BIOC530 AU 2014
Protein NMR Unit
Homework Assignment

The recent paper published in Science "Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone" used a wide variety of NMR experiments to obtain remarkable insights into how a molecular chaperone works. The senior author of the paper, Prof Charalampos Kalodimos, gave a Biochemistry Dept. seminar in October 2014 that you should have attended. This provided an excellent introduction to the paper and should help to make it more accessible.

Your assignment: Read the paper (attached) and answer the questions below.

1. What information was used to generate Figure 2A?

2. What samples (what proteins were in the NMR tube? Which one(s) has isotopic labels? What were the labels?) were used to generate the following:

- spectra shown in Fig. 3B: gray and pink
- spectra shown in Fig. 3B: blue

3. What experiments provided the exchange rates depicted in Fig. 3K?

4. Draw a cartoon or scheme to explain Fig. 7C, showing:

- what are the molecules that generated the spectra
- what molecule(s) are isotopically labeled and with what isotope(s)/

Also use your cartoon to depict the results/interpretation of the experiments.

Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone

Tomohide Saio *et al.*

Science **344**, (2014);
DOI: 10.1126/science.1250494

READ THE FULL ARTICLE ONLINE
http://dx.doi.org/10.1126/science.1250494

Structural and dynamic data show how molecular chaperones prevent aggregation and misfolding of unfolded proteins.

Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone

Tomohide Saio, Xiao Guan, Paolo Rossi, Anastassios Economou, Charalampos G. Kalodimos*

RELATED ITEMS IN *SCIENCE*

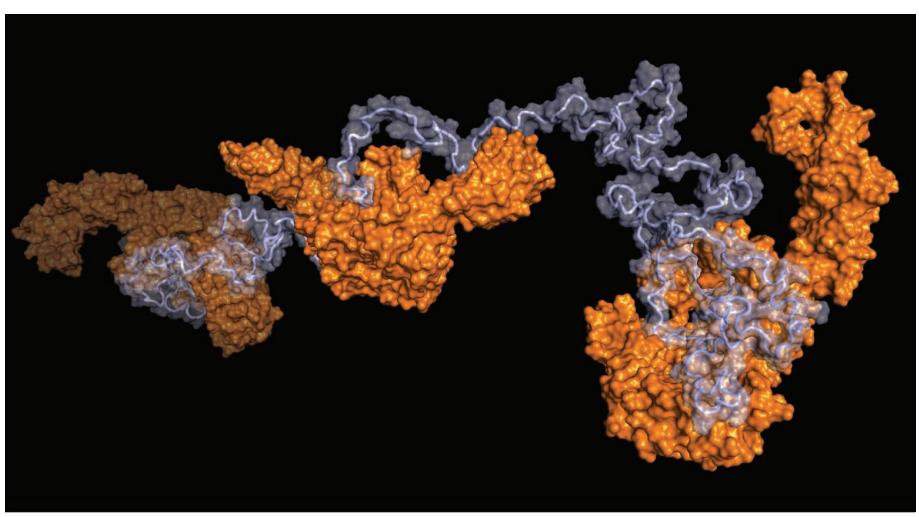
M. Gamerdinger, E. Deuerling, Trigger factor flexibility. *Science* **344**, 590 (2014). DOI: 10.1126/science.1254064

Introduction: Molecular chaperones prevent aggregation and misfolding of proteins in the cellular environment and are thus central to maintaining protein homeostasis. Molecular chaperones are thought to recognize and bind to exposed hydrophobic regions of the unfolded proteins, thereby shielding these regions from the solvent. If unprotected, the proteins would likely aggregate or misfold to bury the hydrophobic residues. Despite the central importance of the binding of chaperones to unfolded proteins, the structural basis of their interaction remains poorly understood. The scarcity of structural data on complexes between chaperones and unfolded proteins is primarily due to technical challenges originating in the size and dynamic nature of these complexes.

Rationale: Recent advances in nuclear magnetic resonance (NMR) and isotope labeling approaches make it possible to study large, dynamic complexes. We used NMR spectroscopy to characterize the binding of the 48-kD unfolded alkaline phosphatase (PhoA) to the 50-kD trigger factor (TF) chaperone. We obtained atomic insight into the dynamic binding and determined the solution structure of PhoA captured in an extended, unfolded state by three TF molecules. Based on our NMR studies, we gained insight into how TF rescues an aggregation-prone protein and how it exerts its unfoldase activity.

Results: We show that TF uses multiple sites, which are located in two different domains and extend over a distance of ~90 Å, to bind to several regions of the unfolded PhoA that are dispersed throughout its entire length. Three TF molecules are required to interact with the entire length of PhoA, giving rise to a ~200-kD complex in solution. The TF-PhoA interactions are mediated primarily by hydrophobic contacts. TF interacts with PhoA in a highly dynamic fashion, giving rise to a rugged landscape for the free energy of interaction. As the number and length of the PhoA regions engaged by TF increases, a more stable complex gradually emerges. The multivalent binding keeps PhoA in an extended, unfolded conformation. Crucially, even the lowest-energy TF-PhoA complex remains rather dynamic with a lifetime of ~20 ms. The structural data of the three TF molecules in complex with different regions of PhoA reveal how the same binding sites within a molecular chaperone can recognize and interact with a large number of substrates with unrelated primary sequences. This promiscuous recognition is further enabled by the notable plasticity of the substrate-binding sites in TF. We finally show that TF in the cytosol prevents aggregation by interacting transiently with the low-populated, aggregation-prone unfolded state of the substrate but acts as a powerful unfoldase when it is bound at the ribosome and thus is colocalized with translating substrate.

Conclusion: The structural data reveal a multivalent binding mechanism between the chaperone and its protein substrate. This mechanism of binding presents several advantages as it enables chaperones to function as holdases and unfoldases by exerting forces to retain proteins in the unfolded state and at the same time protect them from aggregation by shielding their exposed hydrophobic regions. Given the existence of multiple binding sites in other molecular chaperones, this may present a general mechanism for the action of molecular chaperones. The fast kinetics of substrate binding enables chaperones to interact with transiently exposed, aggregation-prone regions of unstable proteins in the cytosol, thereby preventing their aggregation and increasing their solubility.



Structural basis of PhoA binding by TF. PhoA (blue/gray) is captured in an unfolded state by three TF chaperone molecules (orange). Complex formation is mediated by multivalent binding of hydrophobic surfaces, which are shielded from water, thereby preventing folding and, at the same time, aggregation of the substrate protein.

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: babis@rutgers.edu
Cite this article as T. Saio *et al.*, *Science* **344**, 1250494 (2014). DOI: 10.1126/science.1250494

Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone

Tomohide Saio,[1]* Xiao Guan,[1]* Paolo Rossi,[1] Anastassios Economou,[2,3] Charalampos G. Kalodimos[1]†

Molecular chaperones prevent aggregation and misfolding of proteins, but scarcity of structural data has impeded an understanding of the recognition and antiaggregation mechanisms. We report the solution structure, dynamics, and energetics of three trigger factor (TF) chaperone molecules in complex with alkaline phosphatase (PhoA) captured in the unfolded state. Our data show that TF uses multiple sites to bind to several regions of the PhoA substrate protein primarily through hydrophobic contacts. Nuclear magnetic resonance (NMR) relaxation experiments show that TF interacts with PhoA in a highly dynamic fashion, but as the number and length of the PhoA regions engaged by TF increase, a more stable complex gradually emerges. Multivalent binding keeps the substrate protein in an extended, unfolded conformation. The results show how molecular chaperones recognize unfolded polypeptides and, by acting as unfoldases and holdases, prevent the aggregation and premature (mis)folding of unfolded proteins.

Molecular chaperones maintain a functional proteome in the cell by preventing the aggregation of unfolded proteins, assisting with their folding or disassembling aggregates (*1*–*6*). Nascent polypeptide chains emerging from the exit channel of the ribosome expose long hydrophobic regions that are particularly prone to misfolding and aggregation (*7*–*9*). The bacterial trigger factor (TF) (*10*) binds next to the exit channel of the ribosome (*11*–*15*) and directly interacts with the emerging nascent polypeptide (*16*–*19*). With an estimated cellular concentration of ~50 μM (*20*), TF is one of the most abundant proteins in bacteria. TF prevents the aggregation and premature folding of nascent polypeptides and unfolded proteins in vivo and in vitro, likely by interacting with exposed hydrophobic regions, thereby shielding them from the solvent (*18*, *21*–*23*). Long polypeptides (>20 kD), which are the preferred clients for TF in vivo (*24*, *25*), require multiple TF molecules for their protection. TF remains bound to the unfolded polypeptide even after having departed the ribosome (*18*, *21*, *22*, *26*). TF reduces the speed of folding, thereby increasing the yield of biologically active proteins (*21*, *27*–*29*). TF was shown to have both a weak holdase and an unfoldase activity (*30*). Deletion of the TF gene results in the aggregation of many proteins (*25*) and the stimulation of heat shock response (*31*–*33*). Following its interaction with TF, the unfolded polypeptide folds by dissociating, interacting with foldase chaperones such as Hsp70 and GroEL, or enters the posttranslational secretory pathway to interact with the Sec machinery (*2*, *3*, *5*, *34*–*36*).

Despite the importance of chaperone binding to unfolded proteins, the structural basis of this interaction remains poorly understood. The scarcity of structural data on complexes between chaperones and unfolded proteins is primarily due to technical challenges associated with the size and dynamic nature of these complexes (*37*–*42*). We have exploited recent advances in nuclear magnetic resonance (NMR) and isotope labeling approaches (*43*–*48*) to characterize the dynamic binding of unfolded PhoA to TF and to determine the solution structure of the PhoA captured in an extended, unfolded state by three TF molecules.

NMR of TF Chaperone and Unfolded PhoA

Escherichia coli TF consists of 432 amino acids, comprising the ribosome-binding domain (RBD; residues 1 to 112), the peptidyl prolyl isomerase domain (PPD; residues 150 to 246), and a discontinuous C-terminal domain, positioned structurally between the RBD and PPD (*12*), which, on the basis of the current and previous findings (*49*, *50*), we refer to as the substrate-binding domain (SBD; residues 113 to 149 and 247 to 432) (Fig. 1A and fig. S1, A and B). Unliganded TF in solution forms a dimer (*27*) of ~100 kD (fig. S1C) [dissociation constant (K_d) of dimerization is ~18 μM (*51*)]. The quality of the ^{1}H-^{15}N–correlated NMR spectra (figs. S2A and S3) and the ^{1}H-^{13}C–correlated NMR spectra of methyl-bearing (Ala, Ile, Met, Leu, Thr, and Val) and aromatic (Phe, Trp, and Tyr) residues of TF is high (figs. S2B and S4). By exploiting the modular structural architecture of TF (*52*, *53*), we have obtained near-complete assignment of TF (see the Materials and Methods section).

E. coli alkaline phosphatase (PhoA) is an ~50-kD (471 amino acid residues) periplasmic enzyme that requires oxidizing conditions for folding, as do several other periplasmic proteins (*54*). PhoA was shown biochemically to be in an unfolded and thus degradation-prone state in the reducing environment of the cytosol (*54*–*56*) and to interact extensively with TF (*57*, *58*). The NMR spectra of PhoA, as well as relaxation experiments, confirmed that PhoA is unfolded under reducing conditions (figs. S5, B and C, and S6). Assessment of the presence of secondary structure by NMR (*59*) revealed pervasive random-coil and extended-like conformations, although several of the regions that formed a stable secondary structure in the folded PhoA retained a fraction of their fold in unfolded PhoA (Fig. 2, A and B).

The NMR spectra of unfolded proteins typically show very narrow chemical shift dispersion (*60*). To simplify NMR characterization of unfolded PhoA, which is a large protein by NMR standards, we generated a number of PhoA fragments of varying length that spanned its entire sequence (see Materials and Methods). Superposition of the NMR spectra of any of the PhoA fragments with the NMR spectra of the full-length PhoA demonstrated excellent resonance correspondence (fig. S5D). This is expected because each of these fragments is unfolded, as is the full-length PhoA, and thus, the conformational properties of the isolated fragments recapitulate those of the full-length PhoA. Near-complete assignment has been obtained for full-length PhoA.

Interaction Between TF and PhoA

To determine the residues of unfolded PhoA that are recognized and bound by TF, we monitored by NMR the binding of isotopically labeled PhoA by unlabeled TF (fig. S7). In addition to full-length PhoA, several PhoA fragments were used to enhance spectral resolution (fig. S7A). Differential line broadening analysis (*61*) showed that only a relatively small number of residues (~130 of the 471) in PhoA are recognized and bound by TF (Fig. 2C). There are in total seven distinct TF-binding regions in PhoA (labeled using lowercase letters a through g), with lengths ranging from ~8 to ~40 residues (Fig. 2C and fig. S8A). Analysis of the primary sequence of the TF-binding sites showed that they are enriched in hydrophobic and especially aromatic residues (figs. S8, B and C, and S9), in agreement with previous studies (*18*, *62*). Notably, there seems to be a correlation between the hydrophobicity of the PhoA sequence and the binding preference by TF (Fig. 2C).

To identify the sites that TF uses to interact with PhoA, we monitored by NMR the binding of full-length PhoA as well as select PhoA fragments comprising TF-binding sites to labeled TF (figs. S2, C and D, and S10). The combined data demonstrated that TF uses four distinct sites to interact with PhoA (labeled using uppercase letters; A through D). Three of these sites (A, B, and C) are located in the SBD (Fig. 1A), in agreement

[1]Center for Integrative Proteomics Research and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854, USA. [2]Laboratory of Molecular Bacteriology, Rega Institute, Department of Microbiology and Immunology, KU Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium. [3]Institute of Molecular Biology and Biotechnology-FORTH and Department of Biology, University of Crete, P.O. Box 1385, GR-71110, Iraklio, Crete, Greece.

*These authors contributed equally to this work.
†Corresponding author. E-mail: babis@rutgers.edu

with previous data indicating this domain as being the central chaperone module (*49*, *50*). The fourth binding site (labeled D) is located in the PPD and overlaps with its catalytic site (Fig. 1B). All sites are composed primarily of hydrophobic residues (Fig. 1, B and C). NMR titration of additional protein substrates [such as the maltose binding protein (MBP) and the outer membrane protein A (OmpA)] to TF demonstrated that the same sites are used by TF to interact with these substrates (fig. S11). The substrate-binding sites are the best-conserved regions in TF, along with the region in the RBD that binds to the ribosome (Fig. 1D and fig. S12). Residues located in the identified substrate-binding sites were previously reported to cross-link to the emerging nascent chain when TF is bound to the ribosome (*49*, *50*). Although the RBD surface is very hydrophobic, our data show that this domain does not interact with the protein substrates (Fig. 1A and fig. S11). In contrast, the RBD was seen to cross-link to nascent chains in ribosome-bound TF (*49*, *50*). This discrepancy may be due to the physical proximity of the RBD to the exit channel of the ribosome, or RBD may undergo conformational change upon interacting with the ribosome that enables its interaction with the nascent polypeptide (*13*, *18*).

The crystal structure of free TF (*12*) shows that the identified substrate-binding sites (A to D) consist almost exclusively of nonpolar residues (Fig. 1C). Sites A, B, and C are enriched with methyl-bearing residues, and site D is enriched with aromatic residues (Fig. 1B). Thus, the use of hydrophobic sites to interact with unfolded client proteins is shared among several molecular chaperones, such as GroEL, Skp, Hsp40, and Hsp90 (*63*, *64*). Sites A and B are located inside the cradle-like structure formed by the RBD and SBD, with A located at the base of the cradle and B located at the tip of arm 2 (Fig. 1A). Site C forms a rather elongated ridge near the β sheet that connects the SBD and PPD. Sites A, B, C, and D expose ~530, ~680, ~460, and ~650 Å^2, respectively, of hydrophobic surface, thus providing combined up to ~2300 Å^2 of solvent-exposed nonpolar surface available for binding of protein substrates. The four TF sites extend over a distance of more than 90 Å.

Mode of Binding Between TF and PhoA

We characterized by NMR and isothermal titration calorimetry (ITC) the binding and energetics of each of the PhoA sites, individually and in combination to TF. Both chemical shift and intermolecular TF-PhoA nuclear Overhauser effect (NOE) measurements were used to monitor the interaction at the residue level. Starting from the N terminus, the first PhoA region to interact with TF is site *a* (Figs. 2C and 3A). Chemical shift perturbation as well as intermolecular TF-PhoA NOEs (fig. S13) clearly indicated that PhoA site *a* can interact with all four of the binding sites in TF with a moderate affinity (K_d ~25 μM; Fig. 3B). Similarly, PhoA site *b* can interact with any of the four binding sites in TF but with lower affinity (K_d ~200 μM; Fig. 3C). In contrast, PhoA subsite c^1 (the N-terminal region of site *c*; Fig. 3A) binds exclusively to the SBD and preferably to TF sites A and B (on the basis of chemical shift perturbation, subsite c^1 appears to bind to TF site C, but no intermolecular NOEs were detected, indicating a rather transient interaction; Fig. 3D). PhoA subsite c^2 (the middle region of site *c*; Fig. 3A) interacts exclusively with TF site A (Fig. 3E). The binding to TF of a PhoA fragment consisting of subsites c^1 and c^2 is determined by the combined binding preferences of the subsites (Fig. 3F). The PhoA c^1-c^2 fragment binds to TF with higher affinity than the subsites c^1 and c^2 individually because of avidity (Fig. 3, D to F). Notably, the complete PhoA site *c* (consisting of subsites c^1, c^2, and c^3) binds to TF in a specific orientation and arrangement: PhoA subsite c^1 binds to TF site B, subsite c^2 binds to TF site A, and subsite c^3 binds to TF site C, as demonstrated by the measurement of a unique set of characteristic intermolecular NOEs between the TF sites and the PhoA subsites (Fig. 3G). PhoA site *c* binds only to TF sites A, B, and C in the SBD, and with higher affinity (K_d ~10 μM) than PhoA site *a* or *b*. Thus, PhoA sites *a* and *b* can only interact with TF site D in the context of the PhoA$^{a\text{-}c}$ fragment (superscript denotes the TF-binding sites that the PhoA fragment contains). Because PhoA site *a* binds with much higher affinity to TF than PhoA site *b* (Fig. 3, B and C), in the TF–PhoA$^{a\text{-}c}$ complex, PhoA site *a* binds to TF site D, whereas PhoA site *b* does not interact at all with TF (Fig. 3H). The affinity of PhoA$^{a\text{-}c}$ for TF (K_d ~2 μM; fig. S14) is higher than any of the other PhoA sites.

The results clearly show that although some PhoA sites (for example, sites *a* and *b*) bind

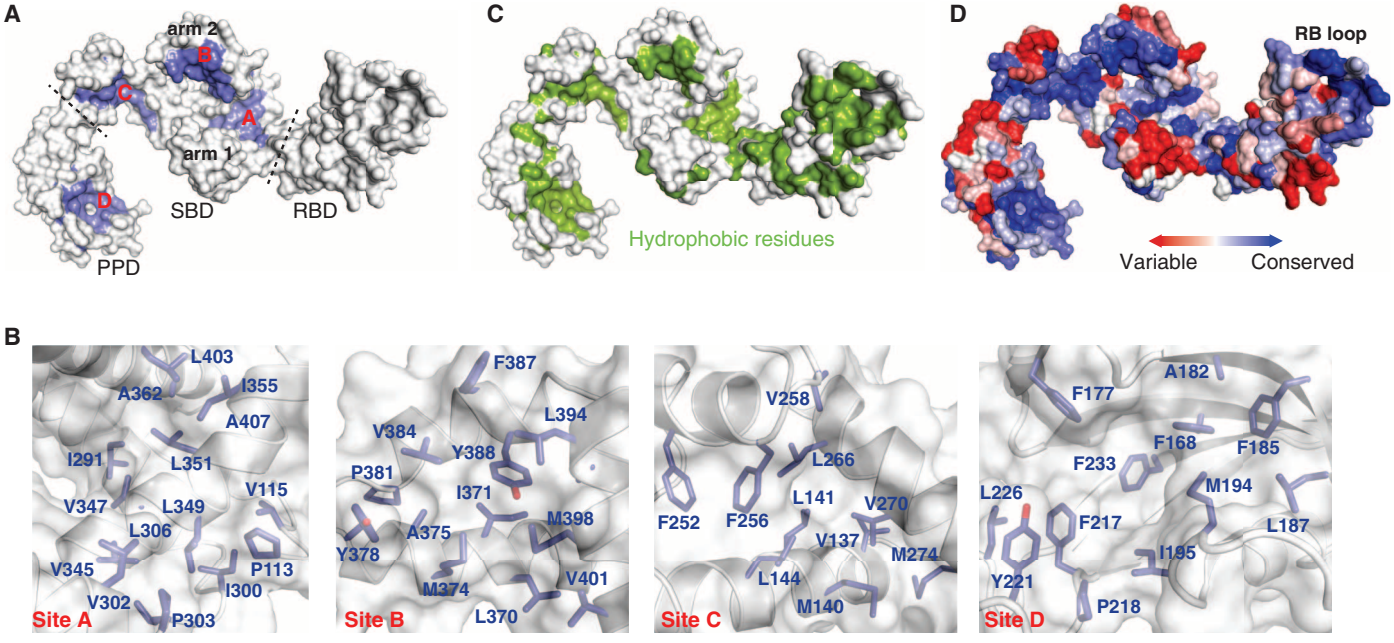


Fig. 1. Substrate-binding sites in TF. (**A**) The TF residues identified by NMR to interact with unfolded protein substrates are colored blue on the crystal structure of free *E. coli* TF [Protein Data Bank (PDB) ID 1W26]. The four main binding sites are labeled A, B, C, and D (in red). The dashed lines indicate the domain boundaries. The two structural protrusions in the SBD are labeled arm 1 and arm 2. An additional binding site, located in arm 1, is used by TF to interact with some of the substrates (see Fig. 7A and fig. S11). (**B**) Expanded view of the four substrate-binding sites in TF identified by the NMR titration experiments. The hydrophobic residues that make up the substrate-binding sites are shown as sticks. (**C**) The hydrophobic residues in TF are colored green, whereas all other residues are colored white. (**D**) TF sequence conservation is color-mapped on the TF structure. The residues that make up the substrate-binding sites are the best-conserved along with the ribosome-binding loop (RB loop) (see fig. S12).

promiscuously to all TF sites, other PhoA sites (for example, subsite c^1) are more selective and some sites (for example, subsite c^2) appear to interact exclusively with specific sites in TF. Therefore, the binding mode of a multisite PhoA fragment to TF will be determined by the combined preferences of the sites (Fig. 3H). Together, the data demonstrate that the interaction of TF and PhoA is dynamic and is best described by a rugged free energy landscape: the various binding interactions shown in panels (B) to (G) of Fig. 3 do not maximize the binding interface and correspond to the higher-energy wells, whereas the maximum-contact arrangement shown in Fig. 3H corresponds to the ground-state, lowest-energy conformation of the complex in which a unique arrangement of the chaperone and the substrate is observed experimentally (Fig. 3H).

The NMR data showed that one TF molecule can accommodate a PhoA fragment encompassing TF-binding sites *a* to *c* ($PhoA^{a-c}$) (Fig. 3 and figs. S14 and S15). Under stoichiometric conditions, the use of a longer PhoA fragment that included site *d* ($PhoA^{a-d}$) resulted in this site not being bound by TF (fig. S16A). Site *d* was bound only when excess TF was added (fig. S16A). Using the experimental approach described in Fig. 3 for

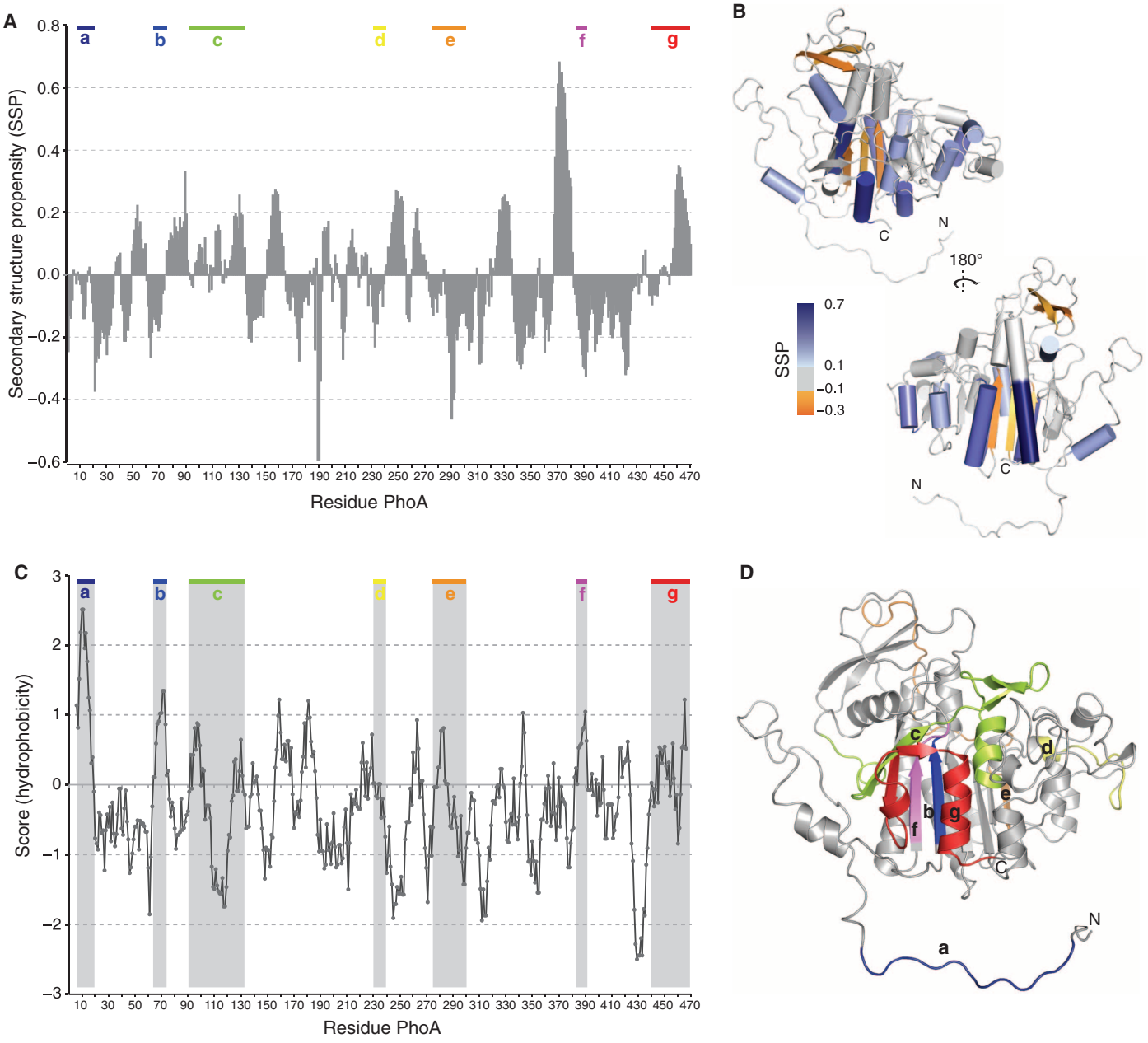


Fig. 2. TF-binding sites and secondary structure in unfolded PhoA. (**A**) Secondary structure propensity (SSP) values (*59*) of unfolded PhoA plotted as a function of the amino acid sequence. A SSP score at a given residue of 1 or −1 reflects a fully formed α-helical or β-structure (extended), respectively, whereas a score of, for example, 0.5 indicates that 50% of the conformers in the native-state ensemble of the protein are helical at that position. The data show that several of the secondary structure elements in the folded PhoA retain some transient structure in the unfolded PhoA. The TF-binding sites (a to g) are indicated. (**B**) The SSP values of unfolded PhoA that correspond to secondary structure elements present in folded PhoA are mapped on the structure of folded *E. coli* PhoA. The signal sequence region (residues 1 to 22) was modeled into the crystal structure of the mature form (PDB ID 1Y6V). Only one subunit of the dimeric PhoA is shown. (**C**) Plot of the hydrophobicity of PhoA as a function of its primary sequence. A hydrophobicity score (Roseman algorithm, window = 9) higher than zero denotes increased hydrophobicity. The seven PhoA regions (labeled a through g) identified by NMR to interact with TF are highlighted in gray. (**D**) Mapping of the TF-binding sites (a to g) on the folded PhoA structure.

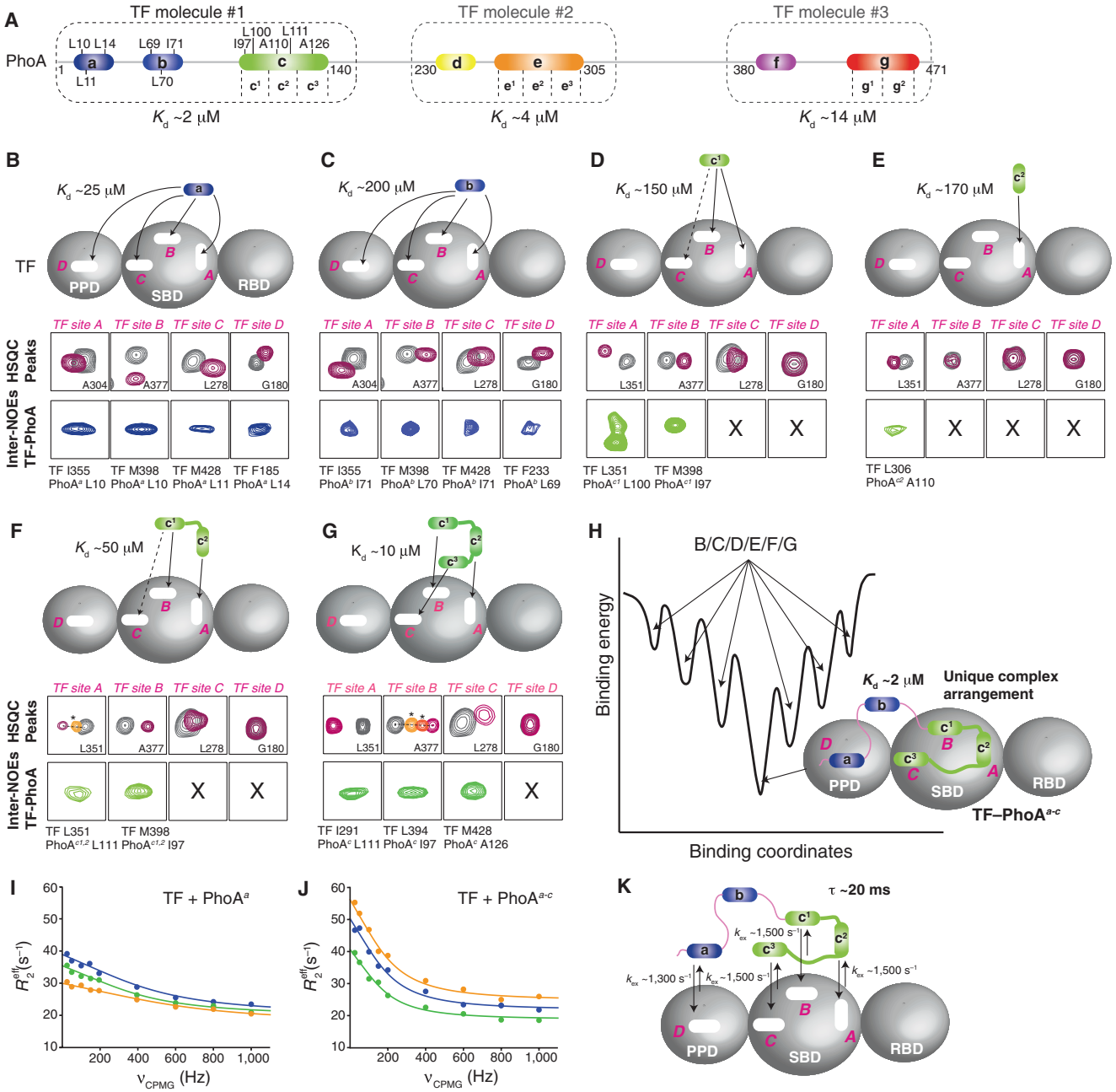


Fig. 3. Dynamic binding and energetics of PhoA by TF. (**A**) TF-binding sites in PhoA colored as in Fig. 2A. Three TF molecules are required for the simultaneous engagement of the entire PhoA. The K_d for each complex is shown. PhoA residues of sites *a*, *b*, and *c* whose intermolecular NOEs to TF residues are depicted in panels (B) to (G) are indicated. (**B** to **G**) The interaction of the independent PhoA site *a* (B), site *b* (C), subsite c^1 (D), subsite c^2 (E), subsites c^1-c^2 (F), and subsites c^1-c^2-c^3 (site *c*) (G) with TF has been characterized by NMR and ITC. The four PhoA-binding sites in TF (A to D) are labeled in pink. In the rectangles labeled "HSQC peaks," representative peaks from ^{1}H-^{15}N–correlated spectra of labeled TF (TF$^{SBD-PPD}$) free (gray) and in complex (magenta) with the indicated PhoA site are shown. In the rectangles labeled "Inter-NOEs TF-PhoA," representative intermolecular NOEs between the TF and PhoA residues are shown. The full NOESY strips are shown in fig. S13. The symbol X indicates the absence of such NOEs. The K_d values were determined by ITC and NMR. Solid arrows indicate experimentally observed binding of the PhoA site to the designated TF site. Broken arrows indicate transient interaction observed by chemical shift perturbation but for which no intermolecular NOEs were detected. The orange cross peak (marked with an asterisk) for TF site A in panel (F) corresponds to the chemical shift observed for the binding of the individual subsite c^2 to TF site A shown in panel (E). Similarly, the orange and dark-orange cross peaks marked with asterisks in panel (G) for TF site B correspond to the chemical shifts for the binding of the individual subsite c^1 (panel D) and the fragment c^1-c^2 (panel F), respectively. All these resonances fall on the same line connecting the free and PhoA site *c*–bound TF peaks (*93*), further supporting the observed synergistic binding between the PhoA sites as the length of PhoA increases, as well as the selectivity of certain PhoA sites for specific TF sites. (**H**) The progressive increase in PhoA length results in increased affinity for PhoA interaction with TF, and when PhoA is sufficiently long to engage all four of the TF sites simultaneously, as is the case with PhoA^{a-c}, a unique binding mode between PhoA and TF is observed as shown. The binding of PhoA to TF can be described by a rugged free energy landscape, with the conformational arrangement of the TF–PhoA^{a-c} complex depicted in this panel corresponding to the lowest-energy, ground-state structure. The binding interactions shown in panels (B) to (G) correspond to higher-energy states. The results for the interaction of PhoA sites *d* to *g* with TF are shown in fig. S17. (**I** and **J**) Representative relaxation dispersion profiles of three TF residues in complex with PhoA site *a* (I) and PhoA^{a-c} (J). R_2^{eff} is the effective transverse relaxation rate, and ν_{CPMG} is the refocusing frequency of the CPMG train pulse. (**K**) Exchange rate constants (k_{ex}) determined experimentally for the complex between TF and PhoA^{a-c}. The residence time of PhoA^{a-c} bound to TF is ~20 ms at 22°C.

the N-terminal region of PhoA, we found that a second TF molecule is required to fully engage PhoA sites *d* and *e* (fragment PhoA$^{d\text{-}e}$) and a third TF molecule to fully interact with PhoA sites *f* and *g* (fragment PhoA$^{f\text{-}g}$) (Fig. 3A and fig. S17). PhoA sites *d* and *f* bound exclusively to the TF PPD (fig. S17, B and E), whereas sites *e* and *g* bound exclusively to the TF SBD in the binding arrangement shown in Fig. 4A and fig. S17, C and F. Similarly to the TF–PhoA$^{a\text{-}c}$ complex, a unique arrangement was observed for the TF–PhoA$^{d\text{-}e}$ complex, as demonstrated by the measurement of a unique set of characteristic intermolecular NOEs between TF and PhoA (fig. S17D). In contrast, two binding arrangements were observed for the TF–PhoA$^{f\text{-}g}$ complex as shown in fig. S17G, but with one of them being the preferred one (populated by more than 70% of the time) (Fig. 4A).

NMR titration experiments showed that three TF molecules are required to bind to all of the seven TF-binding sites in full-length PhoA (Fig. 4A). The stoichiometry of the TF-PhoA complexes was further corroborated by ITC experiments (fig. S14). Thermodynamic characterization showed that TF formed complexes with the three PhoA fragments (PhoA$^{a\text{-}c}$, PhoA$^{d\text{-}e}$, and PhoA$^{f\text{-}g}$) with K_d values of ~2, ~4, and ~14 μM, respectively (Fig. 3A and fig. S14). The binding arrangement shown for the three TF molecules bound to PhoA sites *a* through *g* depicted collectively in Fig. 4A represents the lowest-energy conformational states for the three complexes. The binding avidity among the proximal PhoA sites and subsites, enhanced by the proximity of the A, B, and C sites in TF, further accounts for the higher stability of the observed ground-state conformations. This observation is confirmed by additional NMR titration experiments (fig. S16B).

Dynamic Engagement of PhoA by TF

Rapid binding and release of unfolded proteins by molecular chaperones is likely important for a productive assisted-folding process (*5*). To determine the kinetics of PhoA binding to TF, we used NMR relaxation dispersion (*65*, *66*). This powerful technique allows characterization of the dynamic interaction of protein complexes at the residue level. In the case of the interaction between TF and PhoA, the kinetics are as follows:

$$\mathrm{TF} + \mathrm{PhoA} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \mathrm{TF} - \mathrm{PhoA}$$

where k_{on} and k_{off} are the association and dissociation rate constants, respectively. The exchange rate constant, k_{ex}, of complex formation is given by $k_{ex} = k_{on}[\mathrm{PhoA}] + k_{off}$, where [PhoA] is the concentration of unbound PhoA. We have used ^{15}N relaxation dispersion experiments (*67*, *68*) to determine the k_{ex}, and thus k_{on}, k_{off}, and the residence time τ (=1/k_{off}) (*69*, *70*), of PhoA binding to TF (Fig. 3I). The data show that individual PhoA sites, such as *a* and *d*, interact rapidly with TF, and they stay bound to TF only for a very limited time (τ ~1 ms). As expected, the binding of larger PhoA fragments such as PhoA$^{a\text{-}c}$ or PhoA$^{d\text{-}e}$, which bind to TF with much higher affinity (Fig. 3A), results in increased complex lifetimes. However, even in these cases, the kinetics remain fast with the residence time being only ~20 ms at 22°C (Fig. 3K). The data are consistent with previous findings reporting similar rapid kinetics for the binding of TF to protein substrates (*71*). Thus, even when TF fully engages its unfolded protein substrate by using all of its available binding sites, the complex is still quite dynamic with fast association (k_{on} ~2×10^7 M^{-1} s^{-1}) and dissociation (k_{off} ~50 s^{-1}) rates, as determined by relaxation dispersion.

Structural Basis for the Formation of the TF-PhoA Complex

The three TF molecules in complex with PhoA (Fig. 4A) are linked by long (~80 to 90 amino acids), flexible regions of PhoA that are not bound by TF. Thus, the three PhoA-bound TF molecules tumble and behave as independent entities, corroborated by NMR data of the full-length PhoA in complex with three TF molecules (figs. S7A and S15). We thus determined the high-resolution structure of each of the three complexes (TF–PhoA$^{a\text{-}c}$, TF–PhoA$^{d\text{-}e}$, and TF–PhoA$^{f\text{-}g}$; Fig. 4A) separately. In each case, the structure determined by NMR corresponded to the ground-state, lowest-energy structure. A large number of intermolecular NOEs were collected because of the large interacting surface between TF and PhoA and the high sensitivity provided by the methyl-bearing and aromatic residues of the NMR spectra. The structure and NMR statistics are listed in table S1. Resonance dispersion of PhoA increases markedly in its complexes with TF (fig. S10, A and B) owing to the intimate contacts between the two proteins, thus enabling complete resonance assignment and structure characterization of PhoA.

PhoA binding to TF results in TF monomerization, as demonstrated by multiangle light scattering (MALS) and NMR experiments (figs. S1, C to E, and S18). The data further showed that the putative dimeric interface in TF is dynamic and suggested that substrate-binding regions are partially buried at the dimeric interface (figs. S18A and S19), in agreement with biochemical data (*49*). Thus, full engagement of the unfolded substrate by TF requires monomerization of the chaperone.

Structure determination of the TF–PhoA$^{a\text{-}c}$ complex shows that PhoA$^{a\text{-}c}$ binds simultaneously to all four substrate-binding sites, A to D, in TF (Fig. 4, A and B), confirming the initial binding experiments (Fig. 3H). Specifically, PhoA site *a*, which is ~10 residues long, binds to TF site D located in the PPD, whereas the much longer PhoA site *c* (~40 residues long) binds to all three TF sites located in the SBD (A, B, and C) (Fig. 4, A and B). Other than the regions encompassing the binding sites *a* and *c*, all other regions in PhoA$^{a\text{-}c}$ remain flexible in the complex (fig. S15).

PhoA site *a* uses exclusively nonpolar residues consisting primarily of a cluster of Leu residues to interact through intimate van der Waals contacts with a hydrophobic cluster in TF site D consisting primarily of aromatic residues (Fig. 4B, left). Formation of this subcomplex buries a total of ~1340 Å^2 of nonpolar surface and ~410 Å^2 of polar surface. Following T16 in PhoA site *a*, the next to contact TF is F94, the first residue of PhoA site *c*. PhoA site *c* spans ~40 residues (F94-W131) and binds to all three TF sites in the SBD (Fig. 4, A and B). PhoA site *c* traverses the entire SBD and binds to TF sites B, A, and C using its subsites c^1, c^2, and c^3, respectively (Fig. 4, A and B, right). Thus, each one of the binding sites in TF engages about 6 to 10 PhoA residues. Complex formation between the TF SBD and PhoA site *c* buries a total of ~3820 Å^2 of surface. The majority of the surface (~2580 Å^2) consists of nonpolar residues, whereas ~1240 Å^2 consists of polar residues.

PhoA site *c* (via subsite c^1) is anchored at the tip of arm 2 of the SBD by docking two of its hydrophobic residues, F94 and I97, deep into a hydrophobic pocket formed by TF residues of site B (Fig. 4B, right). Then, using both nonpolar and polar contacts, PhoA twists around the SBD arm 2 to interact with TF site A (via subsite c^2) through contacts involving several aromatic residues. Whereas TF sites A and B are located next to each other, TF site C is located on the other end of the SBD (Fig. 1A). PhoA reaches site C by stretching an ~10-residue-long region (G116-A125) that appears to interact only transiently with TF. The interaction between PhoA site *c* (via subsite c^3) and TF site C is again mediated primarily by nonpolar contacts, with the most prominent being the docking of PhoA W131 into a hydrophobic pocket formed by several residues of TF site C (Fig. 4B). Although the vast majority of the contacts between the TF SBD and PhoA site *c* are mediated by hydrophobic residues, several key hydrogen bonds are also present throughout the complex (Fig. 4B and fig. S20). The structure of the TF–PhoA$^{a\text{-}c}$ complex shows that a total of about 40 to 50 residues of an unfolded client protein can directly interact with the four sites of TF.

A second TF molecule binds to the middle region of PhoA, consisting of sites *d* and *e* (Fig. 4A and fig. S17D). The K_d of the TF–PhoA$^{d\text{-}e}$ complex is ~4 μM (Fig. 3A and fig. S14), similar to that of the TF–PhoA$^{a\text{-}c}$ complex. Structure determination of the TF–PhoA$^{d\text{-}e}$ complex showed that TF recognizes and interacts with the PhoA$^{d\text{-}e}$ region in a similar way as with the PhoA$^{a\text{-}c}$ region. The PPD (site D) binds PhoA site *d* (residues K231-W242) through predominantly nonpolar contacts (Fig. 4C). When compared to PPD-PhoA^{a} (Fig. 4B), two distinct features of the PPD-PhoA^{d} subcomplex are, first, that PhoA site *d* extends over a larger surface, thereby engaging additional PPD residues (for example, T165 and V186), and, second, that recognition of PhoA site *d* is enhanced by three favorable hydrogen bonds (Fig. 4C). Formation of the PPD-PhoA^{d} complex buries a total surface of ~1350 Å^2, with ~880 Å^2 being nonpolar and ~470 Å^2 polar.

PhoA site *e* (residues Leu^{277}-Ile^{301}), like PhoA site *c*, engages all three TF sites of the SBD (Fig. 4A), but the overall architecture is distinct in the two complexes. Whereas PhoA site *c* traverses the SBD

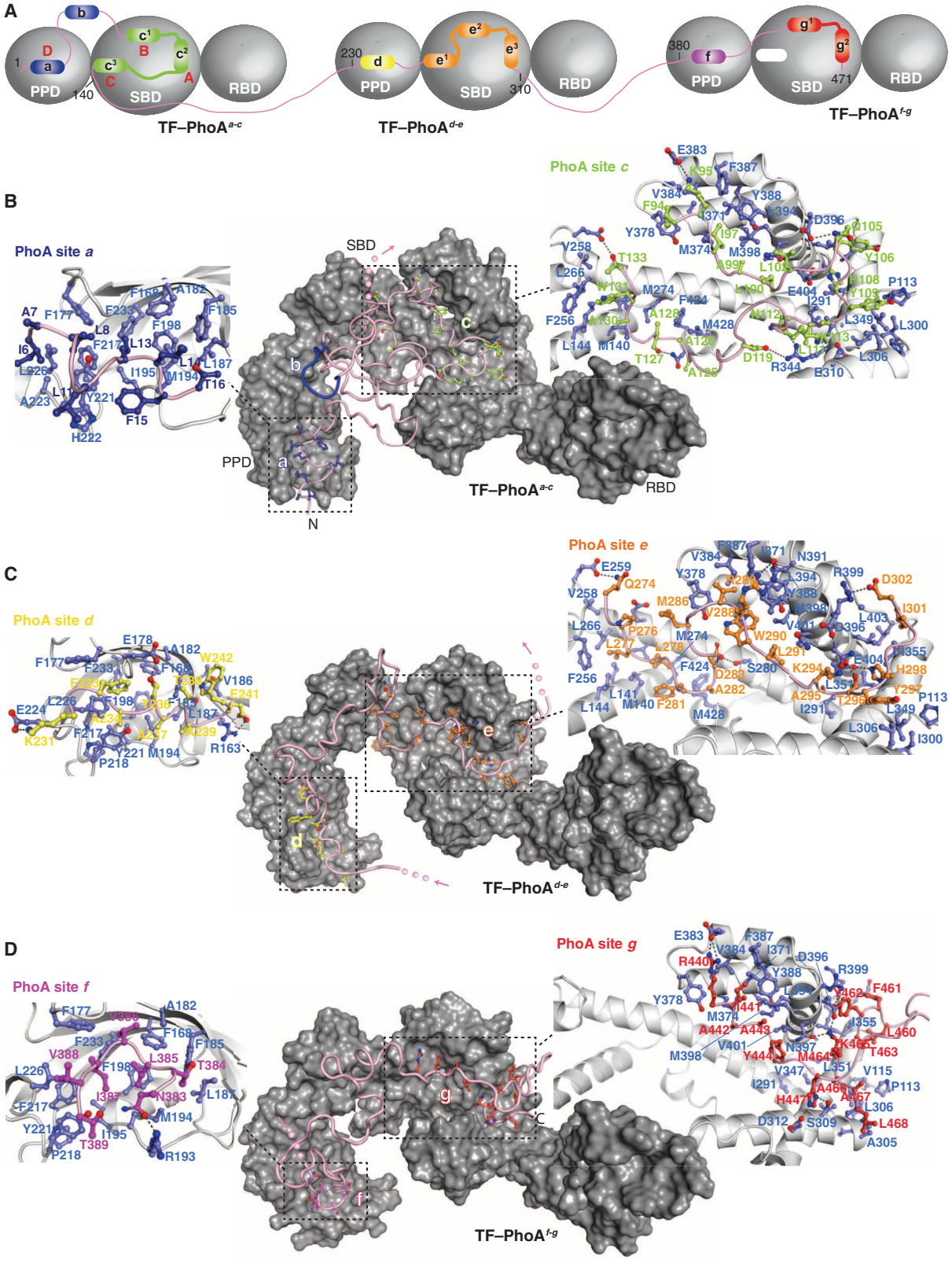


Fig. 4. Structural basis for the formation of the TF-PhoA complexes. (**A**) Schematic for the interaction of PhoA with TF based on the current structural data. The TF-binding sites in PhoA are colored as in Figs. 2 and 3A. (**B** to **D**) Lowest-energy structure of the three TF molecules in complex with the corresponding PhoA regions. TF is shown as a solvent-exposed surface and PhoA as a pink ribbon. The PhoA residues that directly interact with TF are drawn in a ball-and-stick representation, colored as in panel (A). The pink arrows denote the direction of the PhoA chain, from the N to the C terminus. Close-up views of the structures are shown on the left for the PPD (site D) and on the right for the SBD (sites A, B, and C) bound to the corresponding PhoA sites. In the expanded views, the TF backbone is shown as a white cartoon and the PhoA backbone as a pink cartoon. The side chains of TF are colored blue, whereas the side chains of PhoA are colored using the color code for each site. The black broken lines denote hydrogen bonds (present in at least 70% of the conformers of the ensemble).

by interacting first with TF site B, followed by A and C, PhoA site *e* interacts first with TF site C (via subsite e^1), followed by TF site B (via subsite e^2) and TF site A (via subsite e^3) (Fig. 4, A and C, right). Recognition of PhoA site *e* by the TF SBD is mediated predominantly by nonpolar contacts, but several key hydrogen bonds are also present (Fig. 4C and fig. S21). Complex formation between the TF SBD and PhoA site *e* buries a total surface of ~3460 Å^2, with ~2410 Å^2 consisting of nonpolar residues and ~1050 Å^2 consisting of polar residues.

The third TF molecule binds to the C-terminal region of PhoA, consisting of sites *f* and *g* (Fig. 4A and fig. S17G). The K_d of the TF–PhoA$^{f\text{-}g}$ complex is ~14 μM (Fig. 3A and fig. S14), the weakest among the three complexes. Structure determination of the TF–PhoA$^{f\text{-}g}$ complex (the structure of the major conformation out of the possible two was determined; fig. S17G) confirmed that the PPD interacts with PhoA region *f*, and the SBD with PhoA region *g* (Fig. 4, A and D, and fig. S22). PhoA region *f* (residues N383-V388) consisting of only seven residues is the shortest region within PhoA that is recognized and bound by TF (Fig. 2C and fig. S8A). Despite its short length, the nonpolar surfaces between PhoA site *f* and PPD are well juxtaposed, further enhanced by a hydrogen bond, thereby providing sufficient affinity for the formation of a stable complex (Fig. 4D), which buries a total surface of ~1230 Å^2 (~940 Å^2 of nonpolar and ~290 Å^2 of polar surface).

PhoA region *g* (residues Arg^{440}-Leu^{468}) is recognized by the TF SBD. In contrast to PhoA sites *c* and *e*, which both interact with all three TF sites in the SBD (Fig. 4, B and C), PhoA site *g* interacts primarily with TF sites A and B (Fig. 4, A and D, and fig. S17G). PhoA residues Arg^{440}-Tyr^{444} (subsite g^1) interact with TF site B, and PhoA residues L460-L468 (subsite g^2) bind to TF site A (Fig. 4D and fig. S17G). Other than PhoA H447, which forms a bifurcated hydrogen bond with two TF residues (D312 and S309) located in arm 1 of the SBD, none of the other residues of the PhoA region (G445-D459) connecting the g^1 and g^2 subsites appear to interact with TF (Fig. 4D). Complex formation between the SBD and PhoA site *g* buries ~2670 Å^2 of surface (~1840 Å^2 nonpolar and ~830 Å^2 polar), substantially less than the SBD complexes with PhoA sites *c* and *e*, which likely explains the lower affinity of the TF–PhoA$^{f\text{-}g}$ complex (fig. S14).

Amino acid substitutions at the substrate-binding sites in TF resulted in a significant decrease in the affinity for PhoA and a marked decrease of its antiaggregation activity (Fig. 5 and fig. S23).

Recognition of Unfolded PhoA by TF

An important question that remains largely unaddressed is how molecular chaperones recognize unfolded proteins. A common feature shared among molecular chaperones appears to be their interaction with exposed hydrophobic segments of the substrate proteins (*5*). The structural data of TF in complex with PhoA show that TF has a preference for stretches of residues with increased hydrophobicity (Fig. 2C), especially for aromatic residues (Trp, Phe, and Tyr), as well as for large hydrophobic amino acids such as Ile and Leu (fig. S8C). His and Thr residues are also favored (fig. S8C), presumably because they can form both nonpolar and polar contacts. Although there seems to be a good correlation between the sequence hydrophobicity of the substrate protein and its binding preference by TF (Fig. 2C), there are long stretches of hydrophobic residues in PhoA (for example, residues 155 to 185) that are not recognized by TF. This may be due to the complete absence of aromatic residues in this region. The strong preference of TF for aromatic residues may also explain the stronger association of TF with outer membrane proteins (*25*), which are typically enriched in aromatic residues.

The structural data of the three TF molecules in complex with different regions of PhoA (Fig. 4) provide insight into how the same binding sites within a molecular chaperone can recognize and interact with a large number of substrates with unrelated primary sequences. Superposition of the three structures of the TF SBD shows that the binding mode is distinct in the three complexes (Fig. 6 and fig. S24A). The structural data suggest that the two arms of TF are rather dynamic, giving rise to a variable width of the binding crevice in the SBD (fig. S24B), in agreement with NMR dynamic analysis (*53*). In addition, there is notable conformational rearrangement of the side chains in the SBD substrate-binding sites among the three TF complexes with PhoA (Fig. 6E and fig. S24, C and D). Structural analysis of the complexes of the TF SBD with PhoA sites *c*, *e*, and *g* indicated some common binding and recognition patterns (Fig. 6, A to C). For example, TF site B recognizes bulky hydrophobic residues in PhoA that appear to bind to two hydrophobic pockets in TF site B (Fig. 6B). Similarly, there seems to be a pattern, albeit weaker, of bulky and small hydrophobic PhoA residues binding in TF sites A and C (Fig. 6, A and C). A great degree of plasticity is also present in the PPD. For example, PhoA site *f* forms a hairpin-like structure when bound to the PPD, whereas PhoA site *d* binds in an extended conformation (fig. S24, A and F). To accommodate the different PhoA sequences, the residues that make up the PPD binding site appear to undergo significant conformational rearrangement upon complex formation (Fig. 6F).

Structural Basis for the Antiaggregation and Unfoldase Activity of TF

To understand how TF prevents the aggregation of unfolded proteins, we used an aggregation-prone MBP variant as a substrate (Fig. 7A). Such MBP variants have been extensively used as model systems for studying the effect of chaperones on protein aggregation (*72*–*74*). The G32D/I33P double substitution in MBP destabilizes the protein by ~4 kcal mol^{-1} compared to wild-type MBP (MBP^{WT}) (*75*). MBP$^{G32D/I33P}$ (MBP^{mut}) is produced in inclusion bodies in the cell (*75*) and has a high tendency to aggregate in solution (*76*). NMR spectra show that purified MBP^{mut} is folded (Fig. 7B, right, and fig. S25A), in agreement with biochemical (*75*) and crystallographic (*77*) data. However, MBP^{mut} suffers heavy aggregation and precipitation at higher temperatures (above 30°C). Notably, in the presence of TF, MBP^{mut} remains soluble and folded even at temperatures as high as 50°C (fig. S25B).

How does TF prevent aggregation of MBP^{mut}? NMR titration experiments performed at 22°C (Fig. 7B, right, and fig. S25, A and C) show that TF does not interact with the folded state of MBP^{mut}. However, titration of MBP^{mut} to labeled TF at temperatures where MBP^{mut} is prone to aggregation (for example, 38°C) resulted in substantial chemical shift perturbation, indicating complex formation (Fig. 7B, left, and fig. S25, D and E). Thus, unfolded MBP^{mut} is recognized and engaged by TF. In support of this conclusion, titration of

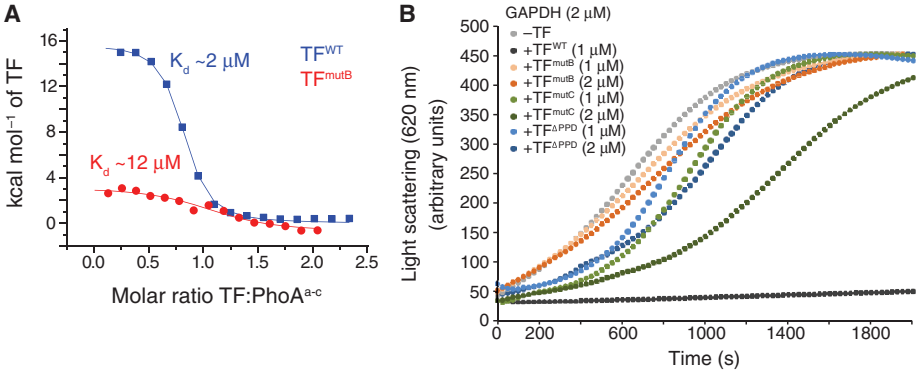


Fig. 5. Effect of mutations in substrate-binding sites on TF activity. (**A**) Binding affinity determined by ITC for PhoA$^{a\text{-}c}$ to TF^{WT} and TF^{mutB}. TF^{mutB} is a mutant in which four hydrophobic amino acids of TF site B were substituted (fig. S23). The mutations result in a sixfold decrease in the affinity. (**B**) Aggregation of chemically denatured glyceraldehyde-3-phosphate dehydrogenase (GAPDH) monitored by light scattering at 620 nm, in the absence or presence of TF^{WT} and its variants. In contrast to TF^{WT}, TF^{mutB} has very poor antiaggregation activity. The antiaggregation activity is also compromised in TF^{mutC}, a mutant in which a single hydrophobic amino acid of TF site C was substituted (fig. S23), as well as in a TF construct wherein the PPD has been deleted (TF$^{\Delta PPD}$).

MBP^{WT} to TF showed no sign of complex formation even at 38°C (fig. S25F). Because of the large difference in their free energy of folding ($\Delta\Delta G_f$ ~4 kcal mol^{-1}, Fig. 7A), the unfolded state will be significantly more populated in MBP^{mut} than in MBP^{WT} at elevated temperatures. NMR analysis (fig. S25, D and E) showed that TF uses all of the possible substrate binding sites to interact with the unfolded state of MBP^{mut} (Fig. 7A and fig. S11C). Together, the results show that multivalent binding between TF and the unfolded state of MBP^{mut} shields and protects the unfolded state from the solvent and results in the prevention of its aggregation (Fig. 7A).

Although TF binds to the unfolded state of MBP^{mut}, the NMR spectra show that TF cannot unfold the protein even at temperatures as high as 50°C (fig. S25B). Because TF lacks adenosine triphosphatase activity, it could, in principle, unfold protein substrates only if the free energy of binding (ΔG_b) of the substrate's unfolded state for TF is much larger than the substrate's intrinsic energy for folding (Fig. 7A). The intrinsic affinity of TF for unfolded substrates is relatively low (ΔG_b ~–7.5 kcal mol^{-1}; fig. S14), and thus, it is expected to unfold only marginally stable proteins. However, the interaction of ribosome-bound TF with the nascent polypeptide is very strong because of their colocalization (*78*). When MBP^{mut} was fused to TF, it was completely unfolded as demonstrated by the disappearance of resonances corresponding to residues located in folded regions and the appearance of resonances in the so-called random-coil region (Fig. 7C). NMR analysis showed that TF makes use of all of its substrate-binding sites to interact and unfold MBP^{mut} (Fig. 7A). Of particular note, maltose binding to MBP^{mut} confers stability (*79*) and prevents its unfolding by TF (Fig. 7C, right). These results are in excellent agreement with elegant biochemical data indicating that TF has a negligible unfoldase activity when free in solution but a strong unfoldase activity when bound to the ribosome (*30*).

Discussion

The TF molecular chaperone engages its unfolded substrate (PhoA) in a highly dynamic manner, giving rise to a rugged free energy landscape (Fig. 3H). It is expected that the features of the binding energy landscape will vary among the various molecular chaperones and even among different substrates binding to the same chaperone. In the case of PhoA binding to TF, we observed a gradual increase in binding affinity as the length of the engaged PhoA increased until all four of the binding sites in TF were fully occupied (Fig. 3 and fig. S17). Because of the topology of the TF-binding sites in PhoA, the selectivity of some of the PhoA sites for specific TF sites (Fig. 3 and fig. S17), and the observed synergistic binding among PhoA sites close to each other (Fig. 3 and fig. S16), the lowest-energy complexes between TF and the three PhoA regions were experimentally observed to exist in one predominant conformation (Fig. 4A and fig. S17, D and G). Notably, the observed binding avidity of the PhoA sites and subsites is relatively weak (Fig. 3 and fig. S17). This is probably because the linker tethering the (sub)sites remains dynamic even in complex with TF, thus allowing for somewhat independent binding of the PhoA sites. The observed weak synergistic binding may be crucial in enabling a dynamic binding and release of the unfolded protein substrate from the chaperone (Fig. 3K).

The excellent quality of the NMR spectra enabled the high-resolution structure determination for each one of the three TF molecules in complex with different regions of PhoA (Fig. 4). Analysis of

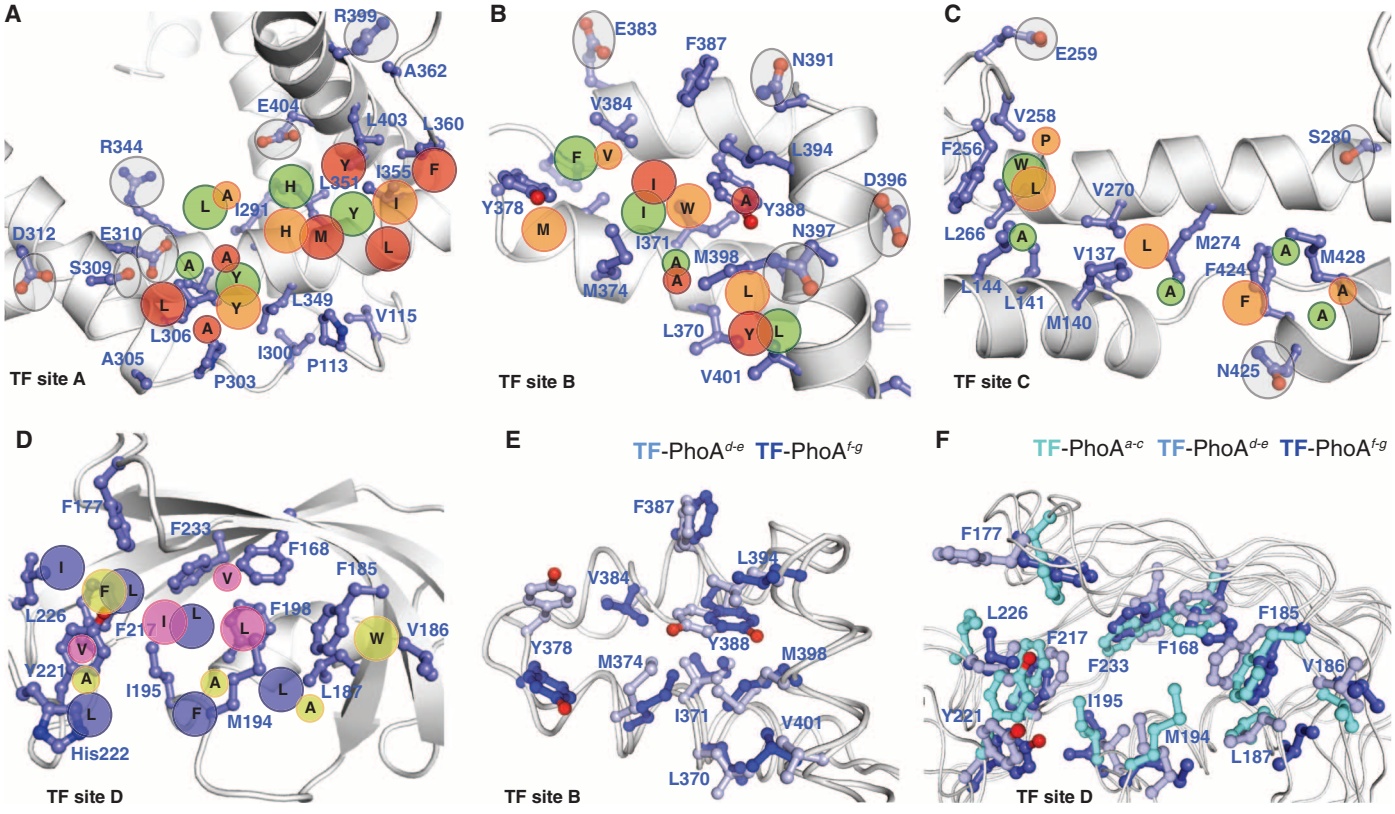


Fig. 6. Substrate recognition conformational plasticity by TF. (**A** to **D**) Interaction and recognition patterns of the PhoA sites in the complexes with the three TF molecules. A close-up view of each one of the four substrate-binding sites in TF (labeled TF site A, B, C, and D) is shown, with the backbone in white ribbon and the interacting residues in blue sticks. The position of hydrophobic residues from PhoA sites interacting with TF is projected on the TF structure and denoted as large (for aromatic or Ile/Leu residues) or small (for Ala/Val residues) circles. For clarity, only one TF molecule is shown. The substrate-binding sites A, B, and C in TF are decorated with polar residues (highlighted with gray circles) that form hydrogen bonds with PhoA. Only PhoA residues involved in nonpolar contacts are shown. The color code of the PhoA residues (in circles) is as follows: site *a*, blue; site *c*, green; site *d*, yellow; site *e*, orange; site *f*, magenta; site *g*, red. (**E** and **F**) Superposition of the TF substrate-binding sites [site B in (E), site D in (F)] in the indicated PhoA-bound complexes. PhoA is not shown for clarity. The TF side chains undergo significant rearrangement to interact with the different PhoA regions.

the structures suggests that the plasticity of the substrate-binding sites in TF, combined with the presence of numerous small hydrophobic pockets on the binding surface and the decoration of these sites by polar residues that can form hydrogen bonds (Fig. 6), is important for the promiscuous recognition of multiple substrate sequences by TF. TF presents a multivalent binding surface to its protein substrate, which binds to TF using several distinct regions. This mechanism of binding presents several advantages because it enables chaperones to function as holdases and unfoldases by exerting forces to retain proteins in the unfolded state, and at the same time protect them from aggregation by shielding their exposed hydrophobic regions. Impairment of the substrate-binding sites in TF results in significant reduction in TF's antiaggregation activity (Fig. 5). Given the existence of multiple binding sites in other molecular chaperones (such as the GroEL), this may present a general mechanism for the action of molecular chaperones (*80*–*82*).

TF appears to have a weak holdase and unfoldase activity in the cytosol but a very strong activity when colocalized with the protein substrate. Thus, in the cytosol, TF is apparently capable of rescuing folded proteins by transiently interacting with the aggregation-prone unfolded population of the substrate, without destabilizing the protein substrate (as seen with MBP^{mut}; Fig. 7, A and B, and fig. S25B). In contrast, when TF is bound to the ribosome, its strong holdase and unfoldase activity (Fig. 7C) will mean that TF can retain the nascent polypeptide in an unfolded state and even actively unfold transiently formed structures. Although PhoA has some fraction of secondary structure present in its unfolded state (Fig. 2, A and B), these regions are devoid of any secondary structure when they are bound to TF (Fig. 4). Together, the data demonstrate that binding of substrate proteins to TF may result not only in global unfolding of the substrate protein but also in melting of secondary structural elements.

The present structural data explain the observation that multiple TF molecules are required to bind to long nascent chains (*18*, *19*, *25*) because a single TF molecule can only accommodate a stretch of about 40 to 50 binding residues of the substrate protein. Moreover, the weak affinity of the first ~70 residues of PhoA for TF, as well as the fact that the predicted TF-binding sites in a number of proteins known to bind strongly to TF (*25*) are located past the first ~70 residues (fig. S26), is in agreement with ribosome profiling data (*25*) that nascent chains with more than ~70 residues protruded from the ribosome exit channel are required for a stable complex with TF. Another interesting aspect of TF is its ability to bind and protect small folded proteins as a dimer (*27*) (fig. S27). The ability of TF to act as a holdase and an unfoldase with unfolded proteins and to encapsulate folded proteins in its dimeric state highlights the versatility of this particular molecular chaperone and the multiple roles it may play in protein folding (*83*).

Materials and Methods

Expression and Preparation of Proteins

The *E. coli* TF was cloned into the pCold vector (Takara Bio). The following TF constructs were prepared: the RBD (residues 1 to 117) and the SBD (residues 113–432Δ150–246) were cloned into pET16b vector (Novagen) and fused to His_6-MBP, including a tobacco etch virus (TEV) protease cleavage site. The PPD (residues 148 to 249), SBD-PPD (residues 113 to 246), and RBD-SBD (1–432Δ150–243) were cloned into pCold vector (Takara Bio). TF mutants were constructed by site-directed mutagenesis using PfuTurbo High Fidelity DNA polymerase (Agilent). All constructs were transformed into BL21(DE3) cells. The following *E. coli* PhoA constructs were prepared (residue numbers of the boundaries are in superscript): $PhoA^{1–471}$ (full length), $PhoA^{1–62}$, $PhoA^{1–82}$, $PhoA^{1–122}$, $PhoA^{1–140}$, $PhoA^{1–150}$, $PhoA^{1–250}$, $PhoA^{23–78}$, $PhoA^{76–150}$, $PhoA^{119–243}$, $PhoA^{1–232}$, $PhoA^{234–349}$, $PhoA^{228–471}$, $PhoA^{349–471}$, $PhoA^{220–310}$, $PhoA^{365–471}$, $PhoA^{228–245}$, $PhoA^{272–327}$, $PhoA^{381–438}$, and $PhoA^{425–471}$. All PhoA fragments were cloned into a pET16b vector and fused to His_6-MBP, including a TEV protease cleavage site. *E. coli* MBP, $MBP^{29–132}$, and $MBP^{G32D/I33P}$ (*75*) were cloned into

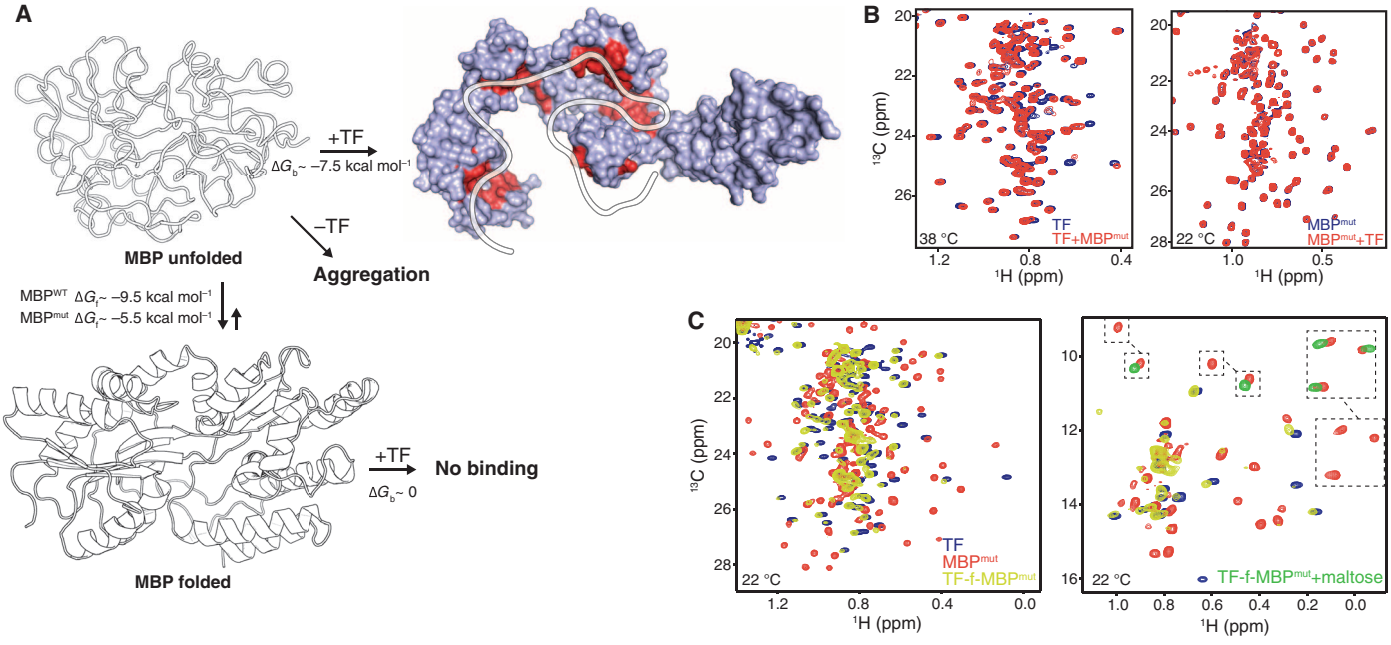


Fig. 7. Antiaggregation and unfoldase activity of TF. (**A**) Schematic of MBP unfolding and binding reactions between TF and MBP^{mut}. ΔG_f is the free energy of folding and ΔG_b is the free energy of binding. The sites that TF uses to interact with the unfolded state of MBP^{mut}, as determined by NMR, are colored red. (**B**) ^1H-^{13}C–correlated methyl NMR spectra of the interaction between TF and MBP^{mut}. The spectrum shown on the left was recorded at 38°C using labeled TF and unlabeled MBP^{mut}, whereas the spectrum shown on the right was recorded at 22°C using the reverse labeling scheme. No interaction was observed at 22°C, where the unfolded population of MBP^{mut} was negligible, whereas binding was observed at 38°C, where the unfolded population of MBP^{mut} was appreciable. For clarity, only the region of the Leu and Val methyl resonances is shown. The full-range spectra are shown in fig. S25. (**C**) ^1H-^{13}C–correlated methyl NMR spectra of the fusion construct between TF and MBP^{mut} (TF-f-MBP^{mut}, yellow). The spectra of free TF (blue) and free MBP^{mut} (orange) are also shown for comparison. The spectrum on the left depicts the region of the Leu and Val methyl resonances, whereas the one on the right depicts the region of the Ile methyl resonances. The disappearance of the peaks belonging to MBP^{mut} residues residing in folded regions and the appearance of peaks in the random-coil region indicates that MBP^{mut} is unfolded when fused to TF. Addition of maltose (green resonances on the Ile spectrum) stabilizes MBP^{mut} and prevents its unfolding by TF.

pET 16b vector (Novagen). The TF-MBP$^{I32D/G33P}$ fusion protein was cloned into pCold vector (Takara Bio) to be fused to TF or TF$^{SBD-PPD}$ separated by a 25-residue linker. The transmembrane region of OmpA (OmpA^{t}; residues 1 to 192) fused with N-terminal His-tag was cloned into pET16b vector (Novagen). For the unlabeled samples, cells were grown in Luria-Bertani (LB) medium at 37°C in the presence of ampicillin (100 μg ml^{-1}). Protein expression was induced by the addition of 0.1 to 0.8 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) at OD_{600} ~0.5, followed by 12 to 16 hours of incubation at 18°C. Expression of MBP$^{G32D/I33P}$ was induced by the addition of 0.01 mM IPTG at OD_{600} ~0.5, followed by 48 hours of incubation at 16°C. Cells were harvested at OD_{600} ~1.5 and resuspended in lysis buffer containing 50 mM tris-HCl (pH 8.0), 50 to 500 mM NaCl, 5 mM imidazole, and 1 mM β-mercaptoethanol (βME) (and 8 M urea in the case of PhoA fragments followed by 20-fold rapid dilution in lysis buffer). For OmpA, the cell pellet was resuspended in buffer containing 50 mM tris-HCl (pH 8.0), 500 mM NaCl, 5 mM imidazole, and 6 M guanidine hydrochloride and incubated for 1 hour at room temperature. Cells were disrupted by a high-pressure homogenizer or sonicator and centrifuged at 50,000*g*. Proteins were purified using Ni Sepharose 6 Fast Flow resin (GE Healthcare), followed by tag removal by TEV protease at 4°C (incubation for 16 hours) and gel filtration using Superdex 75 16/60 or 200 16/60 columns (GE Healthcare). OmpA was eluted with a solution containing 50 mM tris-HCl (pH 8.0), 500 mM NaCl, 400 mM imidazole, and 6 M guanidine hydrochloride. The eluted protein was concentrated and diluted 20-fold by rapid dilution in the presence of TF or TF$^{PPD-SBD}$ in 20 mM potassium phosphate (pH 7.0), 100 mM KCl, 4 mM βME, 0.5 mM EDTA, 0.05% NaN_3. Protein concentration was determined spectrophotometrically at 280 nm using the corresponding extinction coefficient.

MALS Experiments

Multiangle light scattering was measured using DAWN HELEOS-II (Wyatt Technology Corporation) downstream of a Shimadzu liquid chromatography system connected to a Superdex 200 10/300 GL (GE Healthcare) gel filtration column. The typical running buffer was 20 mM KPi (pH 7.0), 100 mM KCl, 4 mM βME, and 0.5 mM EDTA. Two hundred microliters of the sample with a concentration of 100 to 400 μM was injected. The data were analyzed with ASTRA version 6.0.5 (Wyatt Technology Corporation).

GAPDH Aggregation Assay

Aggregation of denatured GAPDH from rabbit muscle (Sigma, G-2267) was measured are described previously (*84*). GAPDH (125 μM) was denatured by 3 M guanidine-HCl in 20 mM KPi (pH 7.0), 100 mM KCl, 4 mM βME, 0.5 mM EDTA, and 0.05% NaN_3 for 12 hours at 4°C. The denatured enzyme was diluted 50-fold in a buffer that does not contain guanidine-HCl, and aggregation was monitored by 90° light scattering at 620 nm on a spectrofluorometer (FluoroMax-4, Horiba) in the presence or absence of TF. The experiment was carried out at 20°C.

Substrate Binding Site Mutations in TF

To assess the effect of reducing the hydrophobicity of the substrate-binding sites of TF on its binding and antiaggregation activity, we sought to mutate the most prominent residues in each one of the binding sites located in the SBD (that is, sites A, B, and C). The following mutants were prepared: L306A, G348E, G352E, I355A (site A); M374A, Y378A, V384A, F387A (site B); M140E, L144A, F256A L266A, M274A, M428A (site C) as single or combined mutants. Although all of the mutants were expressed and purified in a soluble form, NMR characterization indicated that mutations other than the ones shown in fig. S23 are not well tolerated, with the mutated TF proteins showing signs of misfolding and aggregation.

ITC Experiments

Calorimetric titrations were carried out on an iTC200 microcalorimeter (GE Healthcare) at temperatures ranging from 8° to 25°C. All protein samples were extensively dialyzed against ITC buffer containing 20 mM potassium phosphate (pH 7.0), 100 mM KCl, and 1 mM tris(2-carboxyethyl)phosphine. All solutions were filtered using membrane filters (pore size, 0.45 μm) and thoroughly degassed for 20 min before the titrations. For the experiments performed on iTC200, the 200-μl sample cell was filled with ~40 to 100 μM protein (TF or PhoA), and the 40-μl injection syringe was filled with 0.5 to 1 mM solution of protein (PhoA or TF). The titrations were carried out with a preliminary 0.2-μl injection, followed by 15 injections of 2.5 μl each with time intervals of 3 min. The solution was stirred at 1000 rpm. For the experiments performed on VP-ITC, the 1200-μl sample cell was filled with ~40 to 100 μM protein (TF or PhoA), and the 250-μl injection syringe was filled with 0.5 to 1 mM solution of protein (PhoA or TF). The titrations were carried out with 30 injections of 10 μl each with time intervals of 5 min. The solution was stirred at 270 rpm. Data for the preliminary injection, which are affected by diffusion of the solution from and into the injection syringe during the initial equilibration period, were discarded. Binding isotherms were generated by plotting heats of reaction normalized by the modes of injectant versus the ratio of total injectant to total protein per injection. The data were fitted with Origin 7.0 (OriginLab Corporation).

NMR Spectroscopy

NMR samples are prepared in 20 mM KPi (pH 7.0), 100 mM KCl, 4 mM βME, 0.5 mM EDTA, 0.05% NaN_3, and 7% ^2H_2O. Protein concentration was 0.1 to 0.8 mM. NMR experiments were recorded on Agilent UNITY Inova 800 and 600 MHz NMR spectrometers and Bruker Avance III 600 and 700 MHz NMR spectrometers equipped with cryogenic probes. The experiments were run at 10°C for isolated PhoA samples, and at temperatures ranging from 18° to 50°C for the other samples. NMR spectra for the assignment and structure determination of the TF-PhoA complexes were recorded at 22°C. Spectra were processed using the NMRPipe program (*85*), and data analysis was performed with Olivia (fermi.pharm.hokudai.ac.jp/olivia) and SPARKY software.

Binding affinities between TF and PhoA fragments were estimated using ^1H-^{15}N heteronuclear single-quantum coherence (HSQC) titration experiments, where unlabeled PhoA was titrated into labeled TF. Titration curves were obtained by plotting chemical shift perturbations ($\Delta\delta_{ppm}$) against the molar ratio of PhoA and TF. Nonlinear least square fitting calculations were performed in GraphPad Prism (GraphPad Software), using the following equation:

$$\Delta\delta_{ppm} = \Delta\delta_{ppm}^{bound} \times \frac{[P]+[L]+K_d-\sqrt{([P]+[L]+K_d)^2-4[P][L]}}{2[P]}$$

where $[L]$ and $[P]$ are the concentrations of the peptide ligand (PhoA) and protein (TF), respectively, and K_d is the dissociation constant.

Protein Isotope Labeling for NMR Studies

Isotopically labeled samples for NMR studies were prepared by growing the cells in minimal (M9) medium. Cells were typically harvested at OD_{600} ~1.0. U-[^2H,^{13}C,^{15}N]–labeled samples were prepared for the backbone assignment by supplementing the growing medium with $^{15}NH_4Cl$ (1 g liter^{-1}) and $^2H_7$,$^{13}C_6$-glucose (2 g liter^{-1}) in 99.9% ^2H_2O (CIL and Isotec). The methyl-protonated samples (*86*) were prepared as described (*44*, *47*). α-Ketobutyric acid (50 mg liter^{-1}) and α-ketoisovaleric acid (85 mg liter^{-1}) were added to the culture 1 hour before the addition of IPTG. Met-[$^{13}CH_3$]– and Ala-[$^{13}CH_3$]–labeled samples were produced by supplementing the medium with [$^{13}CH_3$]methionine (50 mg liter^{-1}) and [^2H,$^{13}CH_3$]alanine (50 mg liter^{-1}). For Thr labeling, a Thr-auxotrophic cell strain was used, and the medium was supplemented with [^2H,$^{13}CH_3$]threonine (25 mg liter^{-1}) prepared as described (*87*). For Phe, Tyr, and Trp labeling, U-[^1H,^{13}C]–labeled amino acids were used. All precursors and amino acids were added to the culture 1 hour before the addition of IPTG, except Ala, which was added 30 min before induction (*47*).

NMR Assignment of TF

The resonances of the full-length dimeric TF (~100 kD) were assigned by following a domain-parsing approach enabled by the modular architecture of TF. The RBD (*52*), SBD (*53*), and PPD are folded when isolated (fig. S3B). Overlay of the NMR spectra for each one of the domains with the spectra of the full-length TF indicated very good resonance correspondence for the SBD and PPD (figs. S2 to S4) but more significant chemical shift

differences for the RBD (fig. S19). About 70 to 80% of the assignment performed on the isolated domains could be transferred to the full-length TF. Near-complete assignment of TF was achieved by recording nuclear Overhauser effect spectroscopy (NOESY) in full-length TF to resolve ambiguities. In addition, $TF^{SBD\text{-}PPD}$ was prepared and fully characterized by NMR because this fragment is monomeric and contains all of the substrate binding sites. Backbone and side-chain resonances for the domains were assigned using standard triple-resonance NMR pulse sequences. Methyl side-chain and amide resonances for TF and large fragments were confirmed by recording the following NMR experiments [the standard version was used for isolated RBD and PPD and the TROSY (transverse relaxation optimized spectroscopy) version, where applicable, for $TF^{SBD\text{-}PPD}$ and TF]: three-dimensional (3D) (^{1}H)-^{13}C heteronuclear multiple-quantum coherence (HMQC)–NOESY–^{1}H-^{13}C HMQC, ^{13}C-edited NOESY-HSQC, ^{13}C-edited HSQC-NOESY, ^{15}N-edited NOESY-HSQC, 3D (^{1}H)-^{13}C HSQC-NOESY–^{1}H-^{15}N HSQC, and 3D (^{1}H)-^{15}N HSQC-NOESY–^{1}H-^{13}C HSQC.

NMR Assignment of PhoA

Resonance assignment of full-length PhoA was accomplished by first assigning short PhoA fragments followed by assignment transfer. The backbone and side-chain resonances of the following PhoA fragments were fully assigned using standard triple-resonance experiments: $PhoA^{1–122}$, $PhoA^{119–243}$, $PhoA^{234–349}$, and $PhoA^{349–471}$ (the superscript denotes the amino acid residue boundaries of the fragment). Overlay of the spectra of these fragments with the spectra of full-length PhoA indicated excellent resonance correspondence (fig. S5D). This is expected because all of the fragments as well as the full-length PhoA are unfolded. Resonance assignment obtained for the various fragments was transferred to the full-length PhoA, and ambiguities were resolved by the use of 3D NMR spectra. Resonance assignment of PhoA in complex with TF was achieved by tracing the PhoA NMR signals as increasing amounts of unlabeled TF was titrated. Complex formation was on the fast time scale regime, allowing the monitoring of the chemical shift changes in a straightforward manner. Ambiguities were resolved by recording 3D NMR spectra. It should be noted that although resonance dispersion in unliganded PhoA is poor, complex formation with TF alleviates this problem (for the PhoA residues in the TF-binding regions) with the spectra being of high resolution (figs. S2, C and D, and S10).

Structure Determination

In the first step, we used a U-^{12}C,^{15}N–labeled sample that contained specifically protonated methyl groups of Ala, Val, Leu, Met, Thr, and Ile (δ1) and protonated aromatic residues Phe, Tyr, and Trp in an otherwise deuterated background (fig. S4). The high sensitivity and resolution of the methyl region, combined with the high abundance of these eight amino acids in TF (~50% of TF residues), provided a large number of intra- and intermolecular TF-PhoA NOEs sufficient to yield an initial well-defined structure. The initial structure was used to guide the subsequent collection of a much larger number of NOEs, using ^{1}H,^{13}C–labeled samples that resulted in further refinement to precise and accurate coordinates. In the second step, additional NMR restraints were collected for complexes between isolated SBD and PPD, as well as the entire chaperone module $TF^{SBD\text{-}PPD}$, in complex with the corresponding PhoA regions (fig. S10). NMR analysis demonstrated that the structures of the individual SBD and PPD in complex with PhoA are essentially identical to the structure of the full-length TF in complex with PhoA.

The initial structures of TF in complex with PhoA were calculated with CYANA 3.0 (*88*), using NOE peak lists from 3D (^{1}H)-^{13}C HMQC-NOESY–^{1}H-^{13}C HMQC, 3D (^{1}H)-^{15}N HSQC-NOESY–^{1}H-^{13}C HSQC, ^{13}C-edited NOESY-HSQC, and ^{15}N-edited NOESY-HSQC. The ^{13}Cα, ^{13}Cβ, ^{13}C′, Hα, ^{15}N, and NH chemical shifts served as input for the TALOS+ program (*89*) to extract dihedral angles (φ and ψ). Assigned NOE restraints were then used in the second round of calculations using Xplor-NIH (*90*) consisting of two steps. In the first step, the orientation and coordinates of PhoA polypeptide were randomized and then annealed to TF using intermolecular NOEs. TF was kept rigid but PhoA was set as flexible. The best 200 structures were selected and further refined in the second step, where both TF and PhoA were set flexible. Because the RBD does not interact with PhoA, or any other tested unfolded substrate, the coordinates from the *E. coli* TF structure (PDB ID 1W26) were used to model RBD in the current structures. The RBD was held rigid during the entire structure calculation protocol. The structure of the main chaperone module, $TF^{SBD\text{-}PPD}$, was determined using the NMR restraints in each one of the three complexes with $PhoA^{a\text{-}c}$, $PhoA^{d\text{-}e}$, and $PhoA^{f\text{-}g}$. The structures were calculated and refined using all of the restraints, including intra- and intermolecular NOEs, dihedral restraints derived from TALOS+, and hydrogen bond restraints for amides located in helices as determined via NOEs and chemical shift information. Hydrogen bonds between TF and PhoA were assigned when satisfactory geometry was present in at least 70% of the conformers. The 20 lowest-energy structures were finally refined in explicit water with CNS (*91*).

NMR Relaxation Experiments

Three relaxation parameters were measured for backbone amides of full-length PhoA, as well as for $PhoA^{1–232}$ and $PhoA^{228–471}$ for enhanced resolution: the {^{1}H}-^{15}N NOE, the longitudinal relaxation rate R_1, and the transverse relaxation rate R_2. ^{15}N R_1 values were measured from 2D spectra recorded with relaxation delays of 20, 40, 80, 160, 320, 640, 1200, 2500, and 5000 ms; ^{15}N R_2 values were measured from 2D spectra recorded with relaxation delays of 31.7, 47.5, 63.4, 79.2, 95.0, 110.9, 126.7, 158.4, 237.6, 316.8, and 475.2 ms. Data sets were acquired as 100 × 1024 complex points in the $t_1 \times t_2$ time-domain dimensions. Data points were fitted as a function of the length of the parametric relaxation delay to two-parameter decay curves of the form $I(t) = I^0 e^{-Rt}$, where I is the peak intensity. To measure the exchange rate and the kinetics ($k_{ex} = k_{on}[\text{PhoA}] + k_{off}$) of complex formation between TF and PhoA, we used NMR relaxation dispersion experiments (*65*, *66*). An ^{15}N-Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence (*67*, *68*) was used. Two-dimensional data sets were acquired as 100 × 1024 complex points in the $t_1 \times t_2$ time-domain dimensions with a constant relaxation delay of 40 ms. Experiments were recorded at spectrometer frequencies of 600 and 700 MHz. NMR titration experiments showed that formation of all of the TF-PhoA complexes is fast on the NMR chemical shift time scale. k_{ex} was measured by using the Carver-Richards equation (*92*) as applied for fast exchange processes between two sites. Exchange rates were first extracted independently for each residue with appreciable exchange contribution (R_{ex} >5 Hz) for each one of the four binding sites in TF. The similarity of exchange rates extracted from fits of dispersion profiles on a per-residue basis for the four binding sites in TF indeed supports a two-site model of exchange (binding) even for the binding of the larger PhoA sites such as *c* and *e*. k_{on} and k_{off} rates were measured by fitting the relaxation dispersion data obtained for a set of experiments recorded using different PhoA concentrations. Data were analyzed using the programs CPMGFit (palmer.hs.columbia.edu/software/cpmgfit.html) and CPMG_fit provided by L. Kay (University of Toronto). $TF^{SBD\text{-}PPD}$ concentration for the relaxation dispersion experiments was ~0.4 to 0.8 mM in the presence of substoichiometric (~5 to 15%) amount of the corresponding PhoA fragment. It should be noted that the exchange process detected and measured is the binding and release of PhoA from TF as opposed to conformational heterogeneity of PhoA while bound to TF. This is supported by the PhoA concentration–dependent k_{ex} values we measured.

References and Notes

1. B. Bukau, J. Weissman, A. Horwich, Molecular chaperones and protein quality control. *Cell* **125**, 443–451 (2006). doi: 10.1016/j.cell.2006.04.014; pmid: 16678092
2. G. Kramer, D. Boehringer, N. Ban, B. Bukau, The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat. Struct. Mol. Biol.* **16**, 589–597 (2009). doi: 10.1038/nsmb.1614; pmid: 19491936
3. F. U. Hartl, M. Hayer-Hartl, Converging concepts of protein folding in vitro and in vivo. *Nat. Struct. Mol. Biol.* **16**, 574–581 (2009). doi: 10.1038/nsmb.1591; pmid: 19491934
4. A. L. Horwich, Protein folding in the cell: An inside story. *Nat. Med.* **17**, 1211–1216 (2011). doi: 10.1038/nm.2468; pmid: 21989012
5. F. U. Hartl, A. Bracher, M. Hayer-Hartl, Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332 (2011). doi: 10.1038/nature10317; pmid: 21776078
6. H. Saibil, Chaperone machines for protein folding, unfolding and disaggregation. *Nat. Rev. Cancer* **13**, 630–642 (2013). doi: 10.1038/nrm3658; pmid: 24026055

7. F. Hartl, M. Hayer-Hartl, Molecular chaperones in the cytosol: From nascent chain to folded protein. *Science* **295**, 1852–1858 (2002). doi: 10.1126/science.1068408; pmid: 11884745
8. J. Young, V. Agashe, K. Siegers, F. Hartl, Pathways of chaperone-mediated protein folding in the cytosol. *Nat. Rev. Mol. Cell Biol.* **5**, 781–791 (2004). doi: 10.1038/nrm1492; pmid: 15459659
9. S. Preissler, E. Deuerling, Ribosome-associated chaperones as key players in proteostasis. *Trends Biochem. Sci.* **37**, 274–283 (2012). doi: 10.1016/j.tibs.2012.03.002; pmid: 22503700
10. A. Hoffmann, B. Bukau, G. Kramer, Structure and function of the molecular chaperone Trigger Factor. *Biochim. Biophys. Acta* **1803**, 650–661 (2010). doi: 10.1016/j.bbamcr.2010.01.017; pmid: 20132842
11. G. Kramer *et al.*, L23 protein functions as a chaperone docking site on the ribosome. *Nature* **419**, 171–174 (2002). doi: 10.1038/nature01047; pmid: 12226666
12. L. Ferbitz *et al.*, Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* **431**, 590–596 (2004). doi: 10.1038/nature02899; pmid: 15334087
13. D. Baram *et al.*, Structure of trigger factor binding domain in biologically homologous complex with eubacterial ribosome reveals its chaperone action. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12017–12022 (2005). doi: 10.1073/pnas.0505581102; pmid: 16091460
14. F. Schlünzen *et al.*, The binding mode of the trigger factor on the ribosome: Implications for protein folding and SRP interaction. *Structure* **13**, 1685–1694 (2005). doi: 10.1016/j.str.2005.08.007; pmid: 16271892
15. F. Merz *et al.*, Molecular mechanism and structure of Trigger Factor bound to the translating ribosome. *EMBO J.* **27**, 1622–1632 (2008). doi: 10.1038/emboj.2008.89; pmid: 18497744
16. B. Bukau, E. Deuerling, C. Pfund, E. A. Craig, Getting newly synthesized proteins into shape. *Cell* **101**, 119–122 (2000). doi: 10.1016/S0092-8674(00)80806-5; pmid: 10786831
17. R. Maier, B. Eckert, C. Scholz, H. Lilie, F.-X. Schmid, Interaction of trigger factor with the ribosome. *J. Mol. Biol.* **326**, 585–592 (2003). doi: 10.1016/S0022-2836(02)01427-4; pmid: 12559924
18. C. M. Kaiser *et al.*, Real-time observation of trigger factor function on translating ribosomes. *Nature* **444**, 455–460 (2006). doi: 10.1038/nature05225; pmid: 17051157
19. A. Rutkowska *et al.*, Dynamics of trigger factor interaction with translating ribosomes. *J. Biol. Chem.* **283**, 4124–4132 (2008). doi: 10.1074/jbc.M708294200; pmid: 18045873
20. R. Lill, E. Crooke, B. Guthrie, W. Wickner, The "trigger factor cycle" includes ribosomes, presecretory proteins, and the plasma membrane. *Cell* **54**, 1013–1018 (1988). doi: 10.1016/0092-8674(88)90116-X; pmid: 3046750
21. V. R. Agashe *et al.*, Function of trigger factor and DnaK in multidomain protein folding: Increase in yield at the expense of folding speed. *Cell* **117**, 199–209 (2004). doi: 10.1016/S0092-8674(04)00299-5; pmid: 15084258
22. A. Hoffmann *et al.*, Trigger factor forms a protective shield for nascent polypeptides at the ribosome. *J. Biol. Chem.* **281**, 6539–6545 (2006). doi: 10.1074/jbc.M512345200; pmid: 16407311
23. S. Tomic, A. E. Johnson, F. U. Hartl, S. A. Etchells, Exploring the capacity of trigger factor to function as a shield for ribosome bound polypeptide chains. *FEBS Lett.* **580**, 72–76 (2006). doi: 10.1016/j.febslet.2005.11.050; pmid: 16359675
24. E. Deuerling *et al.*, Trigger Factor and DnaK possess overlapping substrate pools and binding specificities. *Mol. Microbiol.* **47**, 1317–1328 (2003). doi: 10.1046/j.1365-2958.2003.03370.x; pmid: 12603737
25. E. Oh *et al.*, Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell* **147**, 1295–1308 (2011). doi: 10.1016/j.cell.2011.10.044; pmid: 22153074
26. S. K. Lakshmipathy, R. Gupta, S. Pinkert, S. A. Etchells, F. U. Hartl, Versatility of trigger factor interactions with ribosome-nascent chain complexes. *J. Biol. Chem.* **285**, 27911–27923 (2010). doi: 10.1074/jbc.M110.134163; pmid: 20595383
27. E. Martinez-Hackert, W. A. Hendrickson, Promiscuous substrate recognition in folding and assembly activities of the trigger factor chaperone. *Cell* **138**, 923–934 (2009). doi: 10.1016/j.cell.2009.07.044; pmid: 19737520
28. E. P. O'Brien, J. Christodoulou, M. Vendruscolo, C. M. Dobson, Trigger factor slows co-translational folding through kinetic trapping while sterically protecting the nascent chain from aberrant cytosolic interactions. *J. Am. Chem. Soc.* **134**, 10920–10932 (2012). doi: 10.1021/ja302305u; pmid: 22680285
29. A. Mashaghi *et al.*, Reshaping of the conformational search of a protein by the chaperone trigger factor. *Nature* **500**, 98–101 (2013). doi: 10.1038/nature12293; pmid: 23831649
30. A. Hoffmann *et al.*, Concerted action of the ribosome and the associated chaperone Trigger Factor confines nascent polypeptide folding. *Mol. Cell* **48**, 63–74 (2012). doi: 10.1016/j.molcel.2012.07.018; pmid: 22921937
31. P. Genevaux *et al.*, In vivo analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep.* **5**, 195–200 (2004). doi: 10.1038/sj.embor.7400067; pmid: 14726952
32. E. Deuerling, A. Schulze-Specking, T. Tomoyasu, A. Mogk, B. Bukau, Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature* **400**, 693–696 (1999). doi: 10.1038/23301; pmid: 10458167
33. S. A. Teter *et al.*, Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains. *Cell* **97**, 755–765 (1999). doi: 10.1016/S0092-8674(00)80787-4; pmid: 10380927
34. G. Calloni *et al.*, DnaK functions as a central hub in the *E. coli* chaperone network. *Cell Rep.* **1**, 251–264 (2012). doi: 10.1016/j.celrep.2011.12.007; pmid: 22832197
35. P. Bechtluft *et al.*, Direct observation of chaperone-induced changes in a protein folding pathway. *Science* **318**, 1458–1461 (2007). doi: 10.1126/science.1144972; pmid: 18048690
36. R. S. Ullers, D. Ang, F. Schwager, C. Georgopoulos, P. Genevaux, Trigger Factor can antagonize both SecB and DnaK/DnaJ chaperone functions in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 3101–3106 (2007). doi: 10.1073/pnas.0608232104; pmid: 17360615
37. S. Rudiger, S. M. V. Freund, D. B. Veprintsev, A. R. Fersht, CRINEPT-TROSY NMR reveals p53 core domain bound in an unfolded form to the chaperone Hsp90. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11085–11090 (2002). doi: 10.1073/pnas.132393699; pmid: 12163643
38. R. Horst *et al.*, Direct NMR observation of a substrate protein bound to the chaperonin GroEL. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12748–12753 (2005). doi: 10.1073/pnas.0505642102; pmid: 16116078
39. T. A. Walton, C. M. Sandoval, C. A. Fowler, A. Pardi, M. C. Sousa, The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1772–1777 (2009). doi: 10.1073/pnas.0809275106; pmid: 19181847
40. S. J. Park, B. N. Borin, M. A. Martinez-Yamout, H. J. Dyson, The client protein p53 adopts a molten globule–like state in the presence of Hsp90. *Nat. Struct. Mol. Biol.* **18**, 537–541 (2011). doi: 10.1038/nsmb.2045; pmid: 21460846
41. A. Zhuravleva, E. M. Clérico, L. M. Gierasch, An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones. *Cell* **151**, 1296–1307 (2012). doi: 10.1016/j.cell.2012.11.002; pmid: 23217711
42. B. M. Burmann, C. Wang, S. Hiller, Conformation and dynamics of the periplasmic membrane-protein–chaperone complexes OmpX–Skp and tOmpA–Skp. *Nat. Struct. Mol. Biol.* **20**, 1265–1272 (2013). doi: 10.1038/nsmb.2677; pmid: 24077225
43. R. Sprangers, L. E. Kay, Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* **445**, 618–622 (2007). doi: 10.1038/nature05512; pmid: 17237764
44. I. Gelis *et al.*, Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* **131**, 756–769 (2007). doi: 10.1016/j.cell.2007.09.039; pmid: 18022369
45. T. L. Religa, R. Sprangers, L. E. Kay, Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. *Science* **328**, 98–102 (2010). doi: 10.1126/science.1184991; pmid: 20360109
46. R. Rosenzweig, S. Moradi, A. Zarrine-Afsar, J. R. Glover, L. E. Kay, Unraveling the mechanism of protein disaggregation through a ClpB-DnaK interaction. *Science* **339**, 1080–1083 (2013). doi: 10.1126/science.1233066; pmid: 23393091
47. N. Popovych, S.-R. Tzeng, M. Tonelli, R. H. Ebright, C. G. Kalodimos, Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6927–6932 (2009). doi: 10.1073/pnas.0900595106; pmid: 19359484
48. L. Chen *et al.*, Structural instability tuning as a regulatory mechanism in protein-protein interactions. *Mol. Cell* **44**, 734–744 (2011). doi: 10.1016/j.molcel.2011.09.022; pmid: 22152477
49. S. K. Lakshmipathy *et al.*, Identification of nascent chain interaction sites on trigger factor. *J. Biol. Chem.* **282**, 12186–12193 (2007). doi: 10.1074/jbc.M609871200; pmid: 17296610
50. F. Merz *et al.*, The C-terminal domain of *Escherichia coli* trigger factor represents the central module of its chaperone activity. *J. Biol. Chem.* **281**, 31963–31971 (2006). doi: 10.1074/jbc.M605164200; pmid: 16926148
51. H. Patzelt *et al.*, Three-state equilibrium of *Escherichia coli* trigger factor. *Biol. Chem.* **383**, 1611–1619 (2002). doi: 10.1515/BC.2002.182; pmid: 12452438
52. S.-T. D. Hsu, C. M. Dobson, ^{1}H, ^{15}N and ^{13}C assignments of the dimeric ribosome binding domain of trigger factor from *Escherichia coli*. *Biomol. NMR Assign.* **3**, 17–20 (2009). doi: 10.1007/s12104-008-9130-8; pmid: 19636937
53. Y. Yao, G. Bhabha, G. Kroon, M. Landes, H. J. Dyson, Structure discrimination for the C-terminal domain of *Escherichia coli* trigger factor in solution. *J. Biomol. NMR* **40**, 23–30 (2008). doi: 10.1007/s10858-007-9207-1; pmid: 18043871
54. J. C. Bardwell, K. McGovern, J. Beckwith, Identification of a protein required for disulfide bond formation in vivo. *Cell* **67**, 581–589 (1991). doi: 10.1016/0092-8674(91)90532-4; pmid: 1934062
55. Y. Akiyama, K. Ito, Export of *Escherichia coli* alkaline phosphatase attached to an integral membrane protein, SecY. *J. Biol. Chem.* **264**, 437–442 (1989). pmid: 2535843
56. S. Kamitani, Y. Akiyama, K. Ito, Identification and characterization of an *Escherichia coli* gene required for the formation of correctly folded alkaline phosphatase, a periplasmic enzyme. *EMBO J.* **11**, 57–62 (1992). pmid: 1740115
57. Q. A. Valent *et al.*, Early events in preprotein recognition in *E. coli*: Interaction of SRP and trigger factor with nascent polypeptides. *EMBO J.* **14**, 5494–5505 (1995). pmid: 8521806
58. H. C. Lee, H. D. Bernstein, Trigger factor retards protein export in *Escherichia coli*. *J. Biol. Chem.* **277**, 43527–43535 (2002). doi: 10.1074/jbc.M205950200; pmid: 12205085
59. J. A. Marsh, V. K. Singh, Z. Jia, J. D. Forman-Kay, Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: Implications for fibrillation. *Protein Sci.* **15**, 2795–2804 (2006). doi: 10.1110/ps.062465306; pmid: 17088319
60. H. J. Dyson, P. E. Wright, Elucidation of the protein folding landscape by NMR. *Methods Enzymol.* **394**, 299–321 (2005). doi: 10.1016/S0076-6879(05)94011-1; pmid: 15808225
61. K. Takeuchi, G. Wagner, NMR studies of protein interactions. *Curr. Opin. Struct. Biol.* **16**, 109–117 (2006). doi: 10.1016/j.sbi.2006.01.006; pmid: 16427776
62. H. Patzelt *et al.*, Binding specificity of *Escherichia coli* trigger factor. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14244–14249 (2001). doi: 10.1073/pnas.261432298; pmid: 11724963
63. W. A. Fenton, Y. Kashi, K. Furtak, A. L. Horwich, Residues in chaperonin GroEL required for polypeptide binding and release. *Nature* **371**, 614–619 (1994). doi: 10.1038/371614a0; pmid: 7935796

64. P. C. Stirling, S. F. Bakhoum, A. B. Feigl, M. R. Leroux, Convergent evolution of clamp-like binding sites in diverse chaperones. *Nat. Struct. Mol. Biol.* **13**, 865–870 (2006). doi: 10.1038/nsmb1153; pmid: 17021621
65. A. Mittermaier, L. E. Kay, New tools provide new insights in NMR studies of protein dynamics. *Science* **312**, 224–228 (2006). doi: 10.1126/science.1124964; pmid: 16614210
66. A. Palmer, NMR characterization of the dynamics of biomacromolecules. *Chem. Rev.* **104**, 3623–3640 (2004). doi: 10.1021/cr030413t; pmid: 15303831
67. J. P. Loria, M. Rance, A. G. Palmer, A TROSY CPMG sequence for characterizing chemical exchange in large proteins. *J. Biomol. NMR* **15**, 151–155 (1999). doi: 10.1023/A:1008355631073; pmid: 10605088
68. F. A. Mulder, A. Mittermaier, B. Hon, F. W. Dahlquist, L. E. Kay, Studying excited states of proteins by NMR spectroscopy. *Nat. Struct. Biol.* **8**, 932–935 (2001). doi: 10.1038/nsb1101-932; pmid: 11685237
69. P. Vallurupalli, D. F. Hansen, E. Stollar, E. Meirovitch, L. E. Kay, Measurement of bond vector orientations in invisible excited states of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18473–18477 (2007). doi: 10.1073/pnas.0708296104; pmid: 18006656
70. K. Sugase, J. C. Lansing, H. J. Dyson, P. E. Wright, Tailoring relaxation dispersion experiments for fast-associating protein complexes. *J. Am. Chem. Soc.* **129**, 13406–13407 (2007). doi: 10.1021/ja0762238; pmid: 17935336
71. R. Maier, C. Scholz, F. Schmid, Dynamic association of trigger factor with protein substrates. *J. Mol. Biol.* **314**, 1181–1190 (2001). doi: 10.1006/jmbi.2000.5192; pmid: 11743733
72. D. Tomkiewicz, N. Nouwen, A. J. M. Driessen, Kinetics and energetics of the translocation of maltose binding protein folding mutants. *J. Mol. Biol.* **377**, 83–90 (2008). doi: 10.1016/j.jmb.2008.01.014; pmid: 18241889
73. N. K. Tyagi, W. A. Fenton, A. L. Horwich, GroEL/GroES cycling: ATP binds to an open ring before substrate protein favoring protein binding and production of the native state. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 20264–20269 (2009). doi: 10.1073/pnas.0911556106; pmid: 19915138
74. K. Chakraborty *et al*., Chaperonin-catalyzed rescue of kinetically trapped states in protein folding. *Cell* **142**, 112–122 (2010). doi: 10.1016/j.cell.2010.05.027; pmid: 20603018
75. J. M. Betton, M. Hofnung, Folding of a mutant maltose-binding protein of *Escherichia coli* which forms inclusion bodies. *J. Biol. Chem.* **271**, 8046–8052 (1996). doi: 10.1074/jbc.271.14.8046; pmid: 8626487
76. S. Raffy, N. Sassoon, M. Hofnung, J. M. Betton, Tertiary structure-dependence of misfolding substitutions in loops of the maltose-binding protein. *Protein Sci.* **7**, 2136–2142 (1998). doi: 10.1002/pro.5560071010; pmid: 9792100
77. F. A. Saul *et al*., Crystal structure of a defective folding protein. *Protein Sci.* **12**, 577–585 (2003). doi: 10.1110/ps.0235103; pmid: 12592028
78. J. Kuriyan, D. Eisenberg, The origin of protein interactions and allostery in colocalization. *Nature* **450**, 983–990 (2007). doi: 10.1038/nature06524; pmid: 18075577
79. V. Novokhatny, K. Ingham, Thermodynamics of maltose binding protein unfolding. *Protein Sci.* **6**, 141–146 (1997). doi: 10.1002/pro.5560060116; pmid: 9007986
80. S. Sharma *et al*., Monitoring protein conformation along the pathway of chaperonin-assisted folding. *Cell* **133**, 142–153 (2008). doi: 10.1016/j.cell.2008.01.048; pmid: 18394994
81. G. W. Farr *et al*., Multivalent binding of nonnative substrate proteins by the chaperonin GroEL. *Cell* **100**, 561–573 (2000). doi: 10.1016/S0092-8674(00)80692-3; pmid: 10721993
82. M. Shtilerman, G. H. Lorimer, S. W. Englander, Chaperonin function: Folding by forced unfolding. *Science* **284**, 822–825 (1999). doi: 10.1126/science.284.5415.822; pmid: 10221918
83. A. Hoffmann, B. Bukau, Trigger factor finds new jobs and contacts. *Nat. Struct. Mol. Biol.* **16**, 1006–1008 (2009). doi: 10.1038/nsmb1009-1006; pmid: 19809489
84. G. Kramer *et al*., Functional dissection of *Escherichia coli* trigger factor: Unraveling the function of individual domains. *J. Bacteriol.* **186**, 3777–3784 (2004). doi: 10.1128/JB.186.12.3777-3784.2004; pmid: 15175291
85. F. Delaglio *et al*., NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995). doi: 10.1007/BF00197809; pmid: 8520220
86. R. Sprangers, A. Velyvis, L. Kay, Solution NMR of supramolecular complexes: Providing new insights into function. *Nat. Methods* **4**, 697–703 (2007). doi: 10.1038/nmeth1080; pmid: 17762877
87. A. Velyvis, A. M. Ruschak, L. E. Kay, An economical method for production of ^{2}H,$^{13}CH_3$-threonine for solution NMR studies of large protein complexes: Application to the 670 kDa proteasome. *PLOS One* **7**, e43725 (2012). doi: 10.1371/journal.pone.0043725; pmid: 22984438
88. P. Güntert, Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* **278**, 353–378 (2004). pmid: 15318003
89. Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223 (2009). doi: 10.1007/s10858-009-9333-z; pmid: 19548092
90. C. Schwieters, J. Kuszewski, G.M. Clore, Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Res. Spect.* **48**, 47–62 (2006). doi: 10.1016/j.pnmrs.2005.10.001
91. A. T. Brunger, Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2**, 2728–2733 (2007). doi: 10.1038/nprot.2007.406; pmid: 18007608
92. J. P. Carver, R. E. Richards, A general two-site solution for the chemical exchange produced dependence of T_2 on the Carr-Purcell pulse separation. *J. Magn. Reson.* **6**, 89–105 (1972). doi: 10.1016/0022-2364(72)90090-X
93. S.-R. Tzeng, C. G. Kalodimos, Protein activity regulation by conformational entropy. *Nature* **488**, 236–240 (2012). doi: 10.1038/nature11271; pmid: 22801505

Acknowledgments: C.G.K. thanks L. E. Kay and B. Bukau for fruitful discussions and comments. We thank M. T. Pai and A. Bhattacharya for contributions during the early phase of the project, S. Patel for the use of a spectrofluorometer, and S. Kim and Y. Xia (University of Minnesota) for their help with setting up NMR experiments. This work is supported by NIH grant GM073854 (C.G.K.), Nonaco (Aristeia-Excellence, #1473), KUL-Spa, RiMembR, and StrepSynth (A.E.). T.S. is the recipient of postdoctoral fellowships from the Japanese Society for the Promotion of Science and the American Heart Association (14POST18600039). We thank Bruker Biospin for the use of the 700-MHz instrument. Coordinates and structural restraints have been deposited with the Protein Data Bank under accession codes 2MLY, 2MLX, and 2MLZ, for TF–PhoA$^{a\text{-}c}$, TF–PhoA$^{d\text{-}e}$, and TF–PhoA$^{f\text{-}g}$, respectively.

Supplementary Materials

www.sciencemag.org/content/344/6184/1250494/suppl/DC1
Figs. S1 to S27
Table S1

10.1126/science.1250494

Supplementary Materials for

Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone

Tomohide Saio, Xiao Guan, Paolo Rossi, Anastassios Economou, Charalampos G. Kalodimos*

*Corresponding author. E-mail: babis@rutgers.edu

Published 9 May 2014, *Science* **344**, 1250494 (2014)
DOI: 10.1126/science.1250494

This PDF file includes:

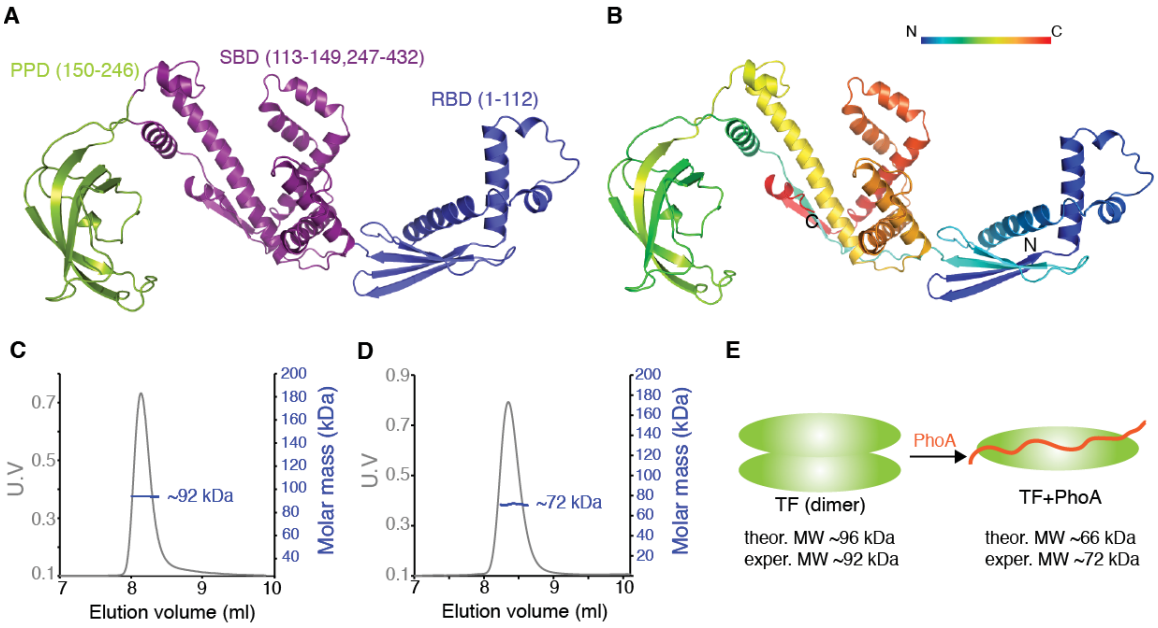


Fig. S1. Domain architecture and oligomerization state of TF. (A) Structure of *E. coli* TF (PDB ID 1W26). The residue boundaries for each one of the three domains are shown in parentheses. SBD is discontinuous and is formed primarily by the C-terminal domain. **(B)** The backbone of TF is colored using a gradient color scheme from the N terminus (blue) to the C terminus (red). **(C)** MALS of unliganded TF shows that the protein forms a dimer in solution. **(D)** MALS of TF in complex with the first 141 residues of PhoA containing TF-binding sites *a*, *b*, and *c* (PhoA$^{a\text{-}c}$) shows that TF binds to PhoA as a monomer. TF is also a monomer in its complex with PhoA$^{d\text{-}e}$ and PhoA$^{f\text{-}g}$, as well as with MBP and OmpA. **(E)** Schematic showing the monomerization undergone by TF upon substrate (PhoA) binding.

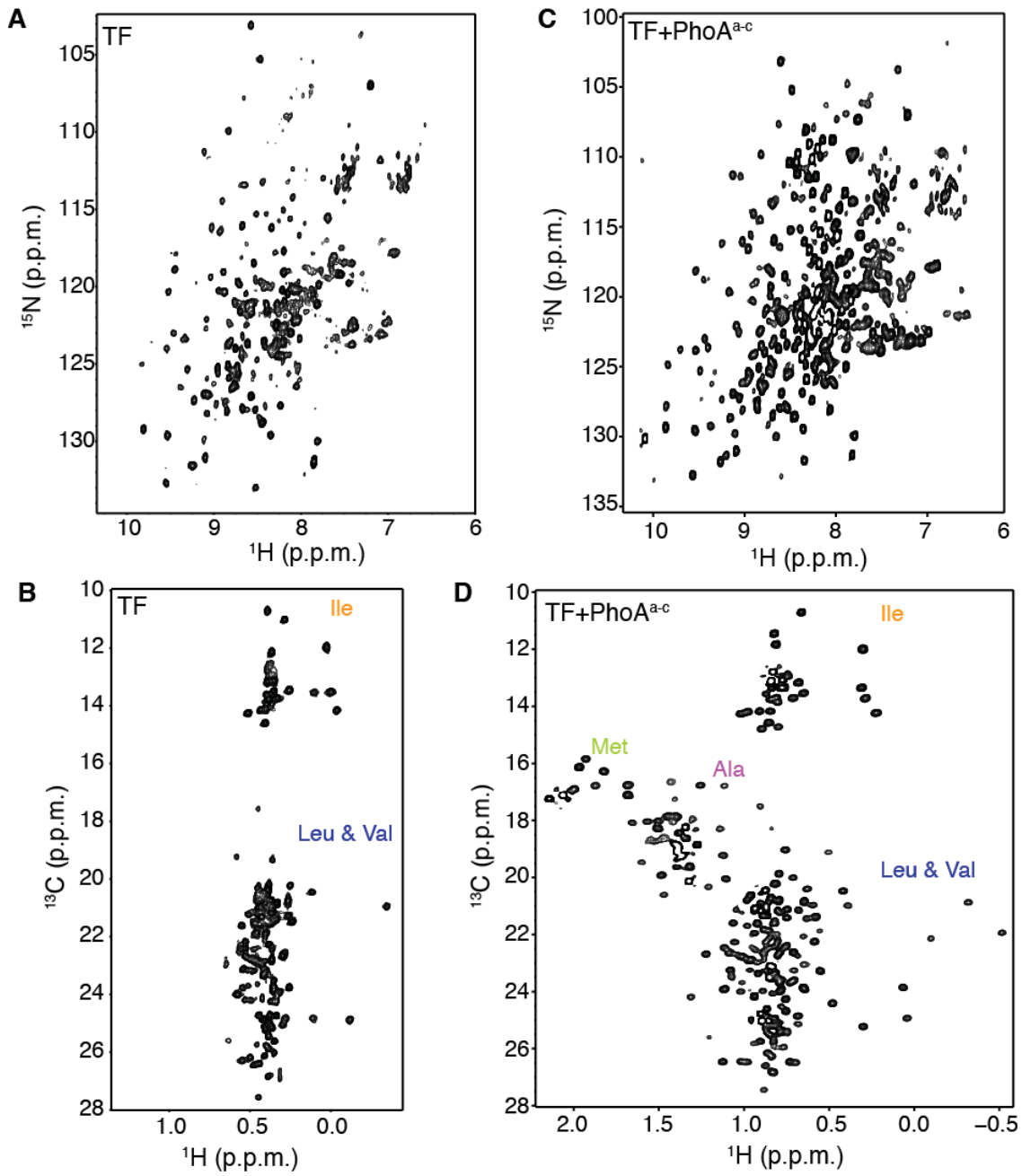


Fig. S2. NMR spectra of unliganded TF and in complex with PhoA. (A) ^{1}H-^{15}N TROSY HSQC of [U-^{2}H,^{15}N]-labeled and **(B)** ^{1}H-^{13}C methyl HMQC spectra of [U-^{2}H; Ile-δ1-$^{13}CH_3$; Leu,Val-$^{13}CH_3$/$^{13}CH_3$]-labeled dimeric TF (unliganded). Analysis of the spectra shows that several resonances are broadened beyond detection, presumably because of slow motions at the dimeric interface (see fig. S19). **(C)** ^{1}H-^{15}N TROSY HSQC of [U-^{2}H,^{15}N]-labeled and **(D)** ^{1}H-^{13}C methyl HMQC spectra of [U-^{2}H; Ala-$^{13}CH_3$; Met-$^{13}CH_3$; Ile-δ1-$^{13}CH_3$; Leu,Val-$^{13}CH_3$/$^{13}CH_3$]-labeled TF in complex with PhoA^{a-c}. Analysis of the spectra indicates that all of the expected resonances are present. The dispersion and sensitivity of these spectra are excellent. Both TF and are PhoA^{a-c} labeled.

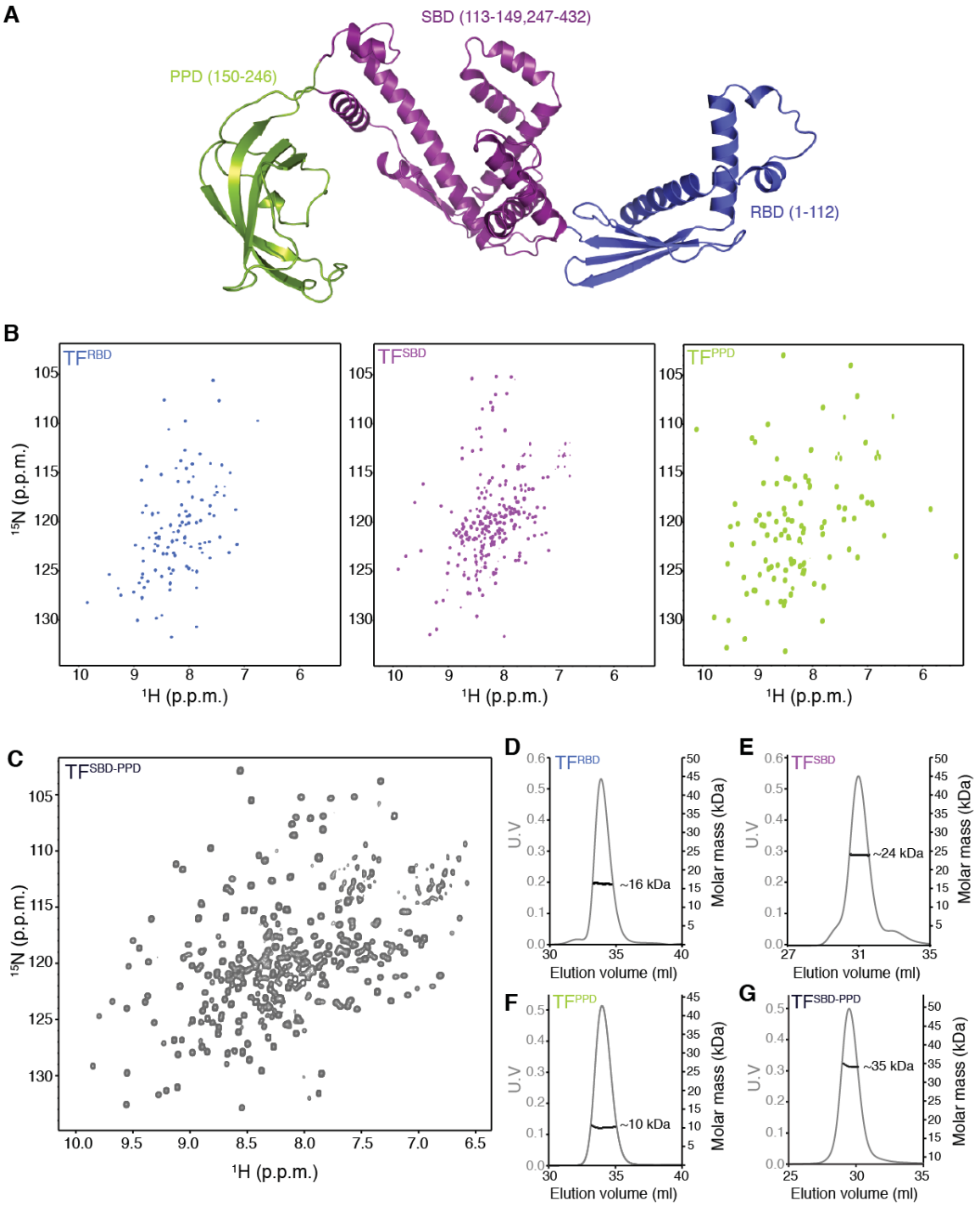


Fig. S3. NMR spectra of TF and its isolated domains. (A) Structure of TF. **(B)** ^{1}H-^{15}N HSQC spectra of the individual domains of TF indicate that they retain their proper fold when in isolation. **(C)** ^{1}H-^{15}N TROSY HSQC spectrum of [U-^{2}H,^{15}N]-labeled TF$^{SBD-PPD}$, a TF construct that consists of the SBD and PPD domains (residues 113-432), the two domains with chaperone activity. **(D** to **G)**, MALS of isolated TF domains RBD (in **D**), SBD (in **E**), PPD (in **F**), and SBD-PPD (in **G**) show that all of these domains or fragments are monomeric in solution.

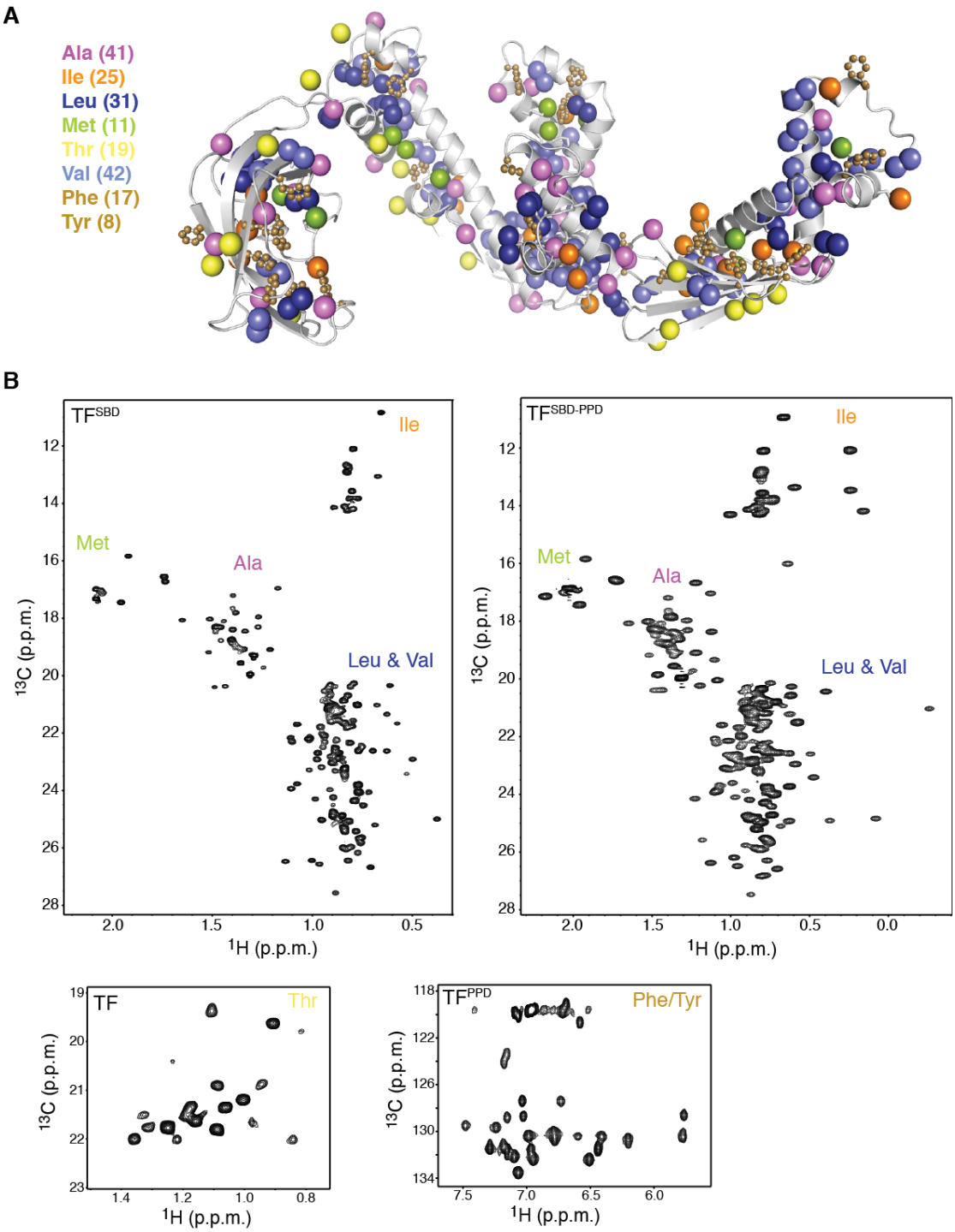


Fig. S4. Side chain NMR spectra of TF. (A) TF is enriched in hydrophobic residues, such as methyl-bearing (Ala, Ile, Leu, Met, Thr, and Val) and aromatic residues (Phe and Tyr). The numbers in parentheses indicate the number of the corresponding residues in TF (per subunit). **(B)** ^{1}H-^{13}C methyl HMQC spectra of [U-^{2}H; Ala-$^{13}CH_3$; Met-$^{13}CH_3$; Ile-δ1-$^{13}CH_3$; Leu,Val-$^{13}CH_3$/$^{13}CH_3$]-labeled TF^{SBD} and TF$^{SBD-PPD}$, and of [U-^{2}H; Thr-$^{13}CH_3$]-labeled dimeric TF, and aromatic ^{1}H-^{13}C HSQC spectra of TF^{PPD}, a domain enriched with aromatic residues.

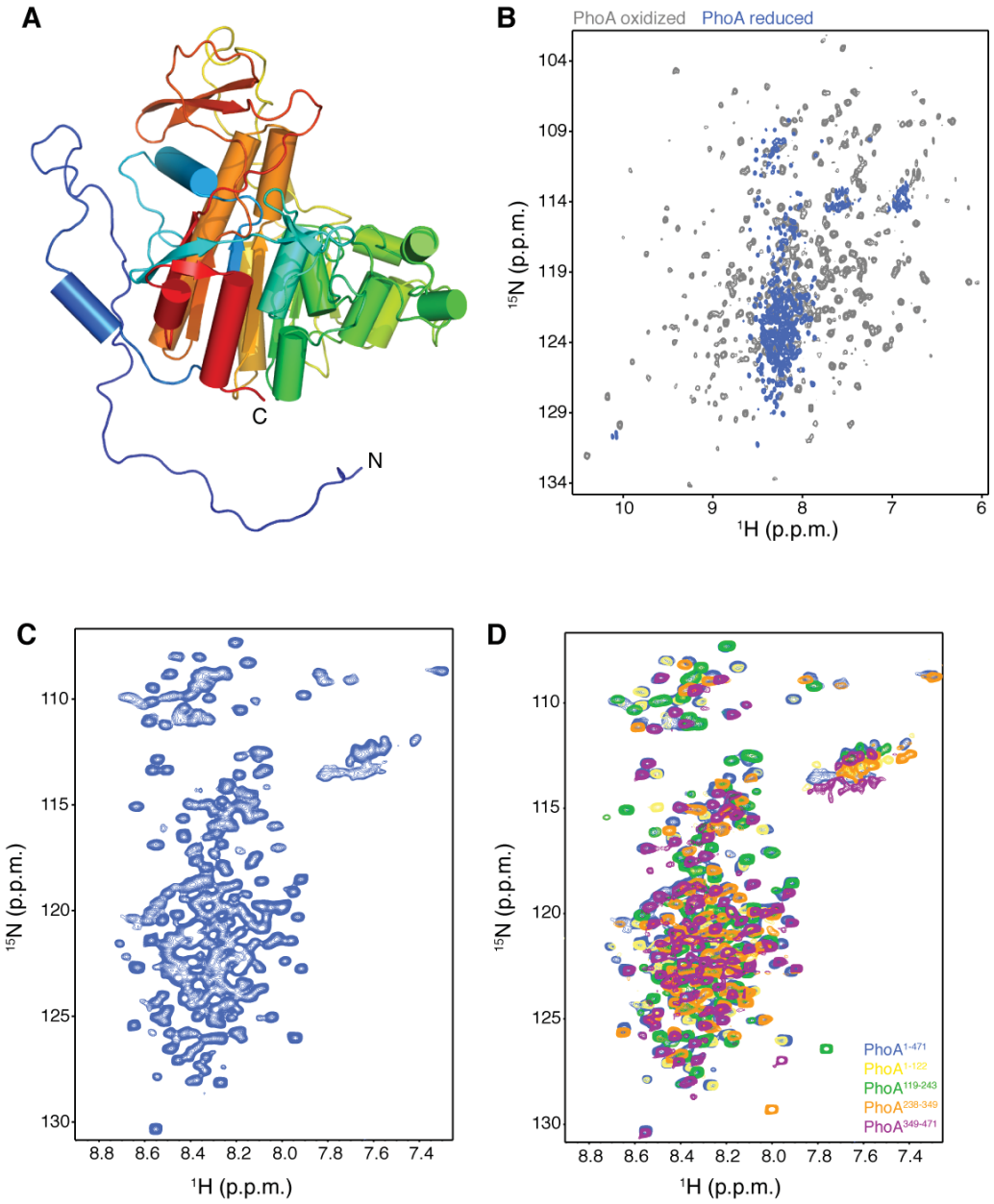


Fig. S5. NMR analysis of PhoA. (A) Crystal structure of PhoA under oxidizing conditions (PDB ID 1Y6V). Folded PhoA forms a dimer but only one subunit is shown here. Only the mature portion of PhoA was solved crystallographically (residues 23-471). The signal sequence (residues 1-22) was modeled as an unstructured N terminal sequence in the crystal structure and is depicted here. The structure is colored using a gradient color from the N terminus (blue) to the C terminus (red). **(B)** ^{1}H-^{15}N HSQC spectra of PhoA under oxidizing (grey) and reducing conditions (blue). The NMR spectra clearly show that oxidized PhoA is folded whereas reduced PhoA is unfolded. **(C)** An expanded view of the ^{1}H-^{15}N HSQC spectrum of reduced, unfolded PhoA. **(D)** Overlaid ^{1}H-^{15}N HSQC spectra of full-length PhoA and select PhoA fragments that span the entire sequence of PhoA (aa 1-471). The spectra of the fragments are exact subsets of the full-length PhoA spectrum. The excellent resonance correspondence is expected since all of the fragments and the full-length PhoA are unfolded. Resonance assignment obtained for the various fragments was transferred to the full-length PhoA using 3D NMR spectra.

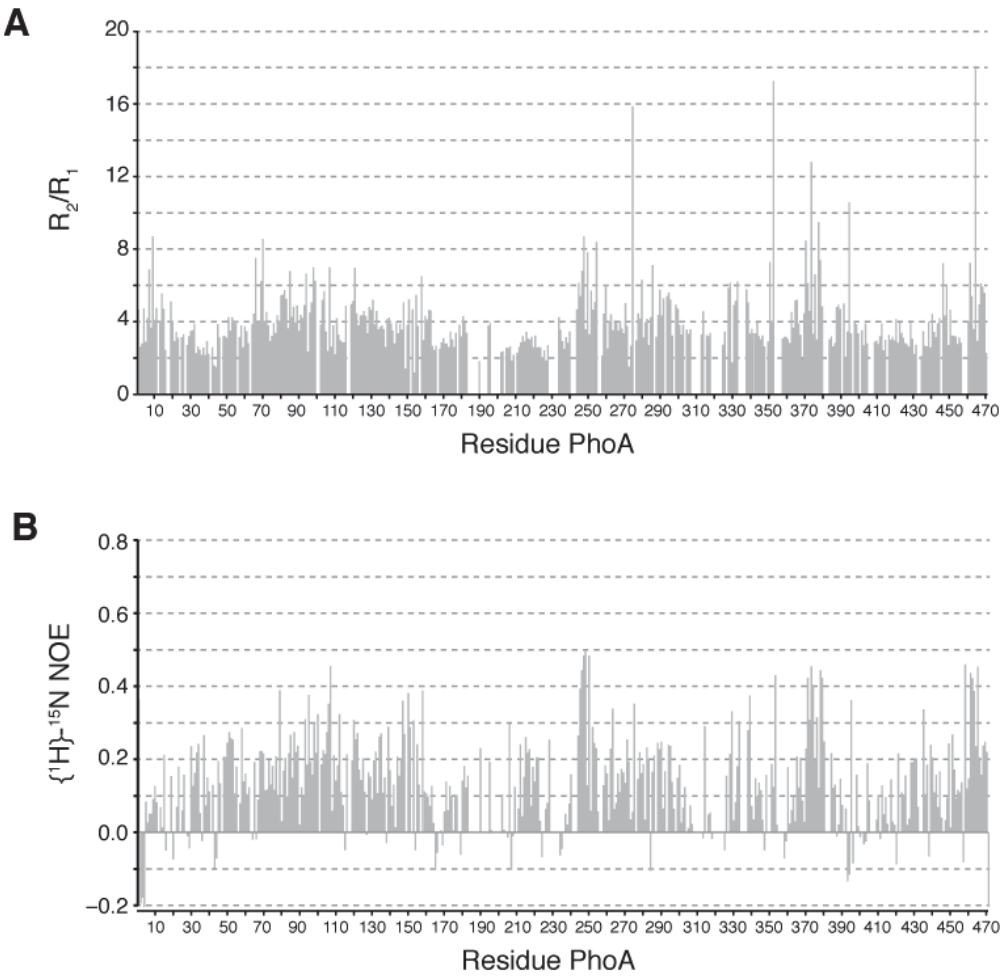


Fig. S6. Relaxation analysis of unfolded PhoA. (A) ^{15}N R_2/R_1 values of unfolded PhoA plotted as a function of the amino acid sequence. R_2 and R_1 are the transverse and longitudinal relaxation rates, respectively. Lower values indicate flexibility whereas higher values indicate rigidity (or the presence of conformational exchange). The low R_2/R_1 values throughout the PhoA sequence suggests that the protein indeed is in a disordered state. Certain regions exhibit higher values suggesting the presence of transiently formed secondary structure. **(B)** Steady-state {^{1}H}-^{15}N values of unfolded PhoA plotted as a function of the amino acid sequence. Values higher than ~0.7 are typically found in regions with secondary structure whereas negative or close to zero are found in completely disordered regions. Values of 0.3-0.4 indicate the presence of transiently formed local secondary structure.

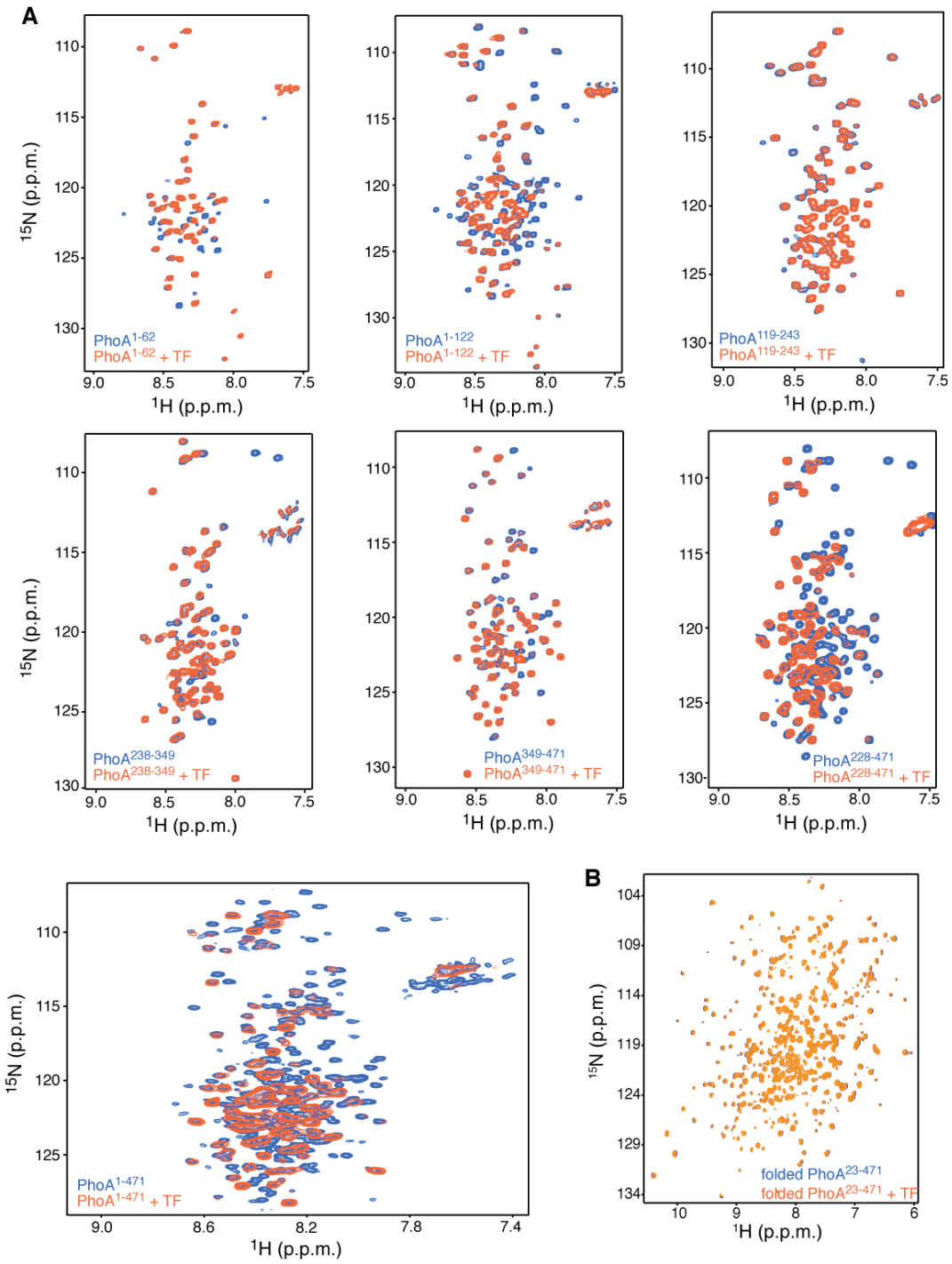


Fig. S7. Identification of the TF-binding sites in PhoA by NMR. (A) ^{15}N labeled full-length PhoA as well as a number of PhoA fragments were titrated with unlabeled TF. Due to the labeling scheme and the size of TF, the intensity of the PhoA residues that are bound by TF decreases dramatically or disappears. The ^{1}H-^{15}N HSQC spectra of PhoA select fragments and the full-length PhoA are shown in the absence (blue) and presence (orange) of TF. The fragments cover the entire sequence of PhoA and the regions identified to interact with TF in the various PhoA fragments are essentially identical to the full-length PhoA. **(B)** Oxidized, folded PhoA (mature form) does not interact with TF.

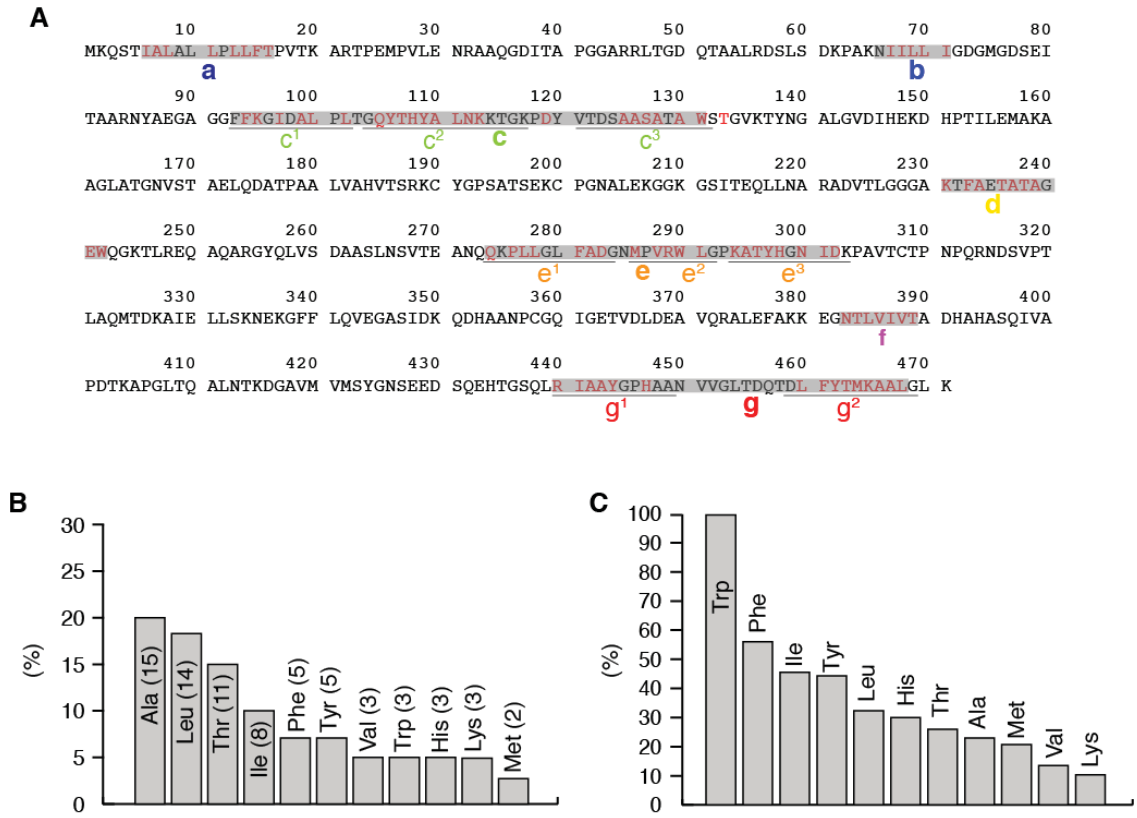


Fig. S8. Amino acid composition of TF-binding sites in PhoA. (A) Primary sequence of PhoA with the TF-binding sites (*a-g*) highlighted. The PhoA residues that form direct contacts with TF are colored red (see Fig. 4). The color code of the sites and subsites are as in Fig. 2C and 3A. **(B)** Percentage values of the PhoA amino acids that directly interact with TF. The number in brackets correspond to the actual number of amino acids. **(C)** Percentage values of PhoA amino acids that interact with TF normalized against the total number of each one of the amino acids in the PhoA sequence. For example, all of the three Trp residues in PhoA interact with TF (that is, 100%).

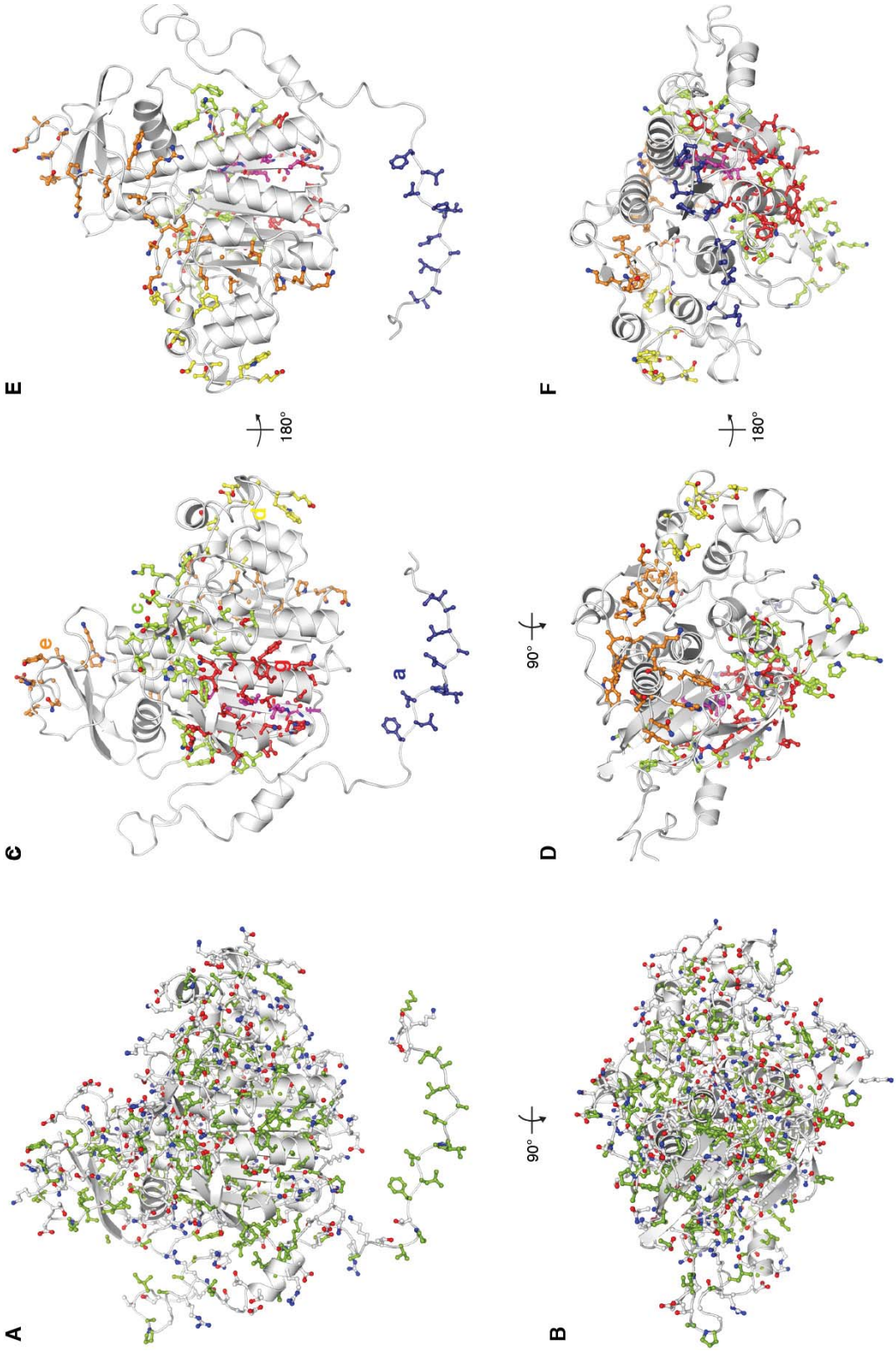


Fig. S9. Distribution of the TF-binding sites in PhoA mapped on its folded structure. (A to **D)** Distribution of nonpolar residues (in green) in the folded structure of PhoA. **(C** to **F)** The residues from the seven distinct regions in PhoA (*a-g*) that interact with TF are shown in sticks.

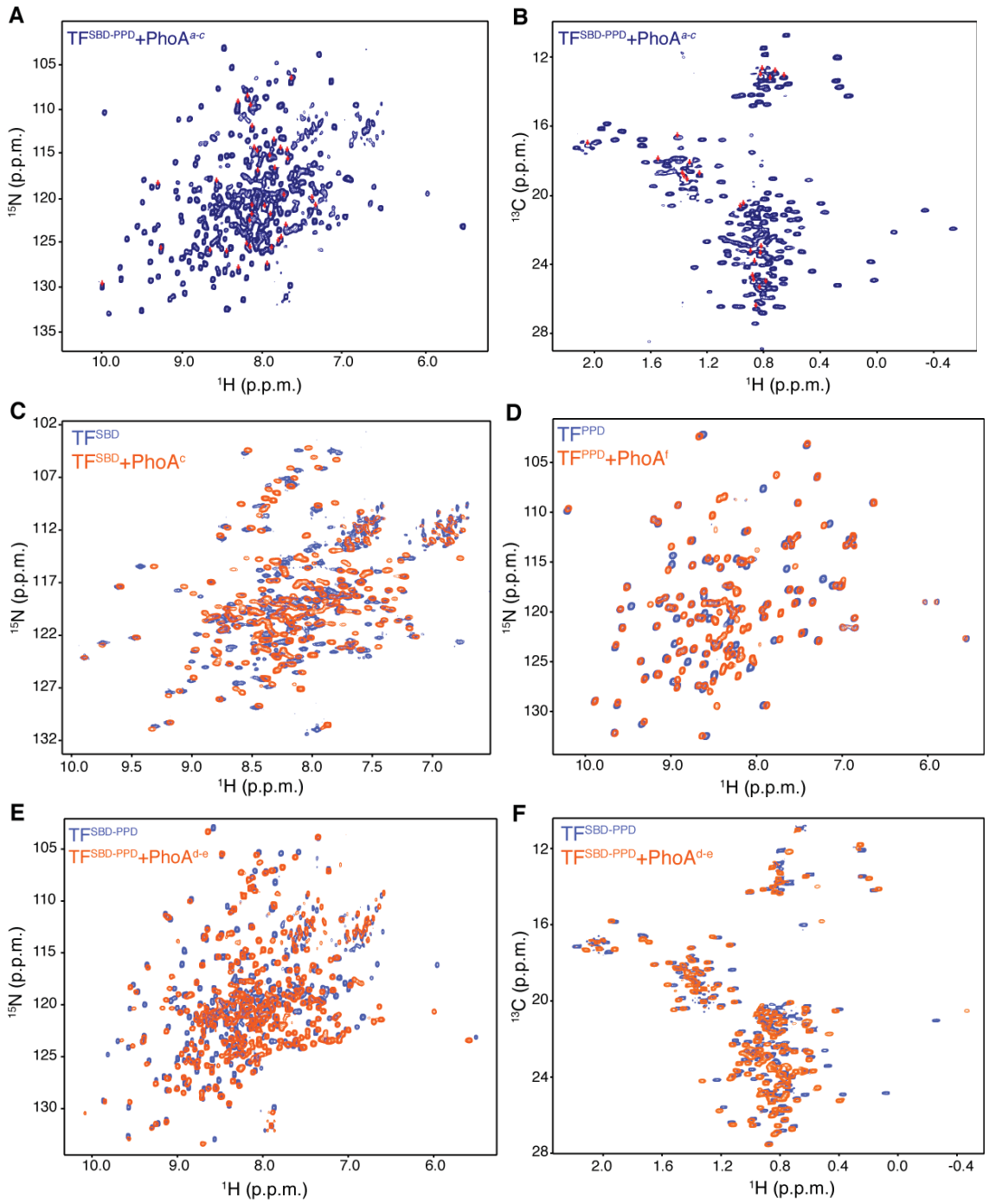


Fig. S10. NMR spectra of TF in complex with PhoA. (A,B) ^{1}H-^{15}N HSQC and ^{1}H-^{13}C methyl HMQC spectra of [U-^{2}H,^{15}N; Ala-$^{13}CH_3$; Met-$^{13}CH_3$; Ile-δ1-$^{13}CH_3$; Leu,Val-$^{13}CH_3$/$^{13}CH_3$]-labeled TF$^{SBD-PPD}$ in complex with PhoA^{a-c}. The arrows indicate the resonances of PhoA that interact with TF and thus exhibit markedly enhanced resonance dispersion. **(C)** ^{1}H-^{15}N HSQC spectra of [U-^{2}H,^{15}N]-labeled TF^{SBD} in the free state (blue) and in complex with PhoA^{c} (orange). **(D)** ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled TF^{PPD} in the free state (blue) and in complex with PhoA^{f} (orange). **(E)** ^{1}H-^{15}N HSQC spectra of [U-^{2}H,^{15}N]-labeled TF$^{SBD-PPD}$ in the free state (blue) and in complex with PhoA^{d-e} (orange). **(F)** ^{1}H-^{13}C methyl HMQC spectra of [U-^{2}H; Ala-$^{13}CH_3$; Met-$^{13}CH_3$; Ile-δ1-$^{13}CH_3$; Leu,Val-$^{13}CH_3$/$^{13}CH_3$]-labeled TF$^{SBD-PPD}$ in the free state (blue) and in complex with PhoA^{d-e} (orange). In all cases both TF and PhoA are similarly labeled.

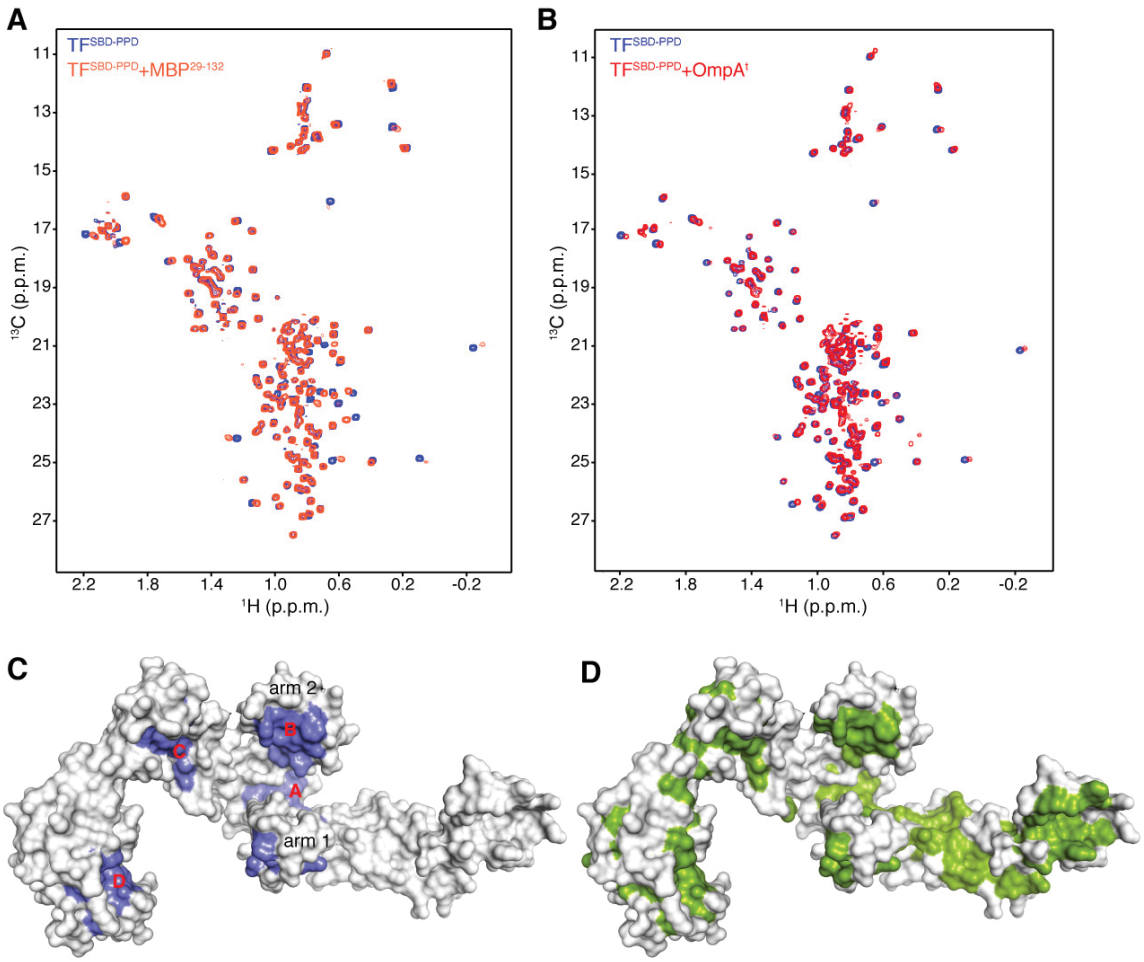


Fig. S11. Interaction of the TF chaperone module ($TF^{SBD-PPD}$) with unfolded fragments of the maltose binding protein (MBP) and the transmembrane region of OmpA. (A) ^{1}H-^{13}C HMQC spectra of [U-^{2}H,^{15}N; Ala-$^{13}CH_3$; Met-$^{13}CH_3$; Ile-δ1-$^{13}CH_3$; Leu,Val-$^{13}CH_3$/$^{13}CH_3$]-labeled $TF^{SBD-PPD}$ in complex with a fragment of unlabeled MBP consisting of residues 29-132, which, based on its hydrophobicity (see fig. S26), is predicted to have several TF-binding sites. **(B)** ^{1}H-^{13}C HMQC spectra of [U-^{2}H,^{15}N; Ala-$^{13}CH_3$; Met-$^{13}CH_3$; Ile-δ1-$^{13}CH_3$; Leu,Val-$^{13}CH_3$/$^{13}CH_3$]-labeled $TF^{SBD-PPD}$ in complex with the transmembrane region of OmpA ($OmpA^{t}$, residues 1-190). **(C)** The TF residues identified by NMR (both from backbone and side chain chemical shifts) to interact with MBP^{29-132} and $OmpA^{t}$ are colored blue. All four binding sites (A-D) that interact with PhoA also interact with MBP and OmpA. In addition, a hydrophobic region located in arm 1 interacts with MBP and OmpA. As with the other substrate-binding sites, this region is hydrophobic, as can be seen in panel **D**, which shows the hydrophobic residues in TF (colored green).

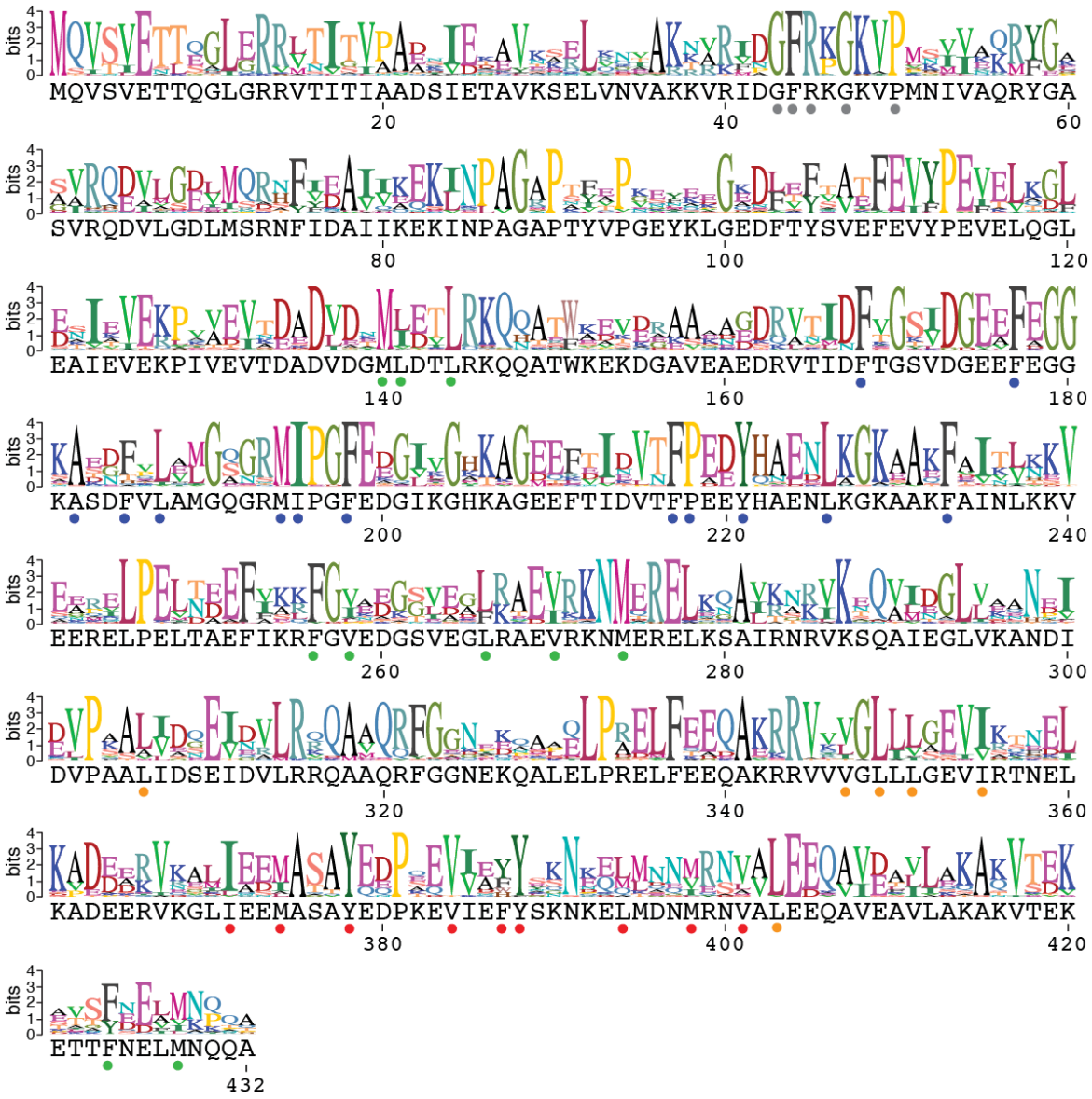


Fig. S12. Sequence conservation of TF. Sequence logo of TF created from a collection of aligned sequences of TF. 60 unique sequences of bacterial TF were used and the sequence of *E. coli* TF is shown. The filled circles denote the most important residues for interacting with PhoA (orange, site A; red, site B; green, site C; blue, site D). Almost all of the interacting residues are very well conserved, with the hydrophobic nature of the residue being absolutely conserved. The grey circles denote the ribosome-interacting residues (G43-F-R-x-G-x-x-P).

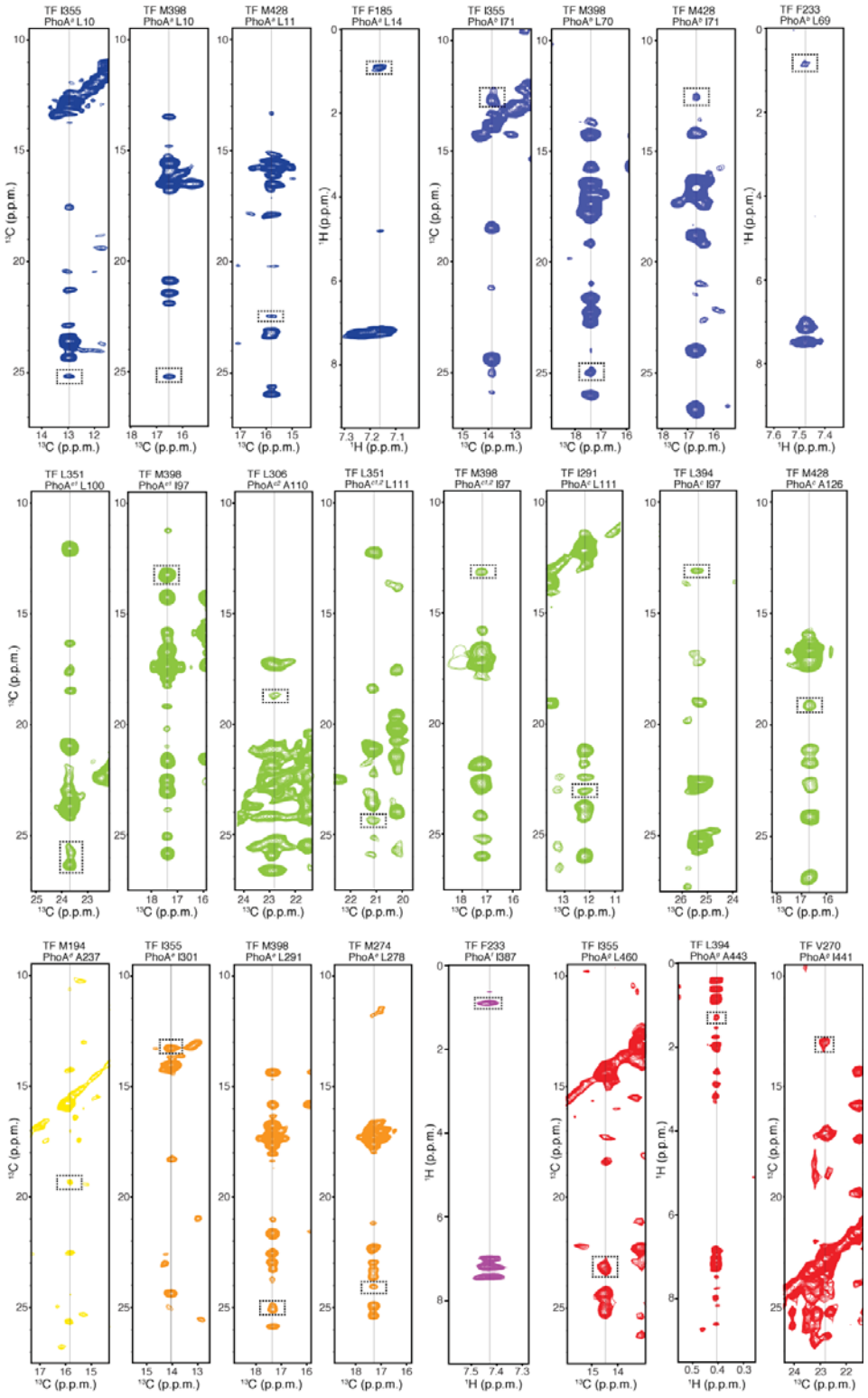


Fig. S13. Inter-molecular NOEs between TF and PhoA. The full strips are shown from ^{13}C-edited NOESY-HSQC and HMQC-NOESY-HMQC NMR experiments corresponding to the inter-NOEs shown in Fig. 3 and fig. S17. The NOE cross-peaks shown in Fig. 3 are designated here by a dashed-line rectangle. The color code of the peaks corresponds to the one used in Fig. 3 and fig. S17.

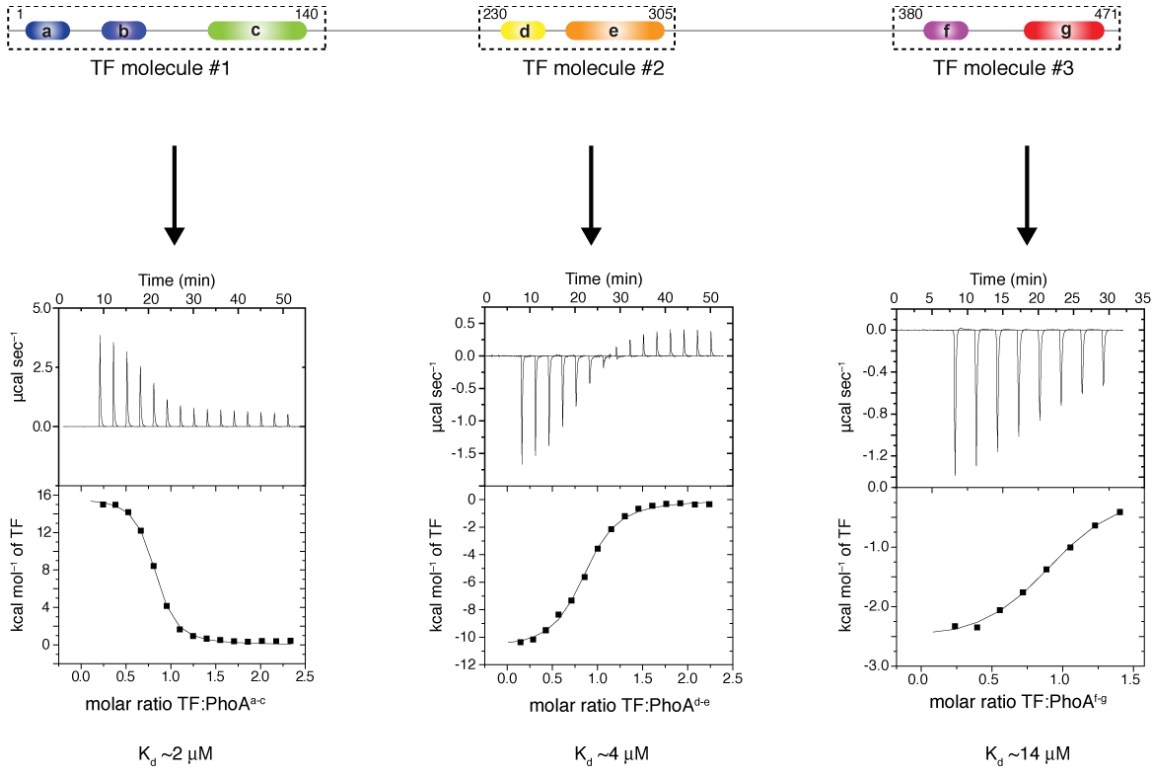


Fig. S14. Thermodynamic characterization of the interaction between TF and PhoA. ITC traces of the titration of $TF^{SBD\text{-}PPD}$ to the three PhoA fragments ($PhoA^{a\text{-}c}$, $PhoA^{d\text{-}e}$, and $PhoA^{f\text{-}g}$). In order to measure pure binding, without contribution from the binding-induced monomerization of TF, the $TF^{SBD\text{-}PPD}$ fragment was used instead of full-length TF. $TF^{SBD\text{-}PPD}$ encompasses all of the PhoA binding sites. Although all three complexes are mediated almost exclusively by hydrophobic contacts, the thermodynamic basis of the binding appears to be different for the different complexes. For example, formation of $TF^{SBD\text{-}PPD}$–$PhoA^{a\text{-}c}$ is entropy driven whereas formation of $TF^{SBD\text{-}PPD}$–$PhoA^{d\text{-}e}$ is enthalpy driven. The different thermodynamic signatures likely reflect differences in the degree of immobilization of the flexible linkers tethering the binding sites in PhoA. Stoichiometry for all complexes is $TF^{SBD\text{-}PPD}$:PhoA 1:1. The experiments shown here were recorded at 10 °C.

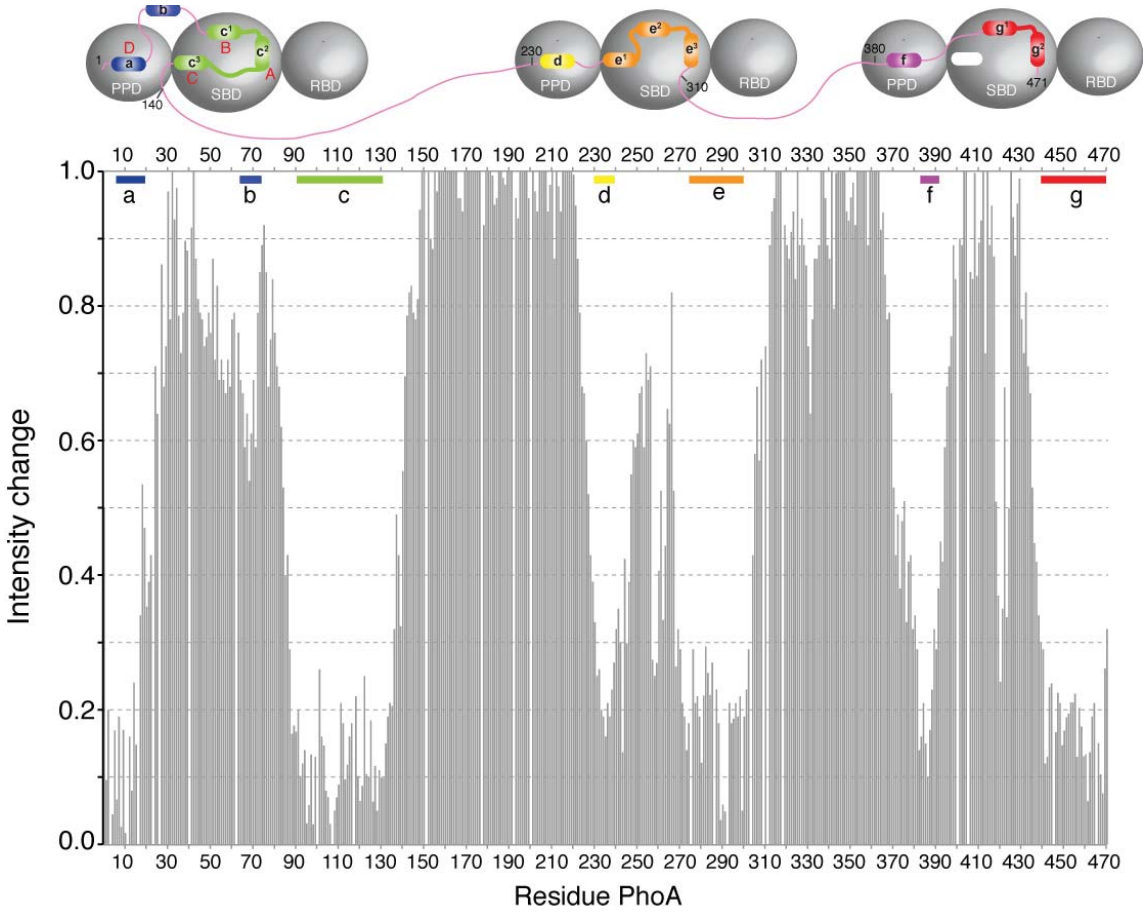


Fig. S15. NMR resonance intensity analysis of the TF−PhoA complex. Plot of resonance intensity changes of PhoA residues as a result of complex formation with TF. Because of the labeling scheme used, PhoA residues that are bound by TF have negligible intensity remaining in the complex, whereas the intensity of residues that are not bound by TF, and are thus motionally not restricted, is not significantly affected. Because of the large size of the TF−PhoA complex, transient interaction of PhoA regions with TF can also result in intensity decrease. The PhoA regions consisting of residues 150-220 and 320-370 do not interact with TF and function as flexible linkers that tether the three PhoA-bound TF molecules.

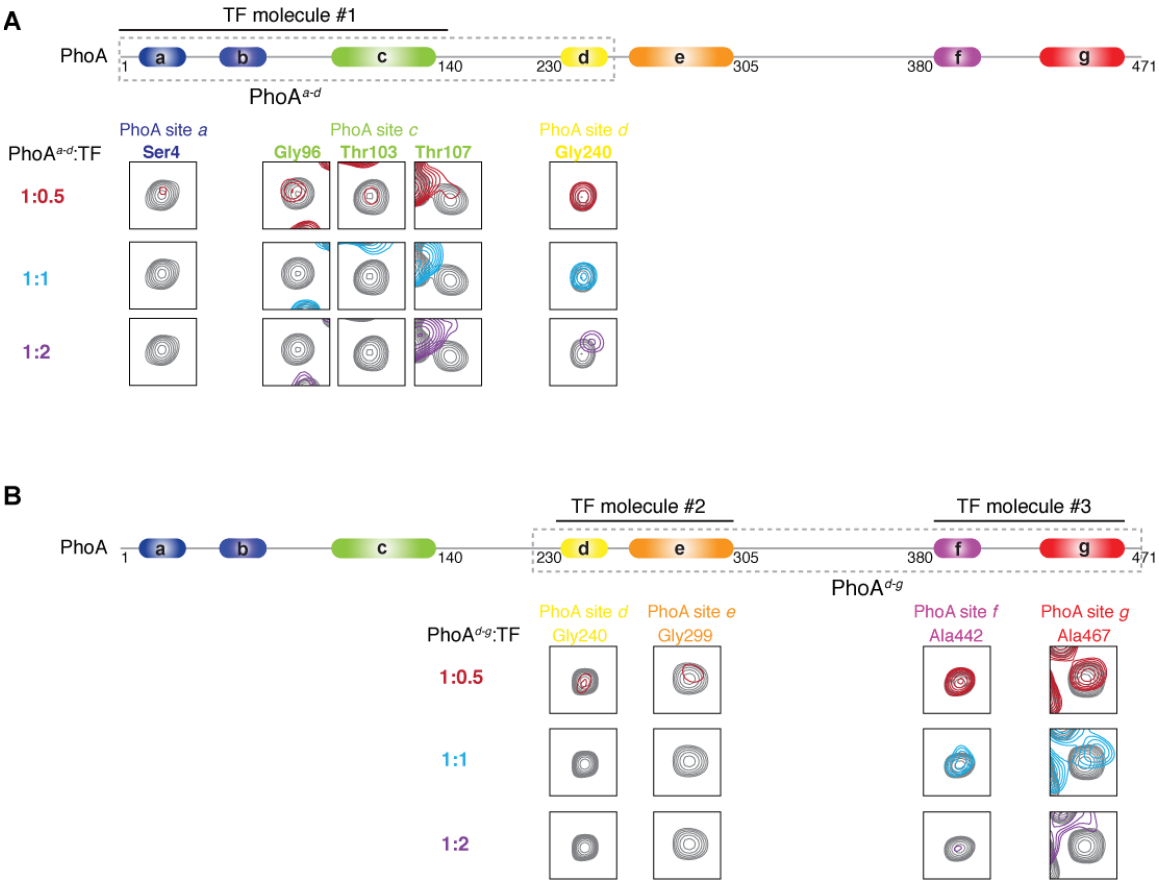


Fig. S16. Synergistic binding and selectivity of PhoA sites for TF. (A) As we explain in detail in Fig. 3, one TF molecule can accommodate sites *a* and *c* (residues 1-140). When a longer PhoA fragment that extends to include PhoA site *d* (PhoA$^{a\text{-}d}$) is used, then at a stoichiometric ratio of 1:1 (PhoA$^{a\text{-}d}$:TF) site *d* is not bound by TF. This is clearly demonstrated by the fact that the chemical shift of Gly240 (located in PhoA site *d*) is not affected at all. When excess of TF is added (PhoA$^{a\text{-}d}$:TF 1:2) then site *d* is bound by TF. **(B)** Direct evidence for synergistic binding among PhoA sites to TF is provided in this experiment. A PhoA fragment consisting of sites *d* through *g* was used (PhoA$^{d\text{-}g}$). Isolated sites *d* and *f* bind exclusively to TF PPD (fig. S17B,E) and with very similar affinities (~240 μM). One TF molecule binds to PhoA sites *d* and *e*, and another TF molecule binds to PhoA sites *f* and *g* (Fig. 4A and fig. S17). At substoichiometric PhoA$^{d\text{-}g}$:TF ratio, only PhoA site *d*, but not PhoA site *f*, is bound to TF because of the higher affinity of the PhoA$^{d\text{-}e}$ fragment, compared to PhoA$^{f\text{-}g}$, for TF (fig. S14). At a PhoA$^{d\text{-}g}$:TF ratio of 1:1, only a small fraction of sites *f* and *g* are bound, as dictated by the Boltzmann distribution. Representative peaks from ^{1}H-^{15}N HSQC spectra of labeled TF (TF$^{SBD\text{-}PPD}$) free (grey) and in complex (red, blue and purple for PhoA:TF ratio 1:0.5, 1:1 and 1:2, respectively) with the corresponding PhoA site are shown. We should note that due to the labeling scheme used for this experiment, and the size of TF, the intensity of the PhoA residues that are bound by TF decreases dramatically or disappears when binding is strong.

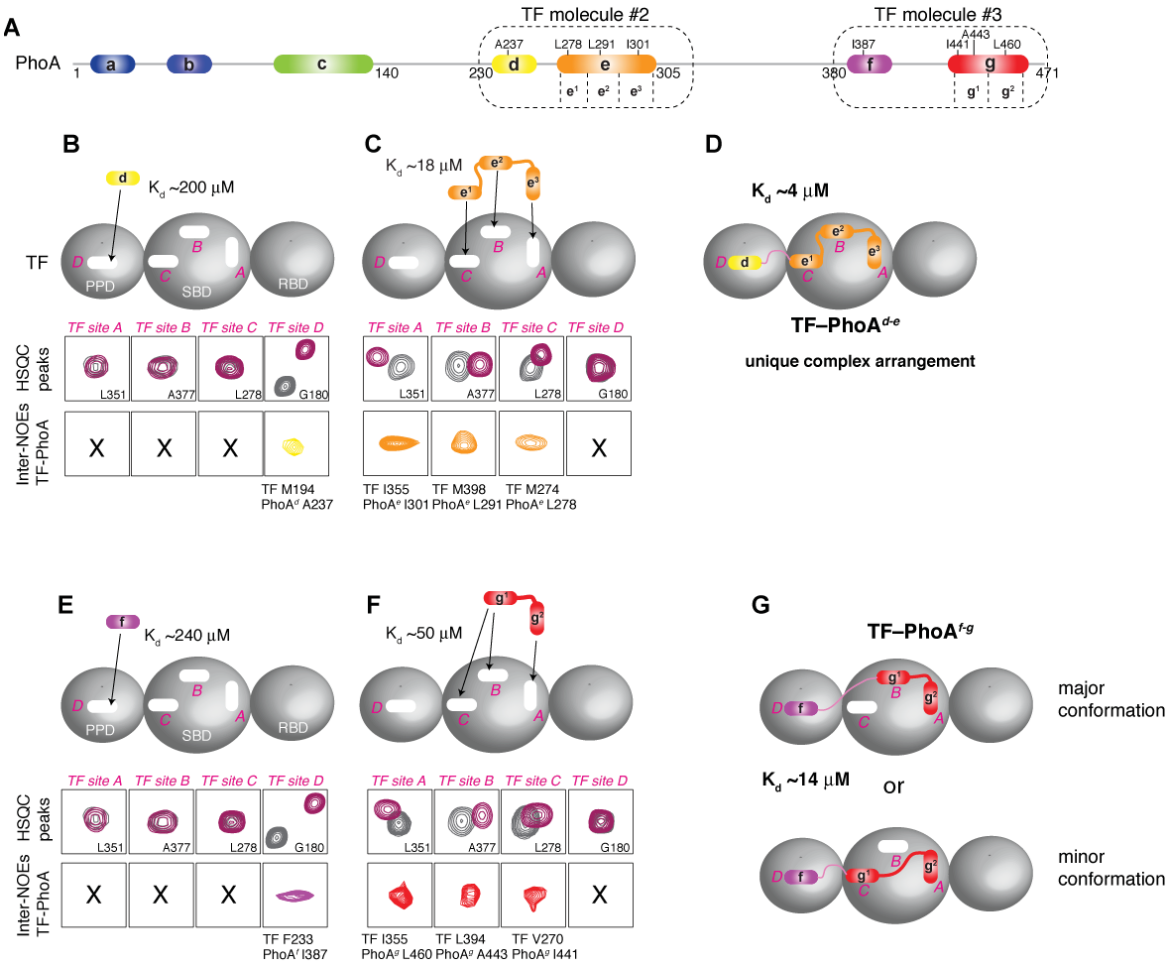


Fig. S17. Binding of the second and third TF molecules to PhoA sites *d* through *g*. This figure is an extension of Fig. 3 where the detailed characterization of the binding of PhoA sites *a*, *b* and *c* (subsites c^1, c^2, and c^3) to TF is described. **(A)** TF-binding sites in PhoA colored as in Fig. 3A. PhoA sites *d* and *g* are engaged by one TF molecule (#2) and sites *f* and *g* by an additional TF molecule (#3). PhoA residues whose inter-molecular NOEs to TF residues are depicted on panels B, C, E and F are indicated. In this figure we provide evidence about the selective binding of PhoA site *d* (in **B**), site e (in **C**), site *f* (in **E**) and site *g* (in **F**). A unique complex arrangement is observed for the TF-PhoA$^{d\text{-}e}$ complex (in **D**), whereas two binding arrangements are observed for TF-PhoA$^{f\text{-}g}$ complex: in the latter case one of them is more populated (major) than the alternative one (minor). The structure of the major conformation was determined. The full strips of the inter-molecular TF-PhoA NOEs are shown in fig. S13.

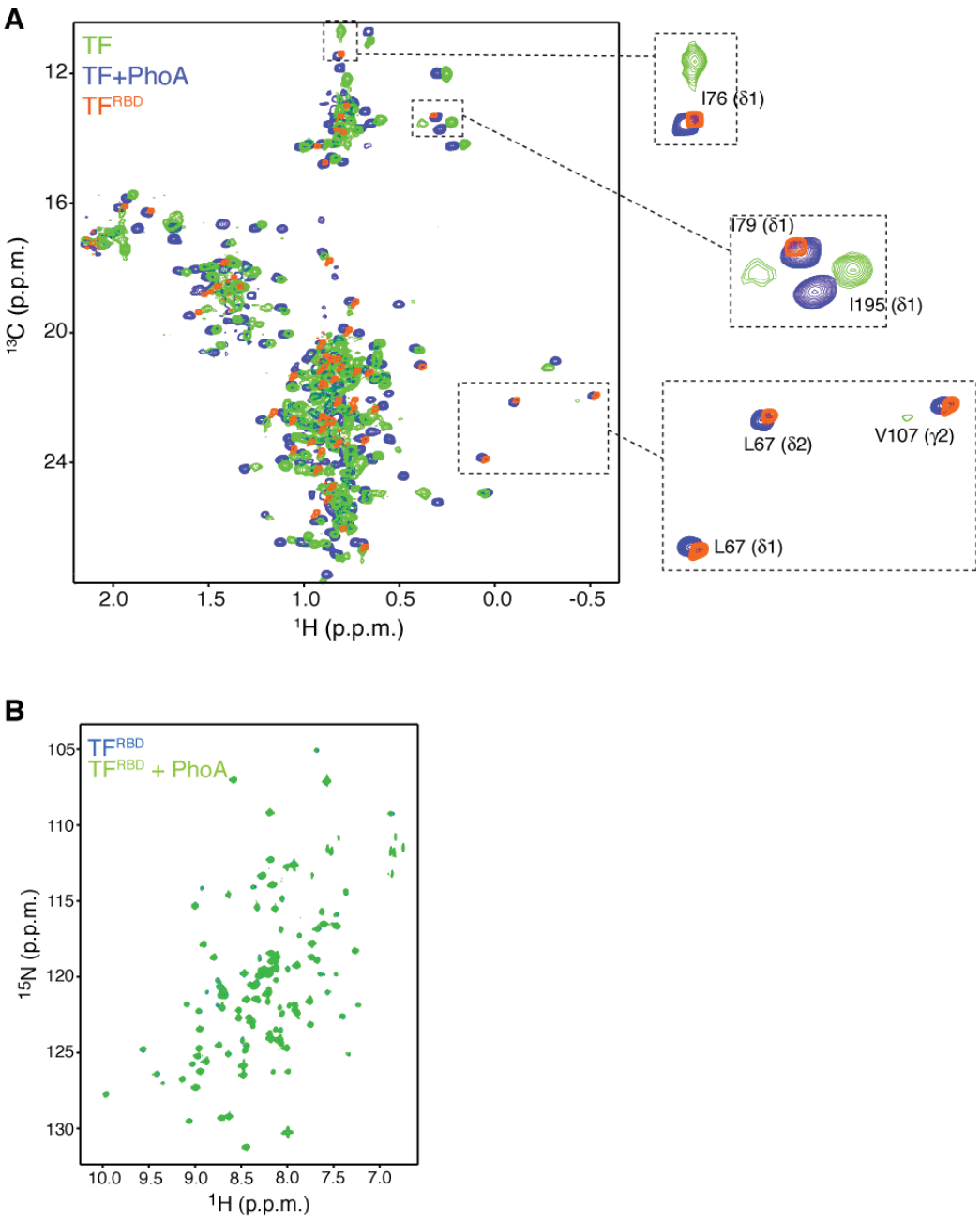


Fig. S18. PhoA binding monomerizes TF. (A) ^{1}H-^{15}N HSQC spectra of TF (green), in complex with unlabeled PhoA (blue), and isolated RBD (orange). TF and TF^{RBD} are [U-^{2}H,^{15}N; Ala-$^{13}CH_3$; Met-$^{13}CH_3$; Ile-δ1-$^{13}CH_3$; Leu,Val-$^{13}CH_3$/$^{13}CH_3$] labeled. Analysis of the spectra shows that several resonances in free TF, which is dimeric (fig. S1C), are broadened beyond detection, presumably because of slow motions at the dimeric interface (fig. S19). NMR analysis demonstrates that TF dimerization is mediated by all three domains (RBD, SBD, and PPD) but RBD resonances in TF are particularly affected (fig. S19). Indeed, monomerization of TF upon PhoA binding (fig. S1C-E) results in the appearance of all of the expected resonances (see for example residues L67, I76, I79 and V107). Furthermore, chemical shift comparison between isolated RBD (orange) and TF in complex with PhoA (blue) suggests that RBD does not bind to PhoA. **(B)** This observation is further corroborated by NMR titration data of ^{15}N labeled RBD with PhoA (green). The ^{1}H-^{15}N HSQC spectra show no perturbation.

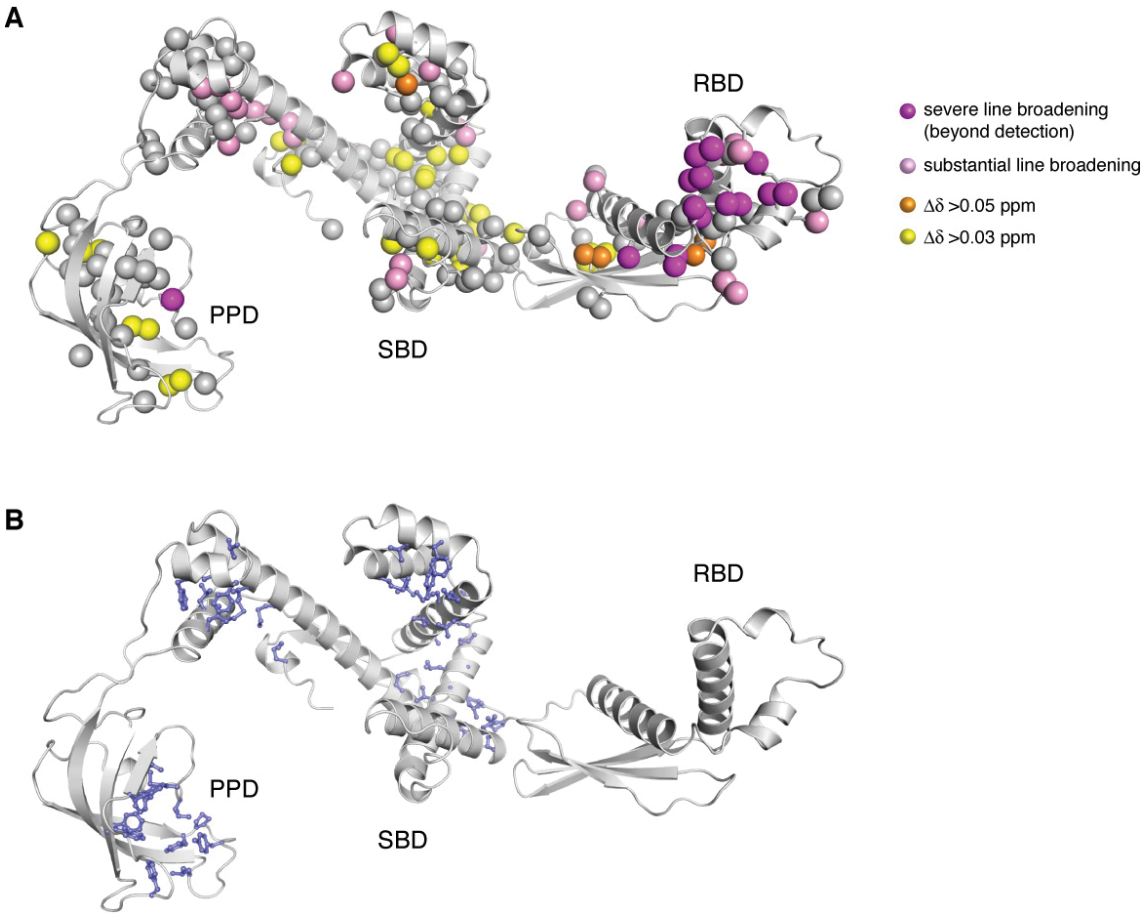


Fig. S19. Changes in chemical shift and slow dynamics between monomeric and dimeric TF. **(A)** Chemical shift changes (Δδ) of methyl resonances between monomeric and dimeric TF. The chemical shifts of monomeric TF were obtained in complex with PhoA or using $TF^{SBD\text{-}PPD}$. Residues whose resonances are significantly broadened are also indicated. Line broadening (see fig. S18A) is a result of the presence of slow (micro-to-millisecond) dynamics. Line broadening is severe for several residues in TF dimer, whereas TF monomerization suppresses these slow dynamics. Therefore, the residues experiencing such slow motions are very likely located at the dimeric interface. Regions most affected are located in RBD, the tips of the SBD arms and the PPD. These results are in agreement with the dimerization interface seen in the crystal structure of dimeric TF in complex with the S7 protein *(27)*. **(B)** Residues that make up the four substrate-binding sites in TF are shown in blue. The combined data suggest that some of the substrate-binding regions may be partially occluded in TF dimer, explaining why PhoA binding to TF results in TF monomerization.

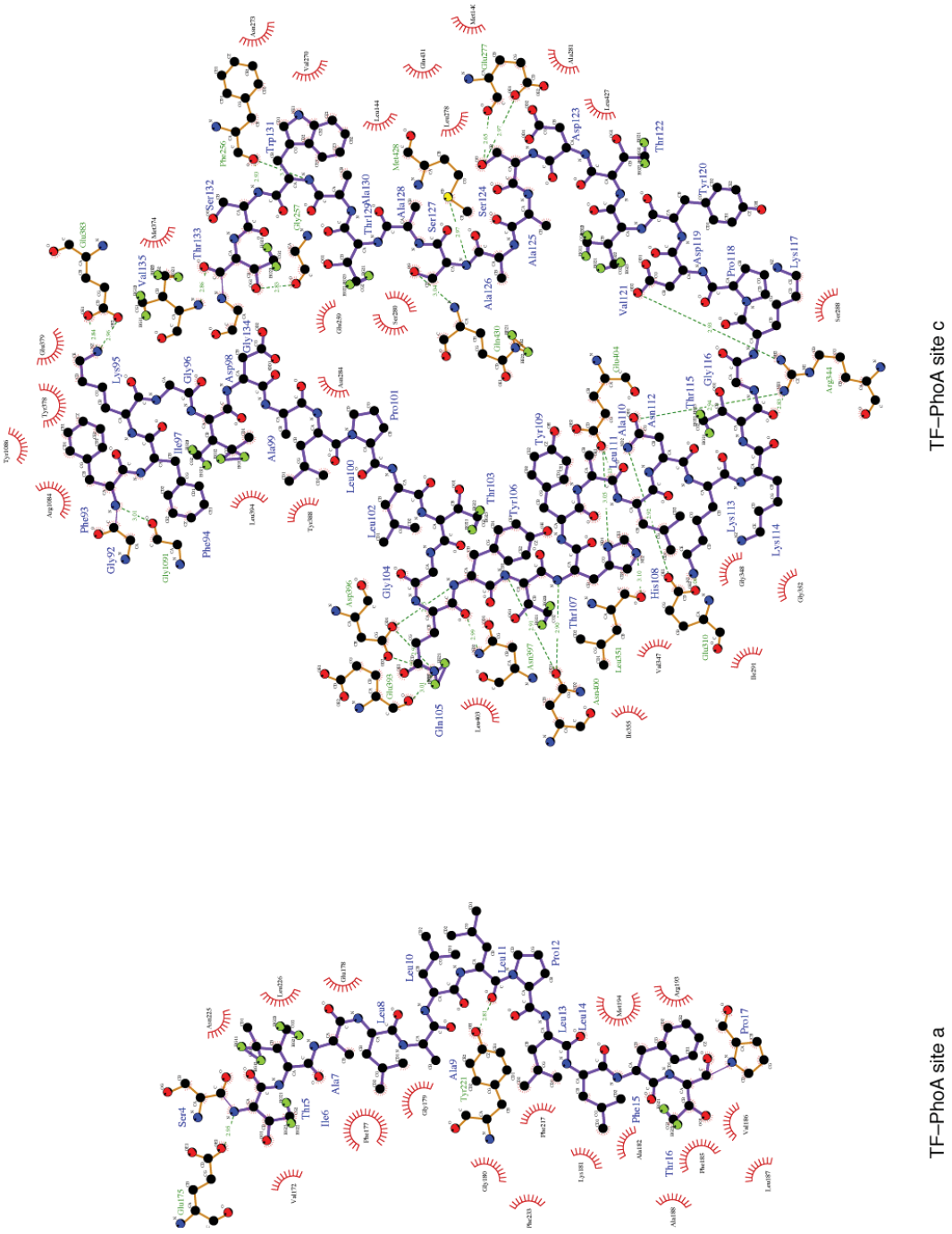


Fig. S20. TF−PhoA$^{a\text{-}c}$ contacts. Inter-molecular contacts between TF and PhoA$^{a\text{-}c}$ generated by Ligplot (http://www.ebi.ac.uk/thornton-srv/software/LIGPLOT/).

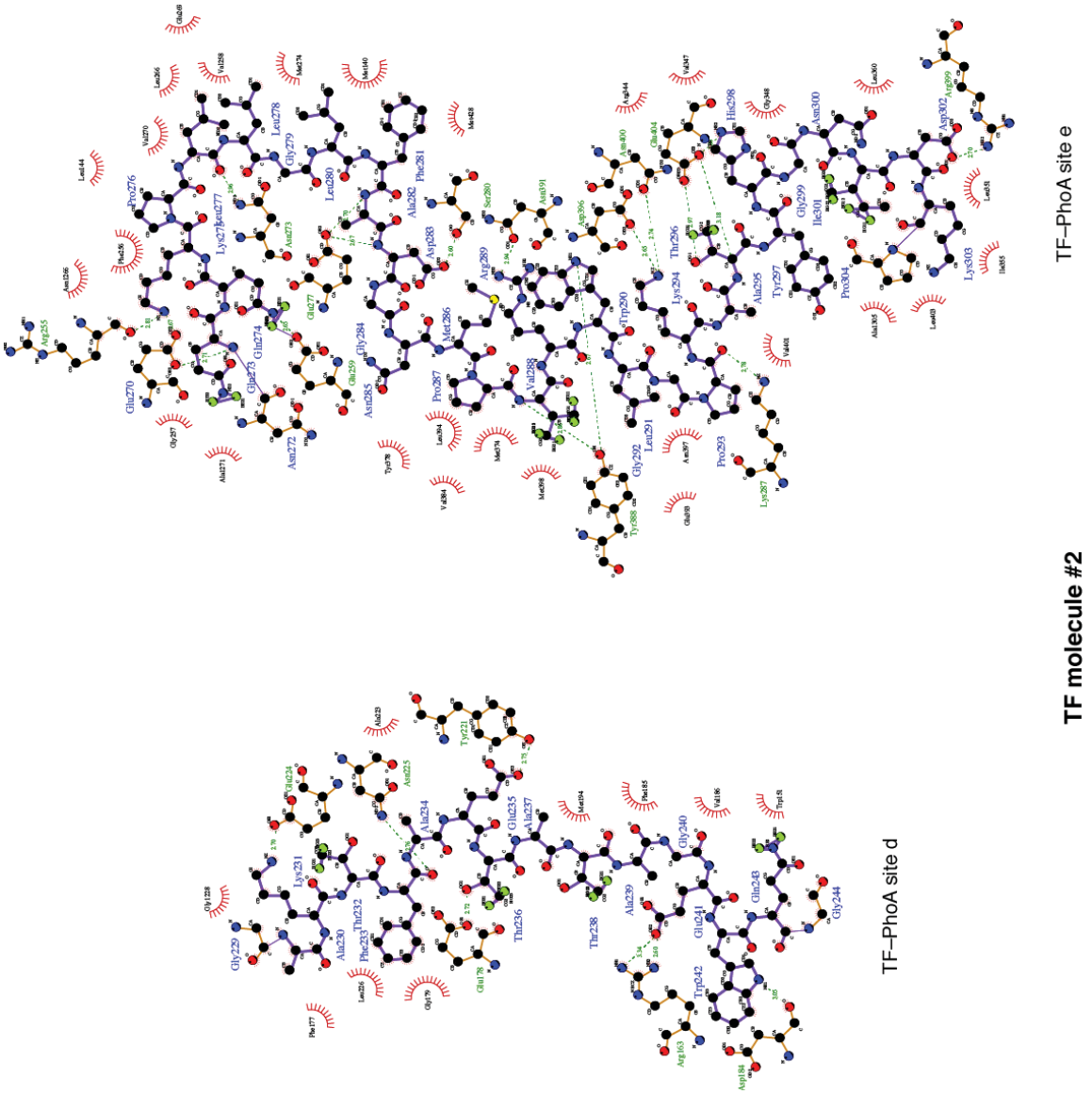


Fig. S21. TF−PhoA$^{d\text{-}e}$ contacts. Inter-molecular contacts between TF and PhoA$^{d\text{-}e}$ generated by Ligplot.

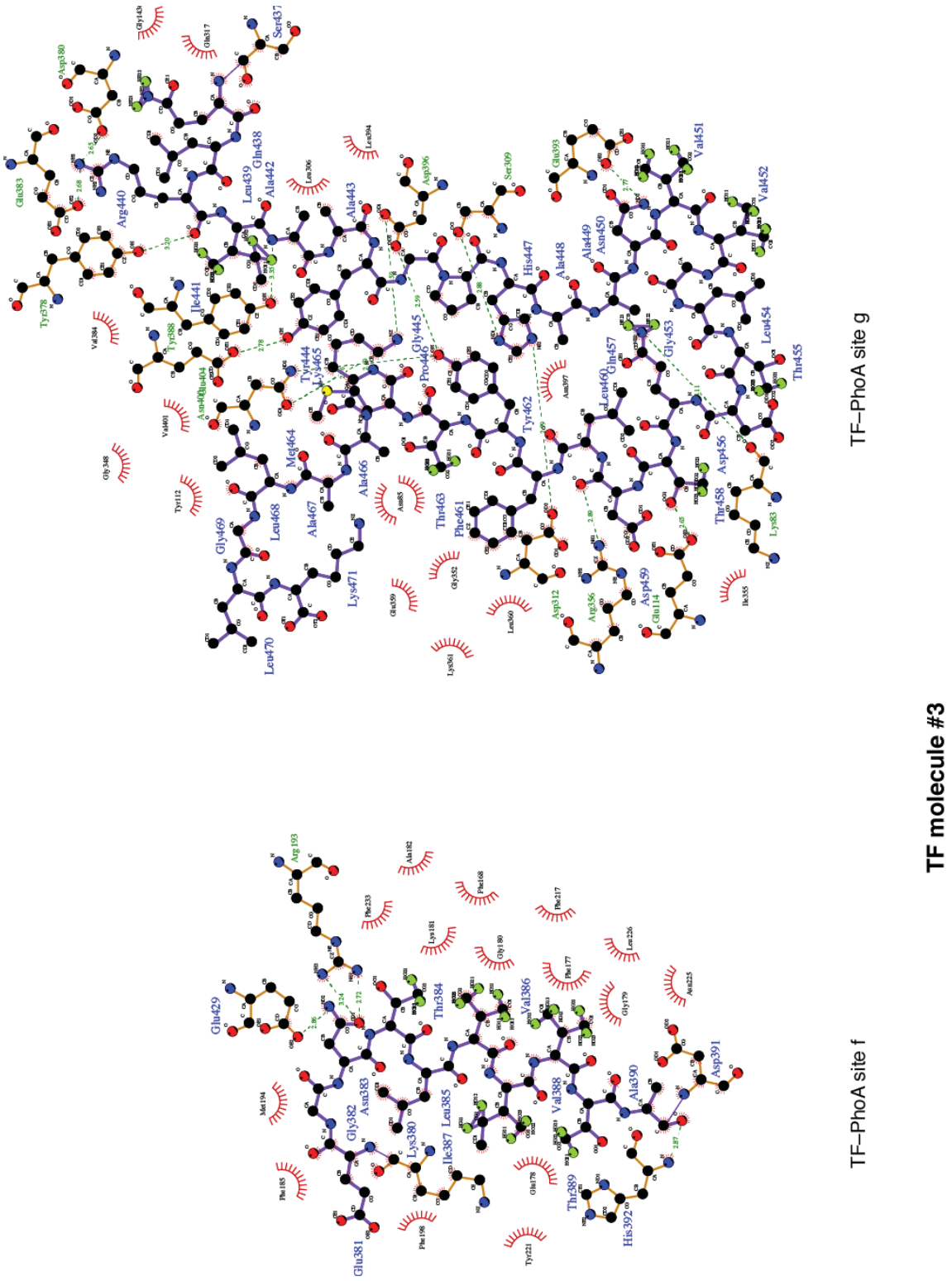


Fig. S22. TF−PhoA$^{f\text{-}g}$ contacts. Inter-molecular contacts between TF and PhoA$^{f\text{-}g}$ generated by Ligplot.

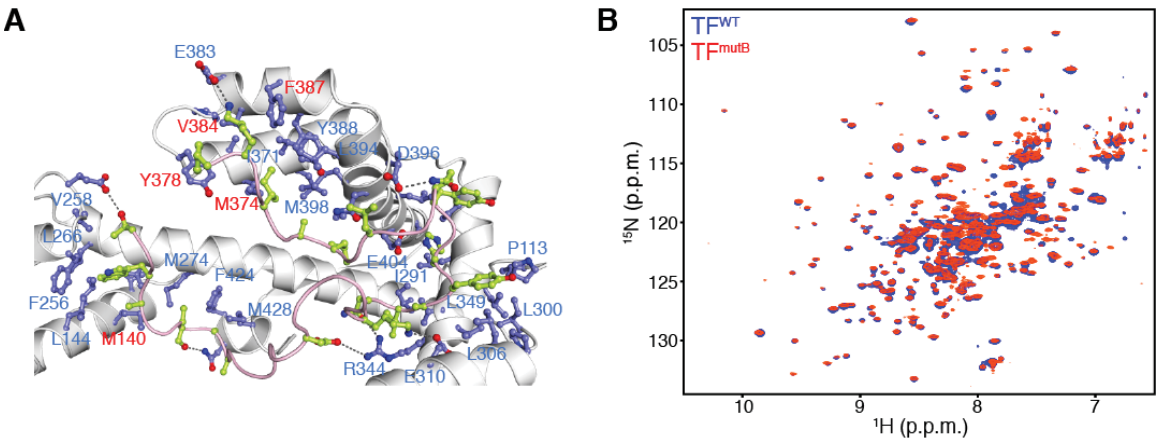


Fig. S23. Mutations in substrate-binding sites of TF. (A) Location of selected amino-acid substitution sites highlighted in red. Four of them (M374A, Y378A, V384A, F387A) are located in the B site, with the corresponding TF mutant referred to as TF^{mutB}, whereas the M140E substitution is located in the C site and thus the corresponding TF mutant is referred to as TF^{mutC} (see materials and methods for more information about the various mutants). **(B)** ^{1}H-^{15}N HSQC spectra of wild type TF (TF^{WT}, blue) and TF^{mutB} (red). The spectra clearly show that TF^{mutB} is well folded. The results of the effect of these mutations on the activity of TF are shown in Fig. 5.

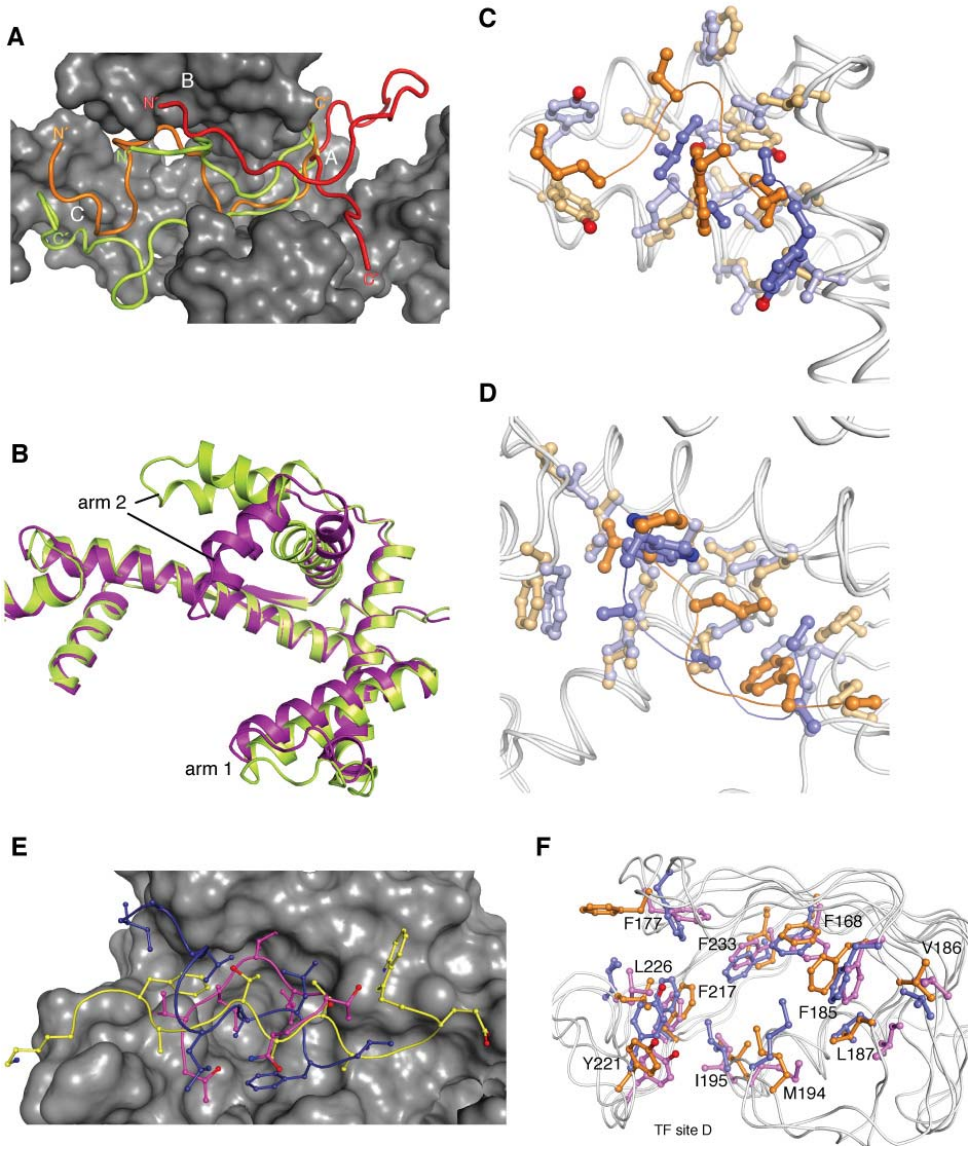


Fig. S24. Conformational plasticity and heterogeneity in TF-PhoA interactions. (A) Superposition of the three structures of TF SBD in complex with PhoA sites *c* (green), *e* (orange), and *g* (red). SBD is shown as dark grey solvent-accessible surface whereas PhoA sites are shown in ribbon. **(B)** Superposition of two TF conformers indicating the dynamic character of arm 2. This flexibility was also seen in the two TF molecules in the crystallographic asymmetric unit *(12)*. **(C,D)** Superimposed structures of TF site B (in **C**) and site C (in **D**) in two of the complexes. Both TF and PhoA are shown. TF side chains are colored light blue and light orange in the two complexes whereas PhoA side chains are colored dark blue and dark orange in the corresponding complexes. The figures demonstrate the different interactions between TF and PhoA in the two complexes. **(E)** Superimposed structures of the three structures of TF PPD in complex with their PhoA sites (site *a* is in blue), *d* is in yellow, and *f* is in magenta. PPD is shown as dark grey solvent-accessible surface whereas PhoA sites are shown in ribbon. The PhoA residues that interact with PPD are shown as sticks. **(F)** Side chain rearrangements in the three complexes between TF and PhoA sites *a*, *d*, and *f*. The TF side chains are shown as sticks and are colored blue, orange and magenta in TF−PhoA^{a}, TF−PhoA^{d}, and TF−PhoA^{f}, respectively. PhoA is not shown for clarity.

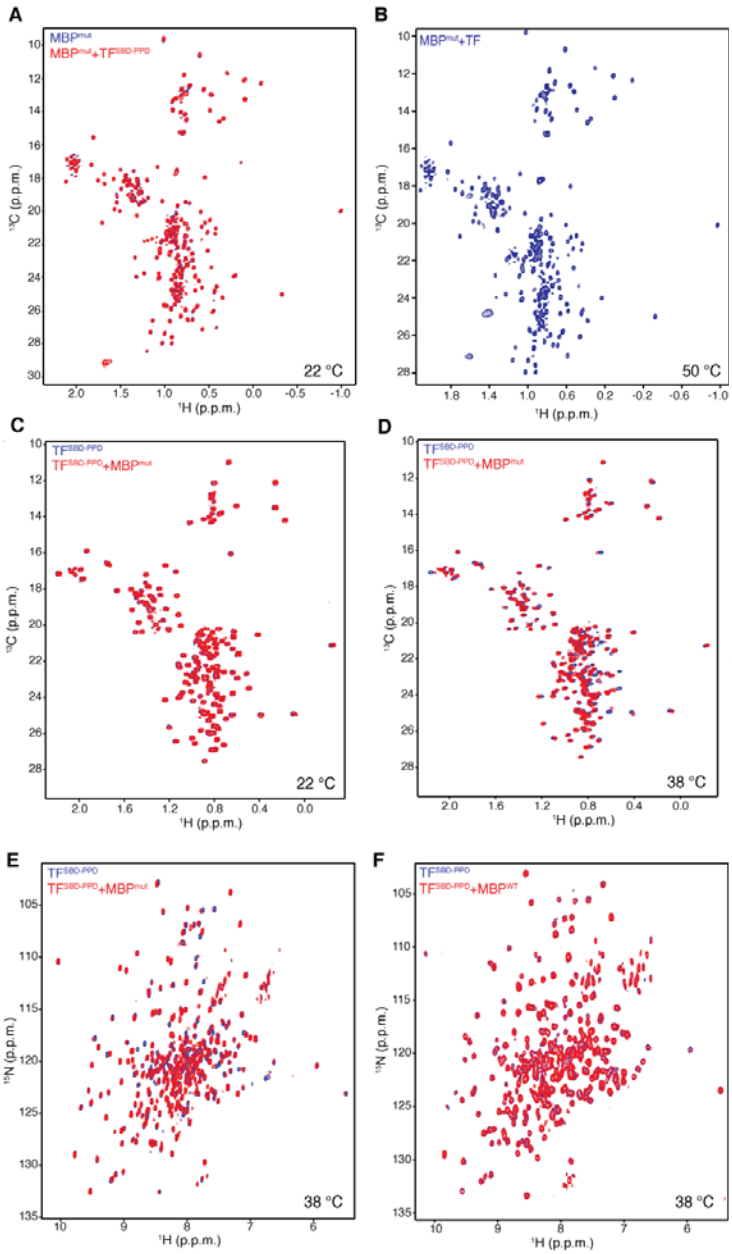


Fig. S25. TF binds to the unfolded state of MBP^{mut}. (A) ^{1}H-^{13}C HMQC spectra of MBP^{mut} free and in complex with TF$^{SBD-PPD}$ recorded at 22 °C. The data show no binding. **(B)** ^{1}H-^{13}C HMQC spectra of MBP^{mut} in complex with TF recorded at 50 °C. MBP^{mut} suffers heavy precipitation and aggregation at temperatures higher than 30 °C, but in the presence of TF it is stable and folded even at temperatures as high as 50 °C. **(C)** ^{1}H-^{13}C HMQC of TF$^{SBD-PPD}$ free and in complex with MBP^{mut}. This experiment is similar to the one shown in panel A but with reverse labeling. No binding was detected at 22 °C between the folded state of MBP^{mut} and TF. **(D)** ^{1}H-^{13}C HMQC of TF$^{SBD-PPD}$ free and in complex with MBP^{mut} at 38 °C. Because of the elevated temperature, a significant unfolded population of MBP^{mut} is present, which binds to TF. **(E)** ^{1}H-^{15}N HSQC spectra of TF$^{SBD-PPD}$ free and in complex with MBP^{mut} at 38 °C. **(F)** ^{1}H-^{15}N HSQC spectra of TF$^{SBD-PPD}$ free and in the presence of MBP^{WT}. No binding was detected between the two proteins. For the entire series of experiments, the effect of MBP^{mut} or MBP^{WT} binding to full-length TF or TF$^{SBD-PPD}$ are essentially identical and so TF and TF$^{SBD-PPD}$ were used interchangeably. Proteins were [U-^{2}H,^{15}N; Ala-$^{13}CH_3$; Met-$^{13}CH_3$; Ile-δ1-$^{13}CH_3$; Leu,Val-$^{13}CH_3$/$^{13}CH_3$] labeled.

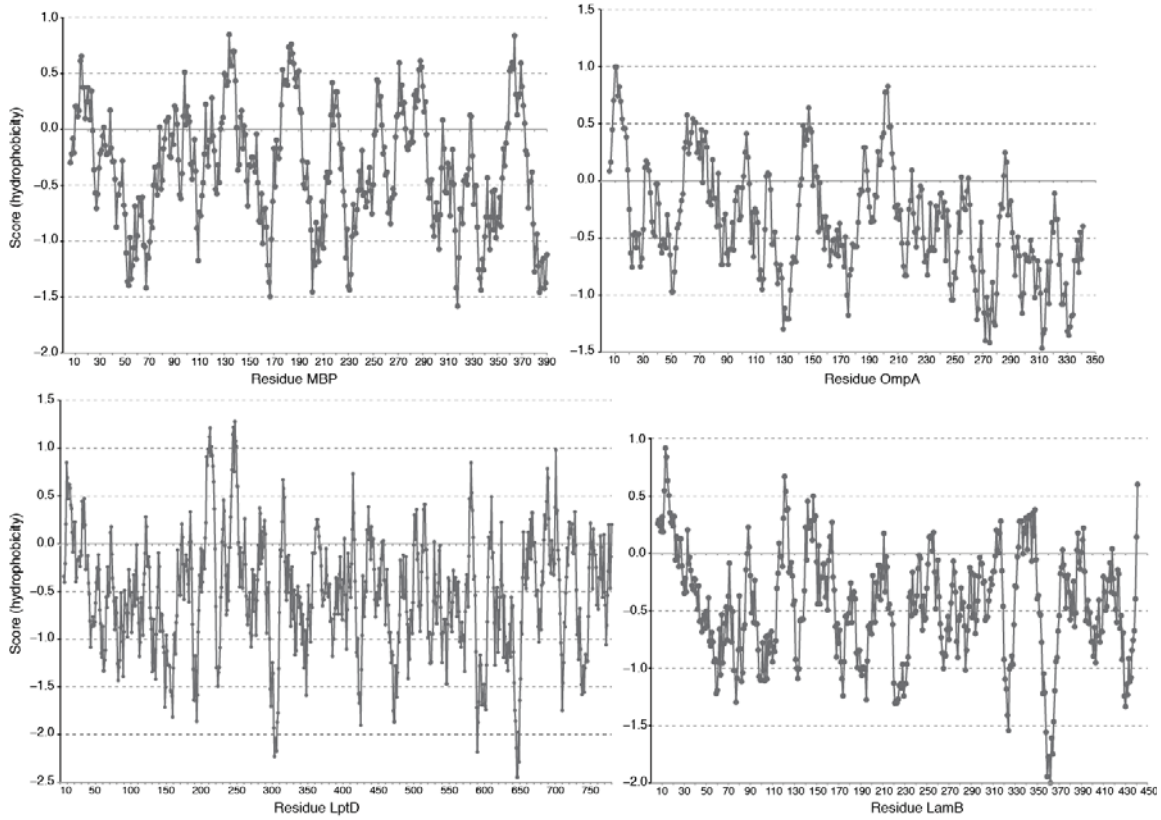


Fig. S26. Sequence hydrophobicity for E. coli proteins previously identified to bind strongly to TF *(25)***.** Plot of the hydrophobicity of MBP, OmpA, LTD and LamB as a function of its primary sequence. A hydrophobicity score (Roseman algorithm, window=9) higher than zero denotes increased hydrophobicity.

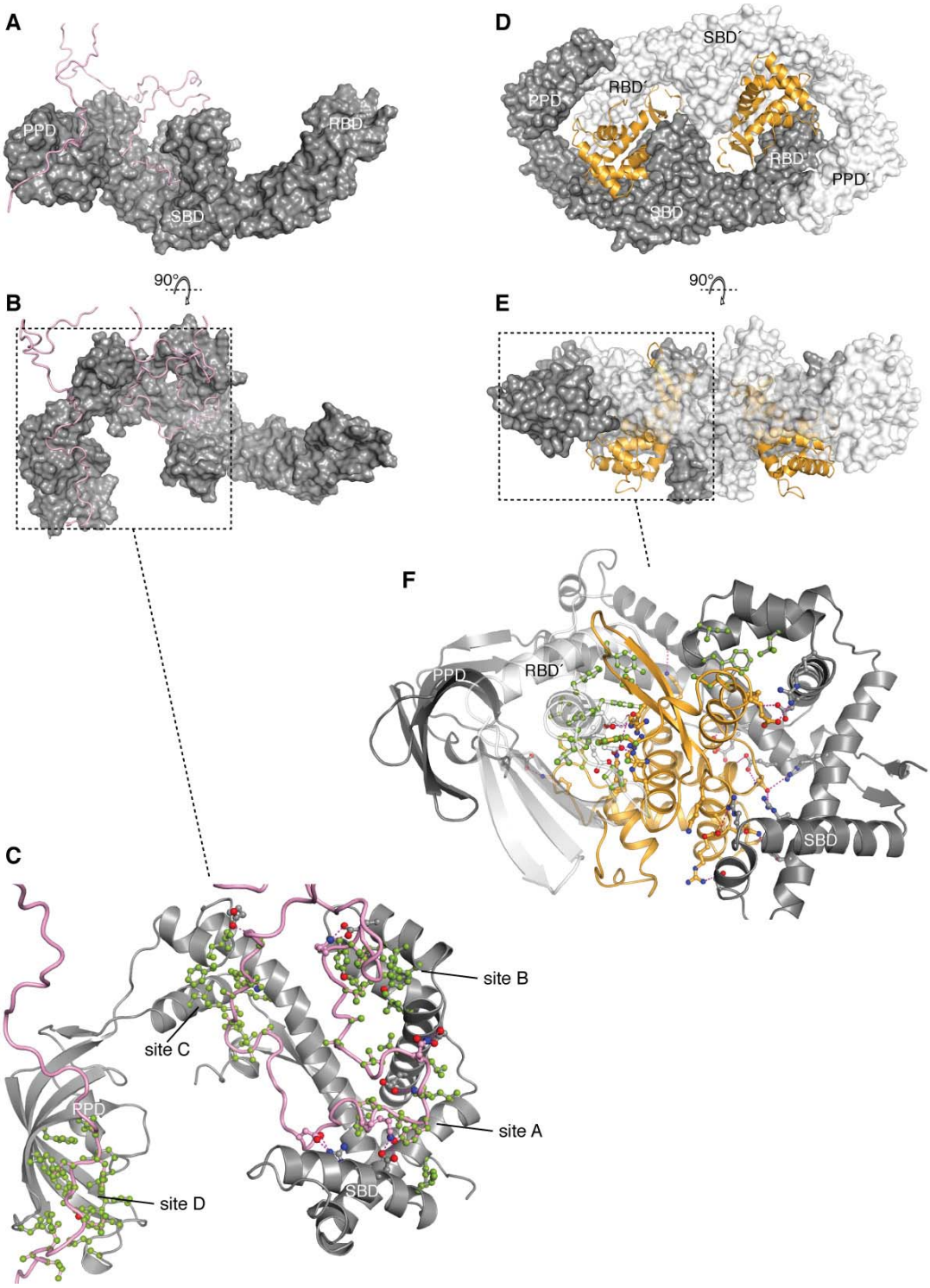


Fig. S27. Functional versatility of TF. In addition to the well characterized binding activity of TF for unfolded proteins, it has been recently observed *(27)* that small proteins can bind in a folded state to a cavity formed in the dimeric TF. **(A** to **C)** Structural basis for the interaction of TF with unfolded proteins, as determined in this work. TF is shown in a grey solvent-accessible surface and PhoA is shown in a pink ribbon. **(D** to **F)** Structural basis for the interaction of TF with small folded proteins as determined previously by crystallography *(27)*. TF is dimeric (the

two subunits are shown in grey and semi-transparent white surface) and the folded protein (ribosomal protein S7) in orange. The TF−S7 complex is from *thermotoga maritima*. Views in **A** and **B** (**D** and **E**) are related by a 90° rotation about the x axis. S7 binds in a cavity that is formed by the SBD and PPD of one subunit and the RBD of the other subunit. The majority of the contacts between S7 and TF are polar (H-bonds and salt bridges, denoted with the magenta broken lines), with only few hydrophobic contacts (hydrophobic residues are shown in green, panel **f**). In contrast, in the TF−PhoA complex, the vast majority of the contacts are hydrophobic (panel **c**). TF dimerization appears to be essential for the stabilization of the TF−S7 complex, whereas TF is forced to monomerize in the TF−PhoA complex. Further experiments are required in order to understand whether TF monomerization is substrate-length dependent and how TF can apparently bind both to unfolded as well as folded proteins. It should be noted that several large (over 30 kDa) folded proteins, such as PhoA and MBP, were tested and were found not to interact with TF. Therefore, it seems that TF can bind only to relatively small folded proteins as they will have to fit inside the cavity formed by the two TF subunits.

Table S1. Structural and NMR statistics of the three TF-PhoA complexes.

Complex	TF-PhoA[a-c]	TF-PhoA[d-e]	TF-PhoA[f-g]
Distance restraints[a]			
NOEs			
Short range (intraresidue and sequential)	858	578	615
Medium range ($2 \leq \mid i\text{-}j \mid \leq 4$)	257	137	211
Long range ($\mid i\text{-}j \mid > 4$)	781	415	587
Intermolecular	63	78	67
Hydrogen bonds	244	250	246
Dihedral angle restraints (φ and ψ)	511	480	500
Violations (mean and SD)[a]			
Distance restraints (Å)	0.013±0.044	0.012±0.051	0.013 ±0.042
Dihedral angle restraints (°)	0.04±0.44	0.04±0.37	0.07±0.50
Structural coordinates rmsd[a]			
PPD			
Backbone atoms	0.41 Å	0.55 Å	0.65 Å
All heavy atoms	0.70 Å	0.85 Å	0.92 Å
SBD			
Backbone atoms	1.22 Å	1.09 Å	1.22 Å
All heavy atoms	1.50 Å	1.25 Å	1.50 Å
Ramachandran plot[a]			
Most-favored regions	86.1%	86.4%	86.6%
Additionally allowed regions	13.7%	13.4%	13.4%
Generously allowed regions	0.2%	0.1%	0.1%
Disallowed regions	0.0%	0.0%	0.0%

[a]The statistics apply to the 20 lowest-energy structures of each of the three complexes. Restraints and analysis applied for the SBD and PPD parts of TF.