DOI: 10.1002/cbic.200800697

Rapid Matrix-Assisted Refolding of Histidine-Tagged Proteins

Tetyana Dashivets,[a] Nichole Wood,[b] Christoph Hergersberg,[b] Johannes Buchner,*[a] and Martin Haslbeck[a]

The formation of inclusion bodies (IBs)—amorphous aggregates of misfolded insoluble protein—during recombinant protein expression, is still one of the biggest bottlenecks in protein science. We have developed and analyzed a rapid parallel approach for matrix-assisted refolding of recombinant His_6-tagged proteins. Efficiencies of matrix-assisted refolding were screened in a 96-well format. The developed methodology allowed the efficient refolding of five different test proteins, including monomeric and oligomeric proteins. Compared to refolding in-solution, the matrix-assisted refolding strategy proved equal or better for all five proteins tested. Interestingly, specifically oligomeric proteins displayed significantly higher levels of refolding compared to refolding in-solution. Mechanistically, matrix-assisted folding seems to differ from folding in-solution, as the reaction proceeds more rapidly and shows a remarkably different concentration dependence—it allows refolding at up to 1000-fold higher protein concentration than folding in-solution.

Introduction

The recombinant expression of genes is of importance both for biotechnology and basic research in protein function and structure. Thousands of DNA targets have been cloned and expressed in bacteria. Although the expression of recombinant proteins in a bacterial host offers several advantages, including short doubling times, well established methods for genetic manipulation, and simple, inexpensive cultivation, the expressed proteins often accumulate in insoluble and inactive inclusion bodies (IBs).[1–3] In a study on human proteins, for example, only 13% of the recombinant proteins could be expressed and purified in a soluble form.[4] Thus, soluble protein expression is a major bottleneck in protein science. Insoluble expression commonly occurs when proteins are produced at unphysiologically high levels, as folding competes with aggregation due to the exposure of hydrophobic surfaces.[5,6] Additionally, the imbalance in the ratio of chaperones and folding catalysts with respect to the folding polypeptides during recombinant protein synthesis can also result in protein aggregation.[7]

The deposition of target protein in insoluble, inactive forms, at first glance, seems to be a failure of the expression strategy. However, IBs also offer advantages as the over-expressed protein is often highly enriched and protected from proteolytic degradation. In addition, the high-level expression of certain proteins in soluble form might be toxic to the host organism. Since the first observation of the formation of IBs more than 30 years ago,[8] a number of folding protocols for different proteins have been described, which overcome problematic steps during the folding process.[9–12] However, to date there exists no common, rational strategy for protein refolding. The highly diverse set of parameters that influence refolding has still to be varied and tested experimentally for each specific protein.

Currently there are three different approaches to refolding: dilution, dialysis and matrix-assisted strategies. So far, less than 20% of the published reports are matrix-assisted, while more than 50% use dilution.[13] The dilution-based refolding strategies, however, are problematic as they produce large volumes of highly diluted ($\mu g\,mL^{-1}$) protein solution. General methods enabling protein refolding at high concentrations are therefore desirable. Matrix-assisted strategies seem to be especially well-suited for this purpose as refolded proteins can be eluted from the matrix at high concentration.[14] Furthermore, the folding pathways of immobilized proteins might be improved.

Here, we describe a stepwise strategy that allows the fast and efficient refolding of immobilized His-tagged proteins. Successful refolding of individual proteins on chelating matrices has been shown previously,[15] although no general strategy has been demonstrated prior to this work. We show that the refolding of five different proteins, on-matrix, displays remarkably similar dependencies on reaction conditions; this suggests that a generic refolding strategy is possible. Also surprising was the observation that refolding and assembly of dimeric and tetrameric proteins appears to be more efficient on-matrix than in-solution.

[a] *T. Dashivets, Prof. Dr. J. Buchner, Dr. M. Haslbeck*
Munich Center for Integrated Protein Science (CIPS^M) and
Chemistry Department, Technische Universität München
85747 Garching (Germany)
Fax: (+49) 89-289-13345
E-mail: johannes.buchner@ch.tum.de

[b] *Dr. N. Wood, Dr. C. Hergersberg*
Biosciences Technologies, GE Global Research Center
Niskayuna, NY 12309 (USA)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.200800697.

Results and Discussion

In vitro refolding of proteins is still a case-specific problem. The efficiency of in vitro folding of proteins from IBs is influenced by many parameters, such as pH, ionic strength, additives or protein concentrations and has to be determined for each protein empirically.[9–11,16] Due to the lack of generic refolding strategies and the need for intensive screening of reaction conditions, refolding continues to be a major bottleneck in protein science. In addition to dilution and dialysis refolding strategies,[17–19] matrix-assisted strategies are attractive because of the spatial separation of matrix-bound proteins during the folding process, which could limit aggregation processes. Furthermore, this strategy allows the optimization of multiple refolding parameters in a fast and efficient manner. Although there are examples of the refolding of His_6-tagged proteins immobilized on chelating materials,[15] there is little information on general strategies and specific mechanisms. We set out to systematically analyze the matrix-assisted refolding performance of five different proteins. To determine the efficiency of the strategy we used proteins from different organisms with a wide range of molecular weights, quaternary structures and enzymatic activities that could be produced either as IBs or in soluble form (Table 1).

To minimize the number of buffer conditions in the optimization process, we established a new, stepwise optimization strategy (Figure 1). After solubilization, the unfolded proteins were bound to the matrix and refolded, on-matrix, in various buffers. The first step was the variation of the basic buffer component and pH value. This was followed by optimization of the ionic strength and additives in further steps. Depending on the results obtained after every step, two to six conditions (in exceptional cases even more) resulting in the highest refolding efficiencies were used as the basis for buffers used in the next screening step. During the initial steps of optimization, the number of selected buffers was usually higher compared to later steps, when specific additives were screened. Here, usually two optimized buffer conditions were sufficient to determine which additives increased the refolding yields.

Initial candidates tested were monomeric proteins, namely enhanced green fluorescent protein (eGFP) and ferredoxin–NADP(+) reductase (FNR). As demonstrated in Figure 2 and Figure S1 in the Supporting Information, this strategy allowed the matrix-assisted refolding with efficiencies of up to 100% for eGFP and ~80% for FNR. The refolding efficiencies were rather low in the initial screening steps of the procedure and

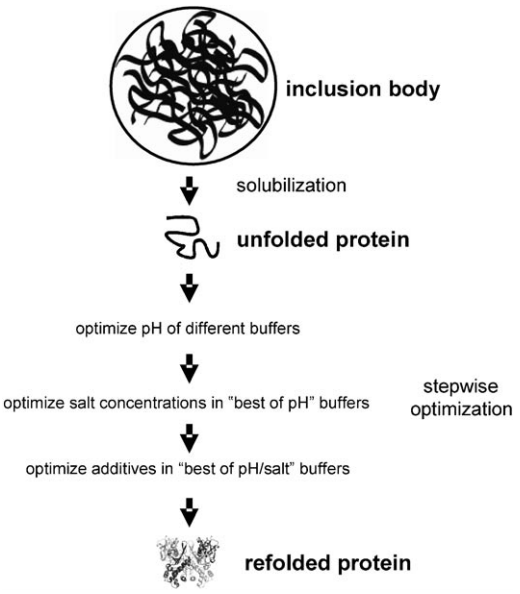

Figure 1. Schematic description of the stepwise optimization of matrix-assisted refolding. Solubilized, unfolded protein was applied to 96-well format HisMultiTrap plates. After loading of the matrices under denaturing conditions, the proteins were incubated in various refolding buffers on-matrix. In a stepwise manner, the respective best buffer components (pH, salt concentration and additive) were determined. After every step, the conditions resulting in the highest refolding efficiencies were selected and used as the basic buffer in the next screening step. The indicated protein structure represents GLK.

increased when additives were screened at already optimized pH and ionic strength conditions (Figure 2, Figure S1). Interestingly, in the case of eGFP, the additive screening indicated that especially a suitable redox potential is necessary to further enhance the refolding efficiency while other additives even decreased the refolding yields. It should be notes that the chromophore of eGFP has to be oxidized correctly to be able to emit fluorescence. It seems that especially buffers containing reducing agents are suitable to reduce unproductive oxidized states or to allow fast oxidation during refolding. Due to variations between individual assays, the standard deviation was on average 8% (Table S1). Thus, triplicates of experiments are advisable for screening.

To compare the quality of the new matrix-assisted strategy versus a classical dilution-based approach, we analyzed refolding by dilution in the respective buffers, also in a stepwise manner (Figure 3, Figure S2). Again, the buffers showing the highest refolding efficiencies were selected and used as the

Table 1. Analyzed proteins.

Protein	Organism	M_w [kDa]	pI	Activity	QS[a]
eGFP[30]	*Aequorea victoria*	28	5.67	intrinsic fluorescence	monomer
FNR[7]	*Arabidopsis thaliana*	35	6.18	catalyzes the final step in the photosynthetic electron transport chain	monomer
CS[34]	*Sus scrofa*	49	8.12	catalyzes the first step in the citric acid cycle	dimer
GLK[35]	*E. coli*	35	6.1	facilitates phosphorylation of glucose to glucose-6-phosphate	dimer
β-Gal[36]	*E. coli*	116	5.28	hydrolysis of terminal nonreducing β-D-galactose residues in β-D-galactosides	tetramer

[a] Quaternary structure.

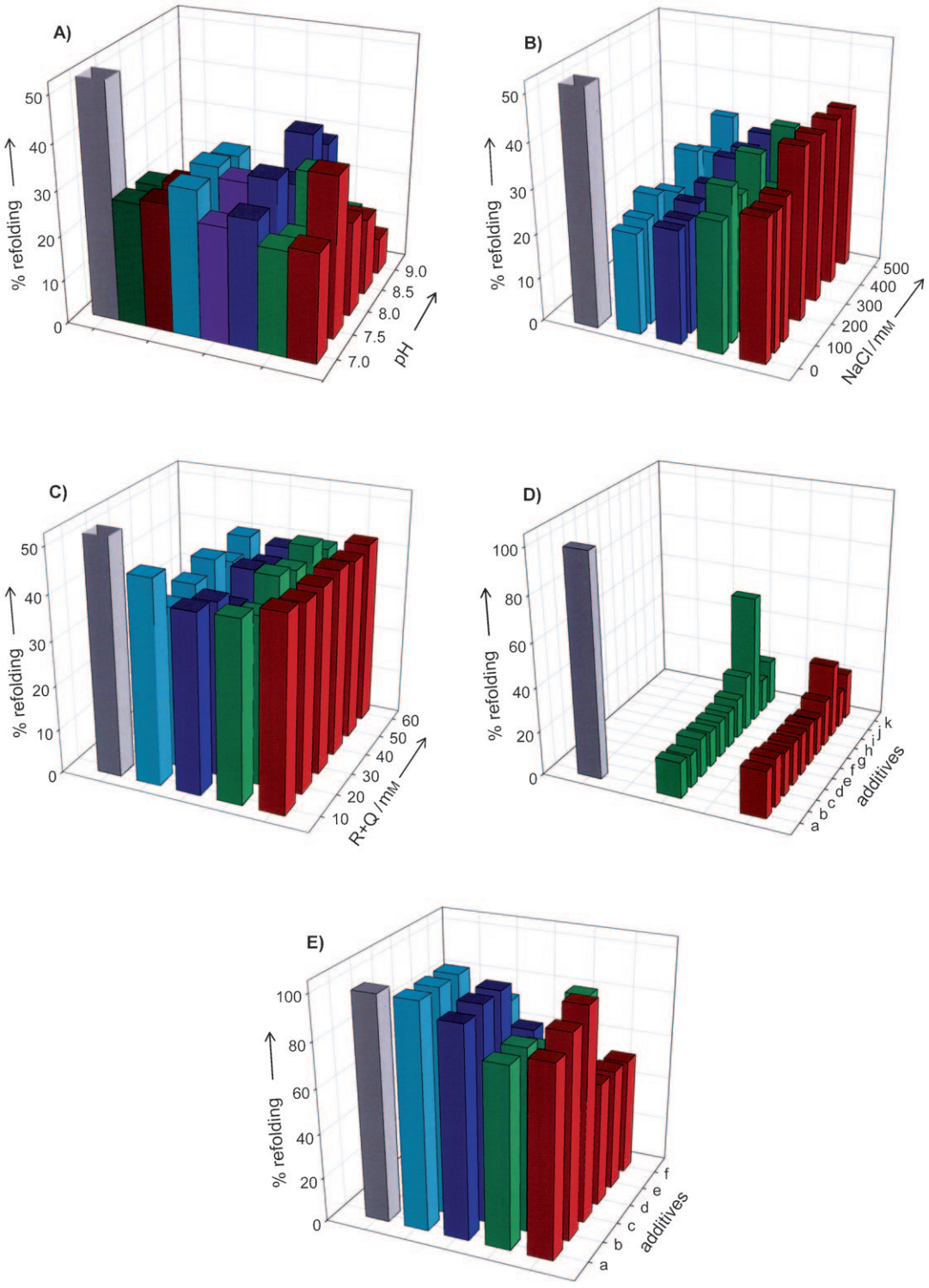

Figure 2. Stepwise optimization of matrix-assisted refolding of eGFP. Efficiencies of refolding after incubation for 1 h at 20 °C in the respective buffers were determined according to the intrinsic GFP fluorescence at 503 nm. Equally treated native GFP, which was used as reference, is depicted in gray. The fluorescence of the native protein was cut at 50% (A–C) for better visualization. A) First step: initial screening for optimal buffer substances and pH. Buffers: 100 mM HEPES (red), 100 mM MOPS (green), 40 mM Na-P (blue), 100 mM Tris-Ac (pink), 100 mM Tris-HCl (cyan), 100 mM Na-P (dark red), 100 mM K-P (dark green). B) Second step: screening of salt concentrations in the best buffers selected in the previous step. Buffers: 40 mM Na-P, pH 7.5 (red) or pH 8.0 (green), 100 mM Tris-HCl, pH 7.5 (blue) or pH 8.0 (cyan). C) Third step: screening of different concentrations of an L-Arg/L-Gln (R + Q) mixture. Buffers: 40 mM Na-P, 200 mM NaCl, pH 7.5 (red) or pH 8.0 (green), 100 mM Tris-HCl, 300 mM NaCl, pH 7.5 (blue) or pH 8.0 (cyan). D) Fourth step: screening of different refolding additives; glycerin: 2.5% (a), 5% (b), 10% (c); sucrose: 100 mM (d), 200 mM (e), 300 mM (f); TCEP: 1 mM (g), 2 mM (h); PEG: 0.01% (i); $CaCl_2$: 1 mM (j); $MgCl_2$: 1 mM (k) in 40 mM Na-P, 200 mM NaCl, 40 mM Arg, 40 mM Gln, pH 7.5 (red) or 100 mM Tris-HCl, 300 mM NaCl, 40 mM Arg, 40 mM Gln, pH 7.5 (green). E) Fifth step: screening of reducing conditions. Buffers: 40 mM Na-P, 200 mM NaCl, 40 mM Arg, 40 mM Gln, pH 7.5 (red) or pH 8.0 (green), 100 mM Tris-HCl, 300 mM NaCl, 40 mM Arg, 40 mM Gln, pH 7.5 (blue) or pH 8.0 (cyan). Reducing agents: DTE: 1 mM (a), 2 mM (b), 5 mM (c); TCEP: 2 mM (d), 5 mM (e), 10 mM (f).

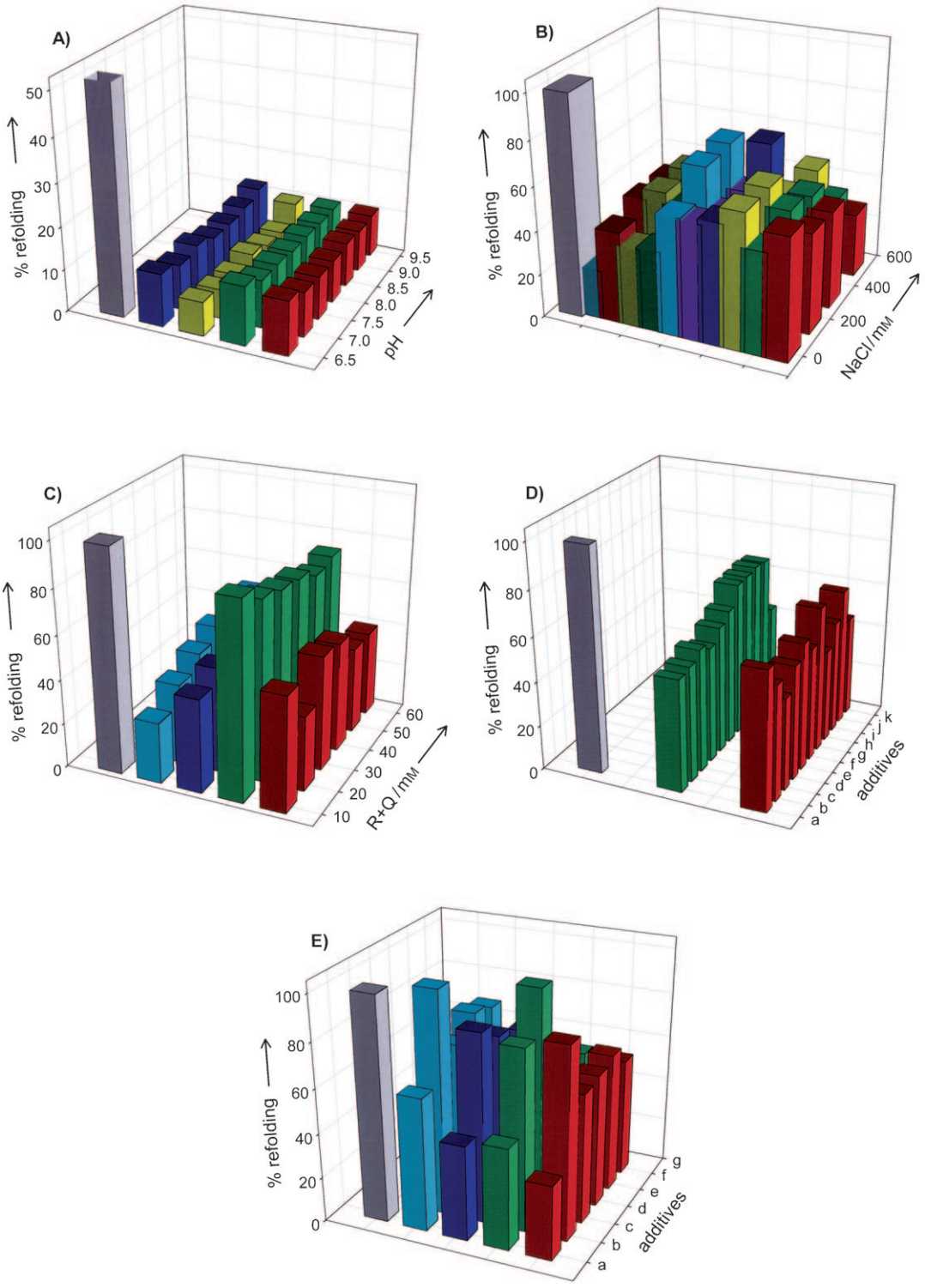

Figure 3. In-solution refolding of eGFP. Efficiencies of refolding, after incubation for 16 h at 20 °C in the respective buffers were determined according to the intrinsic GFP fluorescence at 503 nm. Equally treated native GFP, which was used as reference, is depicted in gray. Fluorescence of native protein was cut at 50% (A) for better visualization. A) First step: initial screen of buffer substances and pH; 100 mM Tris-HCl (red), 100 mM Tris-Ac (green), 40 mM Na-P (yellow), 100 mM HEPES (blue). B) Second step: screening of salt concentrations in the best buffers selected in the previous step. Buffers: 40 mM Na-P, pH 7.0 (red), pH 7.5 (green) or pH 8.0 (yellow), 100 mM Tris-HCl, pH 7.0 (blue), pH 7.5 (pink) or pH 8.0 (cyan), 100 mM MOPS, pH 7.0 (dark green) or pH 8.0 (dark yellow), 100 mM HEPES, pH 7.0 (dark red) or pH 8.0 (dark cyan). C) Third step: screening of different concentrations of an L-Arg/L-Gln (R + Q) mixture. Buffers: 40 mM Na-P, 200 mM NaCl, pH 7.5 (red) or pH 8.0 (green), 100 mM Tris-HCl, 300 mM NaCl, pH 7.5 (blue) or pH 8.0 (cyan). D) Fourth step: screening of different refolding additives; glycerin: 2.5% (a), 5% (b), 10% (c); sucrose: 100 mM (d), 200 mM (e), 300 mM (f); TCEP: 1 mM (g), 2 mM (h); PEG: 0.01% (i); $CaCl_2$: 1 mM (j); $MgCl_2$: 1 mM (k) in 40 mM Na-P, 200 mM NaCl, 40 mM Arg, 40 mM Gln, pH 7.5 (red), or 100 mM Tris-HCl, 300 mM NaCl, 40 mM Arg, 40 mM Gln, pH 7.5 (green). E) Fifth step: screening of reducing conditions. Buffers: 40 mM Na-P, 200 mM NaCl, 40 mM Arg, 40 mM Gln, pH 7.5 (red) or pH 8.0 (green), 100 mM Tris-HCl, 300 mM NaCl, 40 mM Arg, 40 mM Gln, pH 7.5 (blue) or pH 8.0 (cyan). Reducing agents: DTE: 1 mM (a), 2 mM (b), 5 mM (c); TCEP: 2 mM (d), 5 mM (e), 10 mM (f).

basis for the next optimization step. Especially in the case of eGFP (Figure 3A), the first screening step did not allow discrimination between different buffers and pH values. In this case, all conditions were screened with various salts, and the selection of the optimal buffers was performed after this second screening step. Nevertheless, also here, the stepwise-optimization strategy proved to be applicable and resulted in overall high refolding efficiencies (Figure 3, Figure S2). In Figure 3B, the ten buffer setups giving the highest refolding yields after this second screening step are shown. In comparison to the matrix-assisted approach, similar refolding efficiencies were obtained for eGFP (100%). However, for FNR the folding efficiency in-solution (55%) was significantly lower than on-matrix (80%). Comparable to other refolding studies,[7,20] the standard deviation for refolding in-solution was relatively high (on average 15%; Table S2).

On the basis of these initial results, we set out to test whether dimeric proteins like citrate synthase (CS) and glucokinase (GLK) are refoldable using this approach. As demonstrated in Figures 4 and S3, high refolding yields of up to 70–80% were achieved. Again, in comparison to the respective in-solution refolding approach, the matrix-assisted refolding yields were significantly higher (Figure S4). For CS, the optimum refolding yields were up to approximately fourfold higher than reported previously.[21–23]

Finally, we tested whether a large, oligomeric protein, such as β-galactosidase (β-Gal), is refoldable on-matrix. Previously, β-Gal was shown to be recalcitrant with respect to refolding after denaturation, and more or less no refolding was detected in-solution for β-Gal when standard protocols were applied.[24] According to the literature, some refolding in-solution can only be obtained when the molecular chaperones GroEL/ES or Hsp70/Hsp90 are present during refolding.[25,26] Following our matrix-assisted strategy (Figure 1), we were able to optimize β-Gal refolding and we obtained up to 20–30% refolding yield in a set of eight selected buffers (Figure 5). Similar to published reports, we were unable to obtain significant refolding of β-Gal by stepwise optimization in-solution (Figure 5). Interestingly, the obtained matrix-assisted refolding yields for β-Gal are even ~twofold higher than the previously reported optima for chaperone-assisted refolding.[25]

For better comparison, the optimized refolding yields obtained for both matrix-assisted and in-solution refolding are summarized in Figure 6 for all proteins tested. In this context it should be noted that for the individual proteins, additive effects on the refolding efficiency were visible in nearly all optimization steps; this indicates that a stepwise strategy is generally applicable. In addition, the matrix-assisted optimization strategy results in protein samples in identical buffer conditions after elution. In contrast, for the in-solution strategies, the samples are in different screening buffers after the refolding step, which can affect subsequent protein activity assays. Therefore, an additional dialysis step is typically required after the in-solution refolding step in order to obtain samples in identical buffers.

Unexpectedly, for all five proteins tested the optimization of matrix-assisted refolding resulted in very similar optimum

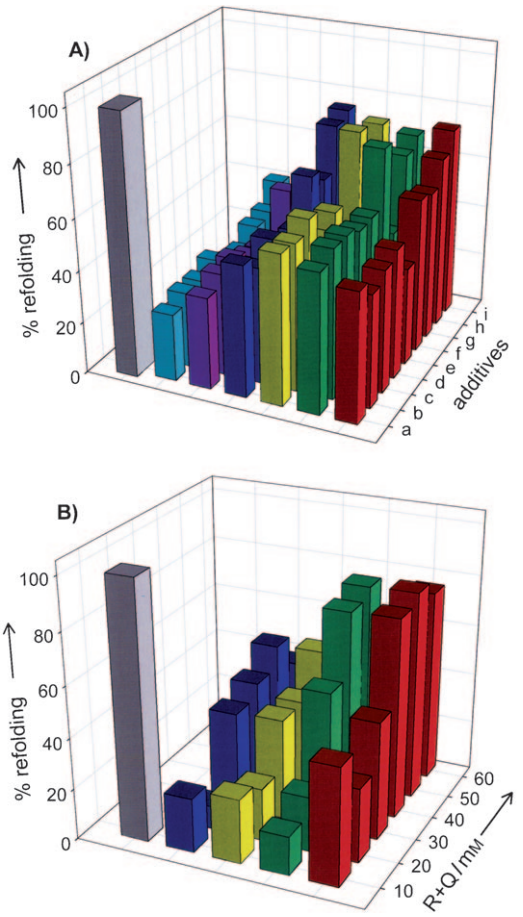

Figure 4. Matrix-assisted refolding of dimeric proteins. A) The final step of the optimization of matrix-assisted refolding of GLK is shown. Buffers: 100 mM Tris-HCl, 300 mM NaCl, 50 mM Arg (R), 50 mM Gln (Q), pH 7.5 (red) or pH 8.0 (green); 40 mM Na-P, 200 mM NaCl, 50 mM Arg, 50 mM Gln, pH 7.5 (yellow) or pH 8.0 (blue); 100 mM K-P, 300 mM NaCl, 50 mM Arg, 50 mM Gln, pH 7.5 (pink) or pH 8.0 (cyan). Additives; sucrose: 100 mM (a), 200 mM (b); PEG 6000: 0.01% (c); glycerin: 5% (d); cyclodextrin: 5 mM (e), 10 mM (f); DTE: 2 mM (g); TCEP: 2 mM (h), 5 mM (i). B) The final step of the optimization of matrix-assisted refolding of CS. Buffers: 100 mM Tris-HCl, 300 mM NaCl, pH 7.5 (red) or pH 8.0 (green); 40 mM Na-P, 200 mM NaCl, pH 7.5 (yellow) or pH 8.0 (blue).

buffer conditions. This suggests that the chromatography conditions might have a greater influence on the buffer requirements than expected. The minimum set of buffers identified in this study is summarized in Table 2 together with the respective refolding yields for all five analyzed proteins.

A further limitation of in-solution refolding strategies includes the variation in optimum reaction times, which might depend on the protein and/or the respective refolding buffers. These can vary from minutes to many hours.[23,27,28] We, therefore, compared the kinetics of refolding in-solution and on-matrix for CS. In solution, incubation times >1 h were needed to reach maximum refolding yields (Figure 7A). For the matrix-assisted strategy, however, the kinetic analysis demonstrated that shorter (~20 min) refolding times (as already used during the optimization screen) seem to be sufficient. This surprising observation suggests that matrix-assisted refolding results in accelerated renaturation compared to in-solution approaches.

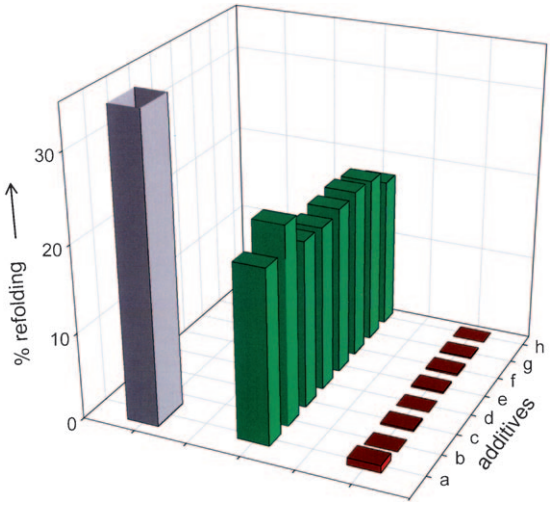

Figure 5. Matrix-assisted refolding of tetrameric β-Gal. Comparison of β-Gal refolding in-solution (red) and on-matrix (green) at 20 °C under identical buffer conditions. Buffers: 40 mM Na-P, 200 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, pH 7.5, with DTE: 2 mM (a), 5 mM (b); TCEP: 2 mM (c), 5 mM (d) or 100 mM Tris-HCl, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, pH 7.5, with DTE: 2 mM (e), 5 mM (f); TCEP: 2 mM (g), 5 mM (h). Equally treated native β-Gal, which was used as reference, is depicted in gray. The activity of the native protein was cut at 35% for better visualization.

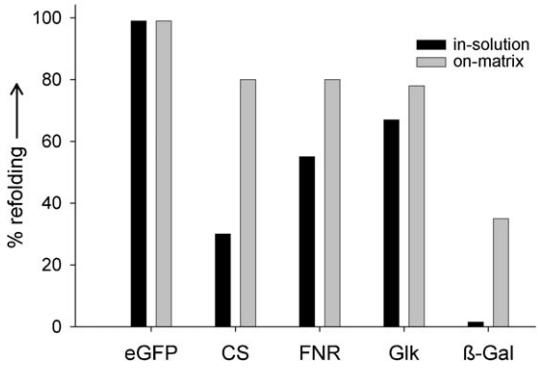

Figure 6. Comparison of refolding efficiencies. The respective refolding efficiencies on-matrix and in-solution are compared for the five proteins tested. The yields for the respective best refolding buffer are shown.

At the moment, it is unclear how this influence on the refolding kinetics can be explained. It might be that binding to the matrix is similar to the binding of substrates to chaperones.[12] Classically, chaperones bind hydrophobic residues that would otherwise be accessible to the solvent and lead to incorrect interactions that can slow down folding.[29] Binding to the matrix might also inhibit protein aggregation in a similar manner by decreasing interactions with other protein chains.

Finally, the most striking difference between folding in-solution and on-matrix is the influence of protein concentration on refolding. For solution-based refolding strategies low protein concentration (ng mL^{-1}–μg mL^{-1}) are generally needed for effective refolding.[11] Consistent with previous studies, the refolding of CS in-solution was most efficient at ng mL^{-1} concentrations and decreased dramatically at higher concentrations with no refolding observed above 15 μg mL^{-1} (Figure 7B). When we analyzed the effect of protein concentration on the matrix-assisted refolding of CS, we found that CS refolding yields were not significantly influenced by concentrations up to 10 mg mL^{-1} (Figure 7C). For β-Gal, the protein concentrations during matrix-assisted refolding could be increased to a similar extent (data not shown). Thus, the matrix-assisted refolding strategy allows the refolding of proteins at up to 1000-times higher concentrations compared to refolding in-solution.

Table 2. Minimum set of buffers for matrix-assisted refolding.

Buffer	eGFP	CS	GLK	FNR	β-Gal
100 mM Tris-HCl, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, pH 7.5	40–45%	80–85%	40–45%	35–40%	10–15%
100 mM Tris-HCl, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, pH 8.0	40–45%	80–85%	35–40%	35–40%	10–15%
100 mM Tris-HCl, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 2 mM DTE, pH 7.5	90–95%	–[a]	50–55%	40–45%	20–25%
100 mM Tris-HCl, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 2 mM DTE, pH 8.0	95–99%	–	55–60%	40–45%	–
100 mM Tris-HCl, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 5 mM DTE, pH 7.5	55–60%	–	–	–	20–25%
100 mM Tris-HCl, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 5 mM DTE, pH 8.0	40–45%	–	–	–	–
100 mM Tris-HCl, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 2 mM TCEP, pH 7.5	60–65%	–	60–65%	60–65%	20–25%
100 mM Tris-HCl, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 2 mM TCEP, pH 8.0	55–60%	–	60–65%	55–60%	–
100 mM Tris-HCl, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 5 mM TCEP, pH 7.5	60–65%	–	70–75%	75–80%	20–25%
100 mM Tris-HCl, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 5 mM TCEP, pH 8.0	55–60%	–	70–75%	75–80%	–
40 mM Na-P, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, pH 7.5	40–45%	45–50%	45–50%	35–40%	10–15%
40 mM Na-P, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, pH 8.0	45–50%	45–50%	45–50%	35–40%	10–15%
40 mM Na-P, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 2 mM DTE, pH 7.5	85–90%	–	70–75%	45–50%	15–20%
40 mM Na-P, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 2 mM DTE, pH 8.0	80–85%	–	60–65%	50–55%	–
40 mM Na-P, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 5 mM DTE, pH 7.5	50–55%	–	–	–	20–25%
40 mM Na-P, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 5 mM DTE, pH 8.0	55–60%	–	–	–	–
40 mM Na-P, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 2 mM TCEP, pH 7.5	50–55%	–	60–65%	50–55%	15–20%
40 mM Na-P, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 2 mM TCEP, pH 8.0	50–55%	–	70–75%	55–60%	–
40 mM Na-P, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 5 mM TCEP, pH 7.5	55–60%	–	70–75%	75–80%	20–25%
40 mM Na-P, 300 mM NaCl, 50 mM L-Arg, 50 mM L-Gln, 5 mM TCEP, pH 8.0	55–60%	–	75–80%	75–80%	–

[a] Yields were not determined; this indicates that the respective buffer was not selected during the optimization process.

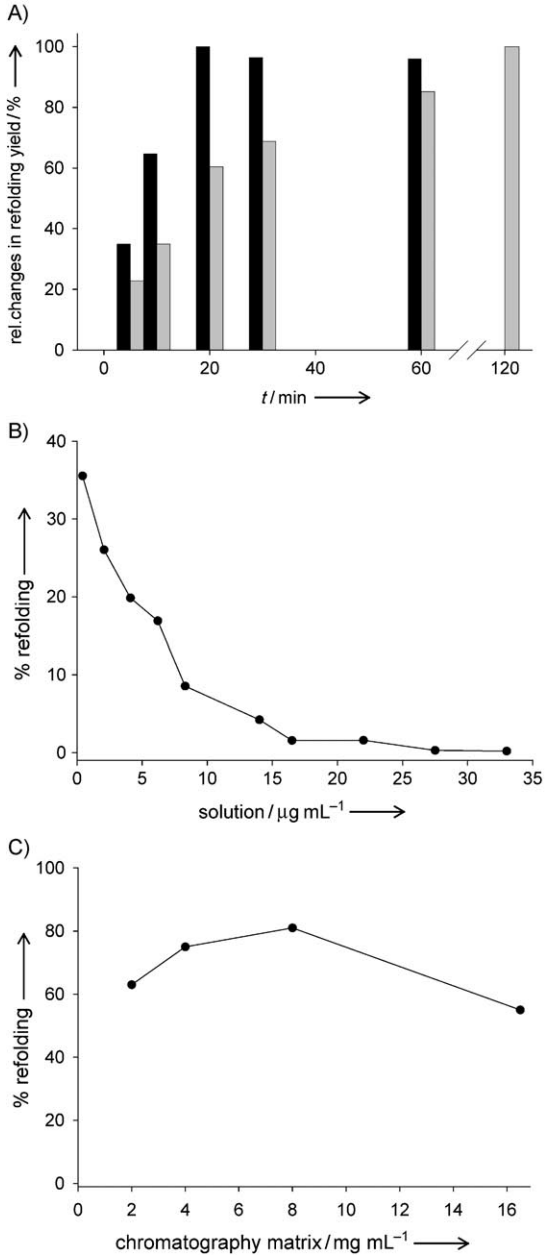

Figure 7. Differences in the mechanism between matrix-assisted folding and folding in-solution. A) Comparative analysis of the refolding kinetics of CS. Refolding on-matrix (black) and in-solution (gray) was analyzed in identical buffers (100 mM Tris-HCl, 300 mM NaCl, 50 mM Arg, 50 mM Gln, pH 7.5). B) Concentration dependence of the refolding efficiency of CS in solution. Denatured CS was diluted 100-fold into the refolding buffer (100 mM Tris-HCl, 300 mM NaCl, 40 mM Arg, 40 mM Gln, pH 8.0) to yield the respective protein concentration. Refolding efficiencies after 16 h of refolding time are compared. C) Concentration dependence of the matrix-assisted refolding efficiency of CS. Denatured CS was loaded onto HisTrap matrix to yield the respective protein concentration per mL of matrix. For refolding, the proteins were incubated in 100 mM Tris-HCl, 300 mM NaCl, 40 mM Arg, 40 mM Gln, pH 8.0. Refolding efficiencies after 1 h of refolding time are compared.

Conclusions

In summary, our systematic analysis of matrix-assisted refolding revealed general reaction principles, which could be applicable for a wide range of proteins. It represents a rational strategy to optimize refolding of insolubly expressed proteins. In the multifactorial space of possible refolding buffers, stepwise optimization allows the efficient minimization of the number of experiments. This offers the possibility to reduce the time of the screening procedure. Additionally, up-scaling of the optimized refolding procedure is directly and easily possible. From the scientific point of view, the striking differences in kinetics and concentration dependence suggest that matrix-assisted folding follows different routes to the native state, which possibly avoid kinetic traps in the energy landscape of folding.

Experimental Section

Materials: All chemicals were p.a. grade and were obtained from Sigma–Aldrich. The expression plasmids for eGFP and FNR were kind gifts from S. Töpell and F. Rohdich, respectively.

Expression and purification of test proteins: GLK und β-Gal were amplified from *E. coli* by PCR and cloned into a pET28b expression vector (Novagen, Madison, US) to generate an N-terminal His_6-tag. The gene for the mitochondrial pig heart CS was obtained as described elsewhere,[7] and cloned, without its leader sequence, into the pIVEX2.3 expression vector (Roche Diagnostics, Mannheim, Germany) to create a N-terminal His_6-tag fusion protein.

All proteins were expressed in *E. coli* BL21 (DE3) COD+RIL cells (Qiagen, Hilden, Germany). Cells were cultured in lysogeny broth (LB) supplemented with ampicillin (100 µg mL^{-1}) or kanamycin (35 µg mL^{-1}) at 37 °C until an $OD_{600} = 0.8$ was reached. Protein expression was subsequently induced with IPTG (1 mM) and the cells were incubated 30 °C for 16 h. Cells were harvested by centrifugation at 3000 *g* for 20 min. The harvested cells were cracked by using a Basic Z cell disruption system (Constant Systems, Warwick, UK). After separation of the soluble and insoluble protein content by centrifugation at 40 000 *g* for 30 min, insoluble target proteins were further isolated by following the IB preparation protocol as described elsewhere.[28] The solubilized proteins were affinity-purified on 5 mL HisTrap columns under denaturing conditions and dialyzed into denaturing buffer (50 mM Tris-HCl, 100 mM NaCl, 10 mM DTE (GLK, FNR, β-Gal), 20 mM DTE (CS) or without DTE (eGFP), pH 8.0 containing 6 M GdmCl) after elution. All chromatography steps were performed by using an Äkta Purifier 10 system (GE Healthcare, Munich, Germany) equipped with an eight column vent system. Native proteins were obtained by purification from the soluble protein fraction by affinity chromatography on 5 mL HisTrap FF columns.

To purify the target proteins to homogeneity, ion-exchange chromatography on a 6 mL Resource Q column, followed by preparative size-exclusion chromatography on a 16/60 Superdex 75 pg column was applied for CS and β-Gal. FNR and GLK were further purified by ion-exchange chromatography on 5 mL HiTrap Q FF columns. After every chromatography step the protein-containing fractions were analyzed by SDS-PAGE.

Matrix-assisted refolding: Solubilized protein samples were applied to HisMultiTrap FF plates (GE Healthcare). HisMultiTrap plates were handled according to the manufacturer's protocols; protein samples were pre-equilibrated in denaturing buffer, solubilized (50 µg) and then applied to each well. The respective refolding buffer (500 µL) was added subsequently to each well and the plates were centrifuged for 5 min at 50 *g*. Addition refolding buffer (500 µL) was added to each well and plates were incubated for 1 h

at 20 °C. Refolded proteins were eluted with Na-P (200μL; 40 mM), NaCl (300 mM) and imidazole (500 mM), pH 7.8.

For reference, in each plate three wells were loaded with the same amount of the respective native protein. These wells were treated under nondenaturing conditions (without GdmCl) and by following identical processing conditions. The activity of the native sample was set as 100% and the refolding yields in the various buffers were normalized accordingly. All data shown represent the average of at least three individual experiments.

Refolding in solution: For each refolding experiment, solubilized protein (5 μg) was diluted 100-fold in the respective buffer. The samples were mixed immediately, gently but vigorously, and incubated at 20 °C for 16 h prior to activity determination. For reference, three dilutions of the respective native protein in Tris-HCl (50 mM), NaCl (100 mM), pH 8.0, were prepared and incubated at 20 °C for 16 h. Data analysis was performed as described for the matrix-assisted procedure.

Determination of enzyme activities: The activity of eGFP (variant F64L and S65T[30]) was determined according to the fluorescence emission of the folded and oxidized protein. Fluorescence was analyzed by using a SPEX II fluorescence spectrometer (Jobin Yvon, Unterhaching, Germany). The excitation wavelength was set to 395 nm and fluorescence emission was scanned from 430–550 nm. After baseline subtraction, the signal intensities at 508 nm were compared with the intensity of equally treated, native eGFP.

To determine FNR activity, reduction of potassium ferricyanide was monitored at 419 nm (decrease in absorbance). The reaction was performed at 25 °C in Tris-HCl (100 mM), pH 8.0, containing potassium ferricyanide (1 mM) and various concentrations of NADPH (200–400 μM). The activity of GLK samples was determined according to Darrow and Colowick.[31]

CS activity was determined as described elsewhere.[32,33] The activity measurements were carried out at 25 °C in Tris-HCl (50 mM), EDTA (2 mM), pH 8.0, by using a Jasco V-550 spectrophotometer (Jasco, Gross-Umstadt, Germany). β-Gal activity was determined as described elsewhere.[24]

Abbreviations: CS: citrate synthase; DTE: dithioeritol; eGFP: enhanced green fluorescent protein; FNR: ferredoxin–NADP(+) reductase; β-Gal: beta galactosidase; GdmCl: guanidinium hydrochloride; GLK: glucokinase; HEPES: *N*-2-hydroxyethylpiperazone-*n*-2-ethanesulfonic acid; His_6-tag: hexahistidine tag; K-P: potassium phosphate buffer; MOPS: 3-(*N*-morpholino)propanesulfonic acid; Na-P: sodium phosphate buffer; PAGE: polyacrylamide gel electrophoresis; PEG: poly(ethylene glycol); SDS: sodium dodecyl sulfate; SEC: size-exclusion chromatography; TCEP: Tris(2-carboxyethyl)-phosphine hydrochloride.

Acknowledgements

We thank Stefanie Heinzlmeier, Bianca Ludwig, Ron Gorham and Amanda Scampini for practical assistance, Reginald Smith, June Klimash and Ake Danielsson for helpful discussion, and Ake Danielsson and GE GRC for supply of materials. This work was supported by grants from the Karl Max von Bauernfeind Verein and GE to T.D., M.H. and J.B. Support by Fonds der Chemischen Industrie to M.H. and J.B. is gratefully acknowledged.

Keywords: protein aggregation · protein expression · protein folding · protein refolding · proteins

[1] D. Christendat, A. Yee, A. Dharamsi, Y. Kluger, A. Savchenko, J. R. Cort, V. Booth, C. D. Mackereth, V. Saridakis, I. Ekiel, G. Kozlov, K. L. Maxwell, N. Wu, L. P. McIntosh, K. Gehring, M. A. Kennedy, A. R. Davidson, E. F. Pai, M. Gerstein, A. M. Edwards, C. H. Arrowsmith, *Nat. Struct. Biol.* **2000**, *7*, 903–909.
[2] R. F. Service, *Science* **2002**, *298*, 948–950.
[3] S. Gräslund, P. Nordlund, J. Weigelt, B. M. Hallberg, J. Bray, O. Gileadi, S. Knapp, U. Oppermann, C. Arrowsmith, R. Hui, J. Ming, S. dhe-Paganon, H. W. Park, A. Savchenko, A. Yee, A. Edwards, R. Vincentelli, C. Cambillau, R. Kim, S. H. Kim, Z. Rao, Y. Shi, T. C. Terwilliger, C. Y. Kim, L. W. Hung, G. S. Waldo, Y. Peleg, S. Albeck, T. Unger, O. Dym, et al., *Nat. Methods* **2008**, *5*, 135–146.
[4] P. Braun, Y. Hu, B. Shen, A. Halleck, M. Koundinya, E. Harlow, J. LaBaer, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2654–2659.
[5] J. Buchner, U. Brinkmann, I. Pasta, *Biotechnologia* **1992**, *10*, 682–685.
[6] J. King, C. Haase-Pettingell, A. S. Robinson, M. Speed, A. Mitraki, *FASEB J.* **1996**, *10*, 57–66.
[7] M. Haslbeck, I. Schuster, H. Grallert, *J. Chromatogr. B* **2003**, *786*, 127–136.
[8] W. F. Prouty, M. J. Karnovsky, A. L. Goldberg, *J. Biol. Chem.* **1975**, *250*, 1112–1122.
[9] M. Mayer, J. Buchner, *Methods Mol. Med.* **2004**, *94*, 239–254.
[10] R. Rudolph, H. Lilie, *FASEB J.* **1996**, *10*, 49–56.
[11] H. Lilie, E. Schwarz, R. Rudolph, *Curr. Opin. Biotechnol.* **1998**, *9*, 497–501.
[12] M. M. Altamirano, R. Golbik, R. Zahn, A. M. Buckle, A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 3576–3578.
[13] A. M. Buckle, G. L. Devlin, R. A. Jodun, K. F. Fulton, N. Faux, J. C. Whisstock, S. P. Bottomley, *Nat. Methods* **2005**, *2*, 3.
[14] M. Li, Z. G. Su, J. C. Janson, *Prot. Expr. Pur.* **2004**, *33*, 1–10.
[15] W. Swietnicki, *Curr. Opin. Biotechnol.* **2006**, *17*, 367–372.
[16] C. Klein, F. Hesse, A. Dehner, R. A. Engh, M. Schwaiger, S. Hansen, *Biol. Chem.* **2004**, *385*, 95–102.
[17] R. Vincentelli, S. Canaan, V. Campanacci, C. Valencia, D. Maurin, F. Frassinetti, L. Scappucini-Calvo, Y. Bourne, C. Cambillau, C. Bignon, *Protein Sci.* **2004**, *13*, 2782–2792.
[18] M. S. Willis, J. K. Hogan, P. Prabhakar, X. Liu, K. Tsai, Y. Wei, T. Fox, *Protein Sci.* **2005**, *14*, 1818–1826.
[19] A. Hofmann, M. Tai, W. Wong, C. G. Glabe, *Anal. Biochem.* **1995**, *230*, 8–15.
[20] R. H. Cowan, R. A. Davies, T. T. J. Pinheiro, *Anal. Biochem.* **2008**, *376*, 25–38.
[21] M. Krause, R. Rudolph, E. Schwarz, *FEBS Lett.* **2002**, *532*, 253–255.
[22] W. Zhi, S. J. Landry, L. M. Gierasch, P. A. Srere, *Protein Sci.* **1992**, *1*, 522–529.
[23] J. Buchner, R. Rudolph, *Biotechnologia* **1991**, *9*, 157–162.
[24] A. Nichtl, J. Buchner, R. Jaenicke, R. Rudolph, T. Scheibel, *J. Mol. Biol.* **1998**, *282*, 1083–1091.
[25] A. Ayling, F. Baneyx, *Protein Sci.* **1996**, *5*, 478–487.
[26] B. Freeman, R. I. Morimoto, *EMBO J.* **1996**, *15*, 2969–2979.
[27] A. Jungbauer, W. Kaar, *J. Biotechnol.* **2007**, *128*, 587–596.
[28] C. Lange, R. Rudolph in *Protein Folding Handbook, Vol. 1* (Eds.: J. Buchner, T. Kiefhaber), Wiley-VCH, Weinheim, **2005**, pp. 1245–1280.
[29] S. Walter, J. Buchner, *Angew. Chem.* **2002**, *114*, 1142–1158; *Angew. Chem. Int. Ed.* **2002**, *41*, 1098–1113.
[30] S. Topell, J. Hennecke, R. Glockshuber, *FEBS Lett.* **1999**, *457*, 283–289.
[31] R. A. Darrow, S. P. Colowick, *Methods Enzymol.* **1962**, *5*, 226–227.
[32] P. A. Srere, *J. Biol. Chem.* **1966**, *241*, 2157–2165.
[33] J. Buchner, H. Grallert, U. Jakob, *Methods Enzymol.* **1998**, *290*, 323–338.
[34] K. Takubo, T. Morikawa, Y. Nonaka, M. Mizutani, S. Takenaka, K. Takabe, M. A. Takahashi, D. Ohta, *Plant Mol. Biol.* **2003**, *52*, 817–830.
[35] D. Meyer, C. Schneider-Fresenius, R. Horlacher, R. Peist, W. W. Boos, *J. Bacteriol.* **1997**, *179*, 1298–1306.
[36] B. W. Matthews, *C. R. Biol.* **2005**, *328*, 549–556.

Received: October 22, 2008
Published online on February 23, 2009

