Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography B, 786 (2003) 247–254
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

A novel strategy for the purification of recombinantly expressed unstructured protein domains

Christoph Kalthoff*

Department of Cell Biology, Center of Anatomy, Hannover Medical School, D-30625 Hannover, Germany

Abstract

We recently found that the larger parts of the endocytic proteins epsin 1 and AP180 consist of an unstructured polypeptide chain. As a result these segments are completely heat-stable without loss of their functional properties. We have taken advantage of this fact and developed a combined heat lysis and pre-purification procedure after expressing the disordered domains in *E. coli*. This results in the irreversible denaturation and precipitation of the majority of bacterial proteins. The bacteria are resuspended in a non-denaturing buffer, heated in a boiling water bath and shock-cooled. We demonstrate that this procedure compared to conventional lysis improves both yield and quality of the purified protein.

Keywords: Purification; Recombinant expression; Intrinsically disordered proteins; Epsin 1; AP180

1. Introduction

Most standard procedures for the purification of recombinantly expressed proteins from bacteria are designed to meet the demands of well-structured polypeptides. Conditions are kept as physiological as possible in order not to compromise the native fold and function of the desired product. After all, proteins are generally assumed to be structured entities even when no information about their secondary and tertiary structure is available.

Recently, we found that segments encompassing about 75% of the endocytic proteins epsin 1 and AP180 are intrinsically disordered. Both are implicated in the formation of plasma membrane-derived clathrin-coated vesicles, which play important roles in membrane trafficking events such as receptor-mediated endocytosis or the recycling of synaptic vesicles (recently reviewed in Ref. [1]). Epsin 1 and AP180 are homologous at their amino termini where they have tightly folded globular α-helical epsin N-terminal homology (ENTH-) domains. They are connected to long, flexible polypeptide chains, which encompass about 75% of the amino acids and lack a regular secondary or tertiary structure under physiological conditions [2].

A literature search reveals that epsin 1 and AP180 are in good company. A growing number of publications reports the identification of intrinsically unstructured proteins and protein segments especially in higher eukaryotes (reviewed in Refs. [3–5]). Most of them are characterized by a combination of low hydrophobicity and high net charge [6]. The following examples show the variety of cellular functions that such proteins are involved in.

First of all, apart from epsin 1 and AP180 other proteins involved in endocytic processes have disordered domains. For example, the light chains of

*Tel.: +49-511-532-3912; fax: +49-511-532-3903.
E-mail address: christoph.kalthoff@zellbiologie.mh-hannover.de (C. Kalthoff).

1570-0232/02/$ – see front matter
doi:10.1016/S1570-0232(02)00908-X

clathrin have been shown to be completely heat stable and are suspected to be at least partly unstructured [7,8]. Furthermore, the cytoplasmic domain of E-cadherin, a transmembrane protein involved in the formation of cell–cell contacts in adherent junctions, is intrinsically disordered but gains structure upon binding to its interaction partner β-catenin [9,10]. Also large parts of the GTPase binding domains (GBD) of the members of the Wiscott–Aldrich syndrome protein (WASP) family are natively unfolded in an aqueous environment but form various α-helical structures when associated either with an intramolecular domain in an autoinhibitory complex or with heterologous activating GTPases [11,12]. The microtubule associated protein tau is an intrinsically disordered protein [13]. In Alzheimer's disease mutated tau protein tends to form β-structures and subsequently polymerizes into the paired helical filaments (PHF) that are suspected to be involved in the development of the neurodegenerative disorders [14]. Also in nucleic acid binding proteins segments without structure are widely spread. As a prominent example the p160 receptor coactivator and the transcriptional coactivator CBP/p300, which are both intrinsically disordered, fold synergistically upon interaction to form a nuclear hormone receptor and transcription activator complex [15].

Interestingly, natively unfolded proteins or protein domains are especially abundant in higher eukaryotes. Recently, several authors have suggested criteria and algorithms upon which it can be judged from the amino acid sequence of a polypeptide chain whether it is unstructured or not and applied these criteria to proteins with unknown structure [5,6]. One of the groups has analyzed the amino acid compositions of the open reading frames of 31 genomes in a systematic study. They predict 41% of all sequences in *Drosophila melanogaster* to contain unstructured segments of 50 amino acids or more and 17% to be wholly disordered. In contrast to that only 2% of the proteins in *E.coli* are predicted to be completely unfolded [5].

A problem upon examining unstructured proteins is that they are particularly sensitive to protease degradation, because the whole polypeptide chains are exposed to the solvent and not protected by a compact fold. Therefore they are easily accessible for proteases. This is especially deplorable since the unstructured regions often contain important properties of the proteins. In the case of epsin 1 and AP180 all known protein–protein interaction sites are located within the unstructured segments. For instance they are able to bind clathrin, the major structural component of clathrin-coated vesicles, and promote the assembly of single clathrin molecules to cage-like structures. So far this is regarded as the most important functional property of both proteins [2].

On the other hand, the extremely high solubility of disordered polypeptides offers opportunities to develop novel strategies for their purification. While almost all folded proteins are sensitive to irreversible heat denaturation, their unstructured counterparts are normally heat stable. This is because upon irreversible denaturation, hydrophobic patches, which are normally hidden in the inside of the folded structure, are exposed to the solvent, enter into uncontrolled inter- and intramolecular interactions, and finally form insoluble aggregates. The high content of hydrophilic and charged amino acid residues keeps unstructured proteins soluble even at high temperatures. For example, it was previously reported that the light chains of clathrin can be purified from a preparation of clathrin triskelia by heat treatment [7,8]. Indeed we have also found that the unstructured parts of epsin 1 and AP180 are stable at 100 °C without loss of their known functional and binding properties [2].

Now we have taken advantage of this fact and developed a heat lysis and prepurification step after expressing the unstructured parts of epsin 1 and AP180 in *E. coli*. This procedure depletes the vast majority of bacterial proteins and therefore also the protease activity before entering the actual purification procedure.

2. Experimental

2.1. Reagents

The affinity purified polyclonal peptide antibody against epsin 1 [2] and the monoclonal antibody mAbAP180.1 against AP180 [16] were described previously. As secondary antibodies we used horseradish peroxidase coupled goat sera against mouse or

rabbit IgG from ICN/Cappel (Aurora, OH, USA). In Western blots, the signals were developed using ECL reagent and films (Amersham Pharmacia Biotech, Freiburg, Germany).

If not stated otherwise all other reagents were purchased from Sigma (Taufkirchen, Germany) in the highest quality available.

2.2. Cloning

The unstructured regions of epsin 1 (amino acids 144–575) and AP180 (amino acids 328–896) were cloned into pQE30 series vectors (Qiagen, Hilden, Germany) as described previously [2] (Fig. 1).

2.3. Expression

All media and solutions were prepared as described [17]. The expression vectors with the inserts were transformed into *E. coli* BL21 CodonPlus (Stratagene, LaJolla, CA, USA). An overnight culture of the bacteria was diluted 1:100 in LB medium containing 50 μg/ml each of ampicillin and chloramphenicol. The cells were grown at 37 °C to an optical density of $A_{600\ nm}$~0.5. Protein expression was induced using 0.5 m*M* isopropylthiogalactoside and performed for 3 h at 25 °C. The bacteria were cooled on ice and washed twice by pelleting them for 15 min at 6000 *g* in a JA-10 rotor (Beckman Coulter, Fullerton, CA, USA) and resuspending them in ice-cold phosphate buffered saline (PBS, 137 m*M* NaCl, 2.7 m*M* KCl, 1.9 m*M* KH_2PO_4, 8.2 m*M* Na_2HPO_4). The bacteria were finally transferred into a 50-ml polypropylene tube (Greiner, Frickenhausen, Germany), pelleted by centrifugation for 20 min at 3300 *g* in a 1.0R centrifuge (Heraeus Instruments, Hanau, Germany), shock-frozen in ethanol with dry-ice and stored at −80 °C until further use. Using this procedure a 1-l culture yielded bacterial pellets with a wet weight of ~3 g.

2.4. Lysis and pre-purification

The bacterial pellets were thawed and resuspended rapidly in 10 ml of lysis buffer (50 m*M* sodium phosphate buffer, 300 m*M* NaCl, 10 m*M* imidazole) under addition of 3 m*M* β-mercaptoethanol. For a pellet obtained from a 1-l culture of bacteria, 10 ml of lysis buffer were used. As soon as a homogeneous suspension was obtained the tube was sealed, placed in a boiling water bath and incubated for 5 min under continuous agitation. Then it was immediately transferred into blended NaCl and ice-water with a temperature of −10 °C where it again was incubated for 5 min under agitation. The bacterial debris and the formed precipitates were pelleted by centrifugation for 15 min at 117,000 *g* in a Type 70Ti rotor (Beckman Coulter). An illustrated flow chart of this procedure is shown in Fig. 2.

A non-denaturing preparation was performed according to the suggestions of the manufacturer of the affinity purification matrix (Qiagen) and served as a control. The bacteria were thawed and resuspended rapidly in lysis buffer. For a pellet resulting from a 1-l culture of bacteria, 9 ml of lysis buffer were used. When the suspension was homogeneous, another 1 ml of lysis buffer with 10% Triton X-100 was added. Lysis was performed for 10 min on ice. Then the lysate was sonicated three times for 10 s using a Sonifier 250C (Branson, Hartford, CT, USA) with an

Fig. 1. Schematic view of the domain structures of epsin 1 and AP180 as well as the recombinantly expressed fragments used in this work. Both full-length proteins have an epsin N-terminal homology (ENTH-) domain at their amino terminal ends. The rest of the proteins is unstructured. The fusion proteins with the fragments encompass most of the disordered parts of both proteins.

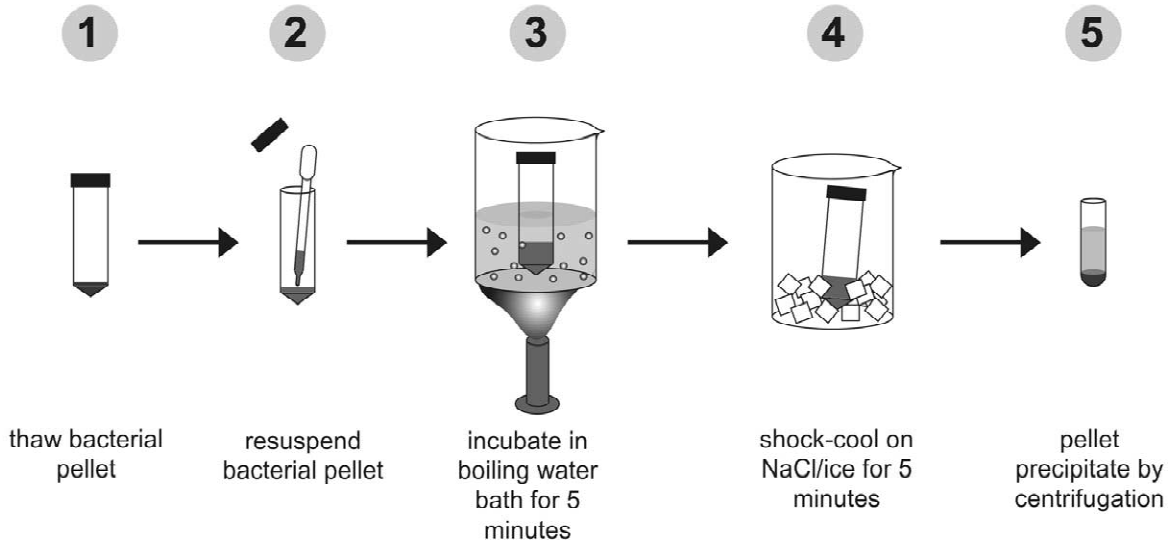

Fig. 2. Flow chart of the lysis procedure. The bacterial pellet is thawed, resuspended in lysis buffer and incubated in a boiling water-bath for 5 min. After that it is immediately shock-cooled in a NaCl/ice-water slurry. Finally the lysate is cleared by centrifugation.

output of 40 W. Finally it was clarified by centrifugation as described above for the heat lysis.

For better comparability of further purification, the buffer system recommended by Qiagen was chosen for both lysis strategies.

2.5. Affinity purification

The 6×His-tagged fusion proteins were purified from the obtained lysates using a Ni-NTA-Agarose matrix (Qiagen). In preliminary tests the maximum capacity of Ni-NTA-Agarose for 6×His-epsin 1-(144–575) and 6×His-AP180-(328–896) was determined to be 4 and 2 mg/ml compact gel, respectively. To determine the quality of the preparations, a lysate prepared from 1 l of bacterial culture was incubated with 1.25 ml of Ni-NTA-Agarose slurry. To determine the maximum yield of a preparation an excess of matrix was used. The Ni-NTA-Agarose was incubated with the lysates for 1 h at 4 °C under continuous rotation. Then the slurry was transferred into an empty gravity flow column and washed with 10 bed volumes of washing buffer (lysis buffer containing 20 m*M* imidazole instead of 10 m*M*). Elution was performed with elution buffer (lysis buffer with 250 m*M* imidazole). The eluate was collected in 0.5-ml fractions, which were checked for protein content. The appropriate fractions were combined.

2.6. Analysis of the purification

From all steps of the purification, samples were taken and adjusted to comparable volumes. Only the final eluates were diluted to an identical protein concentration. All samples were analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and subsequent Coomassie staining as well as Western blot analysis using antibodies against epsin 1 and AP180. The stained gels and developed ECL-films were digitalized using a Duoscan scanner (Agfa, Mortsel, Belgium). Further processing of the images was performed using Photoshop 5.0 software (Adobe Systems, San Jose, CA, USA). The intensity of the bands was measured densitometrically with NIH-Image 1.62 software. For determination of the yield, $A_{280\ nm}$ of the preparations was measured with a DU-640 photometer (Beckman Coulter). The protein content was calculated using an extinction coefficient computed from the amino acid composition with Protean 1.17 software (DNAStar, Madison, WI, USA).

3. Results

For this study we have used previously described constructs [2] with the coding sequences of the disordered regions of epsin 1 and AP180 cloned into pQE30 series vectors (Fig. 1). They provide an amino-terminal fusion tag with a 6×His-cluster for easy purification on metal chelate affinity matrices. The advantage of these vectors is that the tag is very short (10 translated amino acids outside the multiple cloning site) and therefore does not interfere with the heat stability of the protein of interest.

First approaches to benefit from the high solubility and heat stability of disordered proteins for preparative purposes revealed that it is important to remove contaminating proteins as early as possible in the purification procedure, in order to obtain products with the highest achievable purity (not shown). Therefore, we tried to establish a protocol in which heat treatment of the samples is not only used to deplete protein contaminations but rather also to disrupt the bacterial cells. We supposed combining both purposes in one single step would not only improve the quality of the obtained products but rather also simplify the protocol and make the whole procedure very efficient.

3.1. Lysis

To test this conjecture, we incubated a suspension of the washed bacteria in a boiling water bath for 5 min and shock-cooled it afterwards in a slurry of NaCl and icewater (see Fig. 2 for a flow chart). Indeed, this combined lysis and prepurification step was not only fast and easy but also very effective. At least 80% of the total protein content was removed while 72% of the 6×His-epsin 1-(144–575) and 70% of the 6×His-AP180-(328–896) remained in the supernatant (Fig. 3A). In contrast to that, a lysis under native conditions using Triton X-100 and sonication to break up the bacteria released only 29% of the 6×His-epsin 1-(144–575) and 37% of the 6×His-AP180-(328–896). Without any lysis treatment only a negligible amount of protein was found in the supernatant indicating that freezing and thawing alone did not disrupt the bacteria (Fig. 3A and B).

3.2. Purity and affinity purification

After heat treatment the protein content of the lysate is already of astonishing purity. Comparing the SDS-PAGE of the supernatant after clarification of the lysate with the product after affinity purification (see Figs. 3B and 4A) reveals only a small number of additional bands. Apparently besides the elimination of the few heat stable contaminating proteins, the metal chelate chromatography serves mostly to remove the other components of the bacterial lysate, such as nucleic acids, peptides and salts.

Fig. 4A demonstrates that after heat treatment the purity of the final products is significantly increased. Especially 6×His-AP180-(328–896) is of remarkable purity. Most obviously contaminating bands at around 70 kDa are depleted after heat treatment. Even more importantly the share of the desired products in the protein content of the final preparations is considerably higher. Densitometrical analysis of the SDS-PAGE of the eluates from Ni-NTA-Agarose revealed that compared to the purification under non-denaturing conditions 6×His-AP180-(328–896) is 29% and 6×His-epsin 1-(144–575) is 38% enriched after heat treatment. This is presumably either due to improved release of the products from the bacteria upon heat lysis and/or reduced proteolysis after heat denaturation of contaminating proteases.

3.3. Yield

The amount of protein eluted from the affinity purification after heat treatment is 49% higher for 6×His-epsin 1-(144–575) and 21% higher for 6×His-AP180-(328–896) compared to the purification under non-denaturing conditions (see Fig. 4B). Combined with the improved purity of the preparations the recovery of the desired products was increased by 56% and 106%, respectively.

4. Discussion

The presented results show that heat-stable recombinant proteins expressed in bacteria can readily be purified in large yields in a single step by simply

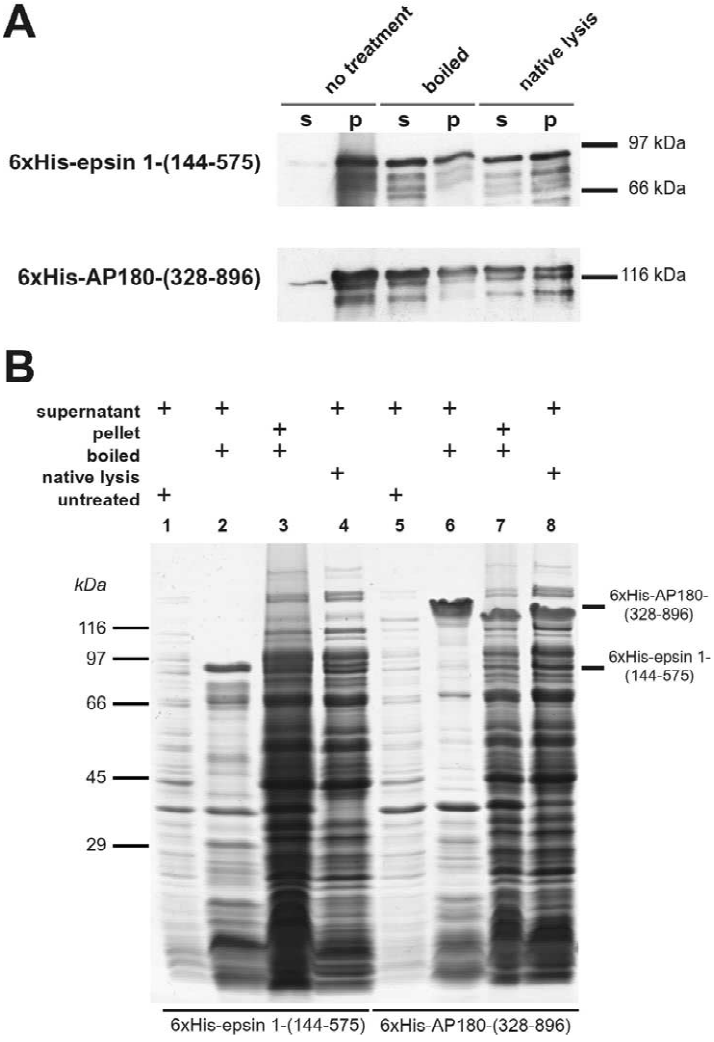

Fig. 3. Comparison of the effectivity of the heat lysis with a conventional lysis procedure under non-denaturing conditions. Bacterial suspension was centrifuged either after no treatment, boiling or native lysis using Triton X-100 and sonication. (A) Western blot of the supernatants (s) and pellets (p) after centrifugation using antibodies against the proteins indicated on the left. While almost no fusion protein is released directly after thawing the bacterial pellet, boiling sets free about 72% of the epsin 1 and 70% of the AP180 fragment. In contrast to that, after native lysis only 29% and 37%, respectively, remain in the supernatant. (B) Coomassie-stained SDS-PAGE of selected samples of (A). Without any treatment the supernatants contain only few proteins (lanes 1 and 5) thus the integrity of the bacteria is hardly influenced by freezing and thawing. After heat treatment the majority of the protein is pelleted by centrifugation (lanes 3 and 7). The lysates are highly enriched of the fusion proteins (lanes 2 and 6) whereas the compositions of the lysates obtained under non-denaturing conditions are much more complex (lanes 4 and 8).

boiling the bacterial suspension. This method is very fast and efficient. In less than 30 min from the moment the bacterial pellet is thawed, a lysate can be obtained from which more than 80% of the contaminating proteins are already removed. This fact is presumably possible due to the previously mentioned low content of intrinsically disordered proteins in procaryotes like *E. coli*.

The technique is suitable for high throughput purification of a series of proteins with similar properties. As already mentioned in the Introduction, many unstructured protein segments contain protein–

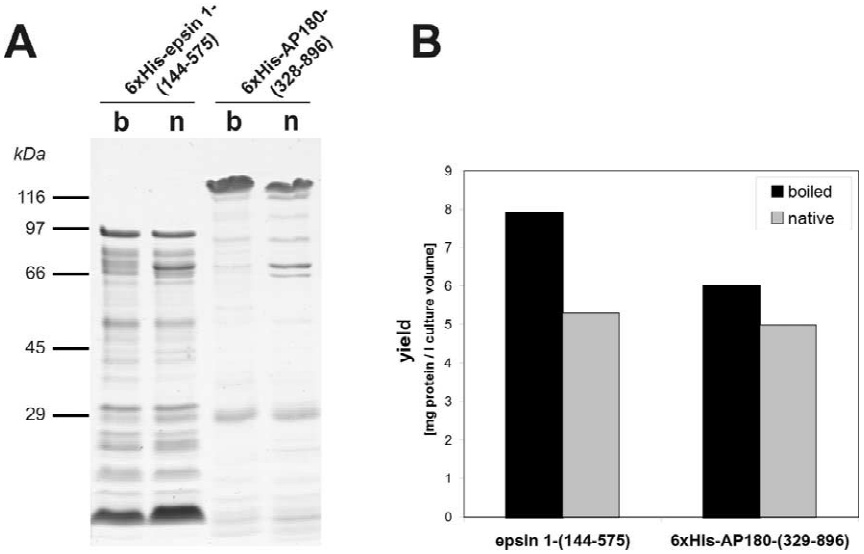

Fig. 4. Comparison of the qualities and yields of the protein preparations after affinity chromatography. (A) Coomassie-stained SDS-PAGE of comparable amounts of protein obtained after boiling (b) and under native conditions (n). Especially in the case of 6×His-AP180-(328–896) the purity of the obtained protein after heat treatment is much better than that obtained after Triton X-100 lysis and sonication. (B) The yields after the different lysis procedures are given as the amounts of total protein eluted from Ni-NTA-Agarose per liter of bacterial culture.

protein interaction sites. For example, the recombinant proteins we used for this study retain the known binding properties of the full-length proteins. To identify their interaction sites, possible motifs can be mutated systematically by site-directed mutagenesis. In such a study, numerous altered proteins need to be purified. If the purification procedure is very time-consuming it can easily be the limiting step in such a study. By having a method like the one we present here at hand, which allows the purification of a product with a higher purity and a greater yield in reduced time, the efficiency of such a study is significantly improved. This is especially the case when the number of necessary purification steps is diminished.

Furthermore, using the heat treatment step not only simplifies and accelerates the purification, it also increases the yield and, what is probably more important, it improves the purity of the preparations. We assume that the increased purity on the one hand is a result of the irreversible denaturation of bacterial proteins containing His-clusters, which would otherwise co-purify upon metal chelate chromatography. On the other hand, it is conceivable that the heat-denaturation of proteases prevents protein degradation during purification and therefore increases the yield and purity of the product in the final eluate. Especially, trace contaminations of proteases can be a problem because unstructured protein domains are particularly sensitive to degradation. This can reduce the quality of the preparation especially upon storage or incubation in experiments. Our experience shows that after heat treatment the long-time stability of the prepared proteins is significantly improved.

Unfortunately one caveat remains. It cannot be excluded that the extraordinary conditions during the heat treatment chemically alter the nature of the amino acid residues. Care should be taken that the solution is well buffered also at higher temperatures (e.g. avoid volatile buffer components) and contains thiol-reagents such as DTT or β-mercaptoethanol to maintain reducing conditions. For the proteins we used here, we compared properties such as SDS-PAGE mobility, UV-spectra as well as functional and binding capabilities and found no differences between samples that were heat treated and those which were not [2]. We recommend to perform a similar initial characterization with other proteins as well, before this method is adopted.

With these precautions in mind we think that the

presented method can be a very useful tool for studying the properties of a great variety of intrinsically unstructured polypeptides.

Acknowledgements

Ernst Ungewickell is thanked for discussions and Christiane Lemke for technical assistance. This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and the Fond der Chemischen Industrie.

References

[1] K. Takei, V. Haucke, Trends Cell Biol. 11 (2001) 385.
[2] C. Kalthoff, J. Alves, C. Urbanke, R. Knorr, E.J. Ungewickell, J. Biol. Chem. 277 (2002) 8209.
[3] H.J. Dyson, P.E. Wright, Curr. Opin. Struct. Biol. 12 (2002) 54.
[4] V.N. Uversky, Eur. J. Biochem. 269 (2002) 2.
[5] A.K. Dunker, J.D. Lawson, C.J. Brown, R.M. Williams, P. Romero, J.S. Oh, C.J. Oldfield, A.M. Campen, C.M. Ratliff, K.W. Hipps, J. Ausio, M.S. Nissen, R. Reeves, C. Kang, C.R. Kissinger, R.W. Bailey, M.D. Griswold, W. Chiu, E.C. Garner, Z. Obradovic, J. Mol. Graph Model 19 (2001) 26.
[6] V.N. Uversky, J.R. Gillespie, A.L. Fink, Proteins 41 (2000) 415.
[7] M.P. Lisanti, L.S. Shapiro, N. Moskowitz, E.L. Hua, S. Puszkin, W. Schook, Eur. J. Biochem. 125 (1982) 463.
[8] F.M. Brodsky, N.J. Holmes, P. Parham, J. Cell Biol. 96 (1983) 911.
[9] A.H. Huber, W.I. Weis, Cell 105 (2001) 391.
[10] A.H. Huber, D.B. Stewart, D.V. Laurents, W.J. Nelson, W.I. Weis, J. Biol. Chem. 276 (2001) 12301.
[11] M.G. Rudolph, P. Bayer, A. Abo, J. Kuhlmann, I.R. Vetter, A. Wittinghofer, J. Biol. Chem. 273 (1998) 18067.
[12] A.S. Kim, L.T. Kakalis, N. Abdul-Manan, G.A. Liu, M.K. Rosen, Nature 404 (2000) 151.
[13] O. Schweers, E. Schonbrunn-Hanebeck, A. Marx, E. Mandelkow, J. Biol. Chem. 269 (1994) 24290.
[14] M. von Bergen, S. Barghorn, L. Li, A. Marx, J. Biernat, E.M. Mandelkow, E. Mandelkow, J. Biol. Chem. 276 (2001) 48165.
[15] S.J. Demarest, M. Martinez-Yamout, J. Chung, H. Chen, W. Xu, H.J. Dyson, R.M. Evans, P.E. Wright, Nature 415 (2002) 549.
[16] S. Ahle, E. Ungewickell, EMBO J. 5 (1986) 3143.
[17] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY, 1989.