Force-Clamp Spectroscopy Detects Residue Co-evolution in Enzyme Catalysis*[S]

Raul Perez-Jimenez[‡1], Arun P. Wiita[‡§], David Rodriguez-Larrea[¶], Pallav Kosuri[‡||], Jose A. Gavira[**], Jose M. Sanchez-Ruiz[¶], and Julio M. Fernandez[‡2]

From the ‡Department of Biological Sciences and the §Graduate Program in Neurobiology and Behavior, Columbia University, New York, New York 10027, the ¶Facultad de Ciencias, Departamento de Quimica Fisica, Universidad de Granada, 18071, Granada, Spain, the ||Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10027, and the **Laboratorio de Estudios Cristalográficos, Instituto Andaluz de Ciencias de la Tierra (Consejo Superior de Investigaciones Cientificas-Universidad de Granada), Parque Tecnologico Ciencias de la Salud, 18100 Granada, Spain

Understanding how the catalytic mechanisms of enzymes are optimized through evolution remains a major challenge in molecular biology. The concept of co-evolution implicates that compensatory mutations occur to preserve the structure and function of proteins. We have combined statistical analysis of protein sequences with the sensitivity of single molecule force-clamp spectroscopy to probe how catalysis is affected by structurally distant correlated mutations in *Escherichia coli* thioredoxin. Our findings show that evolutionary anti-correlated mutations have an inhibitory effect on enzyme catalysis, whereas positively correlated mutations rescue the catalytic activity. We interpret these results in terms of an evolutionary tuning of both the enzyme-substrate binding process and the chemistry of the active site. Our results constitute a direct observation of distant residue co-evolution in enzyme catalysis.

The acquisition of adequate activity by an enzyme is the result of natural selection through mutagenesis (1). Mutations that represent a functional advantage may prevail, reflecting not only the optimization of function but also the emergence of a new enzymatic mechanism (1–4). Disadvantageous mutations can be compensated by mutations in other positions (5–8), as proposed by the concept of co-evolution. This hypothesis suggests that pairs of positions in the sequence evolve together in such a way that mutations in one of the residues induce a coherent response in the other to compensate for a certain functional or structural effect (5, 7). Over the last decades, numerous statistical methods have been applied to detect residue co-evolution in protein sequences (9, 10). Anal-ysis of co-evolving residues has been used to explore functional coupling in processes like protein-protein interaction (11), gating of potassium channels (12), and allosteric communications in proteins (13). Experimental validation has been provided for the identification of networks of energetically coupled residues mediating allostery (14–16). However, in the case of enzyme catalysis, only theoretical studies have been reported (17). The limitation of the available techniques in providing insights into the mechanism of catalysis has made it difficult to experimentally validate co-evolution as a natural means of catalysis optimization. Here, we have combined statistical sequence analysis and single molecule force-clamp spectroscopy to probe the effect of evolutionary correlated mutations in the *Escherichia coli* Trx catalytic mechanism, a ubiquitous disulfide bond reductase.

In a previous study, we developed single molecule force-clamp spectroscopy as a novel sensitive tool to study the mechanism of the catalytic reaction of *E. coli* Trx (18). We showed that Trx reduces disulfide bonds using two distinct force-dependent catalytic mechanisms. The first mechanism is favored at low forces (up to 200 pN),[3] whereas the second mechanism is predominant at high forces (over 200 pN). These two chemical reduction mechanisms are likely to have evolved to allow for the thiol/disulfide exchange reaction to proceed under very different types of stress environments. Understanding the underlying mechanism of disulfide reduction under force is of great interest because force-dependent thiol reactions in cytoskeletal proteins have been shown to occur in cells (19).

Additionally, we also showed that the active site mutation P34H, a natural mutation occurring in some members of the Trx superfamily, had a negative effect in the low force mechanism, whereas the high force mechanism remained unchanged. These studies show that force-clamp spectroscopy is an ideal tool to investigate the sequence-function relationship in Trx because it distinguishes different enzymatic mechanisms and reports precise chemical changes upon sequence alteration.

Considering that the P34H mutation significantly impaired the chemistry of *E. coli* Trx, we have used a simple statistical tool for sequence alignment analysis to investigate the presence

of co-evolving residues (20). We have found that residues Asp[26] and Gly[74] show the highest positive and highest negative values of correlation with residue Pro[34], respectively. Interestingly both residues are situated over 10 Å far away from residue Pro[34]. Our force-clamp experiments, using the single and double mutant forms, show a correlation between the activity of the Trx mutant forms and the co-variance value, which cannot be detected by regular bulk enzymatic assays. Significant changes in the force dependence of the reduction rate for these mutant forms can be detected. Whereas the anti-correlated mutations had a negative effect on the Trx activity, the positively correlated mutations returned the activity to that of the wild type form. We thus corroborate not only that the accumulation of disadvantageous mutations has been avoided by purifying selection to maintain a minimum level of activity but also that mutations that perturb the enzymatic mechanism can and have been reverted by noninteracting mutations. These results are in agreement with the existence of an activity threshold (21). Because residues Pro[34] and Gly[74] are positioned within the catalytic groove and residue Asp[26] influences the protonation state of the reactive cysteine Cys[32], we interpret these results in terms of both an evolutionary optimization of the enzyme-substrate binding affinity and the protonation state of the Trx active site. These results represent an important advance in obtaining information on enzyme catalysis by combining the evolutionary analysis of sequences with the experimental study of catalysis at the single molecule level, a method that could be applied to enzyme design.

EXPERIMENTAL PROCEDURES

Sequence Alignments—We used BLAST 2 to search the non-redundant data base UniProt/Swiss-Prot with the *E. coli* Trx sequence as query and the default options of the search. The sequences found were aligned to the Trx sequence using the Smith-Waterman algorithm (22), and those with similarity with the query higher than 0.25 were selected obtaining a total of 239 sequences. We assumed that most of the retained sequences share the fold of the query protein.

Single Molecule Force-Clamp Spectroscopy and Data Analysis—The details of our custom-made atomic force microscopy apparatus have been previously described (23). Data acquisition cards 6052E and 6703 (National Instruments) were used to control the atomic force microscopy head. Si_3N_4 cantilevers (Veeco) were calibrated in solution using the equipartition theorem. The typical spring constant was ~20 pN/nm. The atomic force microscopy was used in force-clamp mode with an extension resolution of ~0.5 nm and a feedback response of ~5 ms. All of the experiments were performed in 10 mM HEPES buffer with 150 mM NaCl, 1 mM EDTA, 2 mM NADPH, pH 7.2. Before the experiment, Trx reductase was added to a final concentration of 50 nM. The different forms of Trx used were added to yield the desired final concentration. A double-pulse protocol was used to collect data. The first pulse was set at 185 pN of pulling force and held for 0.4 s. The second pulse could be set at different forces and was held for over 10 s. To capture every possible reduction event, the second pulse was maintained for over 25 s in the case of slower kinetics. In addition, we performed the experiments with the double mutant

P34H/G74S at different concentrations to ensure the accuracy of our fitting in the low-force regime. In each force the traces were averaged and fitted to a single exponential. Standard errors in the reduction rates were calculated using the bootstrapping method. The kinetic model used for the interpretation of the experimental data (see Fig. 4A) has been extensively described elsewhere (18). We used software written in IGOR Pro (WaveMetrics) for data collection and analysis.

Protein Expression and Purification—Preparation of $(I27_{SS})_8$ polyprotein was prepared as described previously (18, 24) with a few differences. Two residues were mutated in the I27 module using a QuikChange site-directed mutagenesis kit (Stratagene) to engineer a buried intracellular disulfide bond. Concretely, the mutations are G32C and A75C. Then we followed a multi-step cloning procedure to construct an *N-C*-linked polyprotein gene with eight domains (25), $(I27_{G32C-A75C})_{8-}$. This gene was inserted into the pQE80 vector and transformed and expressed in the recombination-defective *E. coli* strain BLR (DE3) (Novagene). The His$_6$-tagged proteins were purified first in a Co^{2+} column (Clontech) using a Talon resin (BD Biosciences) and second by size exclusion chromatography using a Superdex 200 HR 10/30 column (GE Healthcare). The proteins were then stored at 4 °C in 10 mM HEPES, 150 mM NaCl, and 1 mM EDTA, pH 7.2.

Preparation of mutant forms of *E. coli* thioredoxin was carried out as described elsewhere (26) with some modifications. Thioredoxin gene was cloned in a pET30a vector, and the mutations were carried out in the same way as for I27, using the QuikChange kit. The plasmid in each case was transformed and expressed in BL21(DE3) supercompetent cells and grown in LB medium with 40 μg/ml of kanamycin at 37 °C overnight. The cell pellets were resuspended in 1 mM EDTA, 30 mM Trizma buffer, pH 8.3, and a French press was used for lysis. After centrifugation, the nucleic acids in the supernatant were precipitated by stirring with streptomycin sulfate (10% w/v) at 4 °C overnight. The supernatant was centrifuged at 20,000 rpm for 40 min, filtered, and loaded onto a Sephacryl S-100 size exclusion column (GE Healthcare) equilibrated in the 1 mM EDTA, 30 mM Trizma buffer, pH 8.3. Afterward, a second ion exchange column was used, and Fractogel® EMD DEAE (M) (Merck) equilibrated in 1 mM EDTA, 30 mM Trizma buffer, pH 8.3, was used. The protein was eluted by a linear gradient between 0 and 0.5 M NaCl, and SDS-PAGE gel electrophoresis was used to verify the purity of the proteins. The protein fractions were dialyzed into a buffer of 10 mM HEPES, 150 mM NaCl, 1 mM EDTA, pH 7.2, and stored at −20 °C. Trx concentration was determined spectrophotometrically at 280 nm using an extinction coefficient of 13700 M^{-1} cm^{-1}.

Thioredoxin Bulk Enzymatic Assays—We tested the effect of the mutations by traditional bulk experiments. We carried out two different assays for all the mutant forms and the WT. The first one is the insulin turbimetric assay (27). In this assay thioredoxin catalyzes the reduction of insulin by DTT. The reaction is initiated by combining 0.5 mg/ml of bovine pancreatic insulin, 5 μM Trx, and then adding DTT to a final concentration of 500 μM. The rate of precipitation of the β-chain of insulin can be monitored at 650 nm.
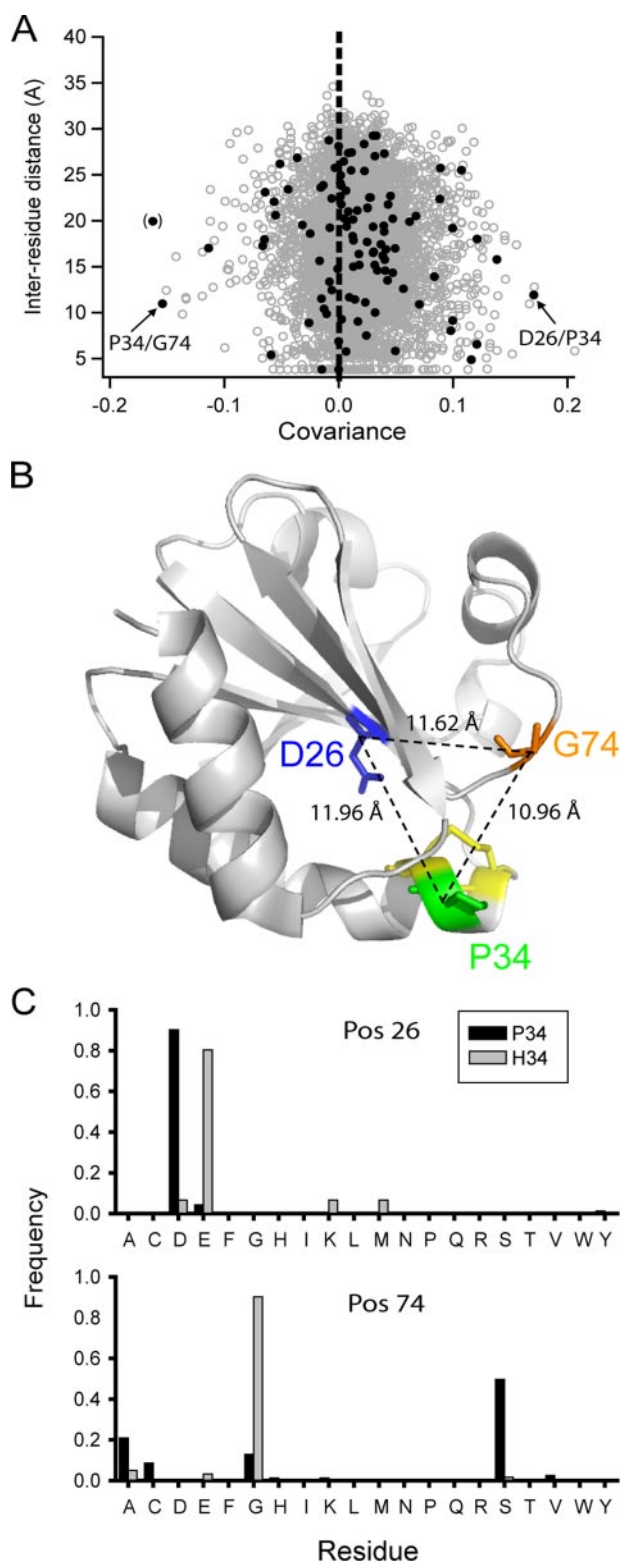
FIGURE 1. **Identification of co-evolving residues in the enzyme Trx.** *A*, plot of inter-residue distance *versus* co-variance for *E. coli* Trx. In particular, the *black circles* mark the co-variance between residue Pro[34] and the rest of the sequence. Residues Asp[26] and Gly[74] present the highest positive and highest negative values of co-variance with Pro[34], respectively. Both of these residues are about a nanometer away from the active site. Because of multiple gaps in the Trx MSA, we did not consider the double mutant in positions 7 and 34 indicated by *parentheses*. *B*, *E. coli* Trx structure (Protein Data Bank code 2trx). Correlated residues Asp[26], Pro[34], and Gly[74] are highlighted. The distance between each pair of residues is indicated. Active site cysteines are colored *yellow*. *C*, residue frequencies from the Trx MSA in position 26 (*upper panel*)

The second test is the 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) assay (28). In this assay the disulfide bond of Trx is reduced in the presence of NADPH in a reaction catalyzed by Trx reductase. The reaction is initiated by adding 50 nM Trx reductase, 0.25 mM NADPH, 0.5 mg/ml of bovine serum albumin, 5 μM Trx, and 500 μM of DTT. The reduced form of thioredoxin can then reduce DTNB giving rise to TNB, the formation of which was monitored at 412 nm. The effect of the mutation P34H had been tested earlier with the DTNB assay (29).

RESULTS

Statistical Basis for Selection of E. coli Trx Mutants—Previous studies have demonstrated the functional importance of residue Pro[34] in *E. coli* Trx (18, 29). We suggest that mutations in this position during evolution should reflect co-evolving alterations in the Trx sequence with a functional meaning. To depict co-evolving residues in the Trx sequence and particularly with residue Pro[34], we have carried out a statistical analysis of the Trx multiple sequence alignment (MSA). We have obtained the distance distribution of correlated mutations using as a correlation measure the co-variance between each pair of positions in the Trx MSA (5778 pairs in total) (20). Considering the distance between residues (C_α-C_α distances from the Protein Data Bank file), we can discriminate between correlations among contact residues (~5 Å) and distant residues. Correlations between proximal residues have been suggested to preserve the structure and stability, whereas correlations between distant sites have a functional significance (9).

The co-variance measures the extent to which two positions (*i* and *j*) vary together and can be defined as follows,

$$\text{Cov}_{i,j} = \sum_{\text{sequences}} \frac{(\delta_i - <\delta_i>) \cdot (\delta_j - <\delta_j>)}{N_S}$$

$$= <\delta_i \cdot \delta_j> - <\delta_i> \cdot <\delta_j> \quad \text{(Eq. 1)}$$

where N_S is the total number of sequences, and the sum is over all sequences. The value of δ_P (with *P* equal to *i* or *j*) is 1 for a given sequence if the residue at position *P* in the sequence is the same as in the Trx sequence and takes a value of 0 otherwise, $<\delta_P>$ is the average value of δ_P (*i.e.* $\Sigma \delta_P / N_S$). The co-variance calculation can be considered itself a measure of the evolutionary pressure because high values of co-variance reveal a high degree of residue co-evolution.

We have found that residue Pro[34] shows the highest positive and the highest negative values of co-variance with residues Asp[26], located at ~12 Å, and Gly[74], located at ~11 Å from Pro[34], respectively (position 7 presents a higher negative value of co-variance with residue 34 that position 74, but we did not consider it because it contains numerous gaps in the MSA) (Fig. 1, *A* and *B*). In other words, a double mutation in positions 34 and 74 is very unlikely, whereas a mutation in position 34 is very

and position 74 (*lower panel*) before (*black bars*) and after (*gray bars*) a perturbation in residue Pro[34]. In position 26 the residue distribution sharply shifts from Asp to Glu, whereas position 74 becomes almost solely occupied by Gly.

often accompanied by a mutation in position 26. The fact that mutations in positions 26, 34, and 74 naturally occur following the pattern captured by the co-variance suggests that those positions are under evolutionary pressure.

The residue Asp[26] plays an important role in the protonation state of the active site acting as an acid in oxidation reactions and a base in reduction reactions catalyzed by Trx (30, 31). In the case of residue Gly[74], there is no evidence of its role in disulfide reduction, but it has been suggested to have importance in the interaction between bacteriophage T7 DNA polymerase and *E. coli* Trx (32), because it may represent an interaction point. In addition, position 74 has been proposed to be determinant in the Trx substrate specificity (33). Because of the individual functional importance of each residue, these correlations suggest an association between these residues that may have driven the functional optimization and/or diversification of Trx through evolution.

A residue frequency analysis in positions 26, 34, and 74 shows that the most common mutation in each case is, respectively, Asp[26] → Glu, Pro[34] → His, and Gly[74] → Ser (supplemental Fig. S1). A perturbation of Pro[34] can be created by extracting from the MSA a subalignment that contains only His at this position (15). The perturbation clearly alters the residue frequency in positions 26 and 74. Position 26 shifts drastically from Asp to Glu and position 74 becomes almost exclusively occupied by Gly (Fig. 1C). Concretely, this analysis corroborates that a double mutation D26E/P34H commonly occurs during evolution, whereas double mutation P34H/G74S is extremely unlikely. We hypothesize that this natural tendency is reflected in Trx activity. Residue co-evolution has been extensively studied in terms of energetic coupling (13, 15, 16, 34). In this work, the co-variance calculation reflects statistical coupling that is not necessarily associated with energetic coupling (35).

In addition to studying P34H Trx, we obtained single mutants D26E and G74S and double mutants D26E/P34H and P34H/G74S and tested their activity. To rule out the possibility of structural effects caused by the mutations, we compared the crystal structure of WT Trx (Protein Data Bank 2trx) with the crystal structures of P34H Trx (Protein Data Bank code 2fd3; resolution, 2.45Å; *R* value, 0.185; R_{free}, 0,305) and G74S Trx (Protein Data Bank code 2fch; resolution, 2.60Å; *R* value, 0.203; R_{free}, 0.276) (supplemental Fig. S2). Even though the crystal structures cannot exclude the possibility of chemical changes on the sub-Ångström level, the basic statement here is that the proteins maintain the thioredoxin fold.

Force-Clamp Experiments—Atomic force microscopy has become a powerful tool to study the mechanical properties of proteins at the single molecule level. In the force-clamp mode, a controlled pulling force can be directly applied to a single molecule, which makes it possible to monitor its unfolding and refolding trajectories as a function of time (36, 37). Although previous works have shown the ability of atomic force microscopy to study single bond ruptures (38, 39), we have recently developed force-clamp as a new tool to study mechanochemistry. By applying a calibrated force to a disulfide bonded substrate and in the presence of a particular reducing agent, single disulfide reduction kinetics can be determined, offering new insights into the chemistry and mechanism of disulfide reduc-

tion either by small molecules (*e.g.* DTT or tris(2-carboxyethyl)phosphine) or by an enzyme as Trx (18, 24, 40). We have used the force-clamp technique to show that Trx regulates the geometry of the participating sulfur atoms in the enzymatic reaction. Our technique is able to capture dynamic sub-Ångström conformational changes during catalysis that are undetectable with any other technique (18).

Given the ability of force-clamp in reporting mechanistic information on enzyme catalysis, we have studied the force-dependent chemical kinetics of all Trx mutant forms to determine whether the co-variance calculation has a functional significance. Similar to previous works, we have first constructed a polyprotein that was used as a substrate for Trx disulfide reduction (18, 40). The polyprotein is composed of eight modules of the 27th domain of the human cardiac titin (I27), which was modified to introduce a disulfide bond between the 32nd and 75th positions in each module (see supplemental data). We applied an initial force pulse to mechanically unfold the I27 modules up to the disulfide bond (Fig. 2A). The individual unfolding events are detected by a series of steps of about 10.5 nm/module (supplemental Fig. S3). This pulse also exposes the previously buried disulfide bond in each module to the solvent environment. This bond remains intact because forces of over 2 nanonewton are needed to break a covalent bond (38). We then applied a second pulse of force to monitor the reduction of single disulfide bonds by the reduced Trx molecules present in the sample (Fig. 2A). These disulfide bond reduction events resulted in a second series of steps of about 13.5 nm/module (Fig. 2, *B* and *C*). In this assay, the reduced Trx is generated by the Trx system composed of Trx, Trx reductase, and NADPH (41). We collected multiple traces that were averaged and fitted with a single exponential with a time constant τ, from which we then obtained the reduction rate ($r = 1/\tau$) at the corresponding force (Fig. 3). Thus, we obtained the reduction rate dependence on force for a given Trx. Considering the number of traces collected per force (15–50 traces), a single exponential is a good approximation and adequately describes the process over the entire range of forces.
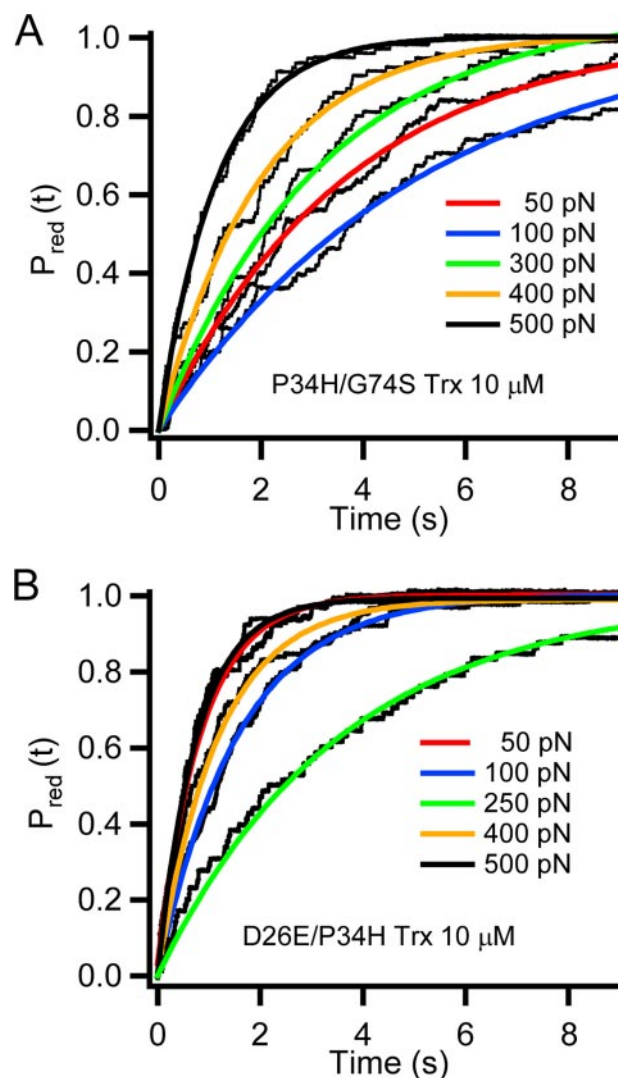
Fig. 4B shows the reduction rate dependence on force for the variant G74S and the double mutant P34H/G74S and also for P34H Trx for comparison. To analyze the data we have used the kinetic model in Fig. 4A. This three-state kinetic model was previously used to describe force-dependent Trx catalysis (18). It was selected from a number of alternate models and found to best describe the experimental data obtained for WT Trx from *E. coli*. In this model, there are two separate mechanisms for Trx catalysis. The first, type I catalysis, is similar to a typical Michaelis-Menten model for enzyme kinetics. This mechanism predominates at low forces, and rate constant k_{12} is inhibited by the applied force. Conversely, type II catalysis, which predominates at high force, is necessary to describe the biphasic force-dependent kinetics found for *E. coli* Trx. In this case, rate constant k_{02} is accelerated by the applied force. In the previous study, we proposed a structural model for Trx catalysis, whereby in type I catalysis the disulfide bond of the substrate polypeptide must be reoriented, on a sub-Ångström scale, within the enzymatic active site for efficient S_N2 reduction of the disulfide bond can occur. We also showed that mechanism
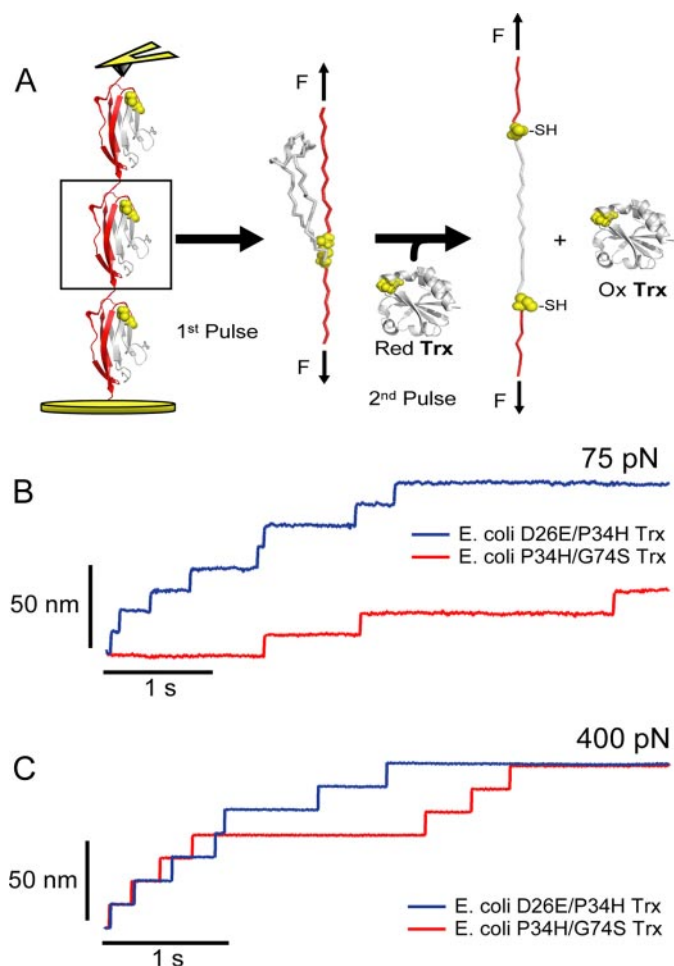
FIGURE 2. **Detecting catalytic co-evolution at the single molecule level.** *A,* schematic representation of the single molecule force-clamp experiment for disulfide reduction by Trx. The first force pulse unfolds and extends each I27 module in the polyprotein exposing the buried disulfide bonds to the solvent. Once unfolded, the second pulse sets the force under which reduction events will take place. The Trx enzymes present in the solution attack and reduce the exposed disulfide bonds, allowing for a stepwise extension of the sequestered residues. *B,* example of traces obtained by stretching $(I27_{ss})_8$ polyproteins at 75 pN of force in the presence of *E. coli* Trx enzymes containing either the D26E/P34H (*blue trace*) or P34H/G74S (*red trace*) double mutations. It is clear that at 75 pN the rate of disulfide bond reduction by the D26E/P34H Trx double mutant is much faster than that obtained with P34H/G74S Trx. *C,* in contrast to the results obtained at 75 pN, a similar experiment done at 400 pN shows little difference in the rate of reduction obtained with either double mutant form of Trx. These observations show that the double mutations control the chemical mechanisms of reduction at low force only.

FIGURE 3. **Ensemble average of single molecule traces measures reduction rates by the double mutants of Trx.** Summed and averaged recordings of disulfide bond reductions at different forces (second force pulse) for: P34H/G74S Trx 10 μM (*A*) and D26E/P34H Trx 10 μM (*B*). The *lines* represent single exponential fits with a time constant τ, where the rate of reduction is $r = 1/\tau$.

II, contrary to mechanism I, is not significantly altered upon mutation in Pro[34], suggesting a different chemical origin for both mechanisms.

Mutation G74S by itself significantly diminishes the reduction rate at low forces. Interestingly, this mutation naturally occurs in numerous bacterial Trxs that may be related to enzyme function diversification (33). The value of α_0, which is related to the initial formation of the enzyme-substrate complex, is 0.12 μM$^{-1}\cdot$s^{-1}. This value is ~45% lower than that of WT and slightly higher to that of the mutant form P34H (Table 1 and Fig. 5). Additionally, the value of Δx_{12}, which is the distance to the transition state along the length coordinate, is 0.86 Å, very close to the WT Trx and P34H Trx values (Table 1). The parameter Δx_{12} represents a shortening of the substrate polypeptide caused by the rotation of the disulfide bond during the reaction to acquire the proper orientation required by a S_N2 reaction (42). Molecular dynamics simulations have shown that a value of Δx_{12} of ~0.8 Å is in agreement with a sampling angle range of the substrate disulfide bond of 30–70° (without considering the starting 20° from the pulling axis) with an average value of ~50° (18).

When mutation G74S is combined with mutation P34H, the value of α_0 sharply decreases to 0.04 μM$^{-1}\cdot$s^{-1}, and the Δx_{12} value increases up to 1.60 Å, which is consistent with a rotation of the substrate disulfide bond angle of ~80°. These values strongly suggest a negative effect in the enzyme-substrate binding affinity when both residues are mutated, making the double mutant less efficient in regulating the orientation of the substrate disulfide bond. Consequently, mainly the conformational dynamics of the disulfide bond (as well as some additional interactions) could give rise to an almost perfect alignment between the three participating sulfur atoms, which may explain the
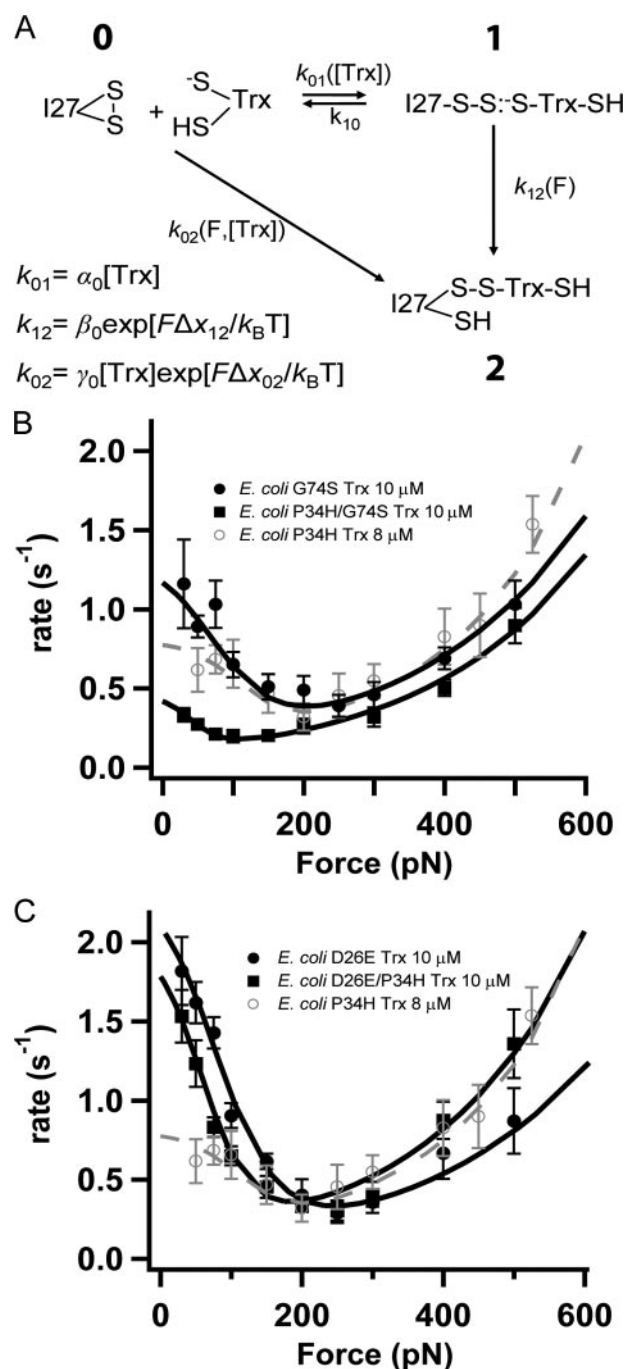
FIGURE 4. **Kinetic analysis of force dependent reduction by Trx double mutants.** *A,* we analyze our data with a kinetic model for force modulated disulfide bond reduction by Trx. The model distinguishes between two catalytic mechanisms: Type I catalysis (follows pathway 0–1-2), favored at low forces, entails the formation of a Michaelis-Menten complex. Type II catalysis is favored at high forces and does not require the formation of a substrate-enzyme complex (follows pathway 0–2). *B,* force-dependent reduction kinetics for 8 μM of P34H Trx (*open circles*), 10 μM of G74S Trx (*closed circles*), and 10 μM of P34H/G74S Trx (*closed squares*). Pathway I is nearly extinguished by the P34H/G74S double mutation, without a significant change in pathway II. *C,* force-dependent reduction kinetics for 10 μM of D26E (*closed circles*), is similar to wild type (not shown). The double mutant enzyme P34H/D26E (10 μM, *closed squares*), rescues the negative effect of P34H (*open circles*), explaining the positive correlation between these two residues. The *solid black lines* in *B* and *C* correspond to fits to the data with the model shown in *A*. The *dashed gray lines* in B and C represent the fit to the data for P34H Trx obtained from Ref. 18. The fitting parameters obtained are summarized in Table 1.

lower reduction rate at low forces for this double mutant. Interestingly, according to the value of the parameters γ_0 and Δx_{02} (Table 1), mechanism II is only slightly affected by these mutations, confirming that the enzymes are similarly active in the high force regime and that both catalytic mechanisms must have a different chemical origin. The unaffected mechanism II can be used as a control for the actual effect of mutations only on type I catalysis.

On the other hand, Fig. 4*C* shows the force dependence for the single mutant D26E and the double mutant D26E/P34H (P34H is also shown for comparison). The single mutation D26E does not seem to affect the activity with respect to the wild type form, according to the parameters obtained from the kinetic model (Table 1). Interestingly, mutation D26E, when combined with mutation P34H, rescued the Trx activity, demonstrating a compensatory effect. The double mutant D26E/ P34H features a value of α_0 of 0.19 $\mu M^{-1} \cdot s^{-1}$ and a value of Δx_{12} of 1.13 Å, which is consistent with a rotation angle of ~60°. This angle value is very close to that of the WT and the single mutants. In this case mechanism II also shows similar kinetics parameters that the WT form (Table 1). It is important to mention that thermal fluctuations can affect our reduction rate resolution. Although we do not provide a model accounting for thermal fluctuations, we assume that the error bars, calculated through bootstrapping of the data set, cover this uncertainty. In addition, the detected changes in the kinetic parameters of the mutants (Table 1) are due to big changes in reduction rate that are beyond the experimental error. Thus, we do not expect thermal motions to impact our sub-Ångström resolution.

Residues Pro[34] and Asp[26] are thought to play a role in the protonation state of the reactive Cys[32]. Although there is controversy about the exact pK_a values of the functional residues in positions 26 and 32 (31, 43), it is known that the pK_a value of the active site residues is a determinant in the function of Trx, and alterations in those values are responsible for changes in the activity of the different Trx superfamily members (44). Our results suggest that mutations P34H and D26E may give rise to a change in the chemistry of the active site of Trx that enhances the binding affinity and eventually the reduction rate, as discussed below.

Bulk Enzymatic Assays—We also tested the sequence-function relationship using two standard methods widely used to test thiol-disulfide oxidoreductase activity. These methods are the insulin turbimetric assay and the DTNB assay (41) (see "Experimental Procedures"). The results of both assays with the single and double mutant forms are summarized in Fig. 6. The only significant result is that mutants containing P34H show an elevated activity. In addition, both double mutants exhibit similar levels of activity, suggesting a neutral effect of the correlated mutations D26E and G74S. These seemingly contradictory results can be explained by the higher redox potential of P34H Trx compared with that of the WT form. This enhances the reduction of Trx by Trx reductase or DTT, thus increasing the rate of the overall process (29). Although these bulk assays are useful as a first approximation to measure reductase activity, it is important to point out that they are coupled assays in which the interaction between Trx reductase and Trx plays a critical role. In contrast, during the single molecule assays we essentially work under constant concentration of reduced Trx,

TABLE 1

Kinetic parameters for all Trx mutants obtained from fits with the kinetic model shown in Fig. 4A

The parameters were obtained by numeric optimization of the global fit to the data using the downhill simplex method (56). The *error bars* are given by the standard error for each parameter obtained with the bootstrap method in combination with the downhill simplex method.

|	α_0	β_0	γ_0	δ_0	Δx_{12}	Δx_{02}
	$\mu_M^{-1} \cdot s^{-1}$	s^{-1}	$\mu_M^{-1} \cdot s^{-1}$	s^{-1}	$\mathring{A}$	$\mathring{A}$
E. coli WT[a]	0.22 ± 0.03	21 ± 6	0.019 ± 0.008	3.2 ± 0.8	-0.79 ± 0.09	0.17 ± 0.02
P34H[a]	0.08 ± 0.02	28 ± 3	0.017 ± 0.003	1.7 ± 0.5	-0.83 ± 0.14	0.21 ± 0.03
G74S	0.12 ± 0.02	21 ± 2	0.017 ± 0.004	3.8 ± 0.5	-0.84 ± 0.16	0.17 ± 0.02
P34H/G74S	0.04 ± 0.01	26 ± 2	0.011 ± 0.001	4.1 ± 0.4	-1.60 ± 0.04	0.17 ± 0.02
D26E	0.23 ± 0.02	28 ± 2	0.018 ± 0.004	4.5 ± 0.4	-0.86 ± 0.05	0.16 ± 0.04
D26E/P34H	0.19 ± 0.02	26 ± 1	0.018 ± 0.004	4.0 ± 0.5	-1.13 ± 0.09	0.19 ± 0.02

[a] These values have been extracted from Ref. 18.
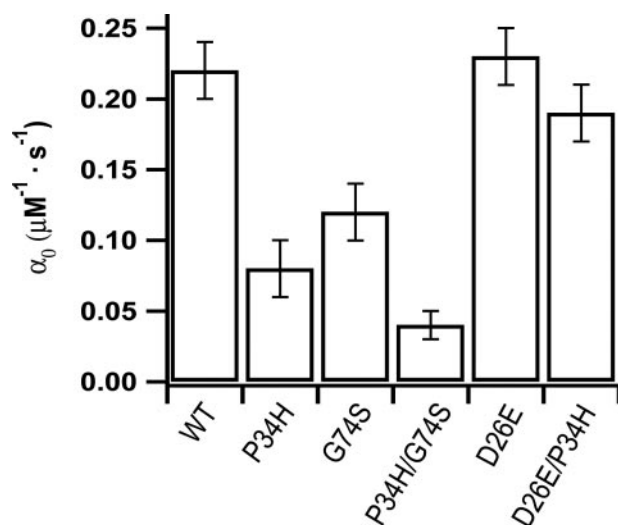
FIGURE 5. **Force spectroscopy captures the functional significance of co-evolved residues.** Shown is a bar graph of the rate of formation of the encounter complex, α_0, obtained from fits of the kinetic model in Fig. 4A. The negatively correlated double mutant P34H/G74S shows a remarkably low value of α_0, whereas the positively correlated mutant D26E/P34H shows a value of α_0 similar to that of WT.
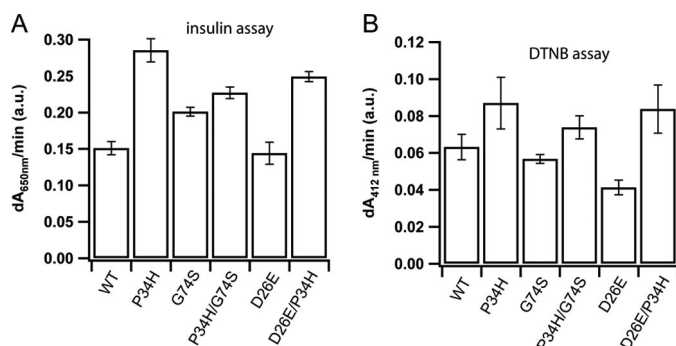

FIGURE 6. **Disulfide reduction by standard bulk assays.** A, rate of disulfide reduction in the bulk insulin assay for different Trx mutants. In this assay, insulin forms a precipitate as its disulfide bonds are reduced by Trx. B, rate of disulfide reduction by Trx in the DTNB bulk assay. In this assay, when the disulfide bond in DTNB is reduced, a leaving group is released, which absorbs strongly at 412 nm. In both assays, the concentration of enzyme is 5 μ_M, and the *error bars* represent the S.D. for three measurements. Note that in both assays the P34H mutation increases the disulfide reduction rate, inconsistent with the co-variance analysis.

and the reduction of the substrate disulfide bond by Trx is the only process detected. Although kinetic measurements using reduced Trx have been reported to yield results similar to those of the single molecule assay (29), the dynamic information on the sub-Ångström scale achieved by force-clamp spectroscopy is still unattainable by any regular assay.

DISCUSSION

Although complex and sophisticated statistical methods have been developed to detect co-evolving residues in protein sequences, understanding the structural or functional effects of co-evolution ultimately requires experimental validation. In the case of protein stability and structure, current techniques are able to detect the energetic and structural effects of detrimental or compensatory mutations (45). However, enzyme catalysis is a complex process that entails several steps: the binding of the substrate, conformational changes of enzyme and substrate, and finally the chemical reaction itself. Standard *in vitro* experiments are limited because they do not provide mechanistic aspects of the process, and NMR studies or x-ray crystallography do not reach the sub-Ångström scale where chemistry takes place (46). Single molecule force-clamp spectroscopy is capable of providing chemical and dynamic insights into the Trx catalysis at the sub-Ångström level. Here we demonstrate that this information is crucial to dissect the functional effects of residue co-evolution.

The negatively correlated residues 34 and 74 belong to hydrophobic surfaces that interact with different proteins (29, 32), in the binding groove of *E. coli* Trx. Binding residues are essential in mediating the correct positioning of the substrate, and mutations in those regions can alter the enzymatic reaction (47). This is in agreement with the value of the parameters obtained from the model for double mutant P34H/G74S (Table 1). The large reduction in α_0 (Fig. 5) suggests a deleterious effect on the binding process, and the increase in the value of Δx_{12} reveals that the double mutant is inefficient in regulating the geometry of the substrate disulfide bond required for a S_N2 reaction. Clearly there is an evolutionary constraint to avoid a double mutation in those positions because it would give rise to a poorly active disulfide reductase activity. Our results also show that the negative effect of mutation P34H can be compensated by the positively correlated mutation D26E, which acts as a neutral suppressor. In the double mutant D26E/P34H, the value of α_0 is returned to that of the WT form (Table 1 and Fig. 5). However, position 26 is in the core of the protein and is not known to interact with the substrate. This suggests that the rescuing of the activity is due to an allosteric chemical change. The mutation D26E may change the pK_a of the reactive Cys[32] by facilitating multiple electrostatic interaction between the residues of the active site (43, 48). These mutations are found, in the corresponding positions, in protein-disulfide isomerase and DsbA, oxidase members of the Trx superfamily (49, 50). Protein-disulfide isomerase and DsbA have a low pK_a for the reac-

tive Cys, which stabilizes the active thiolate form (51, 52). However, if the allosteric rescuing effect of D26E is a consequence of the change in the pK_a of the catalytic cysteine, it is puzzling why the high force pathway II remains relatively unaffected. Although it is known that closely spaced interacting residues can compensate for chemical alterations in enzyme catalysis (53) and protein structure and stability (45, 54), the long distance between residues Asp[26] or Gly[74] and Pro[34], makes our result highly significant.

Here we illustrate that residue Pro[34], known to play an important role in the chemistry of the first mechanism of Trx reduction (type I) (18), sends co-evolutionary signals that can be readily detected with our techniques. Significantly, these same mutations do not alter significantly the second chemical mechanism of reduction (Type II) that operates at high forces. These observations confirm the different chemical origins of both catalytic pathways, suggesting that a different, as yet unidentified set of residues is responsible for type II catalysis. The ability to make these distinctions opens up the possibility of detailed studies of the relationship between sequence evolution and the appearance of the different types of chemical mechanisms in Trx enzymes.

Although in our experiments, P34H represents a loss-of-function mutation, the high number of positive correlations in which residue Pro[34] is involved suggests that accumulating mutations led to enzyme function diversification rather than optimization, as is the case for protein-disulfide isomerase and DsbA. Because the accumulation of mutations is not an all-or-none process, the double mutant D26E/P34H may represent a promiscuous intermediate with enhanced isomerase activity and reductase activity similar to that of the WT Trx.

In this work we show that single molecule force-clamp spectroscopy combined with a straightforward statistical method for sequence analysis can provide experimental validation of the simple concept of co-evolution in enzyme catalysis. Although this concept has been studied in the case of closely spaced interacting residues (55), our experiments directly demonstrate the co-evolution of distant residues in enzyme catalysis.

REFERENCES

1. Patthy, L. (1999) *Protein Evolution*, pp. 80–83, Blackwell Science, Oxford, Malden, MA
2. Jurgens, C., Strom, A., Wegener, D., Hettwer, S., Wilmanns, M., and Sterner, R. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 9925–9930
3. Glasner, M. E., Fayazmanesh, N., Chiang, R. A., Sakai, A., Jacobson, M. P., Gerlt, J. A., and Babbitt, P. C. (2006) *J. Mol. Biol.* **360,** 228–250
4. Norrgard, M. A., Ivarsson, Y., Tars, K., and Mannervik, B. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 4876–4881
5. Altschuh, D., Lesk, A. M., Bloomer, A. C., and Klug, A. (1987) *J. Mol. Biol.* **193,** 693–707
6. Kimura, M. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 5969–5973
7. Neher, E. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 98–102
8. Pritchard, L., and Dufton, M. J. (2000) *J. Theor. Biol.* **202,** 77–86
9. Fares, M. A., and Travers, S. A. (2006) *Genetics* **173,** 9–23
10. Lee, B. C., Park, K., and Kim, D. (2008) *Proteins* **72,** 863–872
11. Pazos, F., Helmer-Citterich, M., Ausiello, G., and Valencia, A. (1997) *J. Mol. Biol.* **271,** 511–523
12. Fleishman, S. J., Yifrach, O., and Ben-Tal, N. (2004) *J. Mol. Biol.* **340,** 307–318
13. Suel, G. M., Lockless, S. W., Wall, M. A., and Ranganathan, R. (2003) *Nat. Struct. Biol.* **10,** 59–69
14. Shulman, A. I., Larson, C., Mangelsdorf, D. J., and Ranganathan, R. (2004) *Cell* **116,** 417–429
15. Lockless, S. W., and Ranganathan, R. (1999) *Science* **286,** 295–299
16. Ozkan, E., Yu, H., and Deisenhofer, J. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102,** 18890–18895
17. Estabrook, R. A., Luo, J., Purdy, M. M., Sharma, V., Weakliem, P., Bruice, T. C., and Reich, N. O. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102,** 994–999
18. Wiita, A. P., Perez-Jimenez, R., Walther, K. A., Grater, F., Berne, B. J., Holmgren, A., Sanchez-Ruiz, J. M., and Fernandez, J. M. (2007) *Nature* **450,** 124–127
19. Johnson, C. P., Tang, H. Y., Carag, C., Speicher, D. W., and Discher, D. E. (2007) *Science* **317,** 663–666
20. Perez-Jimenez, R., Godoy-Ruiz, R., Parody-Morreale, A., Ibarra-Molero, B., and Sanchez-Ruiz, J. M. (2006) *Biophys. Chem.* **119,** 240–246
21. Bershtein, S., Segal, M., Bekerman, R., Tokuriki, N., and Tawfik, D. S. (2006) *Nature* **444,** 929–932
22. Smith, T. F., and Waterman, M. S. (1981) *J. Mol. Biol.* **147,** 195–197
23. Oberhauser, A. F., Marszalek, P. E., Erickson, H. P., and Fernandez, J. M. (1998) *Nature* **393,** 181–185
24. Wiita, A. P., Ainavarapu, S. R., Huang, H. H., and Fernandez, J. M. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 7222–7227
25. Carrion-Vazquez, M., Oberhauser, A. F., Fowler, S. B., Marszalek, P. E., Broedel, S. E., Clarke, J., and Fernandez, J. M. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 3694–3699
26. Perez-Jimenez, R., Godoy-Ruiz, R., Ibarra-Molero, B., and Sanchez-Ruiz, J. M. (2004) *Biophys. J.* **86,** 2414–2429
27. Lundstrom, J., and Holmgren, A. (1990) *J. Biol. Chem.* **265,** 9114–9120
28. Slaby, I., and Holmgren, A. (1975) *J. Biol. Chem.* **250,** 1340–1347
29. Krause, G., Lundstrom, J., Barea, J. L., Pueyo de la Cuesta, C., and Holmgren, A. (1991) *J. Biol. Chem.* **266,** 9494–9500
30. Chivers, P. T., and Raines, R. T. (1997) *Biochemistry* **36,** 15810–15816
31. Chivers, P. T., Prehoda, K. E., Volkman, B. F., Kim, B. M., Markley, J. L., and Raines, R. T. (1997) *Biochemistry* **36,** 14985–14991
32. Himawan, J. S., and Richardson, C. C. (1996) *J. Biol. Chem.* **271,** 19999–20008
33. Masip, L., Klein-Marcuschamer, D., Quan, S., Bardwell, J. C. A., and Georgiou, G. (2008) *J. Biol. Chem.* **283,** 840–848
34. Russ, W. P., Lowery, D. M., Mishra, P., Yaffe, M. B., and Ranganathan, R. (2005) *Nature* **437,** 579–583
35. Chi, C. N., Elfstrom, L., Shi, Y., Snall, T., Engstrom, A., and Jemth, P. (2008) *Proc. Natl. Acad. Sci. U. S. A.* **105,** 4679–4684
36. Fernandez, J. M., and Li, H. B. (2004) *Science* **303,** 1674–1678
37. Garcia-Manyes, S., Brujic, J., Badilla, C. L., and Fernandez, J. M. (2007) *Biophys. J.* **93,** 2436–2446
38. Grandbois, M., Beyer, M., Rief, M., Clausen-Schaumann, H., and Gaub, H. E. (1999) *Science* **283,** 1727–1730
39. Carl, P., Kwok, C. H., Manderson, G., Speicher, D. W., and Discher, D. E. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 1565–1570
40. Szoszkiewicz, R., Ainavarapu, S. R., Wiita, A. P., Perez-Jimenez, R., Sanchez-Ruiz, J. M., and Fernandez, J. M. (2007) *Langmuir* **24,** 1356–1364
41. Holmgren, A. (1985) *Annu. Rev. Biochem.* **54,** 237–271
42. Fernandes, P. A., and Ramos, M. J. (2004) *Chemistry* **10,** 257–266
43. Dillet, V., Dyson, H. J., and Bashford, D. (1998) *Biochemistry* **37,** 10298–10306
44. Mossner, E., Huber-Wunderlich, M., and Glockshuber, R. (1998) *Protein Sci.* **7,** 1233–1244
45. Mateu, M. G., and Fersht, A. R. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 3595–3599
46. Hammes-Schiffer, S., and Benkovic, S. J. (2006) *Annu. Rev. Biochem.* **75,** 519–541
47. Kraut, D. A., Carroll, K. S., and Herschlag, D. (2003) *Annu. Rev. Biochem.*

72, 517–571

48. LeMaster, D. M., Springer, P. A., and Unkefer, C. J. (1997) *J. Biol. Chem.* 272, 29998–30001

49. Gruber, C. W., Cemazar, M., Heras, B., Martin, J. L., and Craik, D. J. (2006) *Trends Biochem. Sci.* 31, 455–464

50. Inaba, K., and Ito, K. (2002) *EMBO J.* 21, 2646–2654

51. Grauschopf, U., Winther, J. R., Korber, P., Zander, T., Dallinger, P., and Bardwell, J. C. (1995) *Cell* 83, 947–955

52. Kortemme, T., Darby, N. J., and Creighton, T. E. (1996) *Biochemistry* 35, 14503–14511

53. Zhang, J., and Rosenberg, H. F. (2002) *Proc. Natl. Acad. Sci. U. S. A.* 99, 5486–5491

54. Wilson, K. P., Malcolm, B. A., and Matthews, B. W. (1992) *J. Biol. Chem.* 267, 10842–10849

55. Choi, S. S., Li, W., and Lahn, B. T. (2005) *Nat. Genet.* 37, 1367–1371

56. Nelder, J. A., and Mead, R. (1965) *Comput. J.* 7, 308–313