elution would protect from oxidation but would inhibit the metalloproteinase and thus leave the protein in its polydisperse form.

IMAC was subsequently replaced by conventional ion exchange chromatography using Q Sepharose High Performance. Use of this medium resulted in a significantly enriched, but not pure, kinase pool (Fig 6A); this pool could be further purified using a Mono Q anion exchange column (Fig 6B). Sequence analysis at this point showed the protein to be N-terminally homogeneous, more easily autophosphorylated, and non-oxidized. Unfortunately, NMR spectroscopy was unable to demonstrate ligand binding.

The protein was subsequently recloned without the His-tag and repurified using anion exchange chromatography as before. The protein was now observed

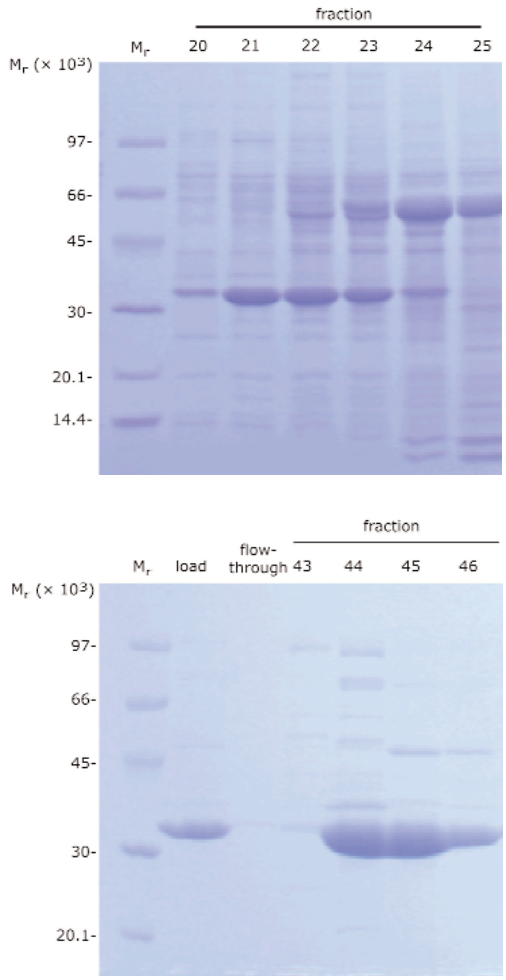

Fig 6. Purification of a receptor tyrosine kinase. Gels are 4–20% polyacrylamide (Novex) run according to the manufacturer's instructions. A. Ion exchange chromatography using Q Sepharose HP. M_r = LMW-SDS Marker Kit (Amersham Biosciences, 17-0446-01); remaining lanes, fractions 20–25 from Q Sepharose HP column. B. Anion exchange chromatography using Mono Q column. M_r, as in A; load = protein loaded on Mono Q HR 10/10 column; flowthrough = unbound material; remaining lanes, fractions 43–46 from Mono Q column.

to autophosphorylate more rapidly, to be more stable during autophosphorylation, and to show ligand binding by both NMR and reverse phase HPLC. Several crystals have been obtained (Fig 7), and the process is currently being optimized.

For this receptor tyrosine kinase, the presence of an N-terminal His-tag caused the protein to aggregate and prevented ligand binding. The use of IMAC chromatography resulted in oxidation of the protein and proteolysis by a contaminating metalloproteinase. Recloning to remove the tag followed by conventional anion exchange chromatography led to a protein that bound ligand and from which crystals could be obtained.

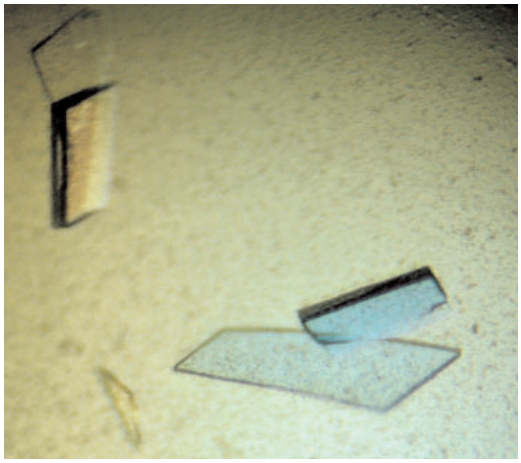

Fig 7. Different crystal forms of a receptor tyrosine kinase. Crystals were grown under oil in complex with an inhibitor. Protein concentration: 12 mg/ml. Diffraction was high in the first two dimensions, but poor in the final dimension (4 Å).

Summary

The four examples above illustrate some of the problems we have seen with His-tagged proteins in our laboratories. Particularly when a protein is being produced for structural studies, great care should be taken in designing constructs. Although affinity tags simplify purification, they are often an unnecessary aid to skilled purifiers. There are instances where tags are useful—for example where only low-level expression is possible, for detection purposes in eukaryotic expression systems, and even to boost expression. The introduction of high-throughput cloning and expression approaches will result in an increased use of affinity tags to facilitate generic purification approaches. The His-tag is deservedly popular because of cloning convenience, the robustness of the affinity purification step, and the ease of elution. We continue to use it, where appropriate. However, care must be taken, and the resulting protein must be carefully watched for unusual or unexpected behavior. As shown above, the His-tag or the IMAC matrix can have a number of different adverse effects on a protein.

ORDERING INFORMATION

HiLoad™ 16/10 Q Sepharose HP	1	17-1064-01
HiLoad 26/10 Q HP	1	17-1066-01
Mono Q HR 5/5	1	17-0546-01
Mono Q HR 10/10	1	17-0556-01
Blue Sepharose 6 Fast Flow	50 ml	17-0948-01
Superdex 75 HR 10/30	1	17-1047-01
Mono S HR 5/5	1	17-0547-01
Mono S HR 10/10	1	17-0557-01
PreScission Protease	500 units	27-0843-01