

Molecular Dynamics Simulations of a Binding Intermediate between FKBP12 and a High-Affinity Ligand

Lilian Olivieri and Fabrice Gardebian*

DSIMB, INSERM, U665, Paris, F-75015, France, DSIMB, INSERM, U665, Faculté des Sciences et Technologies, Université de la Réunion, 15, avenue René Cassin, BP 7151 97715 Saint Denis Messag Cedex 09, La Réunion, France, INTS, Paris, F-75015, France, and Laboratoire de Biochimie et Génétique Moléculaire, Université de la Réunion, 15, avenue René Cassin, BP 7151 97715 Saint Denis Messag Cedex 09, La Réunion, France

 Supporting Information

ABSTRACT: We characterized a binding intermediate between the protein FKBP12 and one of its high-affinity ligands by means of molecular dynamics simulations. In such an intermediate, which is expected to form at the end-point of the bimolecular diffusional search, short-range interactions between the molecular partners may play a role in the specificity of recognition as well as in the association rate. Langevin dynamics simulations were carried out to generate the intermediate by applying an external biasing force to unbind the ligand from the protein. The intermediate was then refined by seven independent molecular dynamics simulations performed with an explicit solvent model. We found consistent results both for the structure of the protein and for the position of the ligand in the intermediate. The two carbonyl oxygens O₂ and O₃ of the ligand core region act as two main anchors, making permanent contacts in the intermediate. The transient contacts with the protein are made by the ligand noncore moieties whose structures and mobilities enable many alternative contacts of different types to be formed: $\pi-\pi$ molecular overlap and weak hydrogen bonds NH $\cdots\pi$, CH $\cdots\pi$, and CH \cdots O. Hence, the stability of the ligand at the entrance of the protein binding pocket offers the possibility of fine-tuning a variety of short-range contacts that involve the ligand noncore moieties. Under the hypothesis that the stability of this intermediate is related to the affinity of the ligand, this binding intermediate model comes closest to explaining the role played by the noncore moieties in the affinity of this ligand. Moreover, this model also provides a plausible explanation for how structurally diverse core motifs that all share the carbonyl atoms O₂ and O₃ bind to FKBP12.

1. INTRODUCTION

The protein FKBP12 is a cytosolic enzyme of 12 kDa that catalyzes the peptidylprolyl *cis-trans* isomerization. FK506, a fungal metabolite, can bind tightly to FKBP12 with an inhibition constant of 0.6 nM (FKBP stands for FK506 binding protein).¹ Rapamycin is another ligand that has a similar inhibition constant (0.3 nM).¹ The immunosuppressive properties of rapamycin and FK506 are currently used for the treatment of transplant rejection.^{2,3} In the brain, the expression level of FKBP12 is much higher than in immune tissues, and in neurons, the FK506-bound form of the protein has been associated with neuroprotective properties.⁴ Nerve regenerative properties are also induced by FK506 or analogues and involve the isoform FKBP52, whose N-terminal FK506 binding domain is very similar in sequence and structure to FKBP12.⁵ Much effort is directed toward finding FK506 analogues with neuroprotective and neurotrophic activities but devoid of the undesirable immunosuppressive activity that is functionally associated with the ligand region responsible for calcineurin inhibition. The high-affinity ligand 8,⁶ bound to the protein FKBP12 in Figure 1, is an example of such nonimmunosuppressive ligands (the ligand is labeled as in the work of Holt et al.).

The development of inhibitors of the protein FKBP12 represents a major interest for extending the potential of many therapeutic treatments. To this end, a detailed understanding of the structure–activity relationships of the FKBP12 ligands is critically

important.^{7,8} To address this problem, we have focused on a complex formed by FKBP12 and the ligand 8 that has an inhibition constant of 10 nM, making it a good model for understanding the molecular basis of its high affinity for this protein. Instead of exploring the structure and dynamics of the bound state, however, this study was intended to find a binding intermediate (or encounter complex) that is expected to form at the end-point of the bimolecular diffusional search. In this intermediate state, short-range interactions could play a crucial role for the specificity of recognition as well as for the association rate.⁹ As the association constant for a protein–ligand complex is the ratio between the association (k_{on}) and dissociation (k_{off}) rate constants, the binding properties of a ligand are thus dependent on k_{on} and/or k_{off} . Under the hypothesis that an intermediate is formed along the binding pathway, the overall rate of association k_{on} is increased, i.e., the affinity, whenever the intermediate has a slow dissociation rate or a fast association rate for the binding.^{10,11} In a NMR study of two phosphopeptides that bind to a mutant of the N-terminal SH2 domain of PI3-K,¹² Mittag et al. have shown the existence of an intermediate state along the association pathway. Interestingly, the lifetime of the intermediate is related to the affinity of the ligand: the higher the affinity of the ligand, the longer is the lifetime of the intermediate (in this case, the slow dissociation rate of the

Received: July 16, 2010

Published: January 31, 2011

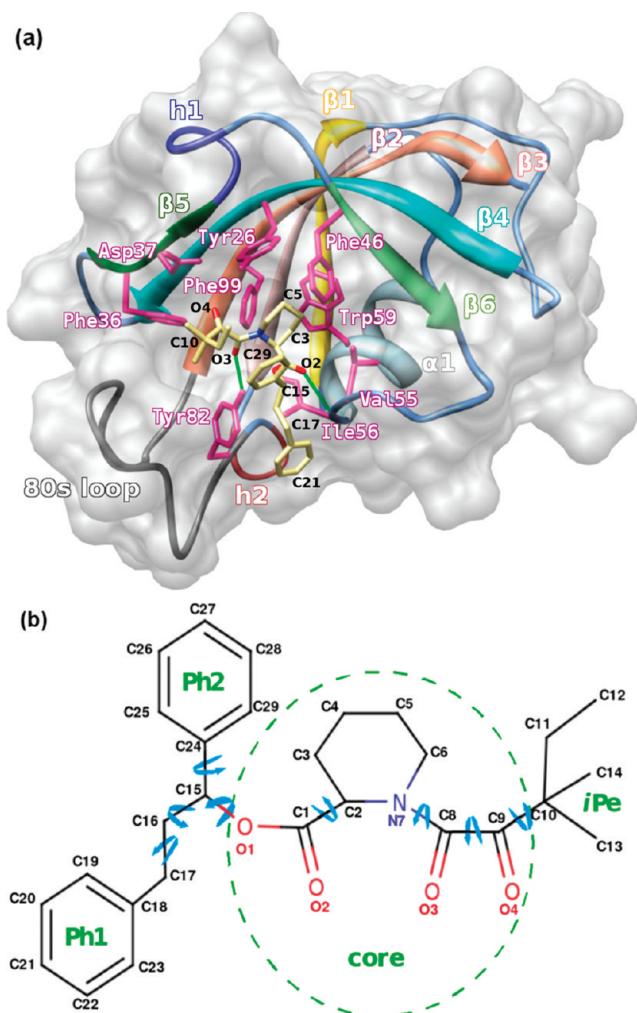


Figure 1. (a) Experimental structure of the complex between FKBP12 and the high-affinity ligand 8 (compound numbering is from ref 6). The β sheet and the α_1 helix are highlighted. The two native hydrogen bonds Ile56-NH \cdots O2 and Tyr82-OH \cdots O3 are also represented by thick green lines. (b) Sketch of the ligand 8 that indicates all of the atom numbers and the eight dihedral angles that were parametrized (highlighted by blue arrows). The core and the noncore regions iPe (C10 up to C14), Ph1 (C18–C23), and Ph2 (C24–C29) are also defined.

intermediate may explain the high affinity). The authors have also stressed that such intermediates may be more common in protein–ligand systems than previously anticipated. In the context of protein–protein association, an experimental kinetic study has correlated the decrease in affinity between Ras and a Ras-binding domain with the increase in the overall dissociation rate constant.¹³ Schreiber et al. have used a computer algorithm to predict the association rate between TEM1 β -lactamase and its inhibitor BLIP.¹⁴ The authors have found that charge mutations in the inhibitor increase the association rate, resulting in higher affinities.

In the case of the complex formed by FKBP12 and the high-affinity ligand 8, the properties of a binding intermediate along the association pathway could help to delineate the essential structural features of this ligand that are important for the recognition step. Two observations made for FKBP12 and its ligands have prompted the search of a binding intermediate. The first observation is the lack of correlation between the structural features

of the various FKBP12 ligands and their measured binding affinities. For example, many core regions of high-affinity ligands, which represent the regions that strongly interact with the protein, are structurally diverse.^{6,15–17} To further complicate the structure–affinity relationships, even for the ligands that are sharing the same diketo-pipercolic acid core (from the bond C9O4 up to the atom O1 in Figure 1b), it is difficult to explain the trends in the affinity data.⁶ The reasons are two-fold: first, the core regions of these ligands in the complexes are nearly superimposable, displaying the same contacts with the protein, and second, the noncore regions of the ligand often make few contacts with the protein.^{6,18} Hence, the structure–affinity relationships of both the core and the noncore regions of the high-affinity ligands of FKBP12 still await clarification. In this work, we address these issues by focusing on the case of the high-affinity ligand 8.

Prior to presenting the approach chosen, let us now consider the second observation that has also motivated this work and that is related to the structure of the binding pocket of FKBP12. In the crystal structure of FKBP12–8, the ligand 8 is deeply buried with only 30% of its surface exposed to the solvent: the pipercolinyl ring is facing the indole ring of Trp59, and the methyl group C13 of the isopentyl moiety (iPe) is lying in a small adjacent cavity lined by the side chains of Tyr82, His87, Ile90, and Ile91. If one considers a rigid docking of ligand 8 to a structure of FKBP12 that is the same as the bound form of the protein, then the binding would be prevented by a large steric hindrance. For a binding mechanism that would occur in such a single step, one may therefore anticipate a large energy barrier of complexation that would make this single step process kinetically inaccessible. Indeed, a positive enthalpic term and a negative entropic contribution would result from the steric hindrance and from the constraints to fit, in only one step, the pipercolinyl ring and the iPe moiety of the ligand into the main pocket and the small side cavity, respectively. The results gathered from the docking of the ligand FK506 to a rigid protein FKBP12 are consistent with the above analysis.¹⁹ In this rigid docking study, the author failed to identify a geometry close to the experimentally reported structure of FKBP12–FK506 because of unfavorable steric overlaps. Moreover, these docking results indicate that the relaxation of the protein in the loop region 82–95, also termed the 80s loop, helps to accommodate the ligand in the binding pocket. Hence, during the binding process of FK506 or related ligands such as 8, a displacement of the 80s loop would facilitate the binding to FKBP12. This hypothesis is also supported by a comparison between solution structures of unliganded FKBP12 and a crystal structure of FKBP12–FK506.²⁰ Taken together, the experimental and the docking studies^{19–21} suggest that, during the binding, the 80s loop undergoes conformational transitions to facilitate the entrance of the ligand by reducing the steric hindrance.

The successive events such as the 80s loop displacement, the first contacts with the ligand, and the formation of the fully bound structure may require only one step, all of these events being concerted; however, these may also proceed in two steps with the formation of a binding intermediate as a first binding event. This intermediate along the association pathway would allow decomposition of the process in a two-step mechanism in order to reduce the overall free energy barrier due both to steric hindrance and to an entropic penalty, as discussed above, thereby increasing the kinetic accessibility to the bound state.¹⁰ Our study has two objectives: (i) to locate a binding intermediate on the free energy surface of FKBP12 and its high-affinity ligand 8 and (ii), by

analyzing this intermediate, to provide a possible guide for understanding the role played by the ligand core and noncore regions in the specificity of recognition. A ligand loosely trapped in a proper orientation at the entrance of the protein binding pocket has translational and rotational motions that are substantially reduced, thereby allowing the ligand to fine-tune its short-range interactions with a more limited region of the protein surface. In particular, the extent and the nature of the intermolecular contacts in the intermediate may be related to its stability (large contact numbers favoring slow dissociation rate), and thus to the binding properties of the FKBP12 ligand. The analysis of this intermediate may also provide clues regarding the diversity of the core structures found in the high-affinity ligands.

To locate this intermediate, Langevin dynamics (LD)²² and an analytical model for the solvation free energy EEF1²³ were used in combination. As a preliminary step toward the search of this intermediate, we assessed the ability of this simulation protocol LD/EEF1 to reproduce accurately the experimental structure of the complex FKBP12–8. LD/EEF1 simulations were then carried out to unbind the ligand. To this end, an external force was applied to pull the ligand out of the binding pocket. This force was chosen as low as possible to ensure a minimal perturbation of the system. The structure of the intermediate state located by the LD simulations was subsequently refined by stochastic boundary molecular dynamics (SBD)²⁴ simulations that focused only on the binding region of the protein.

Under low-force conditions, forced unbinding simulations and spontaneous unbinding likely proceed through similar mechanisms.²⁵ Hence, forced unbinding simulations represent a new tool to investigate the structure–activity relationship.^{7,8,26,27} However, the present study is an original approach that focuses on the structural characteristics of a protein–ligand system in an early stage of recognition. We shall begin by describing the simulation procedure in section 2 (other computational details are provided in the Supporting Information). In section 3.1, the validation step of the LD/EEF1 methodology that was carried out on the bound state is presented. All of the results for the binding intermediate are presented and discussed in sections 3.2 and 4, respectively. These results provide a consistent picture of the binding intermediate with a well-defined position of the ligand that displays both permanent and transient contacts with the protein, thus revealing important aspects of its recognition properties in this early binding stage.

2. METHODS

2.1. Molecular Dynamics Simulations of the Bound State.

The X-ray coordinates of the complex between the protein FKBP12 and the synthetic inhibitor 8 (PDB code: 1FKG;⁶ Figure 1) were used for the simulations of the bound state. The all-atom Charmm22 force field²⁸ was used for the protein and the ligand; however, a few parameters for the ligand were calculated by following the procedure defined by Foloppe and MacKerell²⁹ (see the Supporting Information for details). All of the molecular dynamics simulations that used an implicit or explicit solvation model were performed with the CHARMM program³⁰ using the leapfrog integrator and a time step of 1 ps. SHAKE³¹ restraints were applied to the bonds containing hydrogen atoms. Langevin dynamics (LD)²² were used to simulate the bound state. To reproduce the effects of the water collision frequency at 300 K,³² the friction coefficient was chosen at 60 ps⁻¹. Effective Energy Function 1 (EEF1)²³ was chosen to compute the solvation free energy of the atoms in the LD simulations of both the bound state and

the unbinding process (see the Supporting Information for the parameters).

For the LD simulations of the bound state, the complex was energy-minimized and then heated from 0 to 300 K in steps of 25 K (for 600 ps); different random seeds for the distribution of initial velocities led to four independent simulations. The heating run was followed by an equilibration divided into two periods of 500 ps each (see the Supporting Information for details on the minimization, heating, and equilibration procedures). After the equilibration, the LD production runs were conducted with five NOE restraints between the residue pairs Glu31–Thr96 and Lys34–Ile90 and for a total simulation time of 15.5 ns: 4, 3, 5, and 3.5 ns for the four runs LD1, LD2, LD3, and LD4, respectively. It should be noted that the NOE restraints were used during the LD simulations to avoid a distortion of the 80s loop due to the absence of explicit water molecules (see the Supporting Information).

To compare the implicit solvation model with an explicit model, three independent stochastic boundary molecular dynamics (SBD)²⁴ simulations of the bound state were also carried out. For the SBD simulations of the bound state, the protein and the ligand were immersed in a sphere of water molecules of 25 Å radius (see the Supporting Information for the construction of this sphere). This sphere size is sufficient to provide at least one hydration shell around all of the regions of the complex. For the SBD runs, the complex was energy-minimized and then heated from 0 to 300 K using a step of 10 K (for 30 ps). Three independent heating runs were each followed by an equilibration divided into two periods of 300 and 150 ps, respectively (see the Supporting Information for details). After the equilibration, the SBD production runs were conducted with no restraints for 5 ns (total of 15 ns).

Prior to the unbinding simulations, we checked that the LD simulations of the bound state provided results in agreement with both the X-ray structure and the SBD simulations. The purpose in using LD simulations for the description of the bound state is two-fold: (i) to validate the force field and solvation parameters that are also used in the unbinding simulations and (ii) to obtain initial structures for the unbinding simulations. Implicit solvent simulations are required for the unbinding step, as has been recommended for unfolding or unbinding events.³³

2.2. Unbinding Simulations. The protocol used to pull the ligand out of the binding pocket and to obtain metastable states is shown in Figure 2. This protocol started with 18 structures (labeled “BS from LD”) extracted from the four LD simulations of the bound state. As the unbinding is a long-extended process,³⁴ a pulling force was added to the molecular potential energy function for each unbinding simulation. The time-dependent perturbation implemented in the Biased Molecular Dynamics (BMD)³³ method was used. It is noteworthy that, among the perturbation methods that have been recently compared, BMD is the method that causes the least perturbation in a system.³⁵ In this method, a quadratic time-dependent perturbation is introduced in the system only when the distance between the ligand atom O1 and the C_α atom of Glu5 decreases (d_{RC} ; see the Supporting Information for details). Otherwise, no external perturbation is applied, and the increased distance d_{RC} is taken as the new reaction coordinate for the next time step.

In the initial BMD simulations (step 1 in Figure 2), the constant of the pulling force (parameter α) was chosen at 300 pN/Å (4.32 kcal/mol/Å³), since this value enabled the unbinding of the ligand within a time limit of 5 ns. The unbinding process was

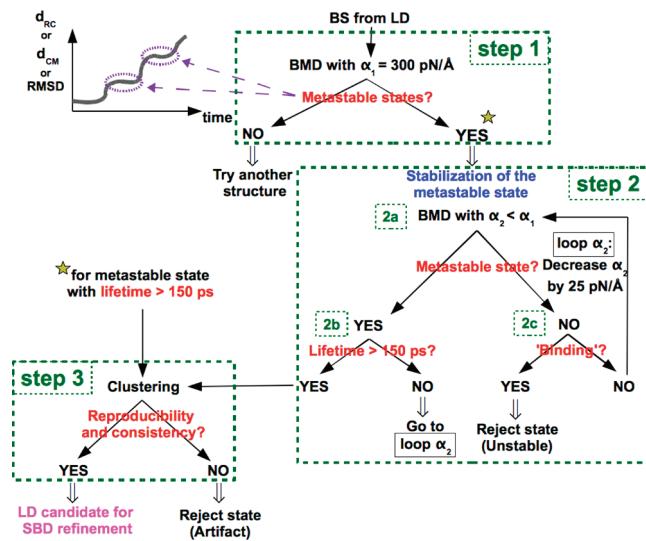


Figure 2. Protocol to find the LD candidates for the binding intermediate. Step 1: Unbinding simulations on each of the 18 structures separated by 1 ns and selected from the four LD simulations of the bound state. The star above “YES” indicates that the metastable states with a lifetime greater than 150 ps are also added to the set of simulations used in the clustering step (step 3). Step 2: In 2a, for each metastable state from step 1, BMD simulations are performed with α_2 lower than α_1 , to extend the lifetime of the metastable state found in step 1. In 2b and 2c, metastable states with lifetime > 150 ps are added to the pool used in step 3; otherwise, a lower α_2 is used for restarting the simulations. “Binding” checks whether the ligand returns to its position in the bound state (BS). Step 3: The metastable states with a lifetime greater than 150 ps are clustered into subsets that exhibit related structural features (see text). Reproducibility and consistency within a cluster are critical in finding the LD candidates for further refinement.

followed by monitoring the time evolutions of d_{RC} and of d_{CM} ; the latter represents the distance between the center of mass of the binding pocket and each of the four centers of mass of the four ligand moieties (core, iPe, Ph1, and Ph2). This procedure was also followed by Curcio et al. in their study of forced unbinding of fluorescein from anti-fluorescein antibody FITC-E2.²⁵ We also monitored the root-mean-square deviations (RMSDs) for both the protein and the ligand moieties (core, iPe, Ph1, and Ph2) with the C_α atoms aligned on those of the X-ray structure of the bound state. From the unbinding simulations (step 1 in Figure 2), many metastable states were found by relying on the time evolutions of d_{RC} , d_{CM} , and RMSDs; constant mean values for these distances and RMSDs for periods of hundreds of picoseconds constitute useful indicators.

In step 2 of the protocol, all of the metastable states found in step 1 were used to perform BMD simulations with smaller pulling forces (stage 2a in Figure 2). The states from step 1 with a lifetime greater than 150 ps were used both in stage 2a and in the clustering step of the protocol (step 3; *vide infra*). The purpose in using smaller values of α (in the range 50–250 pN/Å) in step 2 is to decrease the pulling force, thereby increasing the sampling of the potential energy surface for the ligand coordinates in the vicinity of the binding pocket. Hence, simulations with a longer lifetime were obtained that exhibited relatively stable values of d_{RC} , d_{CM} , and RMSDs for the protein and for the ligand. At stages 2b and 2c in Figure 2, the simulations with short lifetimes (<150 ps) were restarted by gradually decreasing the α_2 value (by steps of 25 pN/Å). The simulations with lifetimes greater than

Table 1. Average RMSDs (\AA) of All of the C_α Atoms of the Protein in the Complex and Those Calculated Only for the 40s, 50s, and 80s Loops (Which Comprise the Residues 39–45, 50–56, and 82–95, Respectively) on the Basis of the Alignment of All of the C_α Atoms

residue	$\langle LD \rangle$	$\langle SBD \rangle$
all	0.86 ± 0.08	1.02 ± 0.08
39–45 (40s)	0.91 ± 0.23	1.68 ± 0.04
50–56 (50s)	0.95 ± 0.17	1.14 ± 0.59
82–95 (80s)	1.38 ± 0.32	1.57 ± 0.01

150 ps were added to the pool of states analyzed in the step 3, and those for which the ligand returned to the binding pocket were discarded.

In step 3 of Figure 2, all of the metastable states with a lifetime greater than 150 ps, obtained from both the unbinding (step 1 in Figure 2) and the simulations performed to increase the sampling (step 2 in Figure 2), were clustered into subsets. The following four criteria were used to cluster and to select the LD simulations that could model the binding intermediate. First, we checked that the trajectories were stable for periods of at least hundreds of picoseconds. Within a subset of metastable states that may represent a binding intermediate, the values derived for d_{RC} , d_{CM} , and RMSDs of the protein and of the ligand were checked to ensure that they remained constant during the course of each simulation as well as consistent among all of the simulations. For the binding intermediate that is presented in this work, the d_{CM} and RMSDs are shown in Table 2 and discussed in section 3.2. The same structural criterion was followed by Li and Daggett in the search of an intermediate state along the unfolding pathway of barnase.³⁶ Second, we checked that the structure of the protein was not significantly altered, except for the 80s loop, as it is known to be highly flexible and, more importantly, surrounds and contributes to the binding pocket. The RMSD values for the protein are presented in section 3.2 and reveal an overall low distortion, except for the 80s loop. Third, we checked that the set of LD simulations sampled a range of conformations that were structurally related and, in particular, the positions of the ligand with respect to the protein were not spread sparsely but rather formed a tight cluster. As seen in the analysis of the position of the ligand in section 3.2, a narrow ensemble of orientation is obtained for the ligand. Central to this last criterion is the reproducibility and consistency that constitute useful indicators that the results have sufficiently converged so as to provide a meaningful average picture of the intermediate. Fourth, as we hypothesize that the intermediate state is an obligatory step from the freely diffusing molecules toward the bound state, we also checked that the intermediate was structurally close to the bound state in relative separation and relative orientation between the two molecular partners. In the intermediate that was identified, the ligand core moiety is lying at 4.6 Å above its position in the native complex (section 3.2).

2.3. Refinement Procedure. Only three sets of LD simulations, the LD candidates, passed the acceptance criteria and were selected for the refinement procedure that used independent SBD production runs with no biasing forces or restraints on the system (see the Supporting Information for details). The purpose in using an explicit solvent representation to further characterize the intermediate state is two-fold: (i) to further probe the stability of the binding intermediate located using LD simulations and (ii) to refine the results since an explicit solvent model

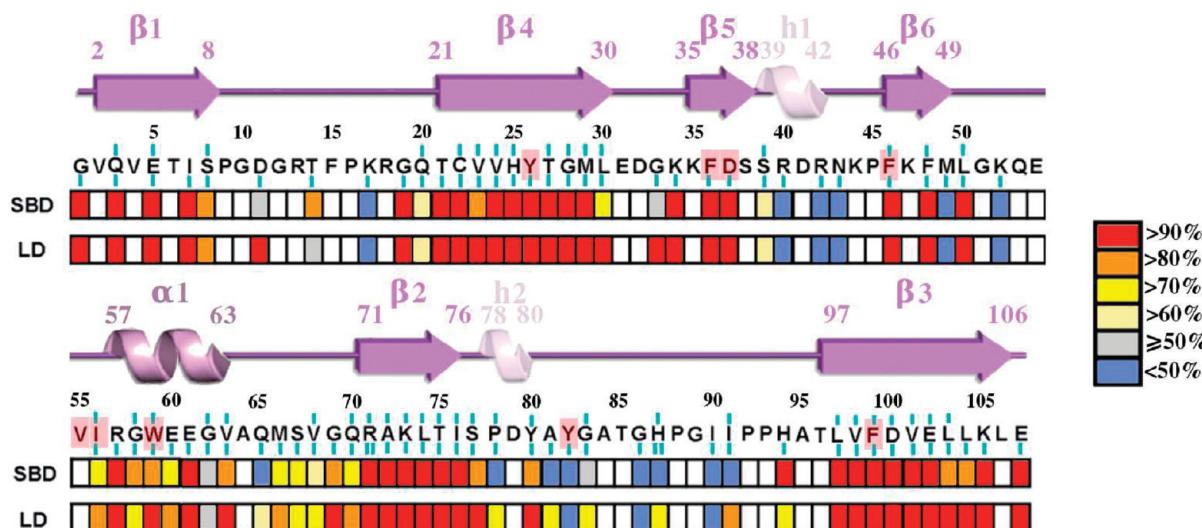


Figure 3. Contact frequencies for all of the backbone hydrogen bonds calculated as ensemble averages for the LD and SBD simulations. In the one-letter code representation of the amino acid sequence, up to three vertical bars are drawn above and below a few residues in the sequence and represent the number of native hydrogen bonds formed by their backbone atoms (through CO and/or NH). When the backbone atoms of a residue are involved in two or three hydrogen bonds, the contact frequency is calculated by averaging over these two or three contacts. For example, the color code adopted for Trp59 results from the average between the contact frequencies calculated for the two native hydrogen bonds Ile56-O \cdots Trp59-HN and Trp59-O \cdots Val63-HN. The one-letter codes of the residues that are in contact with the ligand in the crystal structure are shaded. The secondary structure elements are designated as in the work of Holt et al.⁶ This secondary structure assignment is consistent with a DSSP analysis⁵⁸ except for the regions 39–42 and 78–80, where single turn helices are predicted (helices h1 and h2, respectively).

is known to provide a more realistic description of a system. For the SBD simulations of each candidate for the binding intermediate, the water molecules were added to the selected structures in a sphere of 22 Å in diameter centered on the binding site (see the Supporting Information for the construction of this sphere).

The SBD refinement step enabled us to discard two of the three LD candidates for the binding intermediate. These two states corresponded to more distant positions of the ligand from the binding pocket. And, in contrast to the binding intermediate discussed in this work, the extent of consensus within these two sets of short simulations was rather weak. Despite extensive effort to complete and refine these two sets of metastable states, the overall results did not show sufficient consensus, reproducibility, and stability. This failure to achieve sufficient consistency, either within LD runs or SBD runs or between LD and SBD results, emphasizes the importance of the third criteria used for the clustering step to provide confidence in the intermediate model derived.

It should be mentioned that the analysis of the intermediate rests on the assumption that the results from the unbinding experience are also valid for the reverse process, which is expected to be the case under the chosen conditions, i.e., low forces applied to the ligand and implicit solvent representation.^{25,33} All of the data that are reported for the bound state and the binding intermediate correspond to ensemble averages, except where otherwise indicated.

3. RESULTS

3.1. Bound State. Prior to the unbinding simulations, we tested the ability of our LD/EEF1 protocol to yield stable structures for the bound complex. The stability of the structures generated from the LD approach was evaluated by analyzing (i) the root-mean-square deviations (RMSDs) of the C α atoms from

the experimental complex, (ii) the native hydrogen bonds, and (iii) the position of the ligand inside the binding pocket. Our goal in this validation step is two-fold: first, to check whether the structures generated by LD simulations are consistent with the crystal structure and, second, to compare the LD results with those of the SBD simulations that use an explicit solvent representation.

The average RMSDs of the C α atoms from the experimental complex are reported in Table 1 for the LD and SBD simulations. The ensemble averages are below ~1 Å and are consistent with those typically seen in molecular simulations of X-ray protein structures.³⁷ Though informative, the average RMSDs of all of the C α atoms provide only a global measure of the deviation from the crystal structure. Attention was also focused on the distortion of the three loops closest to the binding site, the 40s, 50s, and 80s loops (the former two loops comprise the residues 39–45 and 50–56, respectively). The RMSDs of the C α atoms of these three loops are also tabulated; their values do not exceed 1.7 Å.

A total of 59 backbone hydrogen bonds were extracted from the experimental structures based on the following contact criteria: a maximum hydrogen-acceptor distance of 2.4 Å and the minimum donor-hydrogen-acceptor angle set to 90°. Figure 3 illustrates the average contact frequencies for all of the backbone hydrogen bonds in the two simulation series. Overall, the two series of LD and SBD simulations give comparable results, especially in the regions of secondary structure. As seen in Figure 3, all of the native hydrogen bonds between the five β strands are well conserved in both series of simulations (contact frequencies are above 80%). These strands form an antiparallel β sheet wrapped around an α helix of seven residues, α_1 . In the two simulation series, the structure of α_1 is preserved owing to contact frequencies higher than 75% (between Arg57-Glu61, Trp59-Val63, Ile56-Trp59, and Ile56-Glu60). In contrast, in the loop regions, four hydrogen bonds between the following pairs of residues are not conserved: Met49–Lys52 in the 50s loop,

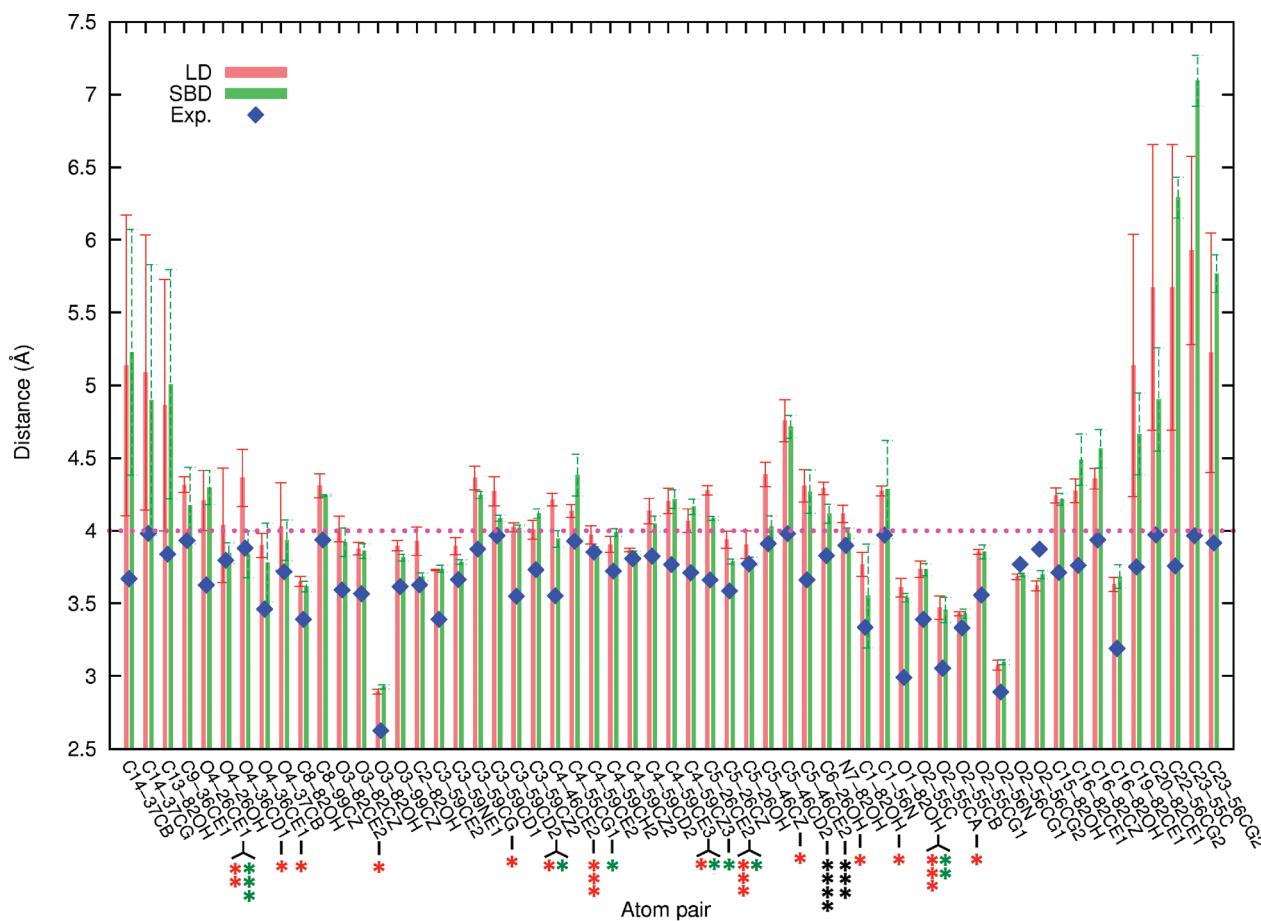


Figure 4. List of 57 separation distances between protein and ligand atoms averaged from the LD and SBD series. This list corresponds to all of the intermolecular distances found in the X-ray structure that are below 4 Å. Any average separation distance that differs from its experimental counterpart by a value found in the ranges]0.4; 0.5[,]0.5; 0.6[,]0.6; 0.7[, and]0.7; 0.8[are indicated by one, two, three, and four stars, respectively. Red and green stars refer to the LD and SBD results, respectively, and the black stars to both.

Tyr82–Gly86 and His87–Ile90 in the 80s loop, and Lys17–Gln20. In the SBD runs only, the hydrogen bond between the residue pair Gly62–Gln65 is weakened, whereas one of the two hydrogen bonds between His87 and Ile91 in the 80s loop is lost (frequencies of 44% and 1%, respectively). Finally, no new hydrogen bonds were formed between backbone atoms in SBD simulations, whereas only one is seen in the LD simulations between Asp11-O and Arg13-NH.

Hence, only four and five of the 59 native hydrogen bonds are not reproduced by LD and SBD simulations, respectively. Two and three of these hydrogen bonds that are disrupted in LD and SBD runs, respectively, are found in the 80s loop. The failure to reproduce these few hydrogen bonds that are assumed in this region may be related to crystal packing effects. For example, in the crystal structure of the protein in complex with FK506, packing effects are known to exist for the segment 82–90.³⁸ In a quantum chemical study of the crystal structure of FKBP12–8, Nakanishi et al.³⁹ have pointed out that ligand 8 experiences packing effects by interacting with adjacent proteins in the crystal, presumably with the same residue segment 82–90 as for FK506. The authors have also derived a RMSD from the crystal structure of 0.80 Å for ligand 8 after a minimization by the quantum mechanical method FMO. This large RMSD has been ascribed to crystal packing effects. In the following, the position of the ligand inside the binding pocket is analyzed for the LD and SBD simulations.

A list of 57 atomic separation distances between the protein and the ligand were monitored to assess the stability of the ligand in the binding site. This list was obtained by considering all of the protein and ligand pairs of heavy atoms within 4 Å of each other in the X-ray structure. In the crystal structure, the diketopiperazine acid core region makes 45 contacts with the protein; the iPe group, three; and the propyl-Ph1 group, nine. The comparison of the 57 average separation distances calculated for the LD and the SBD series (Figure 4) reveals very similar results, especially for the 45 contacts made by the core region. When the separation distances for the ligand core region are compared in the two simulation series, the differences in distances are below ~0.25 Å (except for the distances O4···Phe36-CD1 and C5···Phe46-CZ that are longer in LD runs than in SBD runs by ~0.5 Å). In contrast to the core region of the ligand, large differences between the LD and the SBD results are observed for the propyl-Ph1 region (on the right of Figure 4).

In order to compare the experimental and the calculated separation distances, the magnitude of the errors that are expected in the crystal structure should be first tentatively estimated. On average, the uncertainty on the position of an atom in a X-ray structure is roughly one-fifth to one-tenth of the resolution.⁴⁰ Under this assumption, an estimated error in the range 0.4–0.8 Å can be anticipated for the separation distances. For the core region of the ligand, 16 and 9 separation distances in the LD and SBD

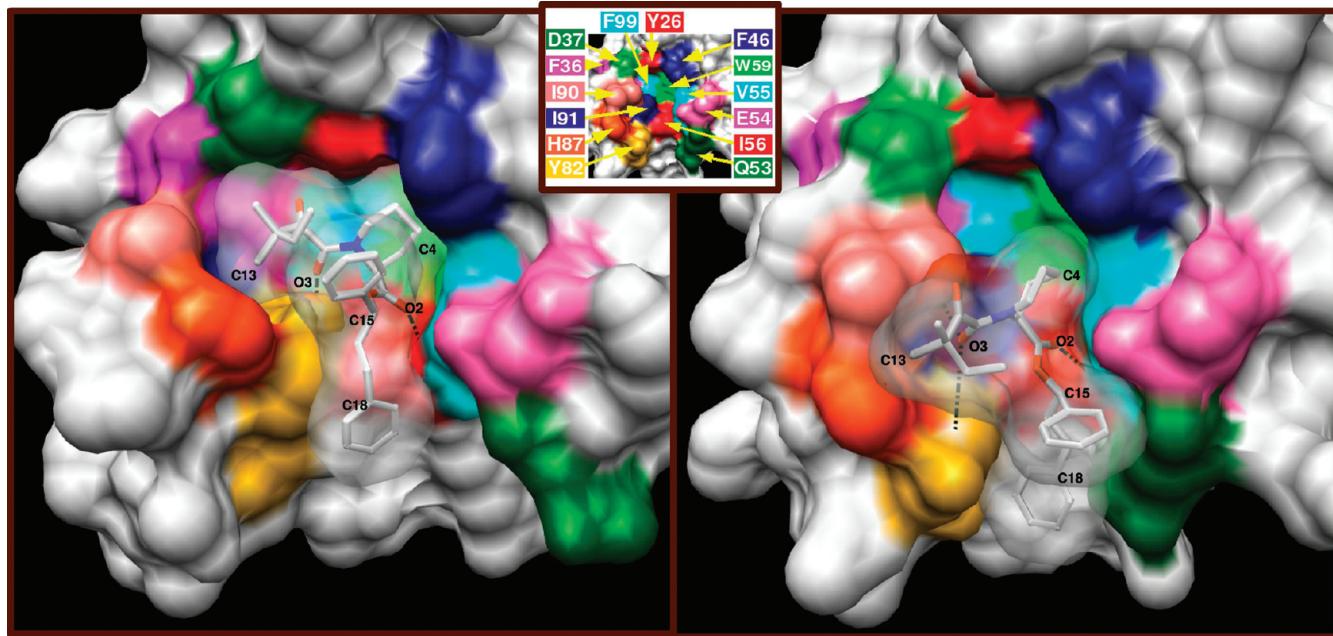


Figure 5. Comparative analysis of the position of the ligand in the binding pocket of FKBP12 in the crystal (left) and in the binding intermediate IS (right). The protein residues are color-coded according to the inset at the top of the figure. The two contacts Ile56-NH···O2 and Tyr82-OH···O3 are represented by dashed lines, along with the contact Ile90-CG2···O3 that is shown only for IS.

simulations, respectively, are longer than their experimental counterparts by more than 0.4 Å; the corresponding 18 distinct distances are highlighted on the x axis in Figure 4. Most of these 18 distances differ from their respective experimental values by, at most, 0.5 Å (marked with one star on the x axis). For the noncore regions iPe and propyl-Ph1, large shifts from the experimental separation distances are observed; these deviations may result from the errors in the experimental positions of these exposed noncore regions that experience packing effects (*vide supra*).^{38,39} Finally, very few new contacts, with averages below 4 Å, are formed between the protein and the ligand: in the SBD simulations, each of the pairs C4···Phe46-CZ and O2···Ile56-CB have a distance of ~3.9 Å; in the LD simulations, the distance C25···Glu54-O is ~3.6 Å.

The results in terms of the RMSDs, of the native hydrogen bonds, and of the native intermolecular contacts for the LD simulations are in good agreement with the experimental data. This provides confidence in the simulation methodology that uses an implicit description of the solvent and an analytical model for the solvation free energy. Previous comparison studies between implicit and explicit solvent representations have also led to the conclusion that the former can replace the latter for many purposes.⁴¹ We have thus applied our methodology to locate a binding intermediate by pulling the ligand out of the binding pocket using a minimal force.

3.2. Binding Intermediate. Only one set of four independent LD simulations satisfied the selection criteria discussed in the section 2. These runs exhibited energetic stability and provided consistent results both for the structure of the protein and for the position of the ligand. As the goal of consistency between independent LD runs was achieved, these runs provided a reasonable guess for the subsequent refinement protocol using SBD simulations. Seven independent SBD trajectories were then obtained for a total of about 12 ns of sampling. A snapshot of the intermediate IS is

shown in Figure 5 along with the crystal structure for purposes of comparison.

In the following comparative analysis between LD and SBD results, we first check the structural integrity of the protein, the overall consistent position of the ligand, and the relative proximity of the ligand to its native binding position. Importantly, this comparative analysis will also reveal a structural consistency between these two approaches, which gives confidence in the prediction of the binding intermediate. Then, in the remainder of this section, we will discuss more specifically the SBD results obtained for the 80s loop (distortion, mobility, and hydrogen bond contact frequencies) and for the protein–ligand contacts.

Table 2 reports the RMSDs of the C_{α} atoms from the experimental complex. The ensemble average RMSDs calculated for the nonloop region of the protein are below ~1.1 Å for both LD and SBD runs, values that are similar to those of the protein in the bound state (Table 1). However, the RMSDs of the 80s loop indicate an overall distortion in the range ~3.2–3.4 Å for the two approaches. This distortion may be expected on the basis of previous works on FKBP12,^{19,20,42} as discussed hereafter. To assess its position with respect to the protein, the ligand was divided into fragments, namely, isopentyl (iPe), a diketo-pipercolinic acid core, and two phenyl rings (Ph1 and Ph2), as shown in Figure 1b, and the separations between the center of mass of each of these fragments and that of the protein binding pocket were checked to ensure they remained constant during the course of each simulation as well as consistent among the two sets of simulations (LD and SBD). Table 2 reports the average distances between the center of mass of the diketo-pipercolinic acid core and that of the binding pocket. The ensemble average position of the core region (d_{CM}) measured from the center of the binding pocket is 7.19 and 6.37 Å in the LD and the SBD simulations, respectively. In particular, for the individual SBD simulations, these values of d_{CM} reveal a stable position of the ligand core in the individual simulation (fluctuations of ~0.3–0.4 Å) and a

Table 2. Average Distances (\AA) between the Center of Mass of the Diketo-Pipecolinic Acid Core and That of the Binding Pocket (d_{CM}) and RMSDs (\AA) of the Four Ligand Moieties and of the C_α Atoms from the Respective X-Ray Structural Positions^a

run	pocket \cdots core d_{CM}	RMSD of ligand				RMSD of protein			
		iPe	core	Ph1	Ph2	nonloop ^b	82–95	82–87/94–95	only 88–93
LD1	6.58 \pm 0.32	8.79 \pm 0.59	6.19 \pm 0.42	4.45 \pm 0.37	7.68 \pm 0.62	1.03 \pm 0.07	3.17 \pm 0.13	1.66 \pm 0.17	4.45 \pm 0.20
LD2	7.09 \pm 0.37	9.29 \pm 0.58	6.69 \pm 0.40	4.38 \pm 0.48	8.25 \pm 0.56	0.90 \pm 0.06	3.27 \pm 0.16	1.39 \pm 0.11	4.73 \pm 0.25
LD3	7.43 \pm 0.31	8.51 \pm 0.50	6.68 \pm 0.31	3.83 \pm 0.46	8.20 \pm 0.54	1.26 \pm 0.10	3.49 \pm 0.22	3.09 \pm 0.25	3.95 \pm 0.24
LD4	7.67 \pm 0.41	10.31 \pm 0.67	7.23 \pm 0.44	4.69 \pm 0.47	8.82 \pm 0.66	1.04 \pm 0.05	3.65 \pm -0.17	1.73 \pm -0.17	5.21 \pm -0.25
$\langle \text{LD} \rangle$	7.19 \pm 0.47	9.22 \pm 0.79	6.70 \pm 0.42	4.34 \pm 0.36	8.24 \pm 0.47	1.06 \pm 0.15	3.40 \pm 0.22	1.97 \pm 0.76	4.58 \pm 0.53
SBD1	6.41 \pm 0.39	8.44 \pm 0.76	6.03 \pm 0.49	4.14 \pm 0.61	7.10 \pm 0.60	0.76 \pm 0.03	3.15 \pm 0.23	1.75 \pm 0.30	4.34 \pm 0.40
SBD2	6.48 \pm 0.45	8.78 \pm 0.68	6.14 \pm 0.53	4.22 \pm 0.61	7.07 \pm 0.70	0.79 \pm 0.11	3.33 \pm 0.14	1.75 \pm 0.30	4.65 \pm 0.24
SBD3	6.36 \pm 0.31	8.50 \pm 0.53	5.93 \pm 0.36	4.09 \pm 0.43	6.95 \pm 0.51	0.83 \pm 0.13	3.27 \pm 0.19	1.74 \pm 0.31	4.56 \pm 0.31
SBD4	6.42 \pm 0.34	8.46 \pm 0.48	6.05 \pm 0.39	4.20 \pm 0.48	7.39 \pm 0.53	0.84 \pm 0.10	3.24 \pm 0.14	1.48 \pm 0.16	4.64 \pm 0.22
SBD5	6.44 \pm 0.43	8.75 \pm 0.63	6.09 \pm 0.50	4.42 \pm 0.53	7.04 \pm 0.55	0.75 \pm 0.09	3.29 \pm 0.18	1.61 \pm 0.22	4.66 \pm 0.29
SBD6	6.36 \pm 0.34	8.50 \pm 0.57	5.95 \pm 0.41	4.22 \pm 0.47	7.14 \pm 0.60	0.79 \pm 0.12	3.29 \pm 0.13	1.47 \pm 0.16	4.73 \pm 0.20
SBD7	6.13 \pm 0.35	8.27 \pm 0.59	5.66 \pm 0.40	3.81 \pm 0.64	6.70 \pm 0.69	0.83 \pm 0.11	3.07 \pm 0.23	1.75 \pm 0.38	4.20 \pm 0.39
$\langle \text{SBD} \rangle$	6.37 \pm 0.11	8.53 \pm 0.18	5.98 \pm 0.16	4.16 \pm 0.18	7.06 \pm 0.21	0.80 \pm 0.04	3.23 \pm 0.09	1.65 \pm 0.13	4.54 \pm 0.20

^a For all of the RMSD calculations, the alignment was based upon the protein C_α atoms. ^b All C_α atoms but those of segment 82–95.

consistent position throughout all of the simulations (standard deviation of 0.11 \AA). The RMSDs of the four ligand moieties from their respective X-ray structural positions after alignment of the protein C_α atoms are also tabulated. These RMSDs are large since they measure the shift in the position of the ligand moieties in IS relative to their positions in the native complex. Most important is the fact that these values are relatively close throughout each set of runs, and taken together, they indicate an overall consistent position of the ligand. For the SBD results, the low standard deviations of ~ 0.2 \AA indicate strikingly similar positions of the ligand moieties in all of the runs. It is noteworthy that the core region of the ligand is 4.64 ± 0.11 \AA away from its position in the bound complex (difference between average distances from the center of the binding pocket of 6.37 \AA in IS and 1.73 \AA in the bound state), which is a distance similar to the ~ 4.5 \AA separating the two protein surfaces in the encounter complex of the barnase–barstar system.¹¹

From Figure 5, the 80s loop clearly has a configuration in IS that differs from that in the native complex. In Table 2, RMSD calculations performed for the protein and the 80s loop are also reported. Their analysis reveals that only a few residues contribute significantly to this distortion. Indeed, the backbone of segment 88–93 exhibits large deviations of ~ 4.5 \AA for both LD and SBD results, whereas the backbone structure of the remaining segments 82–87 and 94–95 of the 80s loop is only moderately distorted (RMSD < 2 \AA). Remarkably, in the short sequence 88–93 at the tip of the 80s loop, the largest structural deviation stems from Gly89 (5.5 \AA). This glycine residue is located in this short sequence at the tip of the 80s loop that also contains the largest proportion of proline residues in FKBP12. In other protein–ligand systems, the location of glycine residues near the active site of proteins has been identified as crucial for ligand recognition by allowing local fluctuations or changes in loop structure that contribute to the entering of the ligand.^{43–45} In this intermediate IS, which is hypothesized to represent an early stage of binding, the distortion facilitated by a glycine residue can thus appear as an effective strategy to favor an interaction between the neighboring residue Ile90 and the diketone group of the ligand, as we will see below in the analysis of the protein–ligand contacts. Previous studies have shown the distortion of the backbone structure

Table 3. Average Separation Distances between the Centers of Mass of the Side Chains of Tyr82, His87, Ile90, and Ile91, along with Those Measured in the Crystal Structure (All Distances Are in \AA ngströms)

distance	reference	average	std dev.
His87 \cdots Ile90	5.17	6.06	0.08
His87 \cdots Ile91	6.22	6.87	0.29
Ile90 \cdots Ile91	5.63	5.73	0.02
Tyr82 \cdots His87	5.50	5.44	0.15
Tyr82 \cdots Ile90	8.43	8.63	0.30
Tyr82 \cdots Ile91	4.97	5.31	0.27

around residue 89 or the importance of this region of the 80s loop for the recognition. For example, in a comparison between unliganded and various bound forms of FKBP12, a systematic difference in the backbone conformations of the region 87–89 of the protein has been pointed out, with RMSDs as large as 2 \AA .⁴⁶ In the context of the formation of the ternary complex FKBP12–FK506–calcineurin, the importance of region 87–91 of the loop in the recognition of calcineurin has also been demonstrated.⁴⁷

Although the 80s loop is distorted in the intermediate IS, the hydrogen bonds found in this loop are well preserved. Six of the 10 native hydrogen bonds found in the 80s loop and in the 3_{10} helix 78–81 have average contact frequencies greater than 55%, including the three interactions between His87, Ile90, and Ile91 at the tip of the 80s loop. For these six hydrogen bonds, the corresponding distances averaged over all of the simulations lie below 2.41 \AA . Contact frequencies that are lower than average are obtained for the four remaining hydrogen bonds of the loop, which each involves a glycine residue: Gly83-O \cdots His94-NH (41%), Ser77-NH \cdots Gly1-O (27%), and values below 3% for both Pro78-O \cdots Gly83-NH and Tyr82-O \cdots Gly86-NH.

By comparing unliganded and FK506-bound conformations of FKBP12, Ivery and Weiler have inferred that the 80s loop must undergo a displacement as a unit in the binding process of the ligand FK506.²⁰ The same conclusion was drawn by Wilson et al. on the basis of a comparative analysis of X-ray structures of unbound, FK506-, and rapamycin-bound forms of FKBP12.⁴²

Table 4. Time Average and Ensemble Average Distances (\AA) for the Most Persistent Contacts between the Protein and the Ligand^a

		atom pair ^b	SBD1	SBD2	SBD3	SBD4	SBD5	SBD6	SBD7	fluctuation range	$\langle \text{SBD} \rangle$	std dev.
protein	ligand											
Ile90	CG2	C8	3.88	3.90	3.86	3.94	3.94	3.93	3.97	0.29–0.46	3.92	0.04
		O3	3.82	3.77	3.82	3.83	3.81	3.87	3.88	0.38–0.59	3.83	0.04
		O4	3.86	4.06	3.79	4.02	4.10	3.95	3.99	0.37–0.53	3.97	0.11
Ile91	CD	CD	4.76	4.05	3.94	4.03	4.17	4.41	4.41	0.49–0.72	4.25	0.29
		O4	4.24	4.25	4.06	4.19	4.26	4.02	4.13	0.38–0.50	4.16	0.10
His87	CD2	C3	4.33	4.82	4.91	4.16	5.53	5.40	3.83	0.59–1.03	4.71	0.63
		O3	4.33	4.82	4.91	4.16	5.53	5.40	3.83	0.59–1.03	4.71	0.63
		C13	4.13	5.32	5.75	3.88	5.82	5.67	6.24	0.79–1.61	5.26	0.90
Tyr82	CG1	CG	4.27	5.48	6.03	4.06	5.91	5.80	6.29	0.78–1.53	5.41	0.88
		NE2	4.35	5.31	5.69	4.07	5.61	5.60	6.20	0.75–1.59	5.26	0.77
		O1	5.74	5.17	4.17	4.36	5.21	4.13	6.19	0.96–2.00	5.00	0.80
Ile56	OH	C8	5.70	5.28	4.37	4.62	5.36	4.43	6.78	0.90–2.13	5.22	0.86
		O3	4.94	4.58	3.58	3.94	4.62	3.66	6.05	0.99–2.21	4.48	0.87
		O1	5.74	5.17	4.17	4.36	5.21	4.13	6.19	0.96–2.00	5.00	0.80
		C16	5.31	4.69	4.08	4.18	4.72	4.06	5.75	0.73–1.69	4.69	0.65
		CE2	4.46	4.21	4.35	3.96	4.35	3.96	5.31	0.54–1.67	4.37	0.46
		CZ	5.11	4.80	4.15	4.37	4.90	4.22	5.99	0.63–1.78	4.79	0.64
Val55	O	O	4.72	5.00	4.85	5.41	5.40	5.42	4.05	0.71–1.68	4.98	0.50
		O2	3.45	3.52	3.45	3.48	3.52	3.40	3.48	0.25–0.34	3.47	0.04
		CA	4.07	4.16	4.05	4.08	4.13	4.01	4.06	0.26–0.37	4.08	0.05
		N	3.34	3.29	3.31	3.30	3.27	3.28	3.32	0.24–0.34	3.30	0.03
Glu54	O	$\angle \text{N}-\text{H}$	88.6	104.1	90.0	96.4	105.2	90.8	86.5	13.6–17.6	94.5	7.6
		CA	3.53	3.44	3.46	3.48	3.45	3.48	3.42	0.22–0.40	3.46	0.04
		C	3.43	3.52	3.40	3.47	3.52	3.40	3.35	0.25–0.37	3.44	0.07
Gln53	C	O	4.08	4.37	4.08	4.24	4.39	4.09	3.99	0.37–0.45	4.17	0.16
		C4	4.29	4.88	4.13	4.49	4.96	4.24	4.08	0.52–0.79	4.44	0.35
		C1	3.97	4.56	3.91	4.18	4.73	3.92	3.83	0.36–0.48	4.16	0.35
		O1	4.04	4.61	3.99	4.21	4.84	3.99	3.97	0.38–0.61	4.24	0.35
		O2	3.74	4.24	3.74	3.96	4.36	3.74	3.62	0.37–0.52	3.92	0.28
		C15	3.84	4.24	3.78	3.91	4.47	3.79	3.70	0.34–0.69	3.96	0.28
		C24	3.94	4.07	3.97	3.98	4.25	3.95	4.14	0.36–0.47	4.04	0.12
		C25	3.75	3.49	4.46	4.66	3.46	3.50	4.21	0.33–0.87	3.93	0.50
		C26	4.54	4.05	5.23	5.30	3.94	4.37	5.18	0.43–0.80	4.66	0.57
		C29	4.80	5.03	4.22	3.92	5.29	5.10	4.99	0.42–0.88	4.76	0.50
Ile91	O	C25	4.26	4.08	5.17	5.44	4.07	4.07	4.77	0.35–0.97	4.55	0.58
		C25	3.91	4.24	4.81	5.18	4.17	3.74	4.19	0.51–1.10	4.32	0.51
		C26	3.82	4.18	5.19	5.42	3.92	3.83	4.65	0.57–1.01	4.43	0.67
		C29	5.67	6.14	4.36	4.26	6.16	5.88	4.87	0.45–1.08	5.33	0.82

^a A persistent contact was considered whenever the corresponding contact frequency was higher than 50% in at least one simulation. The reported fluctuation range (\AA) represents the minimum and the maximum of the fluctuations in the simulation set. One angle value (deg) is reported for the native hydrogen bond between Ile56-NH and ligand atom O2. ^b Except for Ile56 where the angle for the native hydrogen bond it formed with O2 is also reported.

Hence, the conclusions of these studies prompted us to examine whether, in the case of ligand **8**, the 80s loop underwent a displacement as a unit between the native complex and IS. Prior to this analysis, one key observation should be reported for the native complex FKBP12–**8**: the bulky side chains of the 80s loop, namely, Tyr82, His87, Ile90, and Ile91, are clustered together, forming a cylinder-like hydrophobic core whose upper edge defines a small cavity to where atom C13 of the ligand is pointing; this cavity has a horseshoe shape (on the left of Figure 5). We have thus monitored the six center of mass separations between these four side chains in IS for comparison with their values in the crystal structure. The overall averages are

reported in Table 3. All of the distances measured in IS are close to their counterparts in the native complex, with deviations less than $\sim 0.3 \text{ \AA}$, except for the two distances His87···Ile90 and His87···Ile91 that are slightly larger in IS by 0.9 and 0.7 \AA , respectively. Taken together, the analysis of the data for the 80s loop, i.e., RMSDs and separation distances between its bulky side chains, reveals a distortion of the loop backbone conformation near the tip region 88–93 (most severe for the residue 89) and a concerted displacement of all of the bulky side chains of the loop between the bound state and IS.

In Table 4, the distances for the most persistent contacts between the protein and the ligand calculated for every simulation are

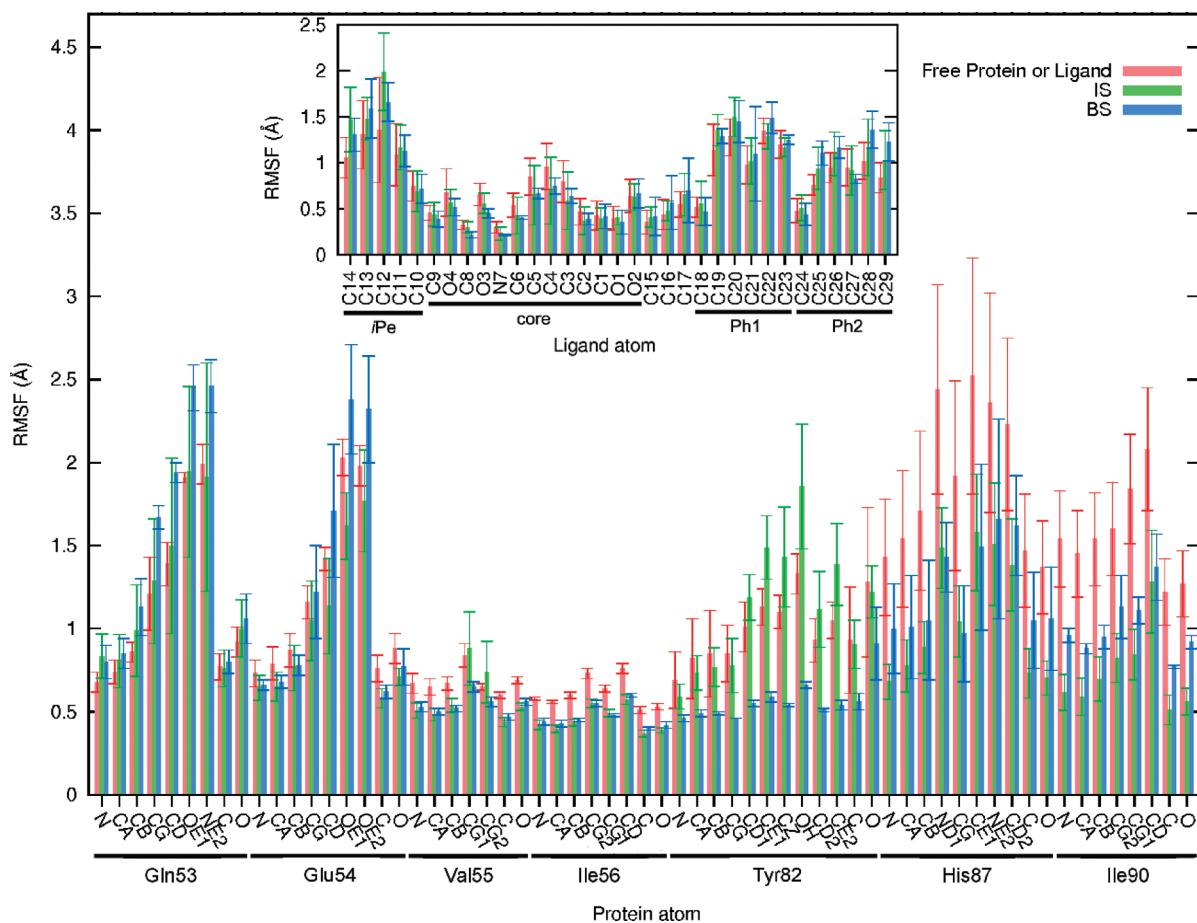


Figure 6. RMSFs calculated for the residues 53–56, 82, 87, and 90 that are in contact with the ligand in IS. The corresponding values for the bound state (BS) are also reported for comparison. The RMSFs for the ligand atoms in IS and the bound state are shown in the inset. For purposes of comparison, the RMSFs of these seven residues in the free protein and those of the free ligand (inset) were also calculated by averaging over four and five 2-ns SBD simulations, respectively.

reported, along with their ensemble average values. It is striking that only a few average contact distances lie below 4 Å: two clusters of interactions are stable, namely, between Ile90-CG2 and the ligand atoms C8, O3, and O4 (average distances are in the narrow range 3.83–3.97 Å) and between the atoms Ile56-(N,CG2) and Val55-(C,CA) of the protein and O2 in the ligand (distance range of 3.30–3.47 Å). The corresponding standard deviations are below ~0.1 Å, indicating overall consistent distances in the simulations (the largest standard deviation is calculated for the contact Ile90-CG2···O4, whose distance lies slightly above 4 Å in three simulations). The backbone oxygen of Glu54 also makes three contacts with distances of ~3.9 Å; however, they are associated with significantly higher standard deviations of 0.28 and 0.50 Å. The native hydrogen bond between O2 and Ile56-NH should more properly be viewed in IS as a van der Waals contact since the corresponding distance and angle are 3.30 ± 0.03 Å and $94 \pm 8^\circ$, respectively. For the other native hydrogen bond between O3 and Tyr82-OH, the average distance is 4.48 ± 0.87 Å. Hence, for the latter contact, a large variability in the seven average distances is observed and also high fluctuations in each simulation; this results from the high mobility of the side chain atoms of Tyr82 (as seen below in the fluctuation analysis) since atom O3 retains a very stable position as stated previously (Table 2). All other average distances in Table 4 lie mainly between 4 and 5 Å with standard deviations less than 1 Å.

In Figure 6, we have reported the RMS fluctuations of the residues that interact with the ligand (except for Ile91, which hardly interacts with the ligand since only one of its contact frequencies is higher than 50%), all of which are listed in the first column of Table 4. For purposes of comparison, the RMSFs of the unbound and bound forms of the protein are also shown, as well as those derived for the ligand itself in an explicit solvent model, in the bound complex, and in IS. The RMSFs for residues 53–56 are very similar in the three states of the protein (except for the amide and the carboxylate groups of Gln53 and Glu54, respectively). However, for the residues Tyr82, His87, and Ile90 in the 80s loop, the RMSFs vary between the three states of the protein. In going from the apoprotein to IS, the fluctuations of the side chain of Tyr82 increase by 20 to 30% (except for CB and CG), whereas those of both backbone and side-chain atoms of His87 and Ile90 are damped by 40 to 50%. Between IS and the bound complex, reversed trends are again observed between Tyr82 and the two residues His87 and Ile90: for the former residue, the RMSFs sharply decrease toward the lower values 0.5–0.7 Å, while no change or a small increase of less than 0.35 Å is observed for the latter residues. The RMSFs derived for the ligand in the three forms are surprisingly similar (with only moderately larger values for the atoms C12 and C14 in IS).

Overall, the convergence achieved in all of the simulations (i) for the ligand core position; (ii) for a few intermolecular distances

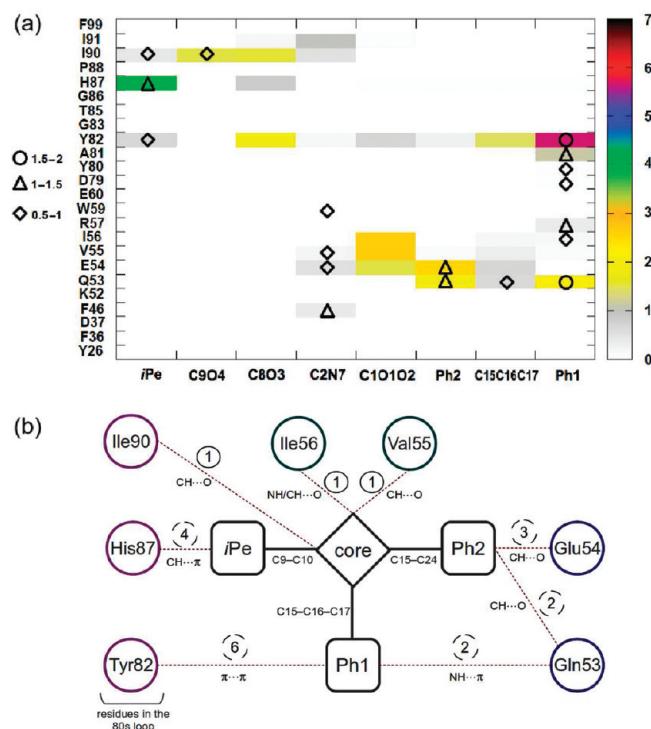


Figure 7. (a) Average number of contacts between each pair of interacting groups of atoms that consists of a ligand moiety and a protein residue (color scheme on the right) and the average number of water molecules in contact with both a ligand moiety and a protein residue (diamond, triangle, and circle on the left are for a number of bridging water molecules in the intervals $]0.5;1]$, $]1;1.5]$, and $]1.5;2]$, respectively). The core region of the ligand is further subdivided into two keto groups ($C9O4$ and $C8O3$), the pipecolinyl ring ($C2N7$), and the ester bond ($C1O1O2$); $C15C16C17$ refers to the propyl group. All types of contacts (cutoff of 4 \AA) were first obtained for each SBD simulation; the average over the seven independent simulations was then derived. (b) Schematic representation of the main protein–ligand contacts. The numbers of permanent and transient contacts are circled in solid and dashed lines, respectively; the type of noncovalent interaction is indicated in each case. The bonds that link the three mobile moieties to the core are also indicated. Positions and distances for all moieties are not realistically represented.

that involve primarily the atom O2 and the bond C8O3 of the ligand and their binding partners Val55, Ile56, and Ile90; and (iii) for the RMSDs of the protein testifies to a convergent picture of the binding intermediate IS. However, a question arises as to how the ligand retains a stable position relative to the protein given the very few persistent contacts between the two molecules. The other intermolecular contacts are indeed showing large distance variations from simulation to simulation and are short-lived. To quantify the amounts of instantaneous short-lived contacts between the two molecules, we have calculated the total number of contacts between each pair of interacting groups of atoms that consists of a ligand moiety and a protein residue. A 4 \AA distance cutoff was used for counting the interatomic contacts between each group of atoms (Figure 7a). In this figure, we can see the two previously reported clusters of contacts between Val55 or Ile56 and the ester bond C1O1O2 (where O2 is engaged in contacts), and between Ile90 and the two bonds C8O3 and C9O4 (of which O4 is responsible for the contacts with the atoms Ile90-(CG2,CD), while C9 is involved in only transient contacts with Ile90-CG2 as all its contact frequencies are well below 50%). About two

contacts are found between C8O3 and Tyr82 because of transient contacts between O3 and the three atoms Tyr82-(OH,CZ,CE2); in SBD3 and SBD6, however, a lasting van der Waals interaction Tyr82-OH···O3 is found (Table 4). It is striking that other groups of atoms also participate in intermolecular contacts, as indicated by the following averages: the iPe moiety makes four contacts with His87; the Ph1 moiety makes two and almost six contacts with the residues Gln53 and Tyr82, respectively; and the Ph2 ring makes about two and three contacts with Gln53 and Glu54, respectively (Figure 7b).

Figure 7a also gives the average number of water molecules that are in contact with both a ligand moiety and a protein residue. Such bridging water molecules are mainly seen in contact with the ligand moieties Ph2, iPe, Ph1, and the pipecolinyl ring ($C2N7$). These contacts are in the range 1–1.5 for iPe and His87, as well as for Ph2 and both Gln53 and Glu54. A slightly larger number of such contacts involve Ph1 (about two with both Gln53 and Tyr82) and the pipecolinyl ring (about one with each of the residues Phe46, Glu54, Val55, and Trp59, respectively).

A detailed analysis of all of the short-lived contacts between the protein and the ligand is useful in identifying the atomic partners and the nature of these contacts. As seen in Table 4, all of the average separation distances between the three mobile moieties of the ligand Ph1, Ph2, and iPe and the residues Tyr82, His87, Ile90, Gln53, and Glu54 are above 4 \AA (except for $C25\cdots\text{Glu54-O}$). The large averages and standard deviations seen for these separation distances result from the large RMSFs obtained not only for the side chains of residues 82, 87, 90, and 53 but also for the ligand moieties Ph1, Ph2, and iPe (Figure 6). The high fluctuations of iPe, Ph1, and Ph2 are due to low dihedral energy barriers associated with the bonds C9–C10 (1.4 kcal/mol), C17–C18 (0.5 kcal/mol), and C15–C24 (1.6 kcal/mol), respectively. The facile rotation of the Ph2 group around the bond C15–C24 enables interchangeable interactions of the segments C25–C26 and C29–C28 with the backbone oxygens of Glu54 and Gln53: these transient interactions are responsible for an average of about two $\text{CH}\cdots\text{O}$ contacts between Ph2 and each of these two residues (Figure 8a). For the iPe moiety, a pseudo-3-fold symmetry can be considered around the bond C9–C10, and similarly, the rotations about this bond exchange the atoms C11, C13, and C14 in the interaction with the imidazole ring of His87 (Figure 8b). Each of the atoms C11, C13, and C14 of the iPe moiety is making an average of four $\text{CH}\cdots\pi$ contacts with the atoms of the imidazole ring.

In contrast to the iPe or Ph2 moiety, where there are only a few easily identifiable contacts with the protein, a higher number of transient contacts is found between the Ph1 moiety and both the amide group of Gln53 and the aromatic ring of Tyr82. The distances between Ph1 and Tyr82 ring centroids as well as between the Ph1 ring centroid and each of the two atoms NE2 and OE1 of Gln53 are reported in Figure 8c. Because of the high mobility of the interacting partners, very short-lived contacts are seen during the course of the simulation (in particular, the propyl “arm” C15–C16–C17 of the ligand is very flexible). With a view to quantitatively analyzing the preferred orientations of the binding partners Ph1 and the side chains of Tyr82 and Gln53, for each simulation, the three instantaneous separation distances, such as those presented in Figure 8c, were collected into equally spaced bins of 0.25 \AA width. This yields a histogram with distance bins each counting the number of occurrences of certain distances within this short distance interval of 0.25 \AA . The same procedure is followed for every simulation, and the resulting averaged

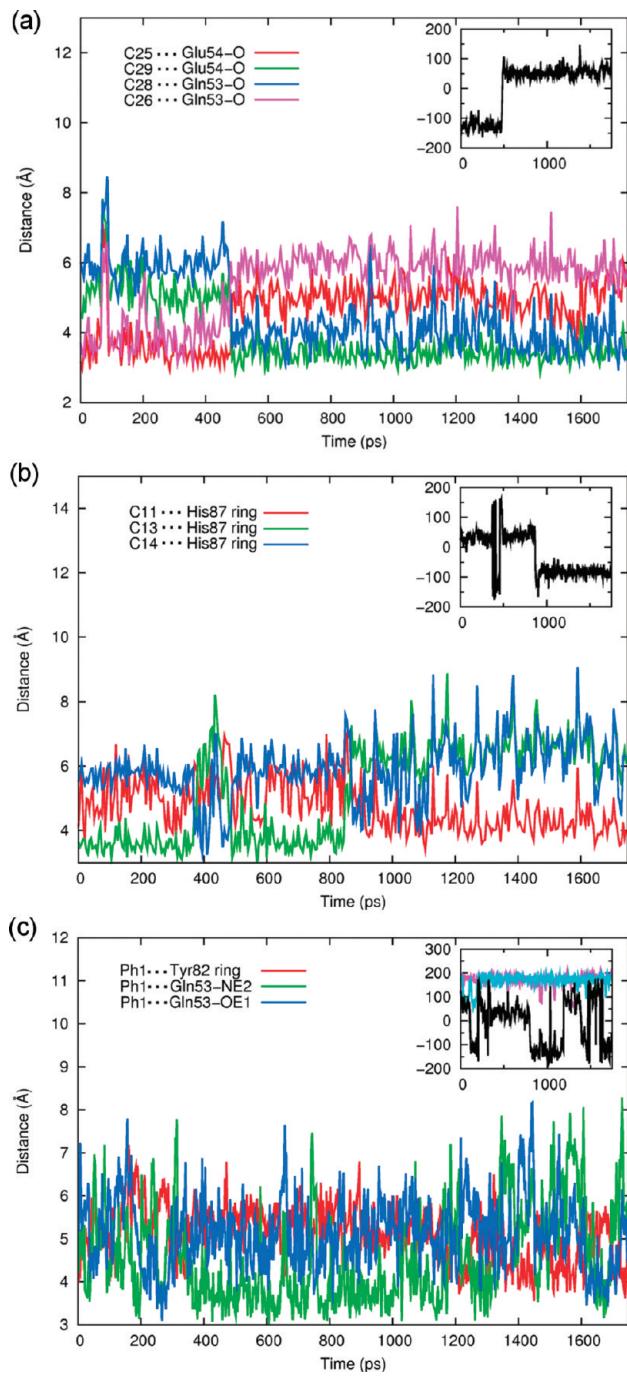


Figure 8. Distances between the noncore regions of the ligand and a few protein atoms as a function of time in the SBD simulations. (a) Distances between the edge atoms C25, C26, C28, and C29 of the ring Ph2 and the backbone oxygen atoms of Gln53 and Glu54. The inset represents the evolution of the dihedral angle around C15–C24. (b) Distances between the atoms of iPe and the center of the imidazole ring of His87. The inset represents the evolution of the dihedral angle around C9–C10. (c) Distances between the center of the Ph1 ring and that of the side chain ring of Tyr82, and between the Ph1 ring centroid and the two atoms NE2 and OE1 of the Gln53 side chain. The inset represents the evolution of the dihedral angles around C15–C16 (purple), C16–C17 (cyan), and C17–C18 (black).

histogram is shown in Figure 9. Between Ph1 and Tyr82 ring centroids, the most populated bin distances are found in the range

4–5.5 Å, a distance range compatible with π – π interactions and consistent with the results obtained for two interacting benzene rings (with benzene ring centroids in the range 3.75–4.02 Å and at 4.96 Å⁴⁸). Regarding the distances between Gln53-NE2 or -OE1 and the Ph1 centroid, the interactions with NE2 are predominant ($\text{NH}\cdots\pi$ type), as reflected by the higher number of contacts than for OE1 at short distance ranges (3–4 Å).

To summarize, one striking feature of the protein in IS is the displacement as a rigid body of its 80s loop, favoring various contacts with the ligand. The segment 88–93 exhibits great flexibility, as previously observed for other FKBP12 complexes.⁴⁶ The ligand core region has a well-defined position with respect to the protein, which enables permanent anchoring contacts of the types $\text{NH}\cdots\text{O}$ and $\text{CH}\cdots\text{O}$ by the ligand atoms O2 and O3. Moreover, since the two native hydrogen bonds are not yet formed in this early binding stage, their formation thus represents a major enthalpic force driving the system toward the native complex. On the other hand, the symmetry and flexibility provided by the noncore regions enable transient contacts of various types (π – π , $\text{NH}\cdots\pi$, $\text{CH}\cdots\pi$, and $\text{CH}\cdots\text{O}$), which is another important aspect of this binding intermediate (*vide infra*). The dynamic aspect of the protein–ligand contacts in IS also helps to avoid a trapped state that could impede the formation of the bound state.

4. DISCUSSION

We have characterized one intermediate state (IS) along the unbinding pathway of ligand 8 to FKBP12 by performing seven independent molecular dynamics simulations with the explicit inclusion of water molecules. The analyses of the separation between the center of mass of the binding pocket and that of the ligand core reveal a stable position of the ligand in IS. The ligand core moiety is restricted to a narrow range of positions, lying at 4.64 ± 0.11 Å above its position in the native complex. The stability of the ligand is further confirmed by the few persistent intermolecular contacts, mainly O2...Val55/Ile56 and O3...Ile90, with corresponding distances varying over limited ranges (Table 4). The ligand core position as well as the protein structure are consistently predicted by the series of simulations. These results, together with the overall agreement with the LD results, provide confidence in the model derived for this binding intermediate.

In contrast to the core moiety, the three peripheral groups (iPe, Ph1, and Ph2) have a high mobility that results mainly from the low dihedral transition barriers around a few bonds. Consequently, no long-lived contacts between these moieties and the protein are consistently observed in all of the simulations. Nevertheless, these groups can form many alternative contacts with the protein, owing to both their intrinsic mobility and their (pseudo)symmetry. Recently, the importance of such alternative contacts has been highlighted in a study of the configurational entropic contributions of the residue side chains upon formation of the PKA/AKAP complex.⁴⁹ The authors found that the affinity between the two proteins is increased with the number of alternative contacts available, which is associated with an increase in configurational entropy. Though their conclusions were drawn from the analysis of a native protein–protein complex, we believe that, in the intermediate IS, such alternative contacts could help the ligand to retain at least a part of the configurational entropy it has in the free state, while the nascent interactions would contribute to enthalpic gain. This assertion is further supported

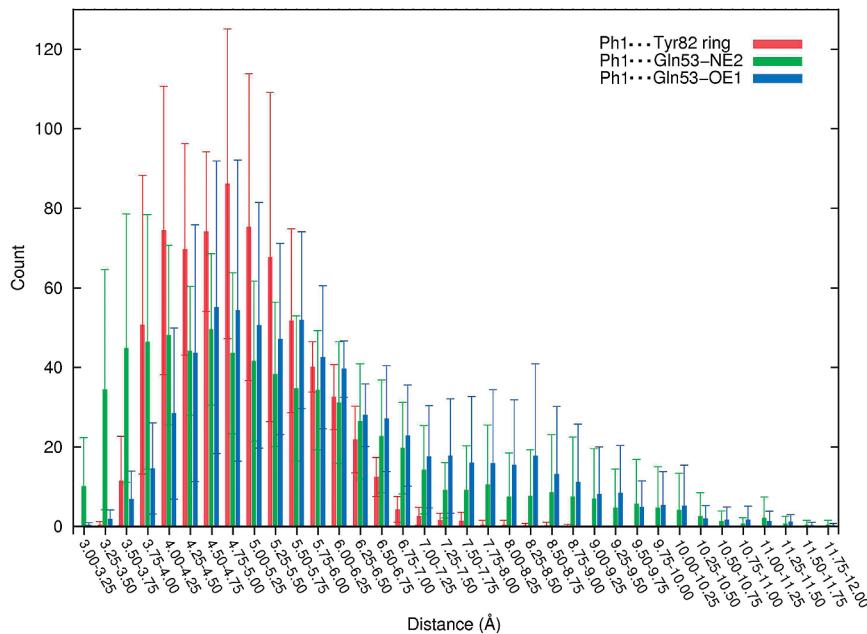


Figure 9. At regular distance intervals of 0.25 Å between 3 and 12 Å, the three distances between the Ph1 ring centroid and the Tyr82 ring centroid, Gln53-NE2, and Gln53-OE1, respectively, that fall into each interval were counted for each trajectory. The resulting numbers were averaged over the whole set of SBD trajectories to yield this histogram. On the x axis, the notation for the distance interval $a-b$ indicates an interval of type $[a;b]$.

by the fluctuations of the ligand calculated in IS that are similar to those in the free state (Figure 6). The dynamic aspect of the protein–ligand contacts also allows avoiding a trapped intermediate along the complexation pathway.

Another important aspect of the binding intermediate is the structure of the 80s loop. As discussed in the Introduction, in the bound state, ligand 8 is largely buried in the binding pocket: the pipecolinyl ring is buried in the main pocket, and the methyl group C13 is buried in the small side cavity defined by residues 82, 87, 90, and 91. Hence, the latter four residues would contribute to hindering the escape or the binding of the ligand if their bulky side chains remained in the position found in the crystal structure. A similar conclusion was drawn by Zacharias in a rigid docking study of FK506 to FKBP12, where the methyl group C35 of the ligand fills the same side cavity as C13 in 8.¹⁹ However, in a binding mechanism that proceeds through the intermediate IS, the concerted motion observed for the bulky side chains of the 80s loop exposes the small side cavity to the incoming ligand, thereby allowing regions of the ligand to initiate the interactions with the protein. Indeed, in IS, atom O3 points to this small side cavity and forms a stable interaction with the side chain of Ile90 as well as weak interactions with the side chain of Tyr82 (Table 4 and right of Figure 5). An important consequence of such a two-step mechanism, as opposed to a one-step one, is the drastic decrease of the energy barrier of association due both to the nascent protein–ligand interactions and to the reduced steric hindrance; the latter two can be regarded as favorable enthalpic and entropic contributions, respectively. In IS, the observed displacement of the 80s loop is also consistent with the conclusion of previous comparative analyses of bound and unbound forms of FKBP12,^{20,42} as well as with NMR data that indicate at least two distinct conformations for the 80s loop of the unbound protein.²¹ The displacement of the 80s loop seen in our model IS is additionally supported by a docking study performed on a set of 20 protein–protein complexes.⁵⁰ Though this study has not dealt with the docking of small molecules to FKBP12, it is worth

mentioning that the authors have found that the regions of the protein that prevent binding because of steric hindrance are undergoing conformational transitions compatible with the recognition of the binding partner.

With a view to testing the transferability of this binding intermediate model to other ligands, we have substituted ligand 8 in IS with FK506. We thus hypothesize here that a similar intermediate exists for the binding mechanism of FK506. This larger ligand was docked to the protein pocket in IS by aligning the atoms of the common core region. FK506 was taken in its bound crystallographic geometry, and no relaxation was further attempted. Figure 10 shows the result of the superposition of the two ligands. Remarkably, ligand FK506 fits well into this model by retaining many of the contact features of ligand 8, with very few short contacts, despite the fact that neither FK506 nor the mobile side chains of His87 and Tyr82 were allowed to relax. In particular, the distances O2···Ile56-N and O3···Ile90-CG2 are both compatible with a van der Waals interaction (3.45 and 3.57 Å, respectively, in the snapshot of Figure 10). The methyl group C35 in FK506 that points to the imidazole ring of His87 has a position equivalent to the methyl or methylene group on C10 in 8 (note that in the respective bound forms of the protein, the methyl C35 of FK506 fills the same small side cavity as the methyl C13 in 8). In Figure 10, the cyclohexyl group (C29–C34) in FK506 has the same role as the Ph1 ring in 8, and, as with Ph1, this group is known to be very mobile.⁵¹ The fact that some configurations taken from the simulations of IS could fit the larger ligand FK506 while still retaining the above-described structural features of ligand 8 suggests a transferability of our model to other ligand structures.

The analysis of this binding intermediate model also sheds some light on the role of the noncore regions iPe, Ph1, and Ph2 in the binding process. Since these three peripheral groups form extensive contacts with a few residues in IS and thus contribute to the stability of this intermediate, we can understand how these groups enhance the binding ability of this ligand. The comparison of

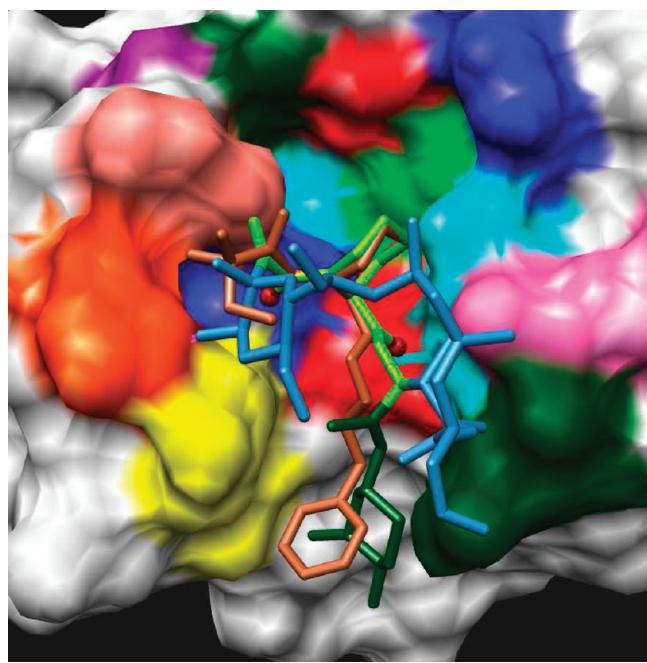


Figure 10. Snapshot that shows a superposition of the FK506 and ligand 8 core regions. The color code for the protein residues is as in Figure 5. Ligand 8 is colored brown, and oxygen atoms O2 and O3 are shown as red spheres. FK506 is colored as follows: core region (C9–C26) in light green, the cyclohexyl region (C27–C34) in dark green, and the nonbinding region or “effector” domain (C25–C10) in blue.

the affinity data for a few close analogues of ligand 8 indicates that the affinity is reduced whenever one of the three peripheral groups of 8 (iPe, Ph1, or Ph2) is missing. The substitution of any of these groups also has important consequences on the affinity data, as discussed below. The iPe moiety seems essential for high-affinity properties since many high-affinity ligands contain this group. Alternatively, for the ligands that have a *tert*-butyl group on C9, the affinity measured is higher than for the group iPe.⁵² Other high-affinity ligands have a phenyl or methoxyphenyl derivatives instead of iPe on C9.⁵³ By relying on our model and by assuming that it is transferable to other ligands, this phenyl group (or phenyl derivatives) linked to C9 would contribute to aromatic stacking with the imidazole ring of His87 in the binding intermediate analogue, thereby stabilizing this species. Regarding the Ph2 group, the close ligand analogue 5 (Figure 11a) exhibits a 10-fold decrease in affinity (110 nM vs 10 nM for 8). It should be stressed here that the role of Ph2 is not only to allow contacts in IS as seen, its presence in 8 also has a consequence on its neighboring group Ph1. Indeed, the analysis of the torsion profile around the bond O1–C15 indicates that Ph2 severely restrains the motion around O1–C15. Thus, the space that is spanned by the Ph1 ring in 8 is narrower than in the case of 5, which makes it more available for the interaction with Tyr82: the presence of the Ph2 ring on C15 in 8 therefore contributes indirectly to enhancing the π – π interactions in IS. For the analogue 9 (Figure 11a), the binding constant remains similar to that of 8 (7 vs 10 nM); this comparison suggests that the aromatic nature of the substituent on C15 is not essential. On the other hand, a substitution of Ph2 by a 1,1-dimethyl-2-propenyl group as in ligand 7 results in a 25-fold lower affinity than that of 8 (250 nM).⁶ By assuming the transferability of our intermediate model to the two previous ligand analogues, we may argue that, in contrast to the case of the

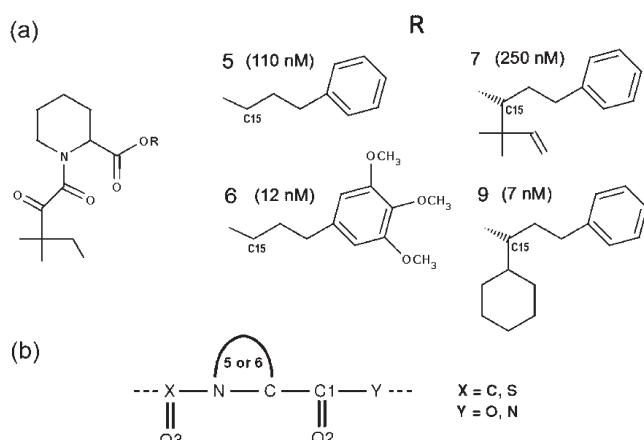


Figure 11. (a) Structures of a few ligands 5, 6, 7, and 9 with structures related to that of the ligand 8. The ligand identifiers and the K_i constants in parentheses are taken from the work of Holt et al.⁶ Atom C15 is also indicated. (b) A five- or six-membered ring is often found at the center of the core binding region of the high-affinity FKBP12 ligands (pyrrolidine or piperoline). A thiazane ring (six-membered ring with a sulfur atom) can also be found as a central motif.¹⁶ Recently, open alkyl forms for the central motif have demonstrated the ability to promote neurite outgrowth, thus suggesting that a central cyclic structure is not a prerequisite for binding to FKBP.¹⁷ An amide bond also replaces the ester bond in the context of the peptidylprolyl *cis*–*trans* isomerization.

ligand 7, in 9 the symmetry of the cyclohexyl substituent allows the formation of alternative contacts after the rotation of this group, as in the case of Ph2. Moreover, as mentioned above, Ph2 in 8 and the cyclohexyl ring in 9 are bulky groups that also indirectly enhance the interactions that Ph1 makes in IS. Hence, the number of contacts seen in the intermediate model, which determines its stability, seems to be related to the affinity of a few analogues of 8.

Regarding the Ph1 group, it is striking that almost all of the high-affinity synthetic ligands of FKBP12 contain an aromatic ring (phenyl, pyridyl, or trimethoxyphenyl ring) linked by a propyl group to the ester bond of the ligand core. The presence of this aromatic group and the length of the alkyl arm that links it to the core therefore seem crucial for the binding properties. The results for our model IS provide a guide for understanding the role played by such an aromatic group in the recognition process. First, in IS the Ph1 moiety makes on average two contacts with Gln53 and almost six others with Tyr82, Ph1 being the ligand moiety that is making the highest average number of contacts with FKBP12 (in comparison, the piperolinyl ring is only weakly interacting with the protein in IS; see Figure 7a). The same analysis performed for Ph1 in the bound state indicates that Ph1 participates in an average of about three contacts both with His87 and with Tyr82, and one with Gly86 (data not shown). Second, in IS the RMSD values of all of the ligand moieties from their respective native positions reveal that Ph1 is the moiety closest to its native position (4.16 ± 0.18 Å in Table 2). Therefore, in IS, where the nascent interactions are formed between the molecular partners, the Ph1 ring appears critical since it not only interacts the most with the protein but is also the moiety closest to its native position. What is more, in tracing the nature of the contacts between Ph1 and Tyr82 by performing semiempirical molecular orbital calculations, a molecular orbital overlap is obtained between the aromatic rings of Tyr82 and Ph1, as illustrated in Figure 12. This molecular overlap is observed whenever the two interacting

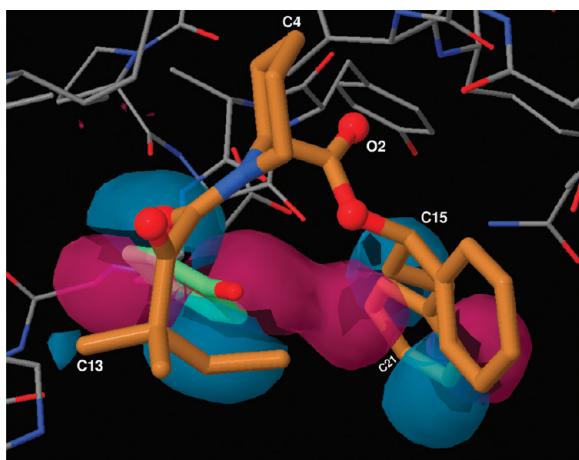


Figure 12. Molecular overlap between the Ph1 moiety of the ligand (shown in brown) and the aromatic ring of Tyr82 (in light green) calculated for IS at the semiempirical level of theory by using the MOZYME method with the PMS Hamiltonian, as implemented in the MOPAC2006 program.⁵⁹ The molecular orbital is the HOMO-40 represented with an isosurface value of 0.008.

rings exhibit a parallel displaced configuration in a series of structures taken from the SBD trajectories. This semiempirical approach has been shown to yield binding enthalpies for protein–ligand systems in nice agreement with experimental values.⁵⁴ To our knowledge, this is the first result showing that short-range electronic interactions (overlap forces) may be important in an early stage of the recognition process in a biomolecular context.

The short-range electronic interactions may also explain the 10-fold difference in K_i measured by Holt et al.⁶ for compounds **5** and **6** (Figure 11a). The two ligands **5** and **6** consist of the iPe and core moieties but differ in the group linked to the ester bond: propylphenyl and (3,4,5)-trimethoxyphenylpropyl, respectively. In **6**, the three methoxy groups act as electron donors on the aromatic ring and are thus responsible for an enrichment in electronic density on both sides of the ring, as compared to the case of the nonsubstituted ring of **5**. Hence, in the hypothetical binding intermediate analogue for **6**, the out-of-plane increased density in the substituted ring could enhance the face-to-face interactions with the Tyr82 ring, as compared to the corresponding interactions in the binding intermediate analogue for **5**. Under the hypothesis that the stability of the intermediate is related to the affinity of the ligand, the stabilization brought by the stronger $\pi-\pi$ interactions in the case of **6** may thus, at least partly, explain the difference in affinity.

From an experimental viewpoint, the importance of the aromatic character of residue 82 in the binding of the ligand can also be seen in the measures of the inhibitory activity against the mutants Y82F and Y82D of the protein. The mutation of Tyr82 to Phe lowers the affinity for FK506 and rapamycin nearly 2-fold.¹ Other experiences performed on the binding of rapamycin to yeast FKBP12 indicate a 95% loss of binding activity for the mutant Y89D;⁵⁵ Y89 corresponds to Y82 in the human form of the protein. Taken together, these results suggest that the hydrogen bond between O3 and Tyr82 is not crucial for the binding but rather the aromatic character of the side chain of residue 82. These observations are consistent with the model IS where we have found short-range electronic interactions that involve the aromatic ring of Tyr82 while the hydrogen bond with this residue is only weak at this early stage. In the hypothetical binding

intermediate analogue for FK506, the interactions of types C–H $\cdots\pi$ and O¹²–H $\cdots\pi$ between the ring C29–C34 and Tyr82 may substitute the $\pi-\pi$ interactions found for **8** (Figure 10). For rapamycin, similar interactions may involve its ring C37–C42 and hydroxyl oxygen O¹³. Interestingly, in a more recent mutational analysis of the FKBP12 enzymatic activity, it was found that, among the 19 substitutions of the 82nd residue, the residues with an aromatic side chain (Phe, Tyr, Trp) exhibited the highest activity, along with Arg (which also has a marked aromatic character) and Pro.⁵⁶ These experimental findings therefore also emphasize the importance of the aromatic character of the 82nd residue.

Finally, the analysis of this binding intermediate also provides clues regarding the diversity of the core structures found in the high-affinity ligands. Our results reveal that atoms O2 and O3 are the only atoms that form stable contacts in the binding intermediate, acting as two anchoring points. This result suggests that the presence of these two atoms in the core region of a FKBP12 ligand may be important for the binding. Interestingly, when the structures of the high-affinity ligands of FKBP12 are compared so as to extract a common structural motif, O2 and O3 are found as recurrent atoms, as seen in Figure 11b. It is noteworthy that, in this motif, the piperolinyl or pyrrolidinyl ring seems not to be a prerequisite for tight binding since a thiazane ring¹⁶ or even open alkyl forms for the central ring¹⁷ are also found in high-affinity ligands. In fact, the precise structure (ring or open alkyl form) of the central region of the motif in Figure 11b appears to be of less importance than that of the peripheral substituents of the core, as concluded by Zhao et al. in their study on the abilities of various FKBP ligand structures to promote neurite outgrowth.¹⁷ Interestingly, these authors have also found that the most important regions of the ligand for promoting neurite outgrowth are the group linked to the ester bond of the core (Ph1 or other substituents).¹⁷ Regarding the bond C9O4 found in the core, Orozco et al. have inferred that this group is not really important for the binding of ligands to FKBP12.⁵⁷ Their finding is also consistent with the observation that some high-affinity ligands contain only one ketone group (C8O3);⁵³ other ligands have one sulfonyl group that replaces the diketone motif.¹⁶ From the above analysis of the core region, O2 and O3 thus appear as the main recurrent atoms. If we hypothesize that the model IS is transferable to other ligand analogues, the prevalence of the atoms O2 and O3 may then be explained because of their role as anchoring points in the respective binding intermediate, as stated for ligand **8**. This result therefore helps to rationalize the observation that structurally diverse core regions that all share the atoms O2 and O3 can bind to FKBP12.

5. CONCLUSIONS

We have characterized one intermediate state (IS) along the unbinding pathway of ligand **8** to FKBP12. The ligand core region is restricted to a narrow range of positions, lying 4.6 ± 0.1 Å above its position in the native complex. Only the two carbonyl atoms of the core regions (O2 and O3) make permanent contacts in the intermediate, acting as two anchoring points. In contrast, the noncore regions have a large mobility that (i) ensures a dynamic aspect of the interactions, thereby avoiding a trapped state along the complexation pathway, and (ii) offers the possibility of fine-tuning the specificity of recognition owing to various types of interaction ($\pi-\pi$ molecular overlap and weak hydrogen bonds N–H $\cdots\pi$, C–H $\cdots\pi$, and C–H \cdots O). Previous results on

the structure of the 80s loop are consistent with our results for the intermediate. Moreover, under the hypothesis that our binding intermediate model is transferable to other related ligands, which seems valid for FK506, this model comes closest to explaining both the diversity found for the core structures of the FKB12 ligands and the role that the noncore regions could play in the recognition and the affinity (if we hypothesize that, for the binding to FKB12, the slower the dissociation rate of the binding intermediate, the higher is the affinity^{10,11}). Indeed, for other high-affinity ligands of FKB12, this model may explain the common occurrence of both the atoms O2 and O3 found in the core and the aromatic ring linked to the core by the same propyl “arm”, the latter being important for the flexibility and the contacts with Tyr82 and Gln53. A striking finding is the important role that the aromatic ring of Ph1 seems to play in the binding process by making short-range electronic interactions with Tyr82. This result is further supported by previous studies that have demonstrated the importance of the aromatic character of the 82nd residue in FKB12.^{1,55,56} We speculate that, in other protein–ligand systems, short-range electronic interactions between an aromatic group of the ligand and an aromatic residue of the protein, as between Ph1 and Tyr82 in our model IS, might play a critical role in an early stage of the recognition process.

■ ASSOCIATED CONTENT

S Supporting Information. More details on the simulation procedure are available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Fax: +262(0)262-93-82-37. E-mail: Fabrice.Gardebien@univ-reunion.fr.

■ ACKNOWLEDGMENT

L.O. was the recipient of a doctoral fellowship from the French Ministry of Research.

■ REFERENCES

- (1) DeCenzo, M. T.; Park, S. T.; Jarrett, B. P.; Aldape, R. A.; Futer, O.; Murcko, M. A.; Livingston, D. J. *J. Protein Eng.* **1996**, *9*, 173–180.
- (2) Fung, J. J.; Starzl, T. E. *Ther. Drug Monit.* **1995**, *17*, 592–595.
- (3) Danovitch, G. M. *Transplant. Proc.* **1999**, *31*, 2–6.
- (4) Herdegen, T.; Fischer, G.; Gold, B. G. *Trends Pharmacol. Sci.* **2000**, *21*, 3–5.
- (5) Davies, T. H.; Sánchez, E. R. *Int. J. Biochem. Cell Biol.* **2005**, *37*, 42–47.
- (6) Holt, D.; Luengo, J.; Yamashita, D.; Oh, H.; Konialian, A.; Yen, H.; Rozamus, L.; Brandt, M.; Bossard, M.; Levy, M. A.; Eggleston, D. S.; Liang, J.; Schultz, L. W.; Stout, T. J.; Clardy, J. *J. Am. Chem. Soc.* **1993**, *115*, 9925–9938.
- (7) Xu, Y.; Wang, R. *Proteins: Struct., Funct., Genet.* **2006**, *64*, 1058–1068.
- (8) Wang, J.; Deng, Y.; Roux, B. *Biophys. J.* **2006**, *91*, 2798–2814.
- (9) Zhang, C.; Chen, J.; DeLisi, C. *Proteins: Struct., Funct., Genet.* **1999**, *34*, 255–267.
- (10) Ubbink, M. *FEBS Lett.* **2009**, *583*, 1060–1066.
- (11) Schreiber, G.; Haran, G.; Zhou, H. X. *Chem. Rev.* **2009**, *109*, 839–860.
- (12) Mittag, T.; Schaffhausen, B.; Günther, U. L. *J. Am. Chem. Soc.* **2004**, *126*, 9017–9023.
- (13) Sydor, J. R.; Engelhard, M.; Wittinghofer, A.; Goody, R. S.; Herrmann, C. *Biochemistry* **1998**, *37*, 14292–14299.
- (14) Selzer, T.; Albeck, S.; Schreiber, G. *Nat. Struct. Biol.* **2000**, *7*, 537–541.
- (15) Lamb, M. L.; Jorgensen, W. L. *J. Med. Chem.* **1998**, *41*, 3928–3939.
- (16) Sun, F.; Li, P.; Ding, Y.; Wang, L.; Bartlam, M.; Shu, C.; Shen, B.; Jiang, H.; Li, S.; Rao, Z. *Biophys. J.* **2003**, *85*, 3194–3201.
- (17) Zhao, L.; Liu, H.; Wang, L.; Li, S. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4385–4390.
- (18) Babine, R. E.; Bender, S. L. *Chem. Rev.* **1997**, *97*, 1359–1472.
- (19) Zacharias, M. *Proteins: Struct., Funct., Genet.* **2004**, *54*, 759–767.
- (20) Ivery, M. T. G.; Weiler, L. *Bioorg. Med. Chem. Lett.* **1997**, *5*, 211–232.
- (21) Rosen, M. K.; Schreiber, S. L. *Angew. Chem., Int. Ed.* **1992**, *31*, 384–400.
- (22) Lemons, D. S.; Gythiel, A. *Am. J. Phys.* **1997**, *65*, 1079–1081.
- (23) Lazaridis, T.; Karplus, M. *Proteins: Struct., Funct., Genet.* **1999**, *35*, 133–152.
- (24) Brünger, A.; Brooks, C. L., III; Karplus, M. *Chem. Phys. Lett.* **1984**, *105*, 495–500.
- (25) Curcio, R.; Caflisch, A.; Paci, E. *Protein Sci.* **2005**, *14*, 2499–2514.
- (26) Jorgensen, W. *Nature* **2010**, *466*, 42–43.
- (27) Colizzi, F.; Perozzo, R.; Scapozza, L.; Recanatini, M.; Cavalli, A. *J. Am. Chem. Soc.* **2010**, *132*, 7361–7371.
- (28) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarty, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (29) Foloppe, N.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86–104.
- (30) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- (31) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (32) Brooks, C.; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312–6325.
- (33) Paci, E.; Caflisch, A.; Plückthun, A.; Karplus, M. *J. Mol. Biol.* **2001**, *314*, 589–605.
- (34) Swegat, W.; Schlitter, J.; Krüger, P.; Wollmer, A. *Biophys. J.* **2003**, *84*, 1493–1506.
- (35) Huang, H.; Ozkirimli, E.; Post, C. *J. Chem. Theory Comput.* **2009**, *5*, 1304–1314.
- (36) Li, A.; Daggett, V. *J. Mol. Biol.* **1998**, *275*, 677–694.
- (37) Caves, L. S.; Evanseck, J. D.; Karplus, M. *Protein Sci.* **1998**, *7*, 649–666.
- (38) van Duyn, G. D.; Standaert, R. F.; Karplus, P. A.; Schreiber, S. L.; Clardy, J. *J. Mol. Biol.* **1993**, *229*, 105–124.
- (39) Nakanishi, I.; Fedorov, D. G.; Kitaura, K. *Proteins: Struct., Funct., Genet.* **2007**, *68*, 145–158.
- (40) Rhodes, G. A User’s Guide to Crystallographic Models. In *Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models*; Academic Press, an imprint of Elsevier Science: San Diego, CA, 2006; pp 159–186.
- (41) Cornell, W.; Abseher, R.; Nilges, M.; Case, D. A. *J. Mol. Graph. Model.* **2001**, *19*, 136–145.
- (42) Wilson, K. P.; Yamashita, M. M.; Sintchak, M. D.; Rotstein, S. H.; Murcko, M. A.; Boger, J.; Thomson, J. A.; Fitzgibbon, M. J.; Black, J. R.; Navia, M. A. *Acta Crystallogr.* **1995**, *D51*, 511–521.
- (43) Peters, G. H.; Bywater, R. P. *Protein Eng.* **1999**, *12*, 747–754.

- (44) Teplyakov, A.; Sebastiao, P.; Obmolova, G.; Perrakis, A.; Brush, G. S.; Bessman, M. J.; Wilson, K. S. *EMBO J.* **1996**, *15*, 3487–3497.
- (45) Hornak, V.; Okur, A.; Rizzo, R. C.; Simmerling, C. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 915–920.
- (46) Burkhard, P.; Taylor, P.; Walkinshaw, M. D. *J. Mol. Biol.* **2000**, *295*, 953–962.
- (47) Rosen, M. K.; Yang, D.; Martin, P. K.; Schreiber, S. L. *J. Am. Chem. Soc.* **1993**, *115*, 821–822.
- (48) Castellano, R. K.; Diederich, F.; Meyer, E. A. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210–1250.
- (49) Chang, C. A.; McLaughlin, W. A.; Baron, R.; Wang, W.; McCammon, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 7456–7461.
- (50) Dobbins, S. E.; Lesk, V. I.; Sternberg, M. J. E. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 10390–10395.
- (51) Lepre, C. A.; Pearlman, D. A.; Cheng, J. W.; DeCenzo, M. T.; Moore, J. M.; Livingston, D. J. *Biochemistry* **1994**, *33*, 13571–13580.
- (52) Hamilton, G. S.; Huang, W.; Connolly, M. A.; Ross, D. T.; Guo, H.; Valentine, H. L.; Suzdak, P. D.; Steiner, J. P. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1785–1790.
- (53) Yang, W.; Rozamus, L. W.; Narula, S.; Rollins, C. T.; Yuan, R.; Andrade, L. J.; Ram, M. K.; Phillips, T. B.; van Schravendijk, M. R.; Dalgarno, D.; Clackson, T.; Holt, D. A. *J. Med. Chem.* **2000**, *43*, 1135–1142.
- (54) Nikitina, E.; Sulimov, V.; Zayets, V.; Zaitseva, N. *Int. J. Quantum Chem.* **2004**, *97*, 747–763.
- (55) Koser, P. L.; Eng, W. K.; Bossard, M. J.; McLaughlin, M. M.; Cafferkey, R.; Sathe, G. M.; Faucette, L.; Levy, M. A.; Johnson, R. K.; Bergsma, D. J.; Livi, G. P. *Gene* **1993**, *129*, 159–165.
- (56) Ikura, T.; Ito, N. *Protein Sci.* **2007**, *16*, 2618–2625.
- (57) Orozco, M.; Tirado-Rives, J.; Jorgensen, W. L. *Biochemistry* **1993**, *32*, 12864–12874.
- (58) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577–2637.
- (59) Stewart, J. J. P. *MOPAC2006*, version 1.0; Fujitsu Limited: Tokyo, Japan, 2006.