VIP Protein Purification

DOI: 10.1002/anie.201108747

The Most Stable Protein–Ligand Complex: Applications for One-Step Affinity Purification and Identification of Protein Assemblies**

*Christoph Giese, Franziska Zosel, Chasper Puorger, and Rudi Glockshuber**

The purification of protein complexes and large-scale investigations of protein–protein interaction networks have been greatly facilitated through the development of a number of affinity tags such as the cmyc, FLAG, and His_6 tags.[2] However, all of the currently available affinity purification systems suffer from dynamic binding equilibria and measurable dissociation rate constants which enable the competitive elution of bound target proteins when an excess of a suitable free ligand is present.[3] This circumstance hampers the quantitative and cost-effective isolation of low-abundance protein complexes. Here, we introduce a new affinity purification system that is derived from type 1 pili of *E. coli*. Type 1 pili are rigid, filamentous supramolecular protein complexes which are anchored in the cell's outer membrane and extend into the extracellular space.[4] They are composed of four structural protein subunits termed FimH, FimG, FimF, and FimA.[4a] In the assembled pilus, these subunits interact noncovalently by a mechanism called donor strand complementation, where the incomplete, immunoglobulin-like fold of one subunit is completed by an N-terminal extension, termed donor strand, of the successive subunit.[5] The complex between FimGt, an N-terminally truncated variant of FimG lacking its own donor strand, and a peptide corresponding to the donor strand of the neighboring subunit FimF (DsF) was found to be the kinetically most stable protein–ligand complex known to date (Figure 1).[5d]

Here, we establish the FimGt/DsF system for use in the affinity purification of heterooligomeric protein complexes from cell extracts. Utilizing the donor strand of FimF as the affinity tag (termed DsF tag) and FimGt as the binding partner, we demonstrate the one-step purification of DsF-tagged *E. coli* tryptophan synthase, a heterotetrameric αββα complex of low cellular abundance. We compare the performance of the DsF tag to that of other commonly used affinity tags and find that, in agreement with the high kinetic stability of the FimGt/DsF complex, enrichment of the tryptophan synthase complex is most efficient for the DsF tag. This result suggests that the DsF tag is most suitable not only for the isolation of low-abundance protein complexes but presumably also for many other technical applications such as, for example, the functional and permanent immobilization of DsF-tagged proteins on surfaces and their detection in cells and on Western blot membranes.

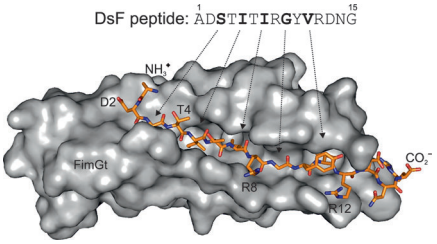

Figure 1. Crystal structure of the FimGt/DsF complex (3BFQ.pdb).[5d] FimGt is shown as a gray surface, the DsF peptide as stick representation. Residues of DsF that point towards FimGt are in bold and their position in the structure is indicated by arrows.

As a prerequisite for the technical application of the FimGt/DsF system we first optimized the production of FimGt by producing it in the cytoplasm of *E. coli* BL21(DE3) cells in the form of inclusion bodies (Figure S1 in the Supporting Information). After solubilization of the inclusion bodies, oxidative refolding, and purification of FimGt by conventional chromatographic techniques, the final yield of purified FimGt was 35 mg per liter of bacterial culture—sufficient amounts for large-scale applications of the FimGt/DsF system. Quantitative formation of the single, structural disulfide bond was verified by the Ellman assay.[6] The identity of FimGt was confirmed by ESI-MS (expected/measured mass: 13656.9/13657.0 Da).

To gain mechanistic insight into the binding reaction between DsF and FimGt, association kinetics were measured for DsF concentrations of 5, 10, 25 and 50 μM while the FimGt concentration was kept constant at 5 μM (Figure 2a). The reaction rates were dependent on the DsF concentration, indicating that binding of DsF is the rate-limiting step of complex formation. The data were globally fit according to an irreversible second-order reaction and yielded a rate constant of association of $(330 \pm 8.9)\,M^{-1}s^{-1}$. Although relatively slow, the binding of DsF to FimGt is fast enough to allow for technical applications on reasonable time scales. We determined the rate constant for spontaneous dissociation/unfolding of the FimGt/DsF complex at pH 8.0 and 25 °C to be

[*] C. Giese, Dr. C. Puorger, Prof. R. Glockshuber
Institut für Molekularbiologie und Biophysik, ETH Zürich
Schafmattstrasse 20, 8093 Zürich (Switzerland)
E-mail: rudi@mol.biol.ethz.ch
F. Zosel
Biochemisches Institut, Universität Zürich
Winterthurerstrasse 190, 8057 Zürich (Switzerland)

[**] We thank B. Bukau (Universität Heidelberg) for the kind gift of *E. coli* strain MC4100 Δ*rplW::kan* [pL23], D. Waugh (National Institutes of Health) for the plasmid pRK793_TEV,[1] and D. Hilvert (ETH Zurich) for the plasmid pAC-PTetT7-HIV. This work was supported by the Swiss National Science Foundation [grants 3100A0-100787 and 310030B_138657/1 (R.G.)] and the Swiss Federal Institute of Technology Zurich within the framework of the National Center for Competence in Research Structural Biology Program.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201108747.

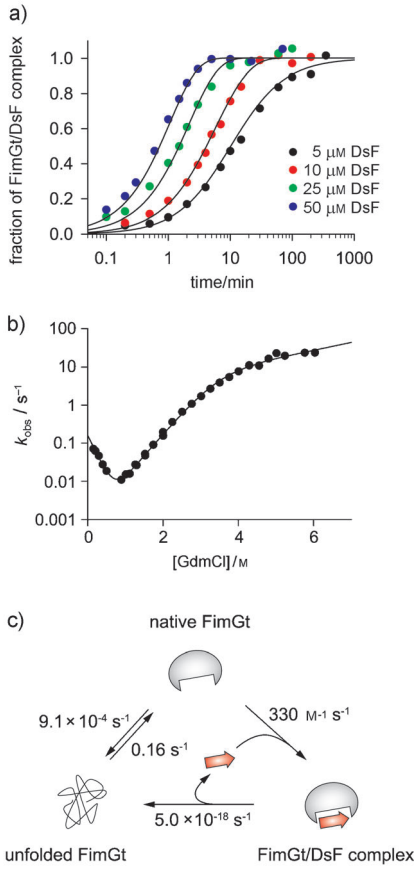

Figure 2. Binding of the DsF peptide to FimGt. a) Binding kinetics of DsF to FimGt (5 μM) at pH 8.0, 25 °C. The solid line is a global fit according to a second-order reaction. b) Kinetic stability of FimGt at pH 8.0, 25 °C. The solid line is a fit according to a three-state model with a metastable on-pathway intermediate. c) Minimal reaction mechanism for the binding of DsF to FimGt. The DsF peptide binds exclusively to native FimGt.

$(5.0 \pm 0.17) \times 10^{-18}\ s^{-1}$ (dissociation half-life: approximately 4 billion years; Figure S2 in the Supporting Information) which is, within experimental error, identical to the value obtained earlier at pH 7.0.[5d] This translates into a dissociation constant (K_D) at pH 8.0 of $(1.5 \pm 0.066) \times 10^{-20}$ M, corresponding to a binding energy of (-113 ± 0.11) kJ mol^{-1}. The FimGt/DsF complex is thus not only the kinetically but, to the best of our knowledge, also the thermodynamically most stable protein ligand complex known to date.

We next determined the kinetic stability of FimGt by measuring the refolding/unfolding kinetics at different concentrations of guanidinium chloride. As refolding of FimGt, when monitored by means of the intrinsic tryptophan fluorescence, includes a spectroscopically silent step,[7] refolding kinetics were obtained by interrupted refolding experiments that directly report on the fraction of native molecules (Figure S3 in the Supporting Information). The half-logarithmic plot of the observed refolding/unfolding rate constants shows a significant deviation from linearity at high denaturant concentrations (Figure 2b). Therefore, the data were fit according to a three-state model with a metastable on-pathway intermediate.[8] From this analysis we obtained the rate constants for the folding and unfolding of FimGt in the absence of denaturant $k_f(H_2O) = (0.16 \pm 0.010)\ s^{-1}$ and $k_u(H_2O) = (9.1 \pm 0.57) \times 10^{-4}\ s^{-1}$, respectively, and the following kinetic *m*-values that define the linear dependence of the logarithm of the rate constant on denaturant concentration: $m_{UI} = d\ln k_{UI}/d[\text{GdmCl}] = (-4.7 \pm 0.18)\ M^{-1}$, $m_{IN} = d\ln k_{IN}/d[\text{GdmCl}] = (-2.2 \pm 0.058)\ M^{-1}$, $m_{NI} = d\ln k_{NI}/d[\text{GdmCl}] = (0.47 \pm 0.016)\ M^{-1}$ (where U denotes the unfolded state, I the intermediate, and N the native state).

With these kinetic values the thermodynamic stability of FimGt can be calculated to be (-12.8 ± 0.2) kJ mol^{-1} and the cooperativity of the equilibrium transition (m_{eq}-value) to be (18 ± 0.47) kJ mol^{-1} M^{-1}, which are in very good agreement with values obtained earlier from equilibrium measurements.[7] The linear dependence of the initial rate of DsF binding to FimGt on DsF concentration (Figure 2a), together with the slow rate of spontaneous FimGt unfolding with a half-life of approximately 13 min ($k_u(H_2O) = (9.1 \pm 0.57) \times 10^{-4}\ s^{-1}$), demonstrates that the DsF peptide directly binds to native FimGt and excludes a mechanism in which unfolding of FimGt would be required for binding. Figure 2c summarizes the resulting DsF binding mechanism, in which only the dissociation of DsF is dependent on unfolding of FimGt. In this context we note that the K_D value for the FimGt/DsF complex calculated above is valid only for concentrations of DsF below 500 μM as otherwise folding of FimGt would become rate-limiting. Moreover, a comparison of $k_u(H_2O)$ of free FimGt to the rate constant of spontaneous dissociation/unfolding of the FimGt/DsF complex shows that donor strand complementation increases the half-life of spontaneous unfolding of FimGt by 14 orders of magnitude.

We further characterized the binding reaction of DsF to FimGt by measuring its NaCl, temperature, and pH dependence (Figure S4 in the Supporting Information). Increasing the concentration of NaCl to 1.0 or 3.0 M led to a twofold and sixfold increase of the rate constant of association, respectively, indicating that complex formation between FimGt and DsF is not primarily driven by electrostatic interactions. Binding at 5 °C occurred three times slower than at 25 °C. From the temperature dependence we obtained the energy of activation to be (34 ± 1.6) kJ mol^{-1}. At pH 4.5 the rate constant of association was ten times higher than at neutral and basic pH. A fit of the data to a single ionization equilibrium gave an apparent pK_a value of 5.6 ± 0.1, suggesting that protonation of either an acidic or a histidine residue is critical for the faster binding of DsF at low pH.

To be able to use the FimGt/DsF system for practical applications such as affinity chromatography, FimGt was covalently linked to succinimidylester-activated sepharose beads through its primary amino groups. Using the purified, DsF-tagged dithiol oxidase DsbA from *E. coli* (22.9 kDa), we determined the dynamic binding capacity of FimGt-sepharose to be 220 nmol or 5 mg of DsbA per milliliter of FimGt-sepharose. To test whether the FimGt/DsF system can be used to purify DsF-tagged proteins and natural binding partners associated with them, the *E. coli* tryptophan synthase complex was chosen; it is composed of two α and β subunits (TrpA and TrpB; αββα tetramer) and catalyzes the last two reactions

in tryptophan biosynthesis. As the complex is of very low abundance in *E. coli* cells,[9] we expected high enrichment factors if the FimGt/DsF system worked efficiently in affinity chromatography. DsF was fused to the N terminus of TrpA and DsF-tagged TrpA was produced at near endogenous levels in *E. coli* BW25113 Δ*trpA768:kan* cells by using the *tetA* promoter (Figure S5a in the Supporting Information). To allow the specific release of DsF-tagged TrpA from FimGt-sepharose under native conditions, a TEV protease cleavage site (TEV = tobacco etch virus) flanked by flexible hexapeptide linkers was introduced between DsF and the N terminus of TrpA (Figure 3). A general scheme for the use of the FimGt/DsF system in affinity chromatography is given in Figure 3b. After incubation of FimGt-sepharose with the soluble cell extract containing DsF-tagged TrpA, the beads were washed to remove impurities and bound TrpA was released by on-column cleavage with His_6-tagged TEV protease. TEV protease was then removed using Ni^{2+}-NTA agarose (NTA = nitrilotriacetate) yielding the equimolar tryptophan synthase complex in the supernatant. Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis displayed two major bands at approximately 28 and 40 kDa, corresponding to TrpA and TrpB, respectively (Figure 4). As a negative control, the experiment was repeated using untagged TrpA and no tryptophan synthase could be detected (Figure S6 in the Supporting Information). Thus, by using DsF-tagged TrpA, we successfully purified the

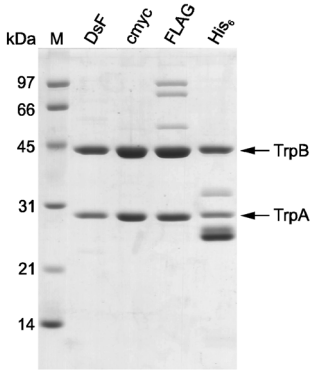

Figure 4. Coomassie-stained 15% polyacrylamide-SDS gel demonstrating the co-purification of the β subunit of tryptophan synthase (TrpB) from cell extracts, using TrpA variants with different affinity tags and the corresponding affinity matrices, followed by TEV protease cleavage and removal of His_6-tagged TEV protease by means of Ni^{2+}-NTA agarose. Of each sample, identical amounts of enzymatic activity (6 mU) were loaded. The highest purity was obtained with DsF- and cmyc-tagged TrpA. Lane M: molecular mass standard.

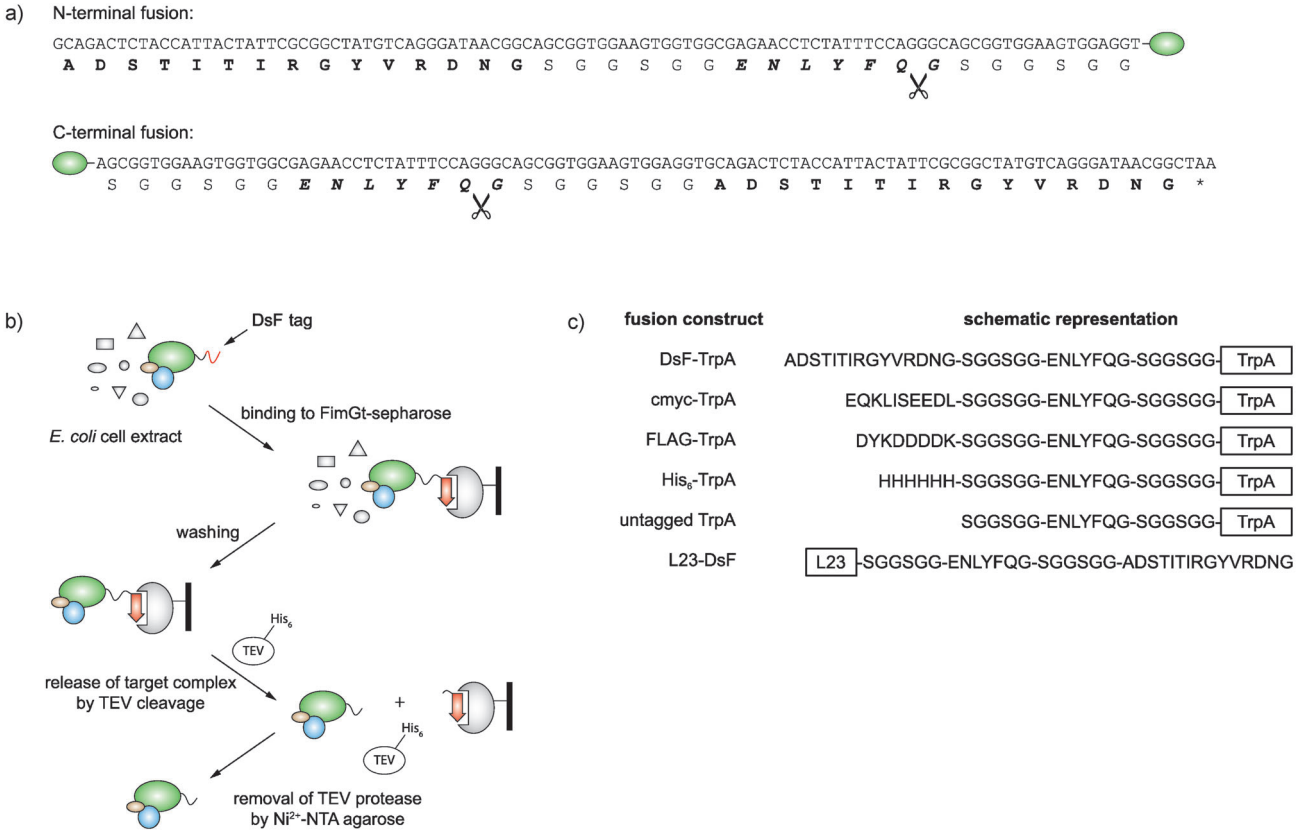

Figure 3. a) Nucleotide and amino acid sequences of the DsF tag as used in this study. The tag can be fused to either the N or C terminus of a target protein. The DsF sequence is highlighted in bold, the TEV protease recognition site is shown in bold and italics, and the cleavable peptide bond is indicated by scissors. To allow access of TEV protease to its cleavage site, two SGGSGG-linker sequences were introduced between the DsF tag and the TEV recognition sequence, and between the TEV site and the N or C terminus of the target protein. The target protein is represented by a green ellipse. b) General workflow for the application of the FimGt/DsF system in affinity chromatography. c) Schematic overview of the protein constructs used in this study for one-step affinity purification of tryptophan synthase (TrpA fusions) or ribosomes (L23 fusion) from *E. coli* cell extracts.

active tryptophan synthase complex from *E. coli* lysate in a single step.

To compare the efficiency of the FimGt/DsF system to that of other commonly used affinity purification systems, we also purified tryptophan synthase using cmyc-, FLAG-, and His_6-tagged TrpA (Figure 3c) and corresponding affinity chromatography materials. Highly pure enzyme preparations were obtained with the DsF and cmyc tags, while the samples obtained using FLAG and His_6 tags contained significant amounts of impurities (Figure 4). To compare the performance of the different tags in a quantitative manner, we calculated the yield and specific activity of purified tryptophan synthase, the enrichment factor, and the overall purification efficiency for each affinity purification system (Figure 5 and Table S1 in the Supporting Information). We find that the DsF tag 1) gives both the highest yield and the highest specific activity of purified enzyme and 2) has a 300-, 5-, and 15-fold higher purification efficiency than the cmyc,

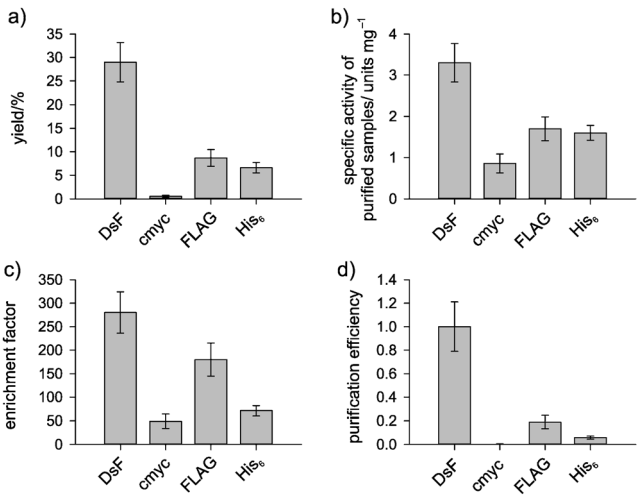

Figure 5. Comparison of different affinity tags regarding the one-step purification of *E. coli* tryptophan synthase. a) Yields of purified tryptophan synthase were calculated by dividing the total activity of tryptophan synthase present in the purified sample by that present in the cell extract. b) Specific activities of purified tryptophan synthase samples were calculated by dividing the enzymatic activity by the protein concentration. c) Enrichment factors were calculated by dividing the specific activity of purified tryptophan synthase by the specific activity of the cell extract. d) Purification efficiencies were obtained by multiplying the yield and enrichment factor and then normalizing. All data are means with standard errors ($n=3$).

was then produced at near endogenous levels in *E. coli* strain MC4100ΔrplW:kan [pL23][10] under conditions at which production of wild-type L23 (encoded on pL23) is reduced to a minimum level (Figure S5b in the Supporting Information). Ribosomes harboring DsF-tagged L23 were purified as described above for tryptophan synthase, that is, FimGt-sepharose chromatography followed by TEV protease cleavage, and were indistinguishable from 70S ribosomes obtained by standard sucrose density gradient centrifugation in terms of their protein and rRNA compositions as well as their morphology as was judged by SDS-PAGE, denaturing agarose gel electrophoresis, and electron microscopy, respectively (Figure 6a–c). Thus, by using DsF-tagged L23 of the

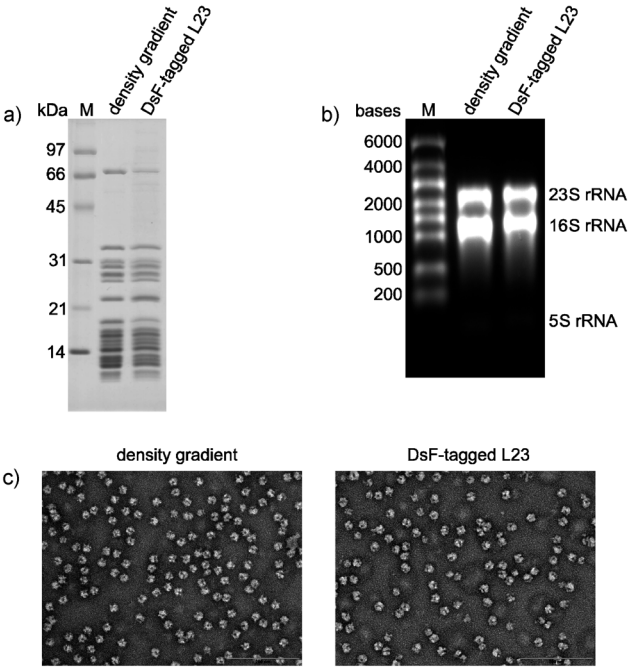

Figure 6. Purification of *E. coli* ribosomes. a) Coomassie-stained 15% SDS-polyacrylamide gel documenting the one-step purification of *E. coli* ribosomes from cell extracts via DsF-tagged L23. Identical banding patterns are obtained for ribosomes purified by sucrose density gradient centrifugation or by using the DsF tag. Lane M: molecular mass standard. b) Denaturing 1% agarose/formaldehyde gel of rRNAs from ribosomes purified by sucrose density gradient centrifugation or by using L23-DsF. Identical banding patterns corresponding to 23S-, 16S-, and 5S-rRNA are visible for both samples. c) Electron micrographs of negatively stained ribosomes purified by sucrose density gradient centrifugation or by using L23-DsF.

FLAG, and His_6 tags, respectively, and is thus superior to established affinity chromatography systems considering purity and yield, in particular.

To demonstrate that the FimGt/DsF system is also applicable for the fusion of the DsF tag to the C terminus of a target protein, as well as for the purification of highly complex macromolecular assemblies and the identification of their molecular components, the DsF tag was fused to the C terminus of protein L23 of the large ribosomal subunit of *E. coli* by the strategy described above for the TrpA fusions (Figure 3c). Using again the *tetA* promoter, DsF-tagged L23 large ribosomal subunit we successfully isolated pure 70S ribosomes in a single step from the crude cell extract.

The new system for single-step affinity purification of protein complexes described in this study has proven to be superior to commonly used systems, in particular in terms of the isolation of low-abundance target protein complexes such as, for example, tryptophan synthase. Moreover, the system should be applicable over wide ranges of pH, NaCl concentration, and temperature and works successfully even for highly complex, supramolecular assemblies such as the ribosome. Also, pure FimGt can be easily obtained recombi-

nantly in high yields from *E. coli*, making the production of FimGt-sepharose efficient and straightforward. Thus, the FimGt/DsF system is generally suitable for the isolation of pure and highly active multi-subunit protein complexes for structure determination by X-ray crystallography. Because of the practically infinite kinetic stability of the FimGt/DsF complex against dissociation under native conditions, the DsF-tagged target, once bound to immobilized FimGt, will survive virtually any washing conditions, which makes the system especially interesting for the permanent and functional immobilization of target proteins under flow conditions where other systems fail owing to their limited kinetic stability against dissociation.

Received: December 12, 2011
Published online: March 21, 2012

Keywords: affinity chromatography · enzymes · kinetics · protein–protein interactions · ribosomes

[1] R. B. Kapust, J. Tozser, J. D. Fox, D. E. Anderson, S. Cherry, T. D. Copeland, D. S. Waugh, *Protein Eng.* **2001**, *14*, 993 – 1000.
[2] a) M. Arifuzzaman, M. Maeda, A. Itoh, K. Nishikata, C. Takita, R. Saito, T. Ara, K. Nakahigashi, H. C. Huang, A. Hirai, K. Tsuzuki, S. Nakamura, M. Altaf-Ul-Amin, T. Oshima, T. Baba, N. Yamamoto, T. Kawamura, T. Ioka-Nakamichi, M. Kitagawa, M. Tomita, S. Kanaya, C. Wada, H. Mori, *Genome Res.* **2006**, *16*, 686 – 691; b) G. I. Evan, G. K. Lewis, G. Ramsay, J. M. Bishop, *Mol. Cell. Biol.* **1985**, *5*, 3610 – 3616; c) C. J. Gloeckner, K. Boldt, A. Schumacher, R. Roepman, M. Ueffing, *Proteomics* **2007**, *7*, 4228 – 4234; d) J. Graumann, L. A. Dunipace, J. H. Seol, W. H. McDonald, J. R. Yates III, B. J. Wold, R. J. Deshaies, *Mol. Cell. Proteomics* **2004**, *3*, 226 – 237; e) E. Hochuli, W. Bannwarth, H. Döbeli, R. Gentz, D. Stüber, *Biotechnology* **1988**, *6*, 1321 – 1325; f) T. P. Hopp, K. S. Prickett, V. L. Price, R. T. Libby, C. J. March, D. P. Cerretti, D. L. Urdal, P. L. Conlon, *Biotechnology* **1988**, *6*, 1204 – 1210.
[3] a) A. G. Atanasov, L. G. Nashev, L. Gelman, B. Legeza, R. Sack, R. Portmann, A. Odermatt, *Biochim. Biophys. Acta Mol. Cell Res.* **2008**, *1783*, 1536 – 1543; b) I. T. Dorn, K. Pawlitschko, S. C. Pettinger, R. Tampe, *Biol. Chem.* **1998**, *379*, 1151 – 1159; c) A. Einhauer, A. Jungbauer, *J. Chromatogr. A* **2001**, *921*, 25 – 30; d) L. Nieba, S. E. Nieba-Axmann, A. Persson, M. Hamalainen, F. Edebratt, A. Hansson, J. Lidholm, K. Magnusson, A. F. Karlsson, A. Pluckthun, *Anal. Biochem.* **1997**, *252*, 217 – 228.
[4] a) E. Hahn, P. Wild, U. Hermanns, P. Sebbel, R. Glockshuber, M. Haner, N. Taschner, P. Burkhard, U. Aebi, S. A. Muller, *J. Mol. Biol.* **2002**, *323*, 845 – 857; b) P. Klemm, G. Christiansen, *Mol. Gen. Genet.* **1990**, *220*, 334 – 338.
[5] a) D. Choudhury, A. Thompson, V. Stojanoff, S. Langermann, J. Pinkner, S. J. Hultgren, S. D. Knight, *Science* **1999**, *285*, 1061 – 1066; b) A. D. Gossert, P. Bettendorff, C. Puorger, M. Vetsch, T. Herrmann, R. Glockshuber, K. Wüthrich, *J. Mol. Biol.* **2008**, *375*, 752 – 763; c) I. Le Trong, P. Aprikian, B. A. Kidd, M. Forero-Shelton, V. Tchesnokova, P. Rajagopal, V. Rodriguez, G. Interlandi, R. Klevit, V. Vogel, R. E. Stenkamp, E. V. Sokurenko, W. E. Thomas, *Cell* **2010**, *141*, 645 – 655; d) C. Puorger, O. Eidam, G. Capitani, D. Erilov, M. G. Grutter, R. Glockshuber, *Structure* **2008**, *16*, 631 – 642; e) C. Puorger, M. Vetsch, G. Wider, R. Glockshuber, *J. Mol. Biol.* **2011**, *412*, 520 – 535.
[6] G. L. Ellman, *Arch. Biochem. Biophys.* **1959**, *82*, 70 – 77.
[7] M. Vetsch, C. Puorger, T. Spirig, U. Grauschopf, E. U. Weber-Ban, R. Glockshuber, *Nature* **2004**, *431*, 329 – 333.
[8] A. Bachmann, T. Kiefhaber, *J. Mol. Biol.* **2001**, *306*, 375 – 386.
[9] Y. Ishihama, T. Schmidt, J. Rappsilber, M. Mann, F. U. Hartl, M. J. Kerner, D. Frishman, *BMC Genomics* **2008**, *9*, 102.
[10] G. Kramer, T. Rauch, W. Rist, S. Vorderwulbecke, H. Patzelt, A. Schulze-Specking, N. Ban, E. Deuerling, B. Bukau, *Nature* **2002**, *419*, 171 – 174.

