Probing Fragment Complementation by Rigid-Body Docking: in Silico Reconstitution of Calbindin D9k

Daniele Dell'Orco,[†] Michele Seeber,[†] Pier Giuseppe De Benedetti,[†] and Francesca Fanelli*[,†]

Department of Chemistry and Dulbecco Telethon Institute, University of Modena and Reggio Emilia, via Campi 183, 41100 Modena, Italy

Received May 17, 2005

Fragment complementation is gaining an increasing impact as a nonperturbing method to probe noncovalent interactions within protein supersecondary structures. In this study, the fast Fourier transform rigid-body docking algorithm ZDOCK has been employed for in silico reconstitution of the calcium binding protein calbindin D9k, from its two EF-hands subdomains, namely, EF1 (residues 1−43) and EF2 (residues 44−75). The EF1 fragment has been used both in its wild type and in nine mutant forms, in line with in vitro experiments. Consistent with in vitro data, ZDOCK reconstituted the proper fold of wild-type and mutated calbindin, locating the nativelike structures (i.e., holding a root-mean-square deviation < 1 Å with respect to the X-ray structure) among the first 10 top-scored solutions out of 4000. Moreover, the three independent in silico reconstitutions of wild-type calbindin ranked a nativelike structure at the top of the output list, that is, the best scored one. The algorithm has been also successfully challenged in reconstituting the EF2 homodimer from two identical copies of the monomer. Furthermore, quantitative models consisting of linear correlations between thermodynamic data and ZDOCK scores were built, providing a tested tool for very fast in silico predictions of the free energy of association of protein−protein complexes solved at the atomic level and known to not undergo significant conformational changes upon binding.

INTRODUCTION

The aim of fragment complementation studies is to achieve the complete reconstitution of a protein from separate fragments, often corresponding to well-defined supersecondary structure elements such as domains or subdomains.[1] This is accomplished without compromising either the 3D structure or the function of the reconstituted protein, with respect to the native system. The increasing picture of protein folding as a hierarchical process,[2−5] hence, involving the assembly of building blocks interacting with one another to achieve the global fold, emphasizes the interrelation between the folding and the binding processes, which appear to be based on the same physical principles.[6,7] Furthermore, this analogy points out the unique advantage of reconstituting systems, which exhibit a dynamic equilibrium of associated and dissociated fragments. Since such an equilibrium can be related to the global folding and unfolding, fragment complementation represents an effective tool for a deep screening of both environmental and direct mutational effects on the global stability, without perturbing the system.[1,8−12]

A wide range of protein systems has been shown to undergo complete reconstitution, either if naturally composed of at least two amino acid chains, such as the dimeric monellin[13] and chimeric G protein coupled receptors,[14,15] or when a chemical cleavage is performed on a single chain, in a region not fundamental for preserving the overall fold.[8−12,16] The calcium binding protein calbindin D9k, the smallest member of the calmodulin family, which is deeply involved in intracellular signaling, belongs to the latter group. The 3D structure of the protein is known at high resolution, and it can be reconstituted with high affinity from its two EF-hand helix-loop-helix subdomains (i.e., EF1, residues 1−43, and EF2, residues 44−75; Figure 1A, B).[11,12,17] The ability of the EF-hand motifs to associate with one another, reconstituting the native fold of a number of calcium-binding proteins, has been widely documented.[18−20] As for calbindin D9k, NMR[11,21,22] and crystallographic[23] determinations show that, in the presence of calcium, the reconstituted protein retains the same structure as the intact calbindin, except for a slight difference at the region connecting the subdomains. The reconstitution capability also appears when single or double point mutations are performed either on the hydrophilic surface or in the hydrophobic core of the EF1 subdomain. This potential allowed for a detailed thermodynamic and kinetic characterization of the reconstitution process for the wild type (WT) and mutants lying either at the solvent-exposed surface of EF1[11] or buried within the EF1−EF2 interface.[12] In vitro experiments pointed out the large effect of point mutations on affinity and kinetics compared to changes in pH or ionic strength of the solvent.[11] Furthermore, mutations involving hydrophobic residues were found to affect both the rate and the equilibrium constants to a notably greater extent than those on the hydrophilic surface.[11,12]

The fact that isolated calbindin fragments retain their structure in the presence of calcium[24] and correspond to well-defined structural motifs, that is, EF-hand domains, makes this system particularly intriguing for in silico studies aimed

* Corresponding Author. Phone: +39-059-2055114. Fax: +39-059-37353. E-mail: ffanelli@dti.unimo.it, fanelli@unimo.it.
[†] University of Modena and Reggio Emilia.

10.1021/ci0501995 CCC: $30.25
Published on Web 09/02/2005

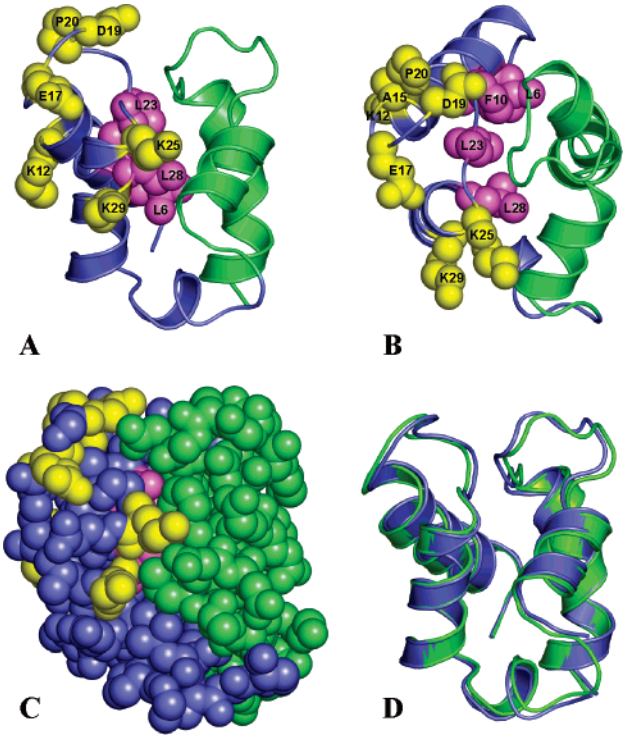

Figure 1. Molecular representation of calbindin D9k's X-ray structure (PDB code 4ICB). (A) EF-hand subdomains are represented by cartoons: EF1 (residues 1−43) is shown in slate blue, whereas EF2 (residues 44−75) is shown in lime green. The amino acid side-chain targets of mutagenesis are represented by spheres. In detail, a yellow color indicates the solvent-exposed residues, whereas a pink color represents residues located in the hydrophobic core. (B) Top view of the same representation as in A. (C) Space-filling representation of the reconstituted calbindin D9k, colored as described above. (D) Cartoon representation of the native complex (yellow-green color) superimposed on the best nativelike EF1−EF2 complex (slate blue color). The Cα RMSD between the two structures is equal to 0.59 Å. Drawings were achieved by means of the PYMOL software.[45]

at investigating the interaction between protein subdomains, in the light of the interrelation between binding and folding, as emphasized also by computational experiments.[25]

Recently, a novel fast Fourier transform protein−protein docking algorithm named ZDOCK2.3 (ZD) was shown to superbly reproduce the nativelike structures of a significant number of protein−protein complexes.[26−28] Results were outstanding for both enzyme−inhibitor and antibody−antigen complexes.[29] The high performance of ZD certainly resides in its docking scoring function, which takes into account shape and electrostatic complementarity, as well as desolvation. In this respect, the pairwise shape complementarity function has particular weight as, instead of relying on protein surface curvature, it computes the total number of target-ligand atom pairs within a distance cutoff, minus a clash penalty. Thus, within such scoring function, the grid-based function enhances pairwise target-ligand contacts, minimizing discontinuity and notably decreasing the number of false positives, whereas the clash penalty accounts for potential bad contacts that may arise from the rigid-body approach.[26]

Under the condition of an exhaustive sampling, the features of the ZD scoring function, taken together, might reasonably represent a first-stage rough estimation of the binding free energy of two interacting proteins for the many cases where experiments demonstrate that both the docked proteins do not undergo major conformational changes upon binding. The latter requirement is due to the fact that the rigid-body approach neglects the "conformational entropy" contribution to the free energy of binding.

In this work, the performance of ZD has been challenged, for the first time, to reconstitute in silico both WT and mutated calbindin D9k, starting from two subdomain fragments. The knowledge, from in vitro experiments, that both WT and mutated calbindin EF-hand fragments associate with each other, reconstituting the native fold of the intact protein, constitutes the foundation of our computational approach. The high performance of ZD in reconstituting calbindin D9k has also prompted us to challenge the algorithm in reconstituting the EF2−EF2 homodimer, reasonably supposed to hold the typical orthogonal bundle architecture shared by the members of the EF-hand homologous superfamily.[30] This would constitute a challenging unbound−unbound docking test, as docking has been performed on two identical copies of the structural motif extracted from the crystallographic structure of calbindin D9k.

Moreover, two different correlative models between ZD scores and in vitro-determined thermodynamic and kinetic data, that is, binding affinities, rate constants, and free energy of association, have been built, which proved a high potential in predicting the thermodynamics of association of novel protein−protein complexes solved at the atomic level and known to not undergo significant conformational changes upon binding.

Ultimately, this study suggests that a rigid-body approach might be a valuable tool for computational investigations of protein domain assembly.

METHODS

In silico Building of the EF1 and EF2 Fragments. Starting from the crystallographic structure resolved at 1.4 Å of the intact protein (PDB code 4ICB), each fragment was built in silico by means of the QUANTA2000 package (www.accelrys.com) by cleaving the peptide bond between M43 and S44 to create the EF1 and EF2 fragments (Figure 1A). The substitution P43M was performed in the WT structure and in the mutants according to the methods previously reported.[11,12,17] The conformations of all the mutated side chains, located in the EF1 fragment, were assigned according to the Karplus rotamer library.[31]

Rigid-Body Docking. Molecular simulations of both WT and mutated calbindin fragment reconstitution were carried out by means of the ZD rigid-body docking program.[28] The two calcium ions, one for each EF-hand motif, were deleted from the structures, as they are expected to play a minor role in the docking simulation, being not exposed to the solvent. The EF2 fragment was kept fixed (i.e., target), whereas the EF1 fragment, carrying mutations, was allowed to rotate and translate around the target (i.e., probe). A rotational sampling interval of 6° was employed, that is, dense sampling, and the best 4000 solutions were retained and ranked according to the ZD score. Docking simulations were also carried out between two identical copies of EF2 to build the homodimer. Three independent sets of docking runs were performed for reconstituting the EF1−EF2 complex, that is, one starting from the X-ray coordinates and the other two randomizing the initial positions of both the

Table 1. Results from the Three Sets of Runs Concerning the WT EF1−EF2 Docking: Solutions with Cα RMSD Lower than 1.0 Å with Respect to the Native Complex

run 1				run 2				run 3			
ZD solution[a]	RMSD[b]	ZD score[c]	3D-profile score[d]	ZD solution[a]	RMSD[b]	ZD score[c]	3D-profile score[d]	ZD solution[a]	RMSD[b]	ZD score[c]	3D-profile score[d]
1	0.74	56.71	40.9	1	0.85	59.09	39.7	1	0.84	60.61	39.8
3	0.84	54.74	38.9	2	0.75	58.40	39.8	2	0.73	60.11	39.2
5	0.81	52.79	40.8	3	0.75	54.79	38.0	3	0.90	58.77	39.0
7	0.79	52.28	39.3	4	0.65	54.21	38.6	6	0.66	54.96	39.6
8	0.78	52.03	40.6	5	0.90	53.48	40.1	7	0.95	52.61	40.5
9	0.79	51.91	40.1	6	0.83	52.98	40.4	9	0.65	51.73	40.0
10	0.96	51.40	38.9	9	0.72	51.70	37.3	10	0.73	51.70	38.6
11	0.64	50.96	39.0	10	0.99	51.16	38.2	11	0.73	51.58	38.6
12	0.69	49.71	39.2	14	0.93	49.23	38.6	12	0.98	49.29	39.8
13	0.94	49.16	38.7	15	0.82	49.14	40.5	13	0.96	49.17	39.8
16	0.63	48.21	40.0	16	0.69	48.81	37.4	19	0.80	46.65	39.5
17	0.77	47.64	41.6	17	0.81	48.42	39.1	20	0.91	46.12	39.3
18	0.97	47.44	43.4	19	0.86	45.26	39.0	24	0.63	45.16	38.2
19	0.99	47.27	40.1	20	0.59	45.21	39.1	28	0.89	43.87	38.6
21	0.99	46.67	39.7	21	0.60	45.06	40.1	29	0.78	43.86	38.2
23	0.78	46.11	39.6	30	0.79	43.17	40.7	30	0.70	43.71	40.1
24	0.90	45.30	39.1	31	0.80	42.94	38.3	33	0.86	42.79	38.7
28	0.78	44.01	39.9	34	0.92	42.51	41.7	34	0.97	42.53	39.2
35	0.97	42.40	39.8	43	0.93	41.00	38.2	37	0.65	42.26	40.2
40	0.86	42.09	39.6	44	0.91	40.92	36.6	39	0.99	41.72	41.6
41	0.91	41.95	40.2	77	0.96	38.08	38.0	56	0.94	39.68	39.7
42	0.70	41.77	39.7	91	0.99	37.49	37.4	58	0.94	39.43	39.6
50	0.78	40.59	40.9	111	0.85	36.88	38.9	86	0.97	37.35	39.1
58	0.96	39.81	41.5	197	0.82	35.27	41.8	133	0.97	35.88	39.4
62	0.95	39.22	38.5	202	0.86	35.18	40.5	359	0.99	33.24	39.3
76	0.84	38.57	39.4	282	0.92	34.26	40.9	581	0.97	32.17	37.0
120	0.98	37.02	40.8	958	0.93	31.49	41.6				
122	0.98	37.00	38.5								
175	0.87	35.98	40.3								
231	0.96	35.11	37.6								

[a] Docking solution numbers ordered according to the ZDOCK ranking; that is, number 1 is the highest scored solution, whereas 4000 is the lowest one. [b] RMSDs (Å) between the Cα atoms of the X-ray structure and each nativelike EF1−EF2 complex. [c] ZD score indicates the score associated with each solution by the docking program. [d] 3D-profile scores calculated according to the Luthy-Eisenberg method.[32] For the X-ray structure, such an index is 40.1.

target and the ligand. On the other hand, only one sampling was done for reconstituting the EF2 homodimer, that is, starting from two EF2 identical molecules in the same orientation held in the X-ray structure.

The predictive ability of ZD concerning the reconstitution of WT and mutated calbindin D9k from the EF1 and EF2 fragments was assessed by computing the root-mean-square deviation (RMSD) between the Cα atoms (i.e., Cα RMSD) of the X-ray structure, that is, native structure, and those of the predicted complexes. We, therefore, selected as nativelike structures all the EF1−EF2 complexes characterized by a Cα RMSD lower than 1.0 Å from the native structure. This criterion was employed both for the complexes involving WT and for those involving mutated EF1, on the basis of the evidence from in vitro reconstitution experiments that the considered EF1 mutants associate with EF2 similarly to the WT.[11,12]

For WT calbindin, the quality of each nativelike complex, expressed as a 3D-profile score,[32] was compared with that of the native structure (Table 1).

For each calbindin form, that is, WT and mutated, the ZD score averaged over all the nativelike complexes resulting from the three independent runs was employed for the correlation analysis with the thermodynamic and kinetic data (Table 2). To evaluate the bias introduced by the bound−bound docking in the simulation of calbindin reconstitution, additional docking runs were carried out by using EF2 fragments extracted from a different X-ray structure, that is, the structure resulting from EF-hand domain swapping[23] (PDB code 1HT9). The Cα RMSD between each of these two EF2 fragments and that extracted from the original X-ray structure of calbindin are 1.3 and 1.6 Å, respectively. Differences arise in both the main-chain and side-chain conformations. Similarly to the bound−bound docking tests, the search for "nativelike" solutions was done by computing the Cα RMSD with respect to the X-ray calbindin structure, from which the EF1 fragment was extracted. In this case, a

Table 2. Wild-Type and Mutated Calbindin D9k: Kinetic and Thermodynamic Data and ZDOCK Scores

	log k^{on}	log k^{off}	log K_A	ZD score[a]
WT[b,c]	5.1[b]/6.6[c]	−5.0[b,c]	10.1[b]/11.6[c]	46.9 ± 0.8
E17Q/D19N[b]	5.4	−4.4	9.8	49.7 ± 0.8
K12Q[b]	4.9	−4.7	9.6	45.6 ± 0.8
K25Q[b]	4.8	−4.0	8.7	40.5 ± 0.7
A15D/P20G[b]	5.1	−3.0	8.1	45.7 ± 0.8
K25E/K29E[b]	4.6	−3.1	7.7	37.5 ± 0.7
L6V[c]	6.0	−4.7	10.6	44.4 ± 0.9
F10A[c]	5.3	−2.5	7.78	41.5 ± 0.8
L23A[c]	5.9	−2.8	8.7	44.1 ± 0.7
L28A[c]	5.9	−3.0	8.8	42.8 ± 0.7

[a] ZD scores averaged over all the nativelike complexes resulting from three independent runs and relative standard errors. [b] These data concern in vitro reconstitution of calbindin D9k as reported by Dell'Orco et al.[11] [c] These data concern in vitro reconstitution of calbindin D9k as reported by Berggard et al.[12]

threshold of 2.0 Å was employed to account for the conformational differences in the EF2 fragment between native and predicted complexes.

Prediction of the most reliable fold for the EF2−EF2 homodimer followed a cluster analysis of the 4000 ZD solutions. This procedure was needed because of the lack of a high-resolution structure (i.e., native structure) for the EF2−EF2 assembly.

Clustering was carried out by using an algorithm from Dr. M. Schaefer (Michael Schaefer, Syngenta Crop Protection AG, unpublished work). The algorithm first calculates the Cα RMSD for each superimposed pair of complexes, and then, it computes the number of neighbors for each complex, using a threshold Cα RMSD. The complex with the highest number of neighbors is considered as the center of the first cluster. All the neighbors of this conformation are removed from the ensemble of conformations to be counted only once. The center of the second cluster is then determined in the same way as for the first cluster, and this procedure is repeated until each structure is assigned to a cluster. To aid the choice of the proper Cα RMSD threshold so as to achieve a restricted number of well-defined clusters, we introduced the dispersion index σ^2, mathematically defined as the variance of the statistical distribution of each structure in the cluster, which has also been used by other researchers,[33] though for a different purpose. Hence, if cluster i groups N_k structures, each indicated by p, its homogeneity may be described by

$$\sigma_i^2 = \frac{1}{N_k}\sum_{p=1}^{N_k}(\mathrm{RMSD}_p - \overline{\mathrm{RMSD}_i})^2$$

where $\overline{\mathrm{RMSD}_i}$ indicates the average Cα RMSD among the structures of cluster i. Clusters with reasonably low σ^2 are, with higher confidence than clusters with a high index, represented by their center, which by definition has a null Cα RMSD. Because of its definition, σ^2 square-rooted can be directly compared with the threshold chosen for running the cluster analysis. Our data suggest that σ values should not significantly exceed 1.0 Å, as an indicative value. An advantage of the dispersion index is that a plot of σ^2 versus the number of clusters, ordered according to decreasing population of the cluster, should result in a roughly linear trend with a barely positive slope. This would be indicative of homogeneity among all the resulting clusters, with a natural increase for the less populated clusters, consistent with the σ^2 definition.

Cα RMSD thresholds ranging from 2.5 to 5.5 Å were probed (Figure 2), leading to the final selection of the lowest value, which flattens the trend and guarantees the high homogeneity of the clusters. The centers of the most populated clusters were, then, subjected to calculations of the 3D-profile scores, to highlight the fold most consistent with the primary sequence. Finally, the best folds were subjected to a visual analysis.

Similarly to the calbindin mutants, three independent sets of docking runs were performed concerning the nine protein−protein complexes used to build a correlation model between ZD scores and in vitro-determined free energy of association (Table 3). The larger molecule of each complex was used as a target, whereas the smaller was considered as a probe. For each complex, the ZD score averaged over all the nativelike solutions was then used to search for correlations with in vitro data.

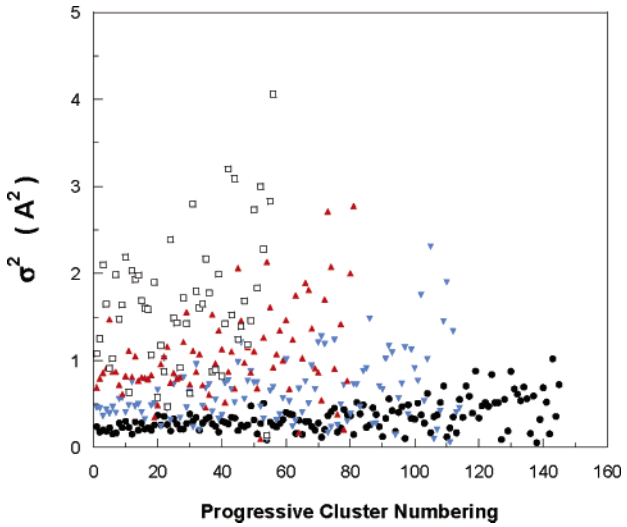

Figure 2. Plot of the dispersion parameter σ^2 (Å^2) versus the cluster numbers for the reconstitution of the EF2 homodimer. Clusters are ordered according to a decreasing population. The results of clustering at different Cα RMSD cutoffs are represented by different symbols. In detail, black open squares refer to a 5.5 Å cutoff, red triangles refer to a 4.5 Å cutoff, blue triangles refer to a 3.5 Å cutoff, and black circles refer to a 2.5 Å cutoff.

Predictions of the free energy of association for all of the calbindin species were also performed by means of the STC 5.0 software,[34] which is based on previous work on empirical parametrization of the binding energetics.[35]

RESULTS

ZD Results: Prediction of the Integral Calbindin Fold. For each docking run, that is, three independent runs for the WT and nine mutants of the EF1 subdomain, ZD was able to predict a correct reconstitution for calbindin D9k. In fact, nativelike complexes, that is, those structures characterized by a Cα RMSD from the native structure lower than 1.0 Å, were found, in all the cases, among the best 10 solutions out of 4000. Furthermore, the first solution in the output list, that is, the one with the absolute highest score, was found to be nativelike in all three of the independent runs of WT reconstitution (Table 1). Among these solutions, the highest and lowest deviations (in terms of Cα RMSD) from the native fold were 0.99 Å and 0.59 Å, respectively (Table 1; Figure 1D). Thus, for each EF1 form, that is, WT or a differently mutated subdomain, the three independent docking runs converged into comparable results, despite the marked differences in the starting positions of the target and probe.

Finally, a further proof of the excellent performance of the docking algorithm comes from calculations of the 3D-profile scores.[32] Indeed, for all the nativelike complexes, this index is comparable to, if not higher than, that of the native structure (Table 1).

The predictive effectiveness of the algorithm is not due to the fact that the two interacting fragments have a structural memory of the interaction, their conformation being exactly the same as that in the native complex (PDB code 4ICB), that is, the bound−bound docking case. Indeed, two ad-

Table 3. Protein−Protein Complexes Used in the Correlative Study: Standard Free Energies of Association and ZDOCK Scores

complex	PDB entry	ZD score[a]	ΔG°_{obs}[b] (kcal/mol)	ΔG°_{cal}[c] (kcal/mol)
barnase-mutant/barstar	1BRS	53.8 ± 0.5	−17.3[d]	−17.0
alfa-chimotrypsin/turkey ovomucouid third domain (OMTKY3)	1CHO	46.3 ± 0.5	−14.4[e]	−14.2
subtilisin Carlsberg/eglin-C	1CSE	43.8 ± 0.4	−13.1[e]	−13.3
subtilisin Carlsberg/genetically engineered eglin-C	2SEC	43.0 ± 0.4	−13.1[f]	−13.0
thermitase/eglin-C	1TEC	47.9 ± 0.4	−14.0[d]	−14.8
subtilisin novo/chymotrypsin inhibitor	2SNI	52.0 ± 0.5	−15.8[d]	−16.3
proteinase B from *Streptomyces griseus*/OMTKY3	3SGB	39.1 ± 0.4	−12.7[e]	−11.6
carboxypeptidase Aα/potato carboxypeptidase A inhibitor	4CPA	38.3 ± 0.2	−10.0[f]	−11.3
human leukocyte elastase/OMTKY3	1PPF	44.9 ± 0.4	−13.5[d]	−13.7

[a] ZD scores averaged over all the nativelike complexes resulting from three independent runs and relative standard errors. [b] Free energy of association as obtained by in vitro measurements. [c] Free energy of association as calculated by using the ZD score and the correlative equation reported in Figure 5A. [d] Data from ref 42. [e] Data from ref 43. [f] Data from ref 44.

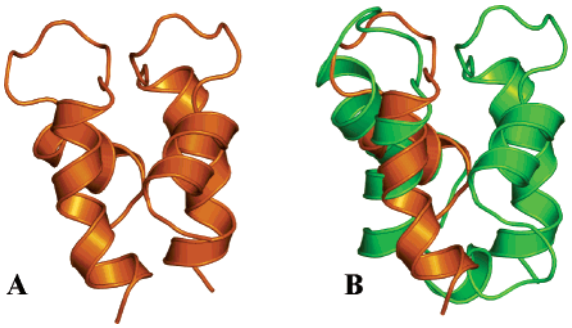

Figure 3. (A) Cartoon representation of one of the most reliable models of the EF2 homodimer. The ZD and 3D-profile scores for this model are, respectively, 41.6 and 33.1. (B) Superimposition of intact calbindin D9k (yellow-green color) on the best-scored complex ranked by ZD (orange color).

ditional independent docking runs, by using the EF2 fragment extracted from each of the two different calbindin structures derived from EF-hand domain swapping (PDB code 1HT9), led to effective reconstitution of calbindin as well. These novel EF2 domains differ in both the main-chain and side-chain conformations from the partner of EF1 in the native complex, providing a sort of bound−unbound docking case. Both of the two independent bound−unbound docking tests provided calbindin-like structures as the top scores in the output list, that is, the first one out of 4000 solutions. These top-scored calbindin-like structures showed Cα RMSDs from native calbindin equal to 1.39 Å and 1.34 Å. This is an excellent result, considering that such Cα RMSD values also include the main-chain conformational differences between the EF2 domains.

ZD Results: Prediction of the EF2−Homodimer Fold. Following cluster analysis with a RMSD threshold of 2.5 Å (Figure 2), EF2−EF2 assemblies in the typical orthogonal bundle architecture shared by the members of the EF-hand homologous superfamily (www.biochem.ucl.ac.uk/bsm/cath; Figure 3) were found in one of the most populated, highest-scored, and extremely homogeneous clusters (i.e., $\sigma^2 = 0.17$ Å^2). The highest 3D-profile score[32] also characterizes the members of this cluster; that is, this index is above 30.0 for the orthogonal bundle architectures, whereas it is below 20 for the members of the other most populated clusters. This result proves the ability of the 3D-profile scores to discriminate correct from incorrect folds.[32]

The model of the EF2−EF2 homodimer, thus, selected as one of the most reliable models on the basis of cluster population as well as ZD and 3D-profile scores, is shown in Figure 3. The buried interface is made of (a) contacts between L46 and both F66 and V70, (b) contacts between L49 and both F66 and I73, (c) contacts between F50 and F63, and (d) interactions between the side chain of V61 from both monomers. Almost all these amino acids also participate in the buried EF1−EF2 interface in the native structure of calbindin D9k. A number of negatively charged and solvent-exposed amino acids approach the interface between the two monomers. Some of them, including E60 and E64 from both monomers, are not involved in the binding of calcium and exert a repulsive effect at the interface. A comparison between the native structure of the EF1−EF2 heterodimer and the computational model of the EF2 homodimer suggests that the electrostatic repulsion at the interface contributes to the lower stability of the homodimer compared to the heterodimer. Phage-display studies on the interaction between the EF3 and EF4 subdomains in calmodulin led to similar conclusions.[36] In line with these observations, the ZD scores properly account for the lower complementarity, which characterizes the association of two EF2 domains compared to the EF1−EF2 heterodimerization. In fact, the best solutions of the EF2 homodimerization show significantly lower ZD scores (at most 41.6) compared to the nativelike heterodimers (at most 60.6), either in their WT or mutated forms (Table 1, Figure 3).

These results are consistent with the evidence, from ion exchange chromatography, agarose gel electrophoresis, and NMR spectroscopy experiments, that EF2 may form homodimers, although with remarkably lower affinity than heterodimers.[17] Consistent with the results of in vitro experiments, these results can be considered a successful example of unbound−unbound docking prediction yet unquantifiable in terms of Cα RMSD because of the lack of a high-resolution structure of the EF2 homodimer.

Structure-Affinity Relationship Analysis of Calbindin Mutants. Figure 1 shows the X-ray structure of WT calbindin D9k. The picture shows, as well, the location of the amino acids targeted by in vitro and in silico mutagenesis. These amino acids, which all belong to the EF1 subdomain, have been divided into two sets, according to previous in vitro experiments.[11,12] One set includes solvent-exposed amino acids in the intact protein (i.e., K12, A15, P20, E17, D19, K25, and K29; yellow color in Figure 1A, B, and C), whereas the other set comprises amino acids buried at the EF1−EF2 interface (i.e., L6, F10, L23, and L28; pink color in Figure 1A, B, and C). Among the solvent-exposed residues, only K25 and K29 participate in the EF1−EF2 interface. In

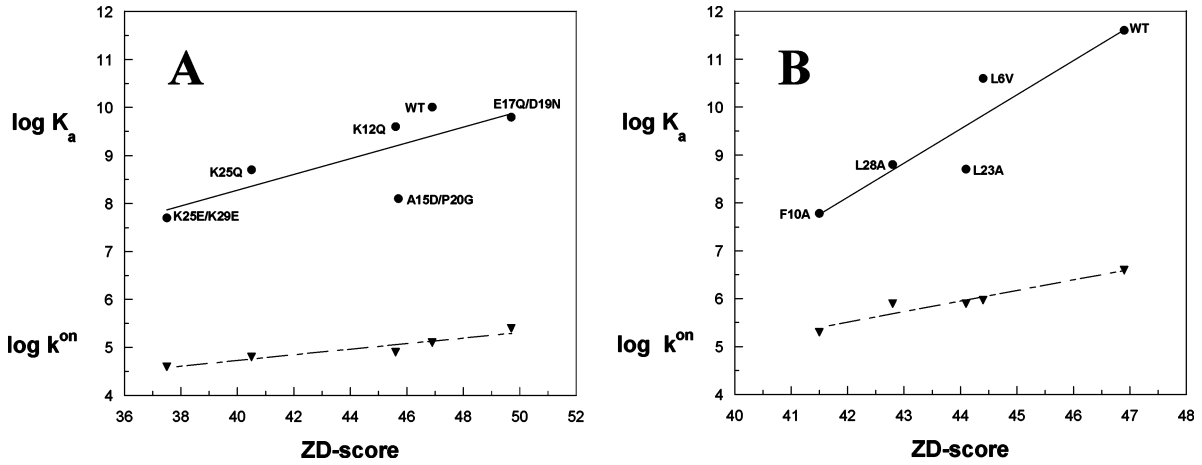

Figure 4. Plot of the calbindin reconstitution kinetic and thermodynamic data versus ZD scores. In each panel, the upper line refers to the equilibrium constant (log K_a) of each calbindin form (i.e., WT or mutated), whereas the lower line refers to the association constant (log k^{on}). (A) Set of solvent-exposed residues:[11] the fitting equations are $y = 1.70 + 0.16x$ (correlation coefficient $r = 0.77$, correlation probability $p = 0.076$, number of points $N = 6$), for log K_a, and $y = 2.40 + 0.06x$ ($r = 0.95$, $p = 0.004$, $N = 6$), for log k^{on}. The first equation becomes $y = 1.10 + 0.18x$ ($r = 0.95$, $p = 0.015$) when A15D/P20G is omitted in the fitting. (B) Set of buried residues:[12] the fitting equations are $y = -22.00 + 0.71x$ ($r = 0.92$, $p = 0.026$, $N = 5$), for log K_a, and $y = -3.7 + 0.22x$ ($r = 0.96$, $p = 0.009$, $N = 5$), for log k^{on}.

fact, K25 is involved in Coulombic interactions with both D47 and E51, in EF2, whereas K29 performs significant Coulombic interactions only with D47. The analysis of the calbindin structure, therefore, suggests that charge reversal or neutralization of K25 and K29 may affect the EF1−EF2 recognition, more than the other solvent-exposed amino acid replacements. In contrast, the D19N mutation might favor the EF1−EF2 association by decreasing the repulsion at the interface.

As for the members of the buried set, (a) L6 lies in a hydrophobic cleft formed by F66, V70, and I73, all in EF2. The conservative L6V mutation is expected to marginally affect the EF1−EF2 binding. (b) F10 is involved in orthogonal $\sigma-\pi$ interactions with both F63 and F66, in EF2. The F10A mutation is, therefore, expected to affect the EF1−EF2 packing. (c) L23 is involved in van der Waals interactions with V61, whereas (d) L28 interacts with F50, L53, and F66. Alanine replacements of either L23 or L28 are, hence, expected to have a negative effect on reconstitution efficiency.

In vitro experiments have shown that the nature of mutations involved in the reconstitution of calbindin D9k affects both the equilibrium and the kinetics of the process. As reported in Table 2, all the solvent-exposed and buried mutants show higher dissociation rates than WT, with a stronger effect observed for the buried set, compared to the solvent-exposed one. The highest effect on k^{off} observed in the solvent-exposed set of mutants is that of the K25E/K29E and A15D/P20G mutations (Table 2). The increase in k^{off} for the K25E/K29E mutant, compared to that of WT, is in line with the inferences from the structural analysis that such double substitution is likely to induce an electrostatic repulsion at the EF1−EF2 interface. Thus, the decrease of k^{on} and increase in k^{off}, which characterize the K25E/K29E mutant, result in a 260-fold decrease in the binding affinity. In all the tested cases, mutations affect the dissociation rate constant much more than the association rate constant (Table 2). Only the E17Q/D19N mutant, which could ameliorate the electrostatic complementarity between EF1 and EF2, shows a 2-fold increased k^{on} with respect to the WT. In contrast, and consistent with a reduction of the electrostatic complementarity at the interface, a 3-fold slower association is observed for the K25E/K29E mutant. The most relevant effect on the dissociation rate is observed for the hydrophobic core mutant F10A, which shows a 300-fold faster dissociation with respect to the WT.

It should be pointed out that important mutational effects on the individual rate constants do not always correspond to significant effects on the equilibrium constant, and, hence, on the thermodynamic stability, as suggested by data reported in Table 2. Significantly broader effects are found within the hydrophobic set (Table 2).

Quantitative Relationships between ZD Scores and Thermodynamic and Kinetic Parameters Concerning WT and Mutated Calbindin D9k. Interesting linear trends have been obtained between the ZD scores, averaged over the scores of all the nativelike solutions from three independent docking runs, and the thermodynamic and kinetic parameters reported in Table 2.

The solvent-exposed and the buried-hydrophobic sets of mutants have been treated separately, since in vitro measurements were performed in different conditions.[11,12] These included the immobilization technique employed for the kinetic measurements, which, contrarily to the case of the solvent-exposed set, for the buried set, was not homogeneous.[11] This might result in the differences observed between the two sets in the association rate constant and binding affinity concerning the WT, which are thought to be overestimated in early measurements, that is, those concerning the hydrophobic set (Table 2).[11,12] Indeed, such overestimations are suspected to concern the whole set of hydrophobic mutants.

Figure 4A reports a plot of log K_a and log k^{on} versus the ZD score concerning the WT and the solvent-exposed mutants. The ZD score gives the best correlation with log k^{on} ($r = 0.95$, $p = 0.004$) and an acceptable correlation with log K_a ($r = 0.77$, $p = 0.076$). The dissociation rate constant k^{off}, instead, does not correlate linearly with the ZD score ($r = 0.5$, $p \approx 0.31$; data not shown). The plot in Figure 4A indicates that the docking program assigns to the A15D/P20G

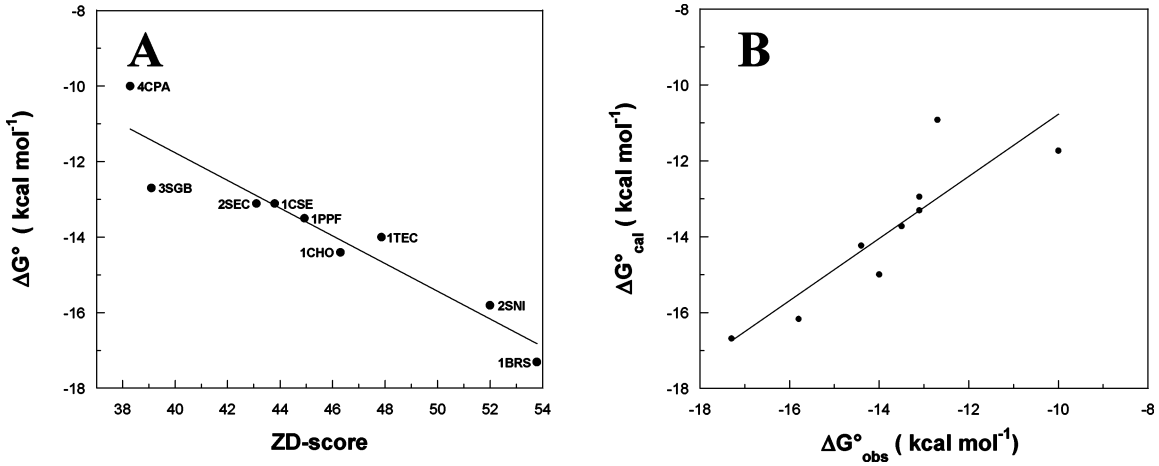

Figure 5. (A) Linear correlation between average ZD scores and in vitro-determined free energies of association concerning a selected set of nine protein−protein complexes (Table 3). Each dot is labeled according to the PDB code of the complex. The linear correlation equation is $y = 2.9-0.37x$ ($r = 0.94$, $p = 0.00015$, $N = 9$). (B) Calculated versus in vitro-determined free energy of association of the same nine complexes. The calculated values refer to a leave-one-out test. The linear correlation equation is $y = -2.6 + 0.82x$ ($r = 0.88$, $p = 0.0018$, $N = 9$; cross-checked $r^2 = 0.77$).

mutant a score similar to that of the WT or of the mutant K12Q, consistent with k^{on} but inconsistent with the binding affinity data. The peculiarity of this mutant from both the kinetic and the thermodynamic points of view has been previously discussed,[11] and the extremely lower affinity, compared to that of WT, might be explained in terms of important structural changes in the mutated fragment compared to the other mutants. It should be pointed out that leaving out the A15D/P20G mutant from the fitting significantly improves the correlation concerning both the log K_a and the k^{on} values ($r = 0.95$, $p = 0.015$ in both cases).

Likewise, Figure 4B reports a plot of the log K_a and log k^{on} values versus the ZD score as obtained from the three sets of independent and randomized docking runs concerning the buried mutants. In this, like in the previous set, a good linear trend can be seen between ZD scores and both log K_a ($r = 0.92$, $p = 0.026$) and log k^{on} (i.e., $r = 0.96$, $p = 0.009$). This suggests that the ZD scores correctly predict that the F10A mutation has the most detrimental effect on the association process and on the binding affinity compared to the other mutants in the set and that all of the mutated forms are expected to interact with lower affinity with respect to the WT, as it is indeed confirmed by in vitro experiments.[12] However, in view of the suspected overestimation of both log k^{on} and log K_a for this set of mutants,[11] the trend is expected to change in an unpredictable way with more accurate in vitro data.

Predictive Model of the Free Energy of Association of WT and Mutated Calbindin D9k based upon the ZD Score of Selected Protein−Protein Complexes. In this study, we have also built a quantitative model consisting of a linear correlation between the experimental free energy of association ($\Delta G°$) and the average ZD scores of nine selected protein−protein complexes with known structures (Table 3). The criteria for selecting the protein complexes include the availability of in vitro-determined $\Delta G°$ and the knowledge that no major conformational changes accompanied complex formation. Interestingly, the ZD scores averaged over all the nativelike structures resulting from three independent runs were highly correlated with $\Delta G°$ ($r = 0.94$, $p = 0.00016$; see Figure 5A). The cross-validated r^2 arising from the leave-one-out procedure was 0.77 (Figure 5B). Other protein complexes, characterized by both known structures and $\Delta G°$, could not be included in this test because of the known conformational changes of at least one of the interacting partners upon binding. For instance, several trypsin-inhibitor complexes (e.g., PDB entries 1TPA, 2TGP, 2PTC, 2KAI, 2TPI, 3TPI, and 4TPI) were tested in the present study with no satisfactory correlation between the ZD score and $\Delta G°$. These results are consistent with the pioneering observations by Katchalski-Katzir and co-workers[37] that the extent of inhibitor conformational changes exceeds the toleration of the grid-based correlation technique, which is employed by the ZD algorithm.[27]

The quantitative model shown in Figure 5 has been, hence, used for estimating the $\Delta G°$ values of WT and mutated Calbindin D9k fragments, by using the computed average ZD scores. The good predictive ability of the model is clearly shown by the linear trend between predicted and measured $\Delta G°$ values (Figure 6A and B). For the solvent-exposed set, the only evident outlier is the above-discussed A15D/P20G double mutant (Figure 6A). In detail, the predicted $\Delta G°$ values hold an average 9% relative error with respect to the experimental value, which becomes 7% if the A15D/P20G mutant (i.e., relative error = 27%) is omitted. In none of the remaining cases does the discrepancy exceed 17% of the experimental value.

The same correlative equation was employed to estimate the free energy of association of the EF2 homodimer, leading to a $\Delta G° = -9.6$ kcal/mol. Despite the lack of an in vitro determination for the homodimer affinity, the estimation appears consistent with the expectations from in vitro biochemical and biophysical experiments.[17,36]

DISCUSSION

There is increasing interest in the development of computational protocols and algorithms able to predict the quaternary structure of protein−protein complexes, for a wide range of issues concerning interacting protein systems.[38] On the other hand, the observation that dissociated subdomains may spontaneously find the way back to restoring the native

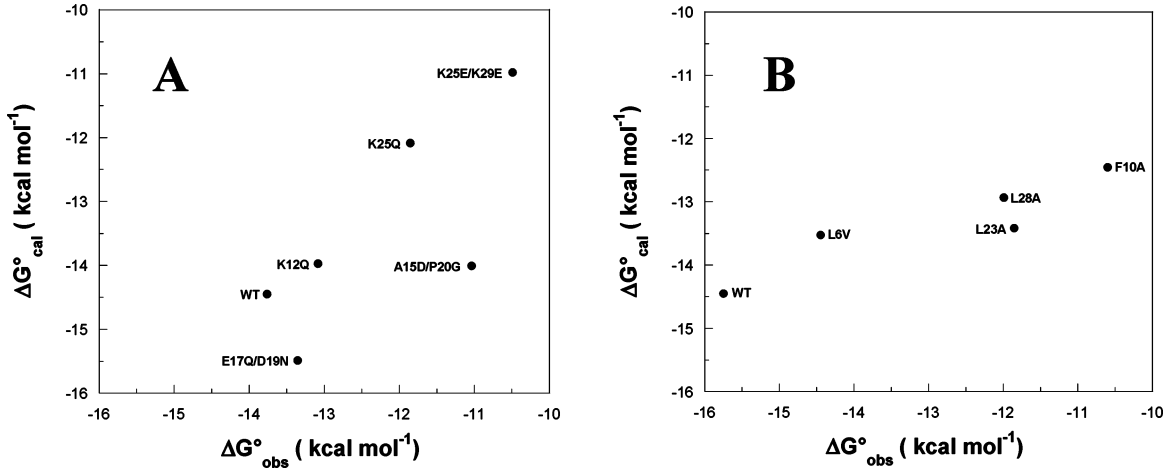

Figure 6. Predicted versus observed free energies of association of calbindin D9k subdomains concerning reconstitution of the solvent-exposed (A) and the buried set of mutants (B). In *Y*, the free energy of association (ΔG°_{cal}, kcal/mol) calculated by means of the correlative equation in Figure 5 is reported, whereas in X, that observed in vitro (ΔG°_{obs}, kcal/mol) is reported. The linear correlative equations are, respectively, $y = 2.7 + 1.3x$ ($r = 0.94$, $p = 0.018$, $N = 5$), omitting the A15D/P20G mutant, and $y = -9.1 + 0.33x$ ($r = 0.92$, $p = 0.026$, $N = 5$).

structure, either when they belong to the same polypeptide chain (i.e., folding) or when they constitute two separate polypeptide chains (i.e., reconstitution), represents the foundation of fragment complementation studies, which have been shown to be valuable nonperturbing tools for investigating the roles of noncovalent interactions in protein stability.[1,11,12]

Recently, in vitro reconstitution experiments on calbindin D9k were carried out, in which the cleavage of the peptide bond between amino acids 43 and 44, leading to separation of the 1−43 (i.e., EF1) and 44−75 (i.e., EF2) domains, was followed by exposure of the immobilized EF2 fragment to the freely diffusing solvated EF1 fragment.[11,12] The latter was considered in its WT and variously mutated forms. The computational study herein conforms to in vitro experiments in that the X-ray structure of calbindin D9k has been cleaved into EF1 and EF2 domains, which have then been subjected to rigid-body docking, following perturbation of their initial positions and employing EF2 as a target (i.e., fixed protein) and EF1 as a probe (i.e., mobile protein). The basic assumption of both in vitro and in silico reconstitution experiments is that the two fragments maintain their structure upon association. This would entitle the use of a rigid-body docking approach for in silico reconstitution of WT and mutated calbindins.

The results of our computational experiments strengthen and widen the performance of the ZD rigid-body docking program. In fact, in all the test cases, ZD has been able to predict nativelike protein assemblies of EF1 and EF2 supersecondary structures, independently of the mutation performed. In particular, the program has found a nativelike complex, that is, with a Cα RMSD lower than 1.0 Å from the X-ray structure, among the first 10 top-scored solutions out of 4000 (Figure 1D; Table 1). For the WT, the lowest Cα RMSD is 0.59 Å and the nativelike sets from the three independent runs always comprise solution number 1, that is, the highest-scored solution out of 4000 (Table 1). Accordingly, the 3D-profile scores of the nativelike solutions were always comparable if not higher than the score of the X-ray structure (Table 1). We have also verified that the high effectiveness of the algorithm in reconstituting calbindin is independent of the potential bias arising from the bound−bound docking.

Another worthy result of computational experiments is that, by using an adequate cluster analysis and ad hoc structure quality checks, we could reliably predict the structure of the EF2 homodimer, that is, an orthogonal bundle architecture typical of the EF-hand homologous superfamily, by unbound−unbound docking. Consistent with the expectations from in vitro experiments, the best scored solutions of homodimerization showed lower scores and estimated lower free energies of association when compared to any of the nativelike reconstitutions of the heterodimer.

Moreover, the challenge for a thermodynamic or kinetic correlative meaning of the ZD scores has led to interesting findings. In this respect, a highly significant correlation between the association rate constant and the docking scores has been observed for the set of the solvent-exposed mutants, including the WT complex, whereas no correlation has been found for the dissociation rate constants (Figure 4A). Also, the association equilibrium constant (log K_a) gives a linear correlation with the ZD scores, though lower than k^{on} (Figure 4A). The goodness of such a linear correlation, considering the fact that the ZD score accounts for almost all of the components of the free energy of binding except for the "conformational entropy", suggests that no major conformational changes occur upon binding of EF1 and EF2, consistent with the evidence from in vitro experiments.[11] In line with these observations, the significant improvement in the correlation between the ZD score and log K_a following deletion of the A15D/P20G mutant suggests that this mutant may have a somewhat considerable conformational effect, unlike the other tested mutants. This hypothesis finds support in the results of isothermal titration calorimetry experiments, which pointed out, for this mutant, an uncommon behavior, suggesting a major intrinsic entropic effect for this substitution compared to all the others.[11] Along the same line, comparisons of the NMR structure of the A15D/P20G mutant[21] with that of the WT calbindin D9k reveals significantly larger RMSDs compared, for example, with the NMR structure of the P43G mutant[39] (results not shown).

The interpretation of the noteworthy correlation regarding ZD scores and association rate constants is even more intriguing. As commonly observed in protein−protein interactions, measurement of the association kinetics, reported in Table 2, shows that bimolecular rate constants for calbindin D9k reconstitution are on the order of 10^6 $M^{-1}s^{-1}$, which is 3 orders of magnitude slower than the diffusion-limited Smoluchowski rate.[40] The evidence for many faster associations ($k^{on} \sim 10^7/10^9$ $M^{-1}s^{-1}$) driven by strong attractive Coulombic forces[40,41] would exclude the possibility that this latter is indeed the case of calbindin reconstitution. However, the combination of structural and correlation analyses done in this study suggests a significant contribution of the long-range electrostatic interactions for calbindin D9k reconstitution as well. In fact, the increase in k^{on} seen for the E17Q/D19N mutant and the decrease in the same constant seen for the single and double mutants of K25, could be, respectively, explained, at least in part, in terms of a reduction and increase of the electrostatic repulsion between the EF1 and EF2 subdomains.

Also intriguing is the linear correlation found between the ZD score and the binding affinity and association constants concerning the set of buried mutants (Figure 4B). This is a further proof of the consistency between in vitro and in silico experiments.

The ability of the ZD score of nativelike complexes to correlate with in-vitro-measured $\Delta G°$ values has also been clearly demonstrated, in this study, for a broad set of protein−protein associations (Figure 5). The high degree of correlation implies that, in the cases where proteins do not undergo major conformational changes upon binding, the ZD scoring function has the potential of an empirically determined free energy function, which does not need the decomposition in enthalpic and entropic contributions, which is often complicated to achieve. The basic requirements with this approach are the native structures of the protein−protein complexes, an exhaustive and dense sampling of the roto-translational space spanned by the probe, and an effective scoring function. In this study, we have demonstrated that the correlative model between average ZD scores and in vitro-determined $\Delta G°$ values has a good capability to estimate the free energy of association of WT and mutated calbindin fragments (Figure 6A and B). This occurs despite the lack of structural similarity between any member of the training set and those of the test set. This makes the model independent of the structural features of the interacting proteins and of general predictive applicability.

The predicted $\Delta G°$ for WT calbindin is close to the one obtained by using an alternative structure/thermodynamics correlation (STC) method.[34,35] In detail, the WT $\Delta G°$ predicted according to the model proposed in this work (Figure 5) and that from the STC method[34] are −14.5 kcal/mol and −13.2 kcal/mol, respectively, whereas the in vitro-determined value is −13.8 kcal/mol. Despite its good performance concerning WT calbindin, the STC method fails in estimating the $\Delta G°$ of the calbindin mutants (results not shown). This may be due, at least in part, to the fact that this method strongly relies on the resolution level of the structural model of the protein−protein complex, which is expected to be low for the computational models of mutant calbindin D9k. In this respect, the strength of our correlative model is that it relies on ZD scores averaged over all the nativelike solutions rather than relative to a unique structure. This might overcome, at least in part, the low resolution of the computational models of the mutants resulting from rigid-body docking.

In conclusion, our results suggest that an extensive use of rigid-body docking with optimized scoring functions may be a useful tool for predicting the architecture of dimeric supersecondary structures, which do not undergo major conformational changes upon association.

Furthermore, computational experiments suggest that ZD may be employed as a first-stage method for a coarse-grained and fast estimation, in a correlative context, of the mutational effects on the energetics and kinetics of binding in those cases where mutations do not affect the conformation of the structural motifs.

ACKNOWLEDGMENT

This study was supported by a Telethon−Italy grant (Grant S00068TELA; to F.F.).

REFERENCES AND NOTES

(1) Hakansson, M.; Linse, S. Protein reconstitution and 3D domain swapping. *Curr. Protein Pept. Sci.* **2002**, *3*, 629−642.
(2) Tsai, C. J.; Ma, B.; Nussinov, R. Folding and binding cascades: shifts in energy landscapes. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9970−9972.
(3) Tsai, C. J.; Maizel, J. V., Jr.; Nussinov, R. Anatomy of protein structures: visualizing how a one-dimensional protein chain folds into a three-dimensional shape. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12038−12043.
(4) Haspel, N.; Tsai, C. J.; Wolfson, H.; Nussinov, R. Hierarchical protein folding pathways: a computational study of protein fragments. *Proteins* **2003**, *51*, 203-215.
(5) Haspel, N.; Tsai, C. J.; Wolfson, H.; Nussinov, R. Reducing the computational complexity of protein folding via fragment folding and assembly. *Protein Sci.* **2003**, *12*, 1177−1187.
(6) Yang, X. M.; Georgescu, R. E.; Li, J. H.; Yu, W. F.; Haierhan. Recognition between disordered polypeptide chains from cleavage of an alpha/beta domain: self-versus nonself-association. *Pac Symp. Biocomput. '99* **1999**, 590−600.
(7) Haliloglu, T.; Keskin, O.; Ma, B.; Nussinov, R. How similar are protein folding and protein binding nuclei? Examination of fluctuations of energy hot spots and conserved residues. *Biophys. J.* **2005**, *88*, 1552−1559.
(8) Kobayashi, N.; Honda, S.; Munekata, E. Fragment reconstitution of a small protein: disulfide mutant of a short C-terminal fragment derived from streptococcal protein G. *Biochemistry* **1999**, *38*, 3228−3234.
(9) Honda, S.; Kobayashi, N.; Munekata, E.; Uedaira, H. Fragment reconstitution of a small protein: folding energetics of the reconstituted immunoglobulin binding domain B1 of streptococcal protein G. *Biochemistry* **1999**, *38*, 1203−1213.
(10) Kobayashi, N.; Honda, S.; Yoshii, H.; Uedaira, H.; Munekata, E. Complement assembly of two fragments of the streptococcal protein G B1 domain in aqueous solution. *FEBS Lett.* **1995**, *366*, 99−103.
(11) Dell'Orco, D.; Xue, W. F.; Thulin, E.; Linse, S. Electrostatic contributions to the kinetics and thermodynamics of protein assembly. *Biophys. J.* **2005**, *88*, 1991−2002.
(12) Berggard, T.; Julenius, K.; Ogard, A.; Drakenberg, T.; Linse, S. Fragment complementation studies of protein stabilization by hydrophobic core residues. *Biochemistry* **2001**, *40*, 1257−1264.
(13) Xue, W. F.; Carey, J.; Linse, S. Multi-method global analysis of thermodynamics and kinetics in reconstitution of monellin. *Proteins* **2004**, *57*, 586−595.
(14) Maggio, R.; Vogel, Z.; Wess, J. Reconstitution of functional muscarinic receptors by coexpression of amino- and carboxyl-terminal receptor fragments. *FEBS Lett.* **1993**, *319*, 195−200.
(15) Scarselli, M.; Armogida, M.; Chiacchio, S.; DeMontis, M. G.; Colzi, A. Reconstitution of functional dopamine D(2s) receptor by coexpression of amino- and carboxyl-terminal receptor fragments. *Eur. J. Pharmacol.* **2000**, *397*, 291−296.
(16) Berggard, T.; Thulin, E.; Akerfeldt, K. S.; Linse, S. Fragment complementation of calbindin D28k. *Protein Sci.* **2000**, *9*, 2094−2108.

(17) Finn, B. E.; Kordel, J.; Thulin, E.; Sellers, P.; Forsen, S. Dissection of calbindin D9k into two Ca(2+)-binding subdomains by a combination of mutagenesis and chemical cleavage. *FEBS Lett.* **1992**, *298*, 211−214.
(18) Permyakov, E. A.; Medvedkin, V. N.; Mitin, Y. V.; Kretsinger, R. H. Noncovalent complex between domain AB and domains CD*EF of parvalbumin. *Biochim. Biophys. Acta* **1991**, *1076*, 67−70.
(19) Durussel, I.; Luan-Rilliet, Y.; Petrova, T.; Takagi, T.; Cox, J. A. Cation binding and conformation of tryptic fragments of Nereis sarcoplasmic calcium-binding protein: calcium-induced homo- and heterodimerization. *Biochemistry* **1993**, *32*, 2394−2400.
(20) Linse, S.; Thulin, E.; Gifford, L. K.; Radzewsky, D.; Hagan, J. Domain organization of calbindin D28k as determined from the association of six synthetic EF-hand fragments. *Protein Sci.* **1997**, *6*, 2385−2396.
(21) Johansson, C.; Ullner, M.; Drakenberg, T. The solution structures of mutant calbindin D9k's, as determined by NMR, show that the calcium-binding site can adopt different folds. *Biochemistry* **1993**, *32*, 8429−8438.
(22) Linse, S.; Johansson, C.; Brodin, P.; Grundstrom, T.; Drakenberg, T. Electrostatic contributions to the binding of Ca2+ in calbindin D9k. *Biochemistry* **1991**, *30*, 154−162.
(23) Hakansson, M.; Svensson, A.; Fast, J.; Linse, S. An extended hydrophobic core induces EF-hand swapping. *Protein Sci.* **2001**, *10*, 927−933.
(24) Julenius, K.; Robblee, J.; Thulin, E.; Finn, B. E.; Fairman, R. Coupling of ligand binding and dimerization of helix-loop-helix peptides: spectroscopic and sedimentation analyses of calbindin D9k EF-hands. *Proteins* **2002**, *47*, 323−333.
(25) Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins* **2002**, *47*, 409−443.
(26) Chen, R.; Weng, Z. A novel shape complementarity scoring function for protein−protein docking. *Proteins* **2003**, *51*, 397−408.
(27) Chen, R.; Weng, Z. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Proteins* **2002**, *47*, 281−294.
(28) Chen, R.; Li, L.; Weng, Z. ZDOCK: an initial-stage protein-docking algorithm. *Proteins* **2003**, *52*, 80−87.
(29) Chen, R.; Tong, W.; Mintseris, J.; Li, L.; Weng, Z. ZDOCK predictions for the CAPRI challenge. *Proteins* **2003**, *52*, 68−73.
(30) Nakayama, S.; Kretsinger, R. H. Evolution of the EF-hand family of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 473−507.
(31) Dunbrack, R. L., Jr.; Karplus, M. Backbone-dependent rotamer library for proteins. Application to side-chain prediction. *J. Mol. Biol.* **1993**, *230*, 543−574.
(32) Luthy, R.; Bowie, J. U.; Eisenberg, D. Assessment of protein models with three-dimensional profiles. *Nature* **1992**, *356*, 83−85.
(33) Paul, N.; Rognan, D. ConsDock: A new program for the consensus analysis of protein−ligand interactions. *Proteins* **2002**, *47*, 521−533.
(34) Lavigne, P.; Bagu, J. R.; Boyko, R.; Willard, L.; Holmes, C. F. Structure-based thermodynamic analysis of the dissociation of protein phosphatase-1 catalytic subunit and microcystin-LR docked complexes. *Protein Sci.* **2000**, *9*, 252−264.
(35) Baker, B. M.; Murphy, K. P. Prediction of binding energetics from structure using empirical parametrization. *Methods Enzymol.* **1998**, *295*, 294−315.
(36) Linse, S.; Voorhies, M.; Norstrom, E.; Schultz, D. A. An EF-hand phage display study of calmodulin subdomain pairing. *J. Mol. Biol.* **2000**, *296*, 473−486.
(37) Katchalski-Katzir, E.; Shariv, I.; Eisenstein, M.; Friesem, A. A.; Aflalo, C. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2195−2199.
(38) Janin, J.; Henrick, K.; Moult, J.; Eyck, L. T.; Sternberg, M. J. CAPRI: a Critical Assessment of PRedicted Interactions. *Proteins* **2003**, *52*, 2−9.
(39) Kordel, J.; Skelton, N. J.; Akke, M.; Chazin, W. J. High-resolution structure of calcium-loaded calbindin D9k. *J. Mol. Biol.* **1993**, *231*, 711−734.
(40) Northrup, S. H.; Erickson, H. P. Kinetics of protein−protein association explained by Brownian dynamics computer simulation. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3338−3342.
(41) Northrup, S. H.; Boles, J. O.; Reynolds, J. C. Brownian dynamics of cytochrome *c* and cytochrome *c* peroxidase association. *Science* **1988**, *241*, 67−70.
(42) Ma, X. H.; Wang, C. X.; Li, C. H.; Chen, W. Z. A fast empirical approach to binding free energy calculations based on protein interface information. *Protein Eng.* **2002**, *15*, 677−681.
(43) Weng, Z.; Delisi, C.; Vajda, S. Empirical free energy calculation: comparison to calorimetric data. *Protein Sci.* **1997**, *6*, 1976−1984.
(44) Horton, N.; Lewis, M. Calculation of the free energy of association for protein complexes. *Protein Sci.* **1992**, *1*, 169−181.
(45) DeLano, W. L. *The PyMOL Molecular Graphics System*, version 0.95; DeLano Scientific: San Carlos, CA.

CI0501995