Functionally Accepted Insertions of Proteins within Protein Domains*

Bruno Collinet†‡, Mireille Hervé, Frédéric Pecorari, Philippe Minard, Olivier Eder, and Michel Desmadril§

From the Laboratoire de Modélisation et d’Ingénierie des Protéines, EP1088 Université de Paris-Sud, F-91405 Orsay Cedex, France

Experiments were designed to explore the tolerance of protein structure and folding to very large insertions of folded protein within a structural domain. Dihydrofolate reductase and β-lactamase have been inserted in four different positions of phosphoglycerate kinase. The resultant chimeric proteins are all overexpressed, and the host as well as the inserted partners are functional. Although not explicitly designed, functional coupling between the two fused partners was observed in some of the chimeras. These results show that the tolerance of protein structures to very large structured insertions is more general than previously expected and supports the idea that the natural sequence continuity of a structural domain is not required for the folding process. These results directly suggest a new experimental approach to screen, for example, for folded protein in randomized polypeptide sequences.

Proteins larger than a few hundred amino acids are almost invariably constituted of several domains. This modularity of protein structure is believed to be the result of an evolutionary process where two or more ancestral genes coding for single domains proteins were at some stage fused together and gave rise to proteins with new biological properties. This model assumes that the putative ancestral proteins were able to fold into stable native structures before the fusion event took place and suggests that the corresponding domains may still be independent folding units in the resultant fused protein. It has indeed been observed in several cases that isolated domains of multidomain proteins can fold independently (1–6). Partially based on these arguments, the folding process of large proteins is often described as a sequential and hierarchical process in which interactions between elements close in sequence position are likely to be formed during early steps, before the interactions involving sequentially distant elements. Consequently, it is often stated that, during the folding process of a multidomain protein, the domains fold first and then interact with each other (7). This conception implicitly associates the continuity of a polypeptide sequence coding for a domain and its folding ability. As first pointed out by Russell (8), this view of the folding process of multidomain proteins is simple and rather satisfying for proteins made by continuous domains but becomes more puzzling for the subset of proteins made of discontinuous domains. Proteins with discontinuous domains can be viewed as proteins in which there is at least one domain, or a fraction of a domain, inserted within another domain. Well known examples of such proteins are dsbA (9), pyruvate kinase (10), and DNA polymerase (11). Although proteins with continuous domains are more common, discontinuous domains made by two segments of the polypeptide chains are not very rare exceptions (12). A recent systematic survey of structural domains showed that 28% of structural domains are not continuous (13).

This suggests that the continuity of the polypeptide sequence within a structural unit, although the most common situation, is not mandatory for a productive folding process. Two other lines of experimental evidence tend to support this somewhat counterintuitive view. First, functional proteins can be reconstituted by the reassocation of two or more polypeptide fragments. It is now clear that this process can occur with pairs of fragments found in many different proteins (14) and that the two complementary fragments need not be able to fold autonomously or correspond to structural domains to reassociate functionally (15). Second, it has been shown that natural proteins made by two continuous domains are still able to fold when one or the other domain is made discontinuous by a circular permutation of the polypeptide sequence (16, 17).

This experimental evidence clearly suggests that the efficiency of the folding process of a continuous structural domain does not strictly depend upon the natural connectivity existing in the native structure. A further consequence of this proposal is that a discontinuity introduced by the artificial insertion of a large sequence, possibly large enough to fold autonomously, could be accepted within a naturally continuous protein domain. This could naturally occur as a consequence of a genetic rearrangement leading to the insertion of a coding sequence within another unrelated coding sequence. Such rearrangement has been observed in natural proteins where domains are inserted or permuted (12). Insertion of a large sequence is generally considered to be incompatible with proper folding of the host protein in which the insertion event took place. However, the observation of accepted natural or engineered discontinuities in structural domains suggests that this intuitive view might be misleading. Large insertion of structural domains could be tolerated as long as the inserted protein is introduced in positions where no steric hindrance occurs. In other words, these types of rearrangement would be more constrained by the stability of the resultant protein than by the folding process of the host protein. We have experimentally tested this assumption and constructed a set of chimeric proteins. In these engineered proteins two different and unrelated proteins of more than 100 amino acids have been inserted in...
several positions of a host protein. The design, expression, and properties of these chimeric proteins are described. Possible functional couplings between inserted and host proteins were investigated.

MATERIALS AND METHODS

Chimera Construction and Expression—The genes coding for Escherichia coli dihydrofolate reductase (DHFR)\(^1\) and \(\beta\)-lactamase (BLA) were amplified by polymerase chain reaction using 100 ng of genomic DNA, 20 units of polymerase, 10 mM of primers, and 10 ng of plasmid pBR322, respectively. An NcoI restriction site was added at the end of the amplified sequences. These fragments were inserted into the phosphoglycerate kinase (PGK) gene carried by the expression vector pYE-PGK (18) in which a unique BspE1 restriction site has been previously introduced by site-directed mutagenesis at the selected positions. The coding sequences of all the chimeras were completely sequenced to check for the absence of unwanted mutations. Wild type (WT) PGK and chimeric proteins were expressed in the Saccharomyces cerevisiae PGK-deficient strain BC3 (19). Crude extracts used for expression efficiency testing were prepared as described (1) and submitted to SDS-polyacrylamide gel electrophoresis (PAGE) (see Fig. 2). Expected molecular masses for PGK/DHFR chimeric proteins, PGK/BLA chimeric proteins, and DHFR (for PGK/DHFR chimeras, WT PGK, and WT DHFR) at various guanidinium chloride concentrations. The gradient was optimized for each particular protein. Fractions containing enzyme activities were pooled and finally purified by gel filtration as described for WT PGK (1). Purified protein used for further characterization appeared as a single band on SDS-PAGE analysis. Protein concentrations were determined spectrophotometrically using the following molar extinction coefficients (20): \(\varepsilon_{280} = 51205 \text{ M}^{-1}\cdot\text{cm}^{-1}\) for PGK/BLA and \(\varepsilon_{280} = 55,375\) for PGK/DHFR constructs. Wild type \(\beta\)-lactamase was produced as described previously (21).

Protein Purification—Crude extracts were precipitated by saturated ammonium sulfate and desalted by chromatography on a Sephadex G-25 column. Protein-containing fractions were purified by HPLC ion exchange on a TSK-DEAE column equilibrated in Tris (20 mM; pH 8) and eluted by a linearly increasing gradient of sodium chloride concentration. The activities were determined following appropriate dilution in activ-

catory products. Enzymatic activity assays were also system-

Equilibrium Unfolding Transitions—Proteins (2 \(\mu\text{g}\), final concentra-

Refolding Kinetic Process—Proteins (70 \(\mu\text{g}\)) were incubated at 22 °C for at least 3 h in 50 mM phosphate buffer, pH 7, containing 10 mM DTT (for PGK/BLA chimeras and WT BLA) or 10 mM 2-mercaptoethanol (for PGK/DHFR chimeras, WT PGK, and WT DHFR) at a final GdmCl concentration of 3.5 M. Refolding was triggered by a 35-fold dilution in the same buffer without GdmCl. Recovery of activity was fitted using a monoeponential term according to the Marquardt algorithm (25).

RESULTS

Design of the Chimeric Protein—The protein used as the host was phosphoglycerate kinase. This protein has been chosen previously by several independent groups (24, 26–31) as a working model for multidomain protein folding. It is a mono-

The abbreviations used are: DHFR, dihydrofolate reductase from E. coli; BLA, \(\beta\)-lactamase from E. coli; DTT, diithiothreitol; GdmCl, guan-
nidinium chloride; PGK, phosphoglycerate kinase from yeast; PAGE, polyacrylamide gel electrophoresis; WT, wild-type.

![Fig. 1. A, structure of yeast PGK. The insertion points are shown in CPK. B, illustrative scheme of the chimeric protein 89BLA.](image-url)
a common problem in bacterial expression systems, could also explain the observed high expression level. However, inclusion bodies are much less frequently observed in yeast expression systems (35). Furthermore, the yeast strain used for expression has a disrupted chromosomal PGK gene and is not able to grow on a glucose-containing growth medium. All the transformed cells expressing the chimeric proteins have recovered the ability to grow on glucose-containing medium. Therefore, the PGK moiety of the chimeric proteins is enzymatically active. A subset of these chimeric proteins has been purified, and the activity of the host as well as of the inserted proteins has been determined. The results reported in Table II reveal that both the refolding rates of these proteins when isolated. The strong effect on their refolding rates shows that the folding process of the inserted proteins is perturbed, in some way, by the connection of their refolding subunits when connected to each other. At least two hypotheses could account for such an effect. First, the folding yield, the observed refolding efficiency remains relatively high in comparison to the in vitro folding efficiency of many natural proteins of similar complexity. There is no clear correlation between the effect on the refolding yield and the point of insertion (type of loop, N-terminal or C-terminal domain) nor with the nature of the inserted proteins.

The reactivation kinetics was followed for the host as well as the inserted proteins. The results reported in Fig. 3 clearly show that the reactivation kinetics of the host protein is slowed down with respect to wild-type PGK. Although their internal connectivity is not modified by their insertion in the host protein, the reactivation rate of the inserted protein (DHFR or BLA) is systematically reduced with respect to the refolding rates of these proteins when isolated. The strong effect on their refolding rates shows that the folding process of the inserted proteins is perturbed, in some way, by the connection of their extremities to the host protein. Although the physical basis for this effect is not clear, this observation suggests strongly that the inserted and host proteins, although constituting autonomous folding units when isolated, are no longer independent folding subunits when connected to each other. At least two hypotheses could account for such an effect. First, the folding process of the inserted polypeptide chain could be affected by transient interactions with other parts of the polypeptide chain outside its own folding unit. Alternatively, during the folding process, the volume statistically occupied by the PGK polypeptide chain fragments linked to each end of the inserted protein, the reactivation rate of the inserted protein (DHFR or BLA) is systematically reduced with respect to the refolding rates of these proteins when isolated. The strong effect on their refolding rates shows that the folding process of the inserted proteins is perturbed, in some way, by the connection of their extremities to the host protein. Although the physical basis for this effect is not clear, this observation suggests strongly that the inserted and host proteins, although constituting autonomous folding units when isolated, are no longer independent folding subunits when connected to each other. At least two hypotheses could account for such an effect. First, the folding process of the inserted polypeptide chain could be affected by transient interactions with other parts of the polypeptide chain outside its own folding unit. Alternatively, during the folding process, the volume statistically occupied by the PGK polypeptide chain fragments linked to each end of the inserted protein could restrict the dynamic possibilities of the inserted polypeptide chain and in this way hinder the folding process of the inserted moiety.

The folding process of BLA and DHFR has been described in detail (33, 34) and is a complex process involving parallel pathways with relaxation times ranging from the millisecond time scale to tens of seconds. In both isolated proteins the formation of a large part of the secondary structure occurs on a time scale shorter than one second. The recovery of enzyme activity used as a folding probe in this work corresponds to the very last step in the folding process and is controlled by slow

---

**TABLE I**

**Sequences of the constructed chimeric proteins**

Numbers are related to the sequences of wild type proteins. Underlined represents sequence of yeast phosphoglycerate kinase (415 residues). Italic font represents linker residues. Bold font represents sequence of the inserted β-lactamase (263 residues) and dihydrofolate reductase (159 residues).

| Chimeric proteins | Denomination | Sequence |
|-------------------|--------------|----------|
| PGK/BLA          | 71BLA        | N11SGGH12P1...H282W283GAGD72... |
|                   | 89BLA        | L17SGGH18P1...H282W283GAGD72... |
|                   | 129BLA       | S17SGGH18P1...H282W283GAGD72... |
|                   | 290BLA       | S17SGGH18P1...H282W283GAGD72... |
|                   | 71DHFR       | N11SGGH12P1...R158R159GAGE72... |
|                   | 89DHFR       | L17SGGH18P1...R158R159GAGE72... |
|                   | 290DHFR      | S17SGGH18P1...R158R159GAGE72... |
| PGK(DHFR + BLA)   | 89DHFR/290BLA| L17SGGH18P1...R158R159GAGD72...S290GGH12P1...H2462W2463GAGA72... |

---

**TABLE II**

**Enzymatic activities of the chimeric proteins**

| Protein  | Host     | Insert    | $k_{cat}$ (s$^{-1}$) | %     | $k_{cat}$ (s$^{-1}$) | %     |
|----------|----------|-----------|----------------------|-------|----------------------|-------|
| wt-PGK   |          |           | 1480                 | 100   | -                    | -     |
| wt-BLA   |          |           | 758                  | 100   | -                    | -     |
| wt-DHFR  |          |           | 8.6                  | 100   | -                    | -     |
| 71BLA    |          |           | 681                  | 46    | 450                  | 59    |
| 89BLA    |          |           | 326                  | 22    | 462                  | 61    |
| 129BLA   |          |           | 296                  | 20    | 87                   | 11    |
| 290BLA   |          |           | 555                  | 38    | 469                  | 62    |
| 71DHFR   |          |           | 503                  | 34    | 3.9                  | 45    |
| 89DHFR   |          |           | 681                  | 46    | 5.4                  | 62    |
| 290DHFR  |          |           | 370                  | 25    | 4.8                  | 56    |
**TABLE III**

Unfolding transition and refolding kinetics parameters for chimeric proteins

| Protein  | Cm (M) | Insert | N | Kinetics refolding | Yield of reactivation |
|----------|--------|--------|---|---------------------|----------------------|
|          | Host   | Insert |   | s⁻¹                 | %                    |
|          |        |        |   |                     |                      |
| wt-PGK   | 0.70   | -      | 19.0⁺ | (5.8 ± 1.0) × 10⁻⁴ | 64 ± 6               |
|          | wt-BLA | 1.0    | -   | (4.0 ± 0.5) × 10⁻⁴ | 53 ± 2               |
|          | wt-DHFR| 2.3    | -   | (4.1 ± 0.5) × 10⁻⁴ | 38 ± 3               |
| 71BLA    | 0.30 ± 0.03 | 0.31 ± 0.02 | 4.2 ± 1.3 | 5.2 ± 1.4 | 64 ± 6 | 80 ± 4 |
| 89BLA    | 0.36 ± 0.01 | 0.39 ± 0.02 | 10.6 ± 1.9 | 8.9 ± 1.7 | 53 ± 2 | 84 ± 3 |
| 129BLA   | 0.22 ± 0.02 | 0.30 ± 0.03 | 3.4 ± 0.7 | 4.4 ± 1.4 | 22 ± 2 | 28 ± 3 |
| 290BLA   | 0.36 ± 0.03 | 0.37 ± 0.02 | 6.9 ± 1.6 | 8.1 ± 1.8 | 27 ± 2 | 31 ± 2 |
| 71DHFR   | 0.35 ± 0.02 | 0.33 ± 0.04 | 7.7 ± 2.4 | 2.6 ± 0.7 | 51 ± 4 | 64 ± 4 |
| 89DHFR   | 0.57 ± 0.04 | 0.31 ± 0.02 | 7.8 ± 2.5 | 2.2 ± 0.4 | 35 ± 2 | 51 ± 3 |
| 290DHFR  | 0.38 ± 0.09 | 0.52 ± 0.11 and | 1.9 ± 1.5 | 3.4 ± 1.7 and | 38 ± 3 | 35 ± 3 |

---

**FIG. 3. Kinetics of refolding monitored by enzymatic activity assay (see “Materials and Methods”).** A, PGK/BLA chimeric proteins. B, PGK/DHFR chimeric proteins. Open symbols, refolding of the host protein (PGK activity). Filled symbols, refolding of the inserted protein (BLA or DHFR activity).

The stability of the chimeric proteins was evaluated by monitoring the loss of each activity induced by increasing guanidinium hydrochloride concentrations. The losses of activity are not fully reversible, and therefore, the transition curves cannot be used to derive a true estimate of each protein’s thermodynamic stability. However, the transition midpoint and the cooperativity of the transition curves (Table III and Fig. 4) clearly reveal that the chimeric proteins are less resistant to denaturation and presumably less stable than are the wild type proteins.

The comparison of the denaturation curves of the chimeras with their parent proteins shows that, in all chimeras but one, the host as well as the inserted proteins are destabilized by the insertion. A clear correlation is observed between the unfolding transition curves and the nature of the inserted protein. In the three BLA-PGK proteins, the host and inserted proteins lose their activity concomitantly as the denaturant concentration increases. This suggests that the stability of the two partners is strongly coupled. This coupling is observed with BLA inserted in PGK at various positions and therefore is not related to particular local interactions between the two partners.

In the DHFR-PGK chimeras, host and insert unfolding curves depend on the position of insertion. In particular, for 290DHFR the host and inserted protein are almost independent. A first transition below 0.4 M corresponds to the unfolding of PGK as shown by the loss of PGK activity and by a strong decrease in the far UV circular dichroism signal. The unfolding of PGK occurring in this first transition does affect DHFR activity but only to a limited extent, suggesting that the DHFR structure is largely unaffected by the unfolding of associated PGK. A second transition corresponding to the inserted DHFR is observed at a much higher denaturant concentration, in the range expected for isolated DHFR. This type of unfolding process is, surprisingly, the only observed example in this set of chimeric proteins of an independent behavior of the inserted protein. In this case, the host protein is destabilized by the discontinuity created by the insertion, whereas the inserted protein remains largely unaffected.

**Functional Interactions between Host and Inserted Modules**—It is generally observed that there is little functional interaction between the two partners in “classical” end-to-end fusion proteins. The presumably stronger structural coupling present in internal-fusion proteins could possibly induce functional interactions between host and inserted proteins (36). The activity either of the host or inserted proteins were measured in the presence of the substrates of the other protein. We observed in three different cases that the presence of the substrates of one protein does modulate positively or negatively the activity of the other one (Fig. 5). Control experiments have shown that there are no such effects for the wild type proteins. These results demonstrate that functional coupling between the two partners of an internal fusion protein is not a rare occurrence, although these interactions were not explicitly designed in these chimeric proteins.
Functional Inserted Proteins within Proteins

**FIG. 4.** Unfolding transition curves assessed by the variation of enzymatic activity. A, PGK/BLA chimeric proteins. B, PGK/DHFR chimeric proteins. Open symbols, refolding of the host protein (PGK activity). Filled symbols, refolding of the inserted protein (BLA or DHFR activity).

**FIG. 5.** Functional interactions observed for three chimeric proteins (89BLA, 290BLA, and 290DHFR). Enzymatic activities are expressed relative to that measured in the absence of the tested ligand 3-PG (3-phosphoglycerate + ADP + Mg^{2+}) or DH (dihydrofolate).

**DISCUSSION**

These experiments reveal that two different unrelated proteins of more than 150 amino acids can be inserted in four different positions of a host protein. The host as well as the inserted protein are able to fold in a functional conformation. The main consequences of the insertions are a loss of stability of the host and inserted proteins and a partial loss of reversibility of the *in vitro* refolding process.

The loss of stability of the host protein is in part due to the perturbation of the local sequence in which the insertion took place, and it seems plausible that the large discontinuity created by the insertion is responsible for the destabilizing effect on the host protein. It is somehow more surprising that in all but one chimeric proteins analyzed (290DHFR), the inserted protein is also strongly destabilized.

The loss of reversibility, although clearly significant, remains relatively moderate in view of the perturbation introduced. No attempts were made to improve the folding reversibility by manipulating the refolding conditions or the temperature refolding protein concentration, which are often necessary to improve the refolding yield of many natural proteins of a similar complexity. This clearly supports the view that, at least with yeast PGK, the sequence continuity of a structural domain is not an essential requirement for a productive folding process. The positions of the insertion points were selected on a very general criterion: They are located on surface loops not directly involved in the catalytic site. These positions were not selected on the basis of a possible folding mechanism or as the result of a genetic screening. An elegant series of genetic experiments on maltose binding protein has shown that several regions of this protein, including secondary structure elements, are permissive regions relative to insertion. These regions have recently been used as insertion sites for β-lactamase, and the two chimeric proteins were shown to be active (21). A chimeric protein made by the insertion of a xylanase domain in a permuted glucanase domain has been recently described (37), showing that the structures of the inserted and host partners are preserved in the fused protein. This result demonstrates that the Jellyroll fold is dominated by long range intramolecular contacts.

The present results further support the idea that large structured insertions can be artificially introduced within a protein. All the designed chimeric proteins, those with two different proteins in four various positions as well as the triple chimera, are biologically active. It should be emphasized that loops involved in the substrate binding sites were not retained as insertion points for practical reasons rather than for folding or structural principles. The idea was simply to preserve the host activity as a convenient way to monitor the proper folding of the host protein. The tolerance observed in this work suggests that insertions in the loops involved in substrate binding would be structurally tolerated. Moreover, as reported by others with maltose binding protein (21), “permissive” regions selected by a genetic screen were shown to occur within secondary structures, showing that the location of an insertion point in a loop is not a mandatory criterion.

Therefore, it seems that tolerance to a large insertion is not restricted to a very limited region of the polypeptide chain (21) nor to a specific topology (37). These results, together with tolerance to circular permutation, clearly support the idea that protein folding does not involve an obligatory step-by-step sequence of structure-building events that cannot tolerate perturbation. It is more suggestive of a process that can lead to the same final structure through different possible sequences of events as originally suggested by Harrison and Durbin (38).

To evaluate the possible evolutionary implications of these results, it must be stressed that the rather general criterion used in this study to select the insertion points does induce restrictions relative to the probability of a random protein insertion event occurring during natural evolution. First, the inserted protein is a folded protein and its extremities are close to one another. Furthermore, if the activity of the two fused partners is under a selective pressure, the insertion points must be compatible with the functional integrity of the two fused partners. Finally, these criteria might be restrictive enough to explain why large insertion and discontinuous do-
mains are not the most common situation observed during protein evolution. New activating or inhibitory effects transmitted between the two fused partners were observed, although this was not intentionally designed. These occurred in three cases in the subset of proteins examined and suggest that insertion of a binding domain within a different protein is a possible evolutionary solution toward new regulatory properties of originally simpler proteins.

Further studies will be necessary to know whether or not the specific properties of the host and inserted proteins used in the present study are responsible for this surprising tolerance. Comparison of natural PGK sequences shows that PGK genes found in Trypanosoma brucei or the related organism Chrytidia fasciculata have insertions in one of the regions selected in this study (position 73). In one of the genes, the insertion is relatively short: 16 amino acids (39). The structure of this protein has been determined and shows that it is folded as a protruding helical region. In another gene of the same organisms and in the same position, there is a much larger insertion of about 100 amino acids. The function and cellular localization of this gene product, which was originally qualified as a PGK-related protein, is not known. Together with the conservation of all the catalytic residues, our results suggest that this protein is a functional PGK. The other positions that were used as insertion sites in the present study do not correspond to positions where natural insertions are observed in other PGK genes.

Relative to many small single domain proteins currently used as folding models, PGK is a relatively large and slow-folding protein (15). Therefore, it could be suggested that the time required for PGK folding would be largely sufficient for the inserted protein to fold first. In other words, the slow folding rate of PGK could be directly responsible for this a priori unusual tolerance to insertions. However, the results show that the folding time range of PGK as well as of DHFR and BLA is considerably slowed down when fused together, suggesting that the folding rates of the proteins taken separately are no longer preserved in the chimeras. The folding rates of isolated proteins might therefore not constitute a relevant criterion to design new properly folded chimeric proteins. The constraint on the relative time range of the inserted and host protein could only be revealed by further experiments on other chimeric proteins. Because the inserted proteins are devoid of any discontinuity in the chimeras, their significant destabilization and slowing down of the refolding process were not predicted by the strictly modular view of the multidomain protein-folding process. Instead, the present results suggest that in terms of folding rate and stability, autonomous folding units do not remain autonomous once their polypeptide chains are connected. These data also show that the folding process of such complex proteins remains relatively efficient in vivo as well as in vitro despite the major changes in the molecular organization and in the folding kinetics of the individual partners. This supports the idea that, if the folded state is stable, the folding process is not highly constrained and does not need to be strictly organized or coordinated.

Remarkably, the extremely efficient yeast PGK expression system used in this work remained so far relatively inefficient to express proteins other than PGK at a very high level. However, the inserted proteins are, in the context of the chimeric proteins expressed, at a very high level and are correctly folded in vivo. It would be useful to verify whether this very high expression efficiency is preserved with other inserted proteins. Then, the yeast expression system could be used to produce large amounts of proteins that are difficult to express or to refold in bacterial systems.

This study raises new questions and opens experimental routes to more systematic studies on the spectrum and properties of the allowed sequences inserted in different sites of a protein. For example, it seems intuitive that a long inserted sequence must necessarily be folded in order to be functionally accepted by the host protein. This property could constitute the basis of a possible and interesting new screening method to select for the folding ability of random amino acid sequences. The fraction of the sequence space giving rise to a folded sequence is indeed an important but still open question.

Acknowledgments—We are very grateful to John O’Neill from C. R. Matthews’ laboratory (Pennsylvania State University) for providing purified WT DHFR for the control experiments.

REFERENCES

1. Minard, P., Hall, L., Betton, J. M., Missiakas, D., and Yon, J. M. (1989) Protein Eng. 3, 55–60
2. Markovic-Housley, Z., Cooper, A., Lustig, A., Flukiger, K., Stoel, B., and Erni, B. (1994) Biochemistry 33, 10977–10984
3. Neck, S., Grillenbeck, N., Ahmadian, M. R., Ribeiro, S., Kreutzer, R., and Sprinzl, M. (1996) Eur. J. Biochem. 234, 132–139
4. Mayr, E. M., Jaenicke, R., and Glockshuber, R. (1997) J. Mol. Biol. 269, 260–269
5. Gill, C. V., Bowers, K. E., and Matthews, C. R. (1997) Protein Sci. 6, 1855–1892
6. Hoszus, L. L., Craven, C. J., Parker, M. J., Lorch, M., Spencer, J., Clarke, A. R., and Waltho, J. P. (1997) Nature Struct. Biol. 4, 801–804
7. Jaenicke, R. (1998) Biol. Chem. 379, 237–243
8. Russell, R. B. (1994) Protein Eng. 7, 1407–1410
9. Martin, J. L., Bardwell, J. C., and Kuriyan, J. (1993) Nature 365, 464–468
10. Levine, M., Murthead, H., Stammers, D. K., and Stuart, D. I. (1978) Nature 271, 626–630
11. Delarue, M., Poch, O., Tordo, N., Moras, D., and Argos, P. (1990) Protein Eng. 3, 461–467
12. Russell, R. B., and Ponting, C. P. (1998)Curr. Opin. Struct. Biol. 8, 364–371
13. Jensen, S., Stewart, M., Michie, A., Swindells, M. B., Orengo, C., and Thornton, J. M. (1998) Protein Sci. 7, 233–242
14. de Prat-Gay, G. (1996) Protein Eng. 9, 843–847
15. Pecorari, F., Guilbert, C., Minard, P., Desmadril, M., and Yon, J. M. (1996) Biochemistry 35, 3465–3476
16. Rito-Vonsiovici, M., Minard, P., Desmadril, M., and Yon, J. M. (1995) Biochemistry 34, 1643–1651
17. Heinemann, U., and Hahn, M. (1995) Prog. Biophys. Mol. Biol. 64, 121–143
18. Minard, P., Bowen, D. J., Hall, L., Littlechild, J. A., and Watson, H. C. (1990) Protein Eng. 3, 515–521
19. Pipper, P. W., and Curran, B. P. G. (1990) Curr. Genet. 17, 119–123
20. Pace, N. C., Vajdos, F., Pee, L., Grimesley, G., and Gray, T. (1995) Protein Eng. 4, 2411–2423
21. Betton, J. M., Jacob, J. P., Hofnung, M., and Bromme-Smith, J. (1997) Nature Biotech. 15, 1276–1279
22. Iwakura, M., and Matthews, C. R. (1992) Nature Struct. Biol. 8, 364–371
23. Ritco-Vonsovici, M., Minard, P., Desmadril, M., and Yon, J. M. (1995) Biochemistry 34, 397–404
24. Ploeg, W. H., Flannery, B. P., Tenkolsky, S. A., and Vetterling, W. T. (1986) in Numerical Recipes, pp. 523–528, Cambridge University Press, Cambridge, UK
25. Garcia, P., Mélola, F., Receveur, Y., Blandin, P., Minard, P., and Desmadril, M. (1988) Biochemistry 37, 7444–7455
26. Grattinger, M., Danelkreiter, A., Schurig, H., and Jaenicke, R. (1998) J. Mol. Biol. 280, 525–533
27. Hoszus, L. L., Craven, C. J., Spencer, J., Parker, M. J., Clarke, A. R., Kelly, M., and Waltho, J. P. (1997) Protein Sci. 6, 333–340
28. Adams, B., Fowler, R., Hudson, M., and Pain, R. H. (1996) FEBS Lett. 385, 101–104
29. Lillo, M. P., Szpikowska, B. K., Mas, T., Sutin, J. D., and Beecham, J. M. (1997) Biochemistry 36, 11273–11281
30. Damaschun, G., Damaschun, H., Gast, K., Misselwitz, R., Müller, J. J., Pfeil, W., and Zirwer, D. (1993) Biochemistry 32, 7739–7746
31. Betton, J. M., Martinez, P., Saurin, W., and Hofnung, M. (1993) Location of tolerated insertions/deletions in the structure of the maltose binding protein.
32. FEBS Lett. 325, 34–38
33. Jennings, F. A., Finn, B. E., Jones, B. E., and Matthews, C. R. (1995) Nature Struct. Biol. 2, 378–389
34. Vanhove, M., Raquet, X., Palzik, T., Pain, R. H., and Frere, J. M. (1996) Proteins 25, 104–111
35. de Cubas-Clark, E., and Georgiou, G. (1991) in Protein Refolding (Georgiou, G., ed.) pp. 1–20, American Chemical Society, Washington, DC
36. Benito, A., Feliu, J. X., and Villaverde, A. (1996) J. Biol. Chem. 271, 21251–21256
37. Ay, J., Goz, P., Borris, R., and Heinemann, U. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 6613–6618
38. Harrison, S. C., and Durbin, R. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 4028–4030
39. Bernstein, B. E., Michels, P. A., and Hol, W. G. J. (1997) Nature 385, 275–278