FEBS Letters 345 (1994) 125–130

FEBS 14015

Refolding proteins by gel filtration chromatography

Milton H. Werner[a], G. Marius Clore[a,*], Angela M. Gronenborn[a,*], Akiko Kondoh[b], Robert J. Fisher[b]

[a]*Laboratory of Chemical Physics, Building 5, National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Building 5, Room 132, 9000 Rockville Pike, Bethesda, MD 20892, USA*

[b]*PRI/DynCorp, National Cancer Institute-Frederick Cancer Research and Development Center, MD 21701, USA*

Received 8 March 1994

Abstract

We have developed a facile means for the refolding of miligram quantities of purified proteins that employs gel filtration chromatography. We demonstrate by electrophoretic mobility shift and NMR spectroscopy that human ETS-1 protein, bovine ribonucelase A and *E. coli* integration host factor can be refolded into the native conformation using this technique. We have extended this strategy to the preparation of miligram quantities of macromolecular complexes suitable for structural analysis by NMR spectroscopy or X-ray crystallography. The diverse challenges to overcome in refolding these proteins illustrates the potential of this technique as a general approach for recovery of recombinant proteins produced as insoluble inclusion bodies.

Key words: Protein refolding; Gel filtration; Technique

1. Introduction

Modern analysis of protein structure by NMR spectroscopy or X-ray crystallography generally requires the overexpression and purification of tens of miligrams of the protein of interest. Frequently, foreign gene products overexpressed in the bacterium *Escherichia coli* are produced in an insoluble form known as inclusion bodies [1]. The isolation of the desired gene product from inclusion bodies requires several washing steps, denaturation of the extracted protein pellet in 6–8 M guanidine hydrochloride (GdmCl) and purification by reversed-phase high pressure liquid chromatography (HPLC) [1]. Finally, the protein must be refolded to the native conformation.

Typical conditions for protein refolding involve dilution of the denatured protein to ≤ 10 μM into the appropriate folding buffer followed by dialysis and concentration. The refolding problem is compounded when disulfide bonds need to be formed as well, often leading to more dilute conditions for efficient pairing of the correct cysteine residues. We have encountered the above problems in the preparation of recombinant human ETS-1 (rETS-1) protein which is expressed as an inclusion body in *E. coli* and contains 9 cysteines.

An alternative methodology of refolding using gel filtration chromatography shows promise in the preparation of samples suitable for structural studies. Gel filtration of the GdmCl denatured protein on a column lacking denaturant allows the refolding of up to 10 mg of protein, even using starting concentrations of protein greater than 10 mg/ml. The protein is diluted only 5–20-fold during elution, depending on the column size used for the renaturation process. This approach has been successfully applied using a variety of salt, buffer and temperature conditions as well as for proteins of different sizes and compositions. Methods for refolding rETS-1 and bovine ribonuclease A (RNase A) are presented. The general applicability of this approach for the large scale preparation of macromolecular complexes, demonstrated by the the refolding of the heterodimeric protein *E. coli* integration host factor (IHF) in the presence of a 30 bp DNA binding site, is also pointed out.

2. Materials and methods

2.1. Gel filtration chromatography

1–10 mg of the protein was solubilized in 1–2 ml of 50 mM Tris-HCl, 50 mM dithiothreitol (DTT), 200–500 mM NaCl and 6–8 M GdmCl, pH 8.5; samples were typically allowed to stand several hours or overnight at ambient temperature under these conditions prior to refolding.

rETS-1 isoform proteins were refolded on a Superdex 75 HR10/30 column (Pharmacia/LKB) equilibrated with 20 mM HEPES, 150 mM NaCl, 3.3 mM Na_2EDTA and 0.1% Tween 20, pH 6.8. RNase A was refolded on a Sephacryl S-100 column (2.6 × 100 cm, Pharmacia) equilibrated with 20 mM sodium phosphate and 200 mM sodium chloride. The Superdex 75 column was run at 25°C and the Sephacryl S-100 column was run at 4°C, although there was little difference in the recovery of folded protein between these two temperatures. Flow rates were 0.2 ml/min for Superdex 75 and 0.5 ml/min for Sephacryl S-100 under the control of a Pharmacia LCC-500 Liquid Chromatography Controller and a Pharmacia P-500 pump. The eluted samples were concentrated in a colloidian membrane apparatus using a 10,000 molecular weight cutoff nitrocellulose membrane (Schleicher & Schuell) or a Micro-ProDiCon apparatus (Spectrum) using a 5000 molecular weight cutoff membrane at 25°C.

For the complex with IHF, uniformly ^{15}N-labelled and unlabelled protein were denatured in 8 M GdmCl and the α and β subunits isolated and purified using reversed phase HPLC. The proteins were reassembled as ^{15}N-labelled α and natural abundance β by dissolving the lyo-

*Corresponding authors. Fax: (1) (301) 496 0825.

SSDI 0014-5793(94)00401-G

philized HPLC fractions in the above denaturation buffer (without DTT) and refolding them as described above in the presence of a 30 bp oligonucleotide comprising the λ *attP* H′ site.

2.2. Nuclear extracts

Nuclear extracts (NE) were prepared from 10 g of CEM cells. The cells were lysed in 100 ml of isotonic low salt lysis buffer (20 mM HEPES, pH 6.8, 0.25 M sucrose, 5 mM $MgCl_2$, 5 mM KCl, 0.05% Nonidet P-40, 1% aprotinin and 0.4 mM phenylmethanesulfonylfluoride (PMSF)) and centrifuged at 1000 × *g* for 10 min. The nuclear pellet was then incubated for 60 min at 4°C with 50 ml of NE buffer (20 mM Tris-HCl, pH 7.5, 20% glycerol, 0.35 M KCl, 0.1 mM EGTA and 0.5 mM EDTA). Then, the nuclear extract was centrifuged at 100,000 × *g* for 60 min and dialyzed for 18 h against NE buffer containing 100 mM KCl, 0.1% aprotinin and 0.1% NaN_3. The dialyzed nuclear extract was then centrifuged at 100,000 × *g* for 30 min and stored at −70°C. Nuclear extracts prepared in this way contained about 5 mg/ml protein as determined by the BCA color reaction (Pierce).

2.3. Electrophoretic mobility shift assay (EMSA)

DNA sequence-specific binding by nuclear extracts from CEM cells, purified renatured ETS1 proteins or by purified rETS1 isoforms were assessed by using EMSA. The nuclear extracts or purified ETS1 isoforms were incubated with ^{32}P-labelled probe (5′-GATCTC-GAGCCGGAAGTTCGA-3′) for 60 min at room temperature in 4% glycerol, 1 mM EDTA, 5 mM DTT, 10 mM Tris-HCl, pH 7.5 and 1.0 mg/ml bovine serum albumin (BSA). Experiments using CEM nuclear extracts contained 100 ng/20 μl poly dI–dC as a non-specific competitor to reduce the background. After incubation, 20 μl were loaded onto 4.5% or 6% native polyacrylamide gels. These gels contained 0.25 × TBE buffer (1 × TBE=89 mM Tris-HCl, 89 mM boric acid, 1 mM EDTA), 5% glycerol, 1 μl/ml TEMED and 14 μl/ml 5% ammonium persulfate. The electrophoresis running buffer was 0.25 × TBE. The EMSA gels were pre-run for 30 min at 250 V. 20 μl of sample was loaded and electrophoresis was continued for an additional 1.5 h. The EMSA gel was dried for 30 min and exposed to X-ray film for an appropriate length of time.

3. Results and discussion

3.1. Refolding of rETS-1

Human ETS1 exists as a 51 kDa protein (full length, p51) and an alternatively spliced isoform of 42 kDa (p42) that lacks the phosphorylation domain [3]. These proteins have been isolated from CEM cells [4] as a mixture of p42 and p51 using an immunoaffinity technique and have been renatured by a 10-fold dilution into a buffer lacking denaturant. The recombinant p51 and p42 ETS1 proteins each contain 9 cysteines and are expressed in *E. coli* as inclusion bodies [3]. While the pattern of disulfide crosslinks is unknown, the large number of potential crosslinks made the refolding of this protein a challenge [3]. Fig. 1 (lanes 3 and 5) demonstrates that the p51 and p42 ETS1 proteins refold to the proper conformation on a Superdex 75 column. In EMSA assays, recombinant p51 and p42 recognize a specific binding site and exhibit the same mobility shift as protein found in CEM cell nuclear extracts (lane 8). The recovery of p42 ETS1 from the Superdex 75 column was 71 ± 15% (six determinations). Renatured ETS1 isoform proteins from CEM cells (Fig. 1, lane 1) show mobility shifts similar to the recombinant ETS1 proteins (lanes 3 and 5) and to the ETS1 isoforms found in CEM nuclear cell extracts (lane 8). The NE protein represents the native form of ETS and has not been exposed to denaturants.

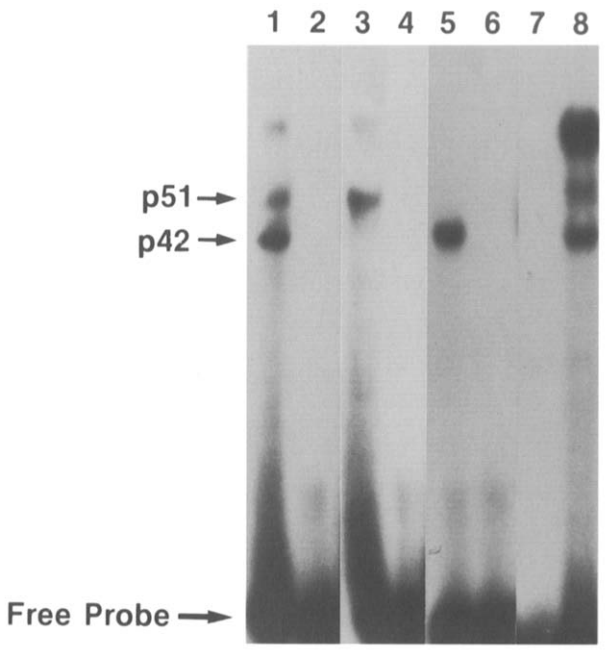


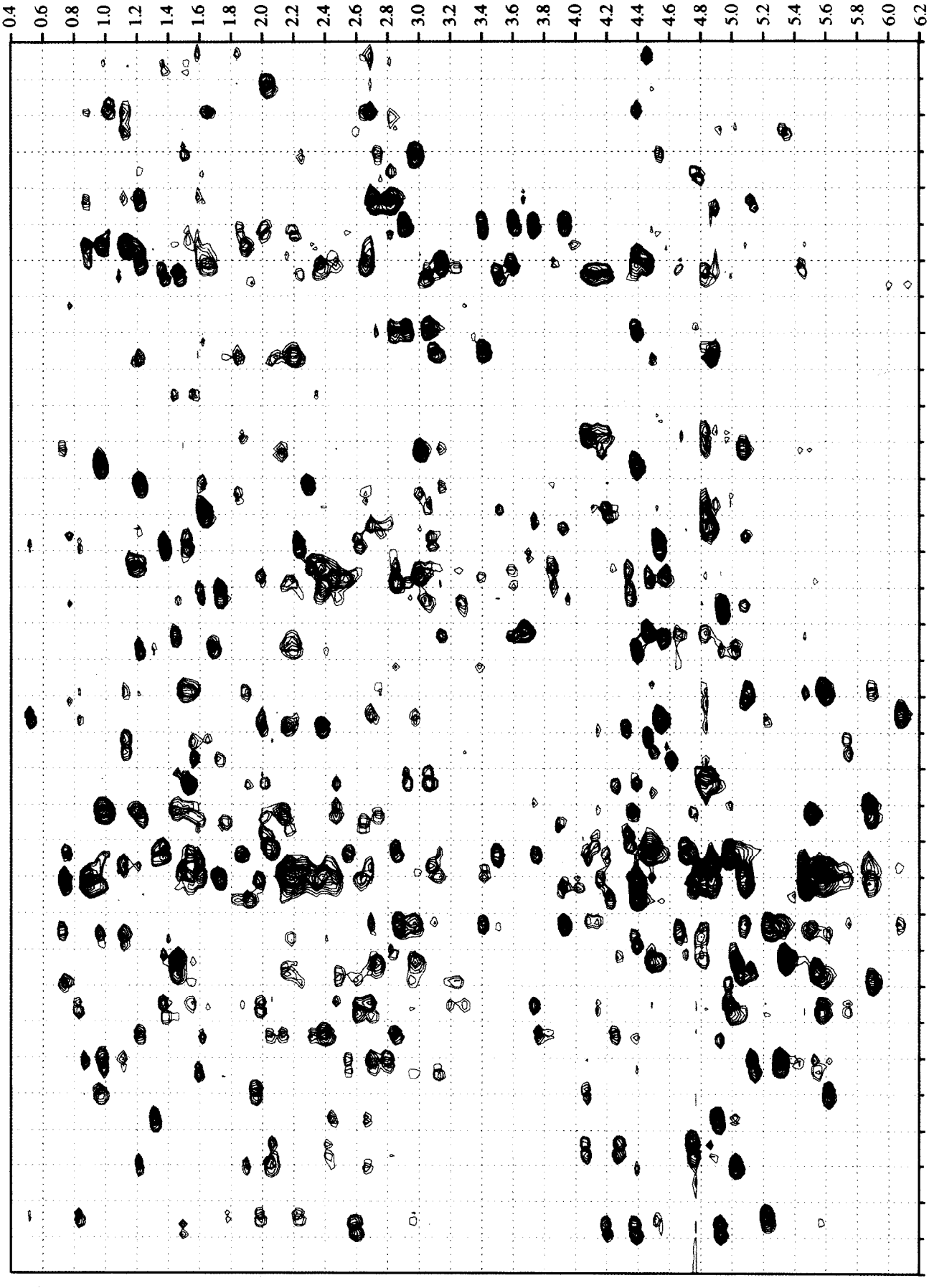
Fig. 1. Mobility shift of renatured human ETS1 isoforms, renatured recombinant human ETS1 isoforms and native ETS1 (CEM nuclear extract). The CEM nuclear extract (1 μl of 5 mg/ml protein extract) (lane 7), 1 ng of the renatured ETS1 (lane 2) or 1 ng of renatured recombinant p42 (lane 4) and p51 (lane 6) isoforms were preincubated 30 min at room temperature with 25 ng unlabelled DNA probe. Subsequently, 1 ng of ^{32}P-labelled DNA probe was added followed by an additional 60 min of incubation. The control CEM nuclear extract (lane 8), control renatured ETS1 (lane 1) or 1 ng of renatured recombinant p42 (lane 3) and p51 (lane 5) isoforms were preincubed 30 min without any additions before the 60 min incubation with 1 ng of ^{32}P-labelled probe. The EMSA gels were then run as described in section 2.

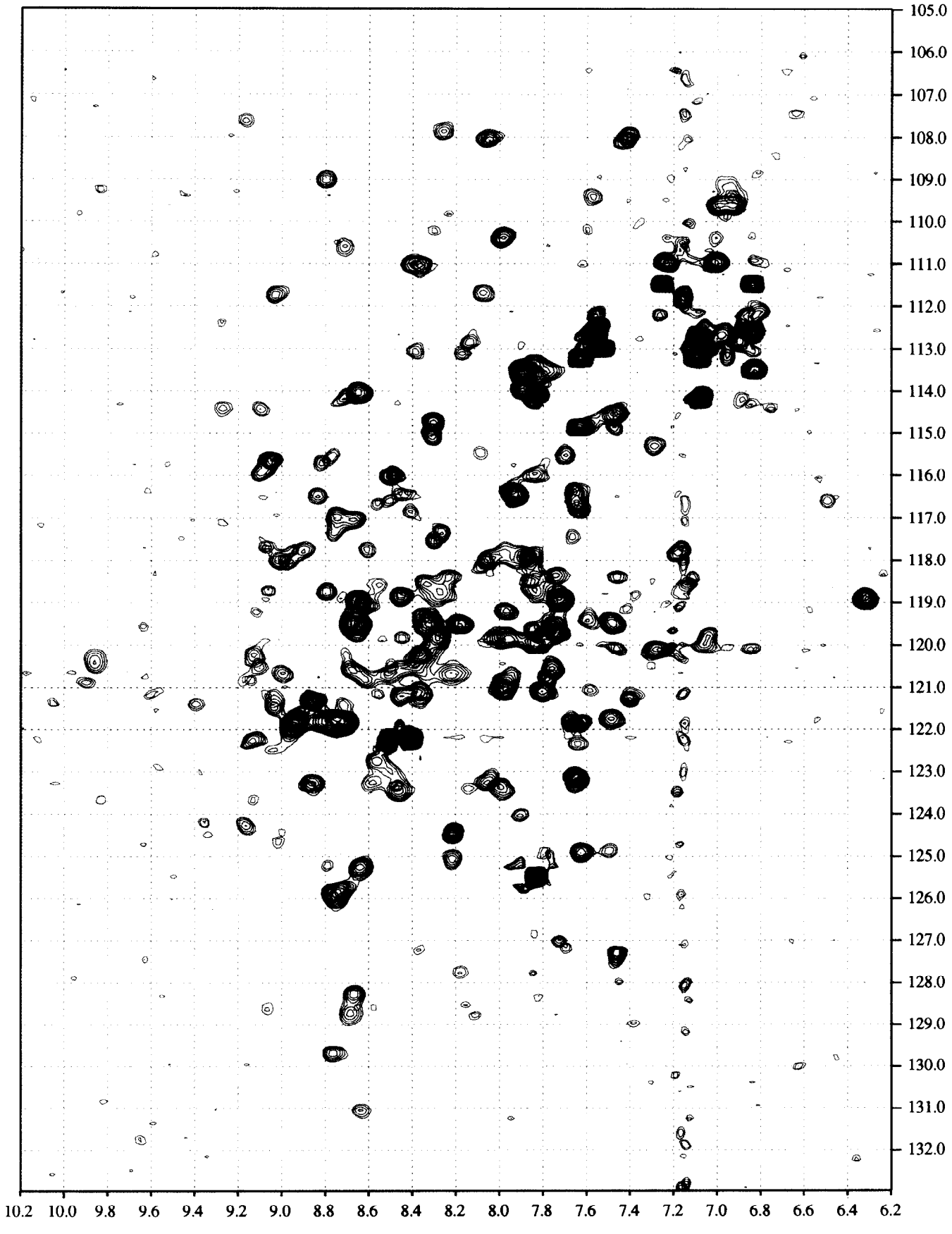
3.2. Refolding of RNase A

To confirm that the protein conformation of the refolded proteins is indeed the same as that of the native protein, we applied this technology to RNase and qualitatively analyzed the protein conformation by NMR spectroscopy. Comparison of NOESY spectra [5] from native and refolded protein demonstrates results obtained by the gel filtration technique (Fig. 2); the pattern and intensity of peaks in the refolded protein spectrum (red) is indistinguishable from that of the native protein spectrum (black) (note that the offset in superposition has been introduced for clarity). The similarity in the NOESY peak pattern and intensities clearly demon-

→

Fig. 2. ^{1}H NOESY spectrum of native and refolded ribonuclease A. The 500 MHz NOESY spectrum shows a portion of the crosspeaks between backbone amide and side-chain protons. Native RNase is shown in black and refolded RNase is shown in red. The uniform shift in the refolded spectrum has been introduced for ease of inspection.


0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.0 4.2 4.4 4.6 4.8 5.0 5.2 5.4 5.6 5.8 6.0 6.2
6.6 6.8 7.0 7.2 7.4 7.6 7.8 8.0 8.2 8.4 8.6 8.8 9.0 9.2 9.4 9.6 9.8


105.0
106.0
107.0
108.0
109.0
110.0
111.0
112.0
113.0
114.0
115.0
116.0
117.0
118.0
119.0
120.0
121.0
122.0
123.0
124.0
125.0
126.0
127.0
128.0
129.0
130.0
131.0
132.0
10.2 10.0 9.8 9.6 9.4 9.2 9.0 8.8 8.6 8.4 8.2 8.0 7.8 7.6 7.4 7.2 7.0 6.8 6.6 6.4 6.2

strates that the refolded protein conformation is basically identical to that of the native protein. Recovery of the refolded protein was ≥ 90% of the denatured protein loaded onto the column.

3.3. Refolding of IHF

Structural analysis by NMR spectroscopy of macromolecular complexes requires the preparation of 1–2 mM complex samples in a final volume of 0.5 ml. In our experience, formation of the proper complex frequently requires diluting the component molecules to 10–50 μM, then mixing and incubating for several minutes to hours, followed by concentration. An easier route involves mixing the denatured protein with its cognate target and refolding by controlled dilution on the column. This is particularly useful if one of the components requires folding from the denatured state. IHF is a moderately sized basic DNA binding protein of *E. coli* with a minimum binding site of 30 bp [6]. IHF is comprised of two subunits, α and β, each of which has a molecular weight of approximately 11 kDa. To simplify data analysis of the 40 kDa complex of IHF bound to a natural DNA target, we prepared IHF in which only the a subunit was labelled with ^{15}N. In this way, is it possible to distinguish the signals of the two subunits by selectively filtering for those signals that arise from nuclei directly bonded to ^{15}N [7]. To this end, we prepared ^{15}N-labelled α-subunits and natural abundance β-subunits by denaturation of purified IHF (labelled and unlabelled) in 6 M GdmCl and separation of the protein subunits by reversed-phase HPLC. The isolated subunits were refolded in the presence of a binding site oligonucleotide, as described in section 2. Fig. 3 illustrates that the oligomeric protein can be successfully refolded and assembled on the specific binding site without preparing and mixing separate dilute solutions of all three components. The pattern of peaks for α-labelled protein (red) is nearly identical to the corresponding peaks in the uniformly labelled native protein (black). Recovery of the folded and assembled complex was 60% of the amount loaded onto the column.

The above examples show that gel filtration permits the preparation of a large quantity of protein under the desired buffer condition(s) with a minimum of dilution and other sample handling steps. We have also tested the method with smaller columns, such as a Pharmacia Superose HR12 (20 ml column volume, data not shown) and have had similar success with protein refolding, albeit with reduced loading capacity (e.g. only 1–2 mg of rEts-1). Moreover, gel filtration also permits the formation of a large quantity of macromolecular complexes with little more effort than dissolving lyophilized protein and its cognate target into a denaturation buffer. This is perhaps the greatest promise of the technique, for it greatly reduces the number of sample handling steps for preparing large quantitites of macromolecular complexes suitable for structural analysis.

One posible explanation for the relatively high success rate using gel filtration may be that refolding and association on the column occurs under essentially irreversible conditions, thereby circumventing a major problem in refolding experiments, namely kinetic competition between folding and aggregation. The folded, native protien is essentially removed from the equilibrium due to its different flow characteristics on the column.

While the recoveries of renatured protein are as good as, or often better than, more traditional gradual dialysis techniques, they do vary from protein to protein. Recently, we have begun to prepare complexes of the DNA-binding domain of rETS1 and have found that different constructs result in recoveries of the renatured domain that range from 30% to 70%. Thus, while we find the technique to be of general importance, as with many techniques, we expect that the method will be more successful with some proteins than with others. The limiting factor seems to be the solubility of the folding intermediates. As the denaturant is diluted away, the partially renatured protein may aggregate and/or crosslink (if many cysteines are present), resulting in significant losses. While gel filtration provides an advantage in that these undesirable forms would be separated from the desired form due to the differences in molecular weight or flow characteristics, these other forms can also block the column when they form in high quantities (depending on the exlusion limit of the resin being used and the bead size of the resin). Thus, it is advisable to test the solubility of the protein under several conditions prior to trying a large scale preparation on a large column. Nonetheless, with these precautions in mind, we find the method to be widely applicable to protein refolding, and, more importantly, to assembly of oliogmeric proteins or macromolecular complexes in large quantities.

Acknowledgements: This work was supported by the AIDS Targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health (to G.M.C. and A.M.G.) The contents of this publication do not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

←

Fig. 3. ^{15}N-^{1}H HSQC spectrum of native and refolded integration host factor complexed with a 30 bp oligonucleotide. Native IHF complex is shown in black, the reconstituted complex is shown in red. Samples were prepared as described in section 2. Only half of the peaks are observed in the refolded spectrum due to only one subunit being labelled with ^{15}N.

References

[1] Schein, C.H. (1993) Curr. Opin. Biotechnol. 4, 456–461.
[2] Nash, H.A., Robertson, C.A., Flamm, E., Weisberg, R.A. and Miller, H.I. (1987) J. Bacteriol. 169, 4124–4127.
[3] Fisher, R.J., Fivash, M., Casas-Finet, J., Erickson, J.W., Kondoh, A., Bladen, S.V., Fisher, C., Watson, D. and Papas, T. (1994) Prot. Sci. 3.
[4] Fisher, R.J., Koizumi, S., Kondoh, A., Mariano, A., Mavrothalassitis, G., Bhat, N.K. and Papas, T.S. (1992) J. Biol. Chem. 267, 17957–17965.
[5] Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley, New York.
[6] Nash, H.A. (1990) Trends Biochem. Sci. 15, 222–227.
[7] Bax, A., Ikura, M., Kay, L.E., Torchia, D.A. and Tschudin, R. (1990) J. Magn. Reson. 86, 304–318.