NMR studies of protein-small molecule and protein-peptide interactions

Guan, J.

3

Structure determination of a protein-ligand complex by NOE

Abstract

This chapter presents the determination of ligand binding pose by the traditional NOE approach. Chemical shift perturbations (CSPs) were used to determine K_D and initially also to determine the binding site. However, CSPs monitor both direct and indirect effects, as observed in the complex. Therefore isotope filtered/edited NOESY measurements were used to determine the ligand binding mode. To characterize the interaction, full assignments on the protein and the ligand were performed. Many intermolecular NOEs were identified from the NOESY spectra and subsequently applied for structural calculation. The resulting structure is used as a reference structure to compare with the structure of the same complex calculated based on paramagnetic NMR data (Chapter 4).

This work has been published as part of

J.-Y. Guan, P. H. J. Keizers, W.-M. Liu, F. Löhr, E. Heeneman, S. P. Skinner, H. Schwalbe, M. Ubbink and G. Siegal. **Small molecule binding sites on proteins established by paramagnetic NMR spectroscopy.** *J. Am. Chem. Soc.*, **2013**, *135* (15), pp 5859–5868

3.1 Introduction

Structure information of protein-ligand complexes is highly valuable in the early stages of structure-based drug design (SBDD) and fragment-based drug discovery (FBDD), where X-ray crystallography plays an important role to provide structure information. In many cases, however, hit compounds that interact weakly with their targets may not readily crystallize, due to a variety of reasons. Due to its high sensitivity in detecting weak interactions, NMR has proved to be a very useful tool when X-ray crystallography cannot be applied.

Ligand-induced chemical shift perturbations (CSPs) are commonly used to monitor ligand binding and, for rigid proteins, to determine ligand binding sites. However, CSPs caused by direct contact with ligand and indirect conformational changes cannot be distinguished by simply mapping CSPs on the protein structure.

NOEs develop due to through-space interactions rather than through-bond interactions. The intensity of an NOE is proportional to the inverse of the sixth power of the distance that separates the two dipolar-coupled spins within 5 Å, and so NOEs are sensitive probes of short-range through-space intramolecular and intermolecular interactions. In standard isotope-edited NOESY spectra, intermolecular NOEs are indistinguishable from intramolecular protein NOEs. Acquiring 2D and 3D edited/filtered NOESY[159,160] on a sample of $^{13}C/^{15}N$-labeled protein saturated with unlabeled ligand can overcome this problem. This type of experiments detects only NOEs between protons directly bound to ^{13}C or ^{15}N (protein) and protons bound to ^{12}C or ^{14}N atoms (ligand) while suppressing all other cross-peaks by isotope filtering and editing.[161] Crosspeaks from the filtered NOESY spectrum can therefore be unambiguously assigned to protons on the small-molecule ligand.

The aim of the work presented here is to characterize the structure of a small molecule ligand bound to a 12 kDa FK-506 binding protein (FKBP12) using CSP and intermolecular NOE restraints. FKBP12 is a peptidyl-prolyl isomerase which belongs to the family of immunophilins and is a drug target for the immuno-suppressants rapamycin and FK506. There have been many structural studies on this protein.[9,10,15,20,162,163] The ligand in this study is a fragment that was identified as a hit against FKBP12 from a screen of a fragment library using Target-Immobilized NMR Screening (TINS).[84] Using high-resolution NMR spectroscopy, we determined the binding mode of the bound ligand. The determined ligand binding mode serves as a reference structure for Chapter 4, in which a paramagnetic NMR approach was applied to determine the binding mode of the same ligand.

3.2 Materials and Methods

Ligand preparation

The ligand **1**, [2-(4,4-dimethyl-5,5-dihydro-1,3-oxazol-2-yl)phenyl]methanol, was purchased from MayBridge (catalog number S13756). Assignment of the ligand resonances was achieved from 2D ^{1}H, ^{13}C-HSQC and ^{1}H,^{13}C-HMBC spectra acquired on a 100 mM solution of the ligand in D_2O.

Protein expression and purification

Uniformly ^{15}N-labeled and ^{15}N, ^{13}C-labeled FKBP12 was purified from *Escherichia coli* cells grown on minimal media containing $^{15}NH_4Cl$ and $^{12}C_6$- or $^{13}C_6$-glucose, respectively. The proteins were essentially expressed and purified as described in Chapter 2.

NMR measurements

All protein NMR samples contained 15 mM Tris-HCl, 25 mM NaCl, pH 7.7, 6% D_2O for ^{15}N-labeled protein and >95% D_2O for non-isotope-labeled protein. The concentration of wild type FKBP12 was 100 μM for titrations, and 1.5 mM for resonance assignment and NOESY experiments. 1D-^{1}H, [^{1}H,^{15}N]-HSQC, HNCA, HNCACB, HN(CO)CA, HN(CO)CACB, HNCO, (H)CCH-TOCSY, HN(CA)CO and HBHA(CBCACO)NH spectra were recorded at 290 K on a Bruker Avance DMX-600 spectrometer equipped with a TCI-Z-GRAD cryoprobe. A ligand to protein ratio of 10:1 was used in the 3D ^{15}N-separated ω1-^{13}C/^{15}N-filtered NOESY-TROSY, 3D NOESY-[^{1}H,^{15}N]-TROSY, 3D ^{13}C-separated ω1-^{13}C/^{15}N-filtered NOESY-[^{1}H,^{13}C]-HSQC, 3D NOESY-[^{1}H,^{13}C]-HSQC, 2D ^{13}C-edited, ^{13}C/^{15}N-filtered NOESY, and a ratio of 3:1 used in 2D NOESY, 2D DQF-COSY, and 2D TOCSY spectra of the complex were recorded on a Bruker Avance 950 MHz spectrometer with a $^{1}H\{^{13}C,^{15}N\}$ cryogenic probe. Data were processed in TopSpin (Bruker) and then spectra were analyzed in Sparky.[164] Dr. Frank Löhr (Goethe University Frankfurt, Germany) is acknowledged for the NMR measurements on the 950 MHz spectrometer.

Calculations of dissociation constants and bound ligand fractions

Ligand binding was observed via the changes of protein resonances in the [^{1}H,^{15}N]-HSQC spectrum upon titration with the ligand.[165] For analysis of the chemical shift perturbations (CSPs) of ^{1}H and ^{15}N backbone resonances, the weighted average chemical shift values were calculated and normalized according to equation 3.1:

$$\Delta\delta_{avg} = \sqrt{\frac{1}{2}[(\frac{\Delta\delta_N}{5})^2 + \Delta\delta_H{}^2]} \quad (3.1)$$

where $\Delta\delta_N$ and $\Delta\delta_H$ are the differences of ^{15}N and ^{1}H chemical shift of an amide group, respectively.

The dissociation constant (K_D) was determined using a two-parameter non-linear regression curve fitting based on a one-site binding model as described in equation 3.2:

$$\Delta\delta_{avg} = \frac{1}{2}\Delta\delta_0(A - \sqrt{A^2 - 4R})$$
$$A = 1 + \frac{1}{R} + \frac{P_0R + L_0}{P_0L_0\left(\frac{1}{K_D}\right)} \quad (3.2)$$

where R is the total [ligand] to [protein] ratio, $\Delta\delta_{avg}$ is the average CSP (equation 3.1) at a given R, $\Delta\delta_0$ is the CSP at 100% bound protein, P_0 is the starting concentration of the protein, L_0 is the stock concentration of ligand and K_D is the dissociation constant.[166] The fraction of bound ligand was calculated using the dissociation constant.

NOE-based structure calculations

Intermolecular NOE crosspeaks were identified in 2D and 3D NOESY spectra with NOE mixing times of 50-70 ms and then converted into distances using the CYANA[167] calibration function. Intermolecular restraints were introduced as ambiguous restraints if degenerate protons were involved or the protein resonance could not be unambiguously assigned.

Protein coordinates were taken from the X-ray and NMR structures of FKBP12 (PDB entries 1FKR,[9] 1FKS,[9] 1FKT,[9] 1D6O[27] and 2PPN[29]) and the backbone was kept fixed throughout the process. Starting structures for the complex were generated by placing the ligand in random orientations with respect to FKBP12. Then the NOE distance restraints were applied to generate ligand orientations which satisfied the intermolecular restraints. Side chain atoms within 8 Å of the ligand were allowed to rotate during subsequent energy minimization. The XPLOR-NIH script is available in Appendix B.

3.3 Results and Discussion

Characterization of the ligand-protein interaction

The ligand **1** (Figure 3.1A) was identified from a TINS screen[84] of a library of commercially available, low molecular weight "drug fragments" for binding to FKBP12. To confirm the binding of **1** to FKBP12 and obtain structural insight into the binding site, we titrated the

ligand into ^{15}N labeled protein and observed CSPs in a series of [^{1}H,^{15}N]-HSQC spectra with increasing ligand concentration. Figure 3.1B shows the titration curves for the ten residues most effected from which the equilibrium dissociation constant, K_D, could be extracted. CSPs occurred throughout the protein, including in two previously defined ligand binding sites, referred to as site-1 and site-2.[30] Five of these residues are site-1 residues and the other five residues belong to site-2. In Figure 3.1C the CSPs have been mapped onto the crystal structure of FKBP12 (PDB entry 2PPN[29]) and color coded according to their magnitude. While CSPs can be caused by direct changes in the electronic environment of spins close to the ligand, many other factors can contribute. For example, it has been shown that, upon ligand binding to FKBP12, perturbations of both main-chain and side-chain dynamics can occur at sites distal to the binding interface.[168] Since such changes in dynamic behavior may also lead to CSPs, it was not possible to define the ligand binding site by CSP mapping alone. Therefore we sought an alternative method to elucidate the structure of the complex.

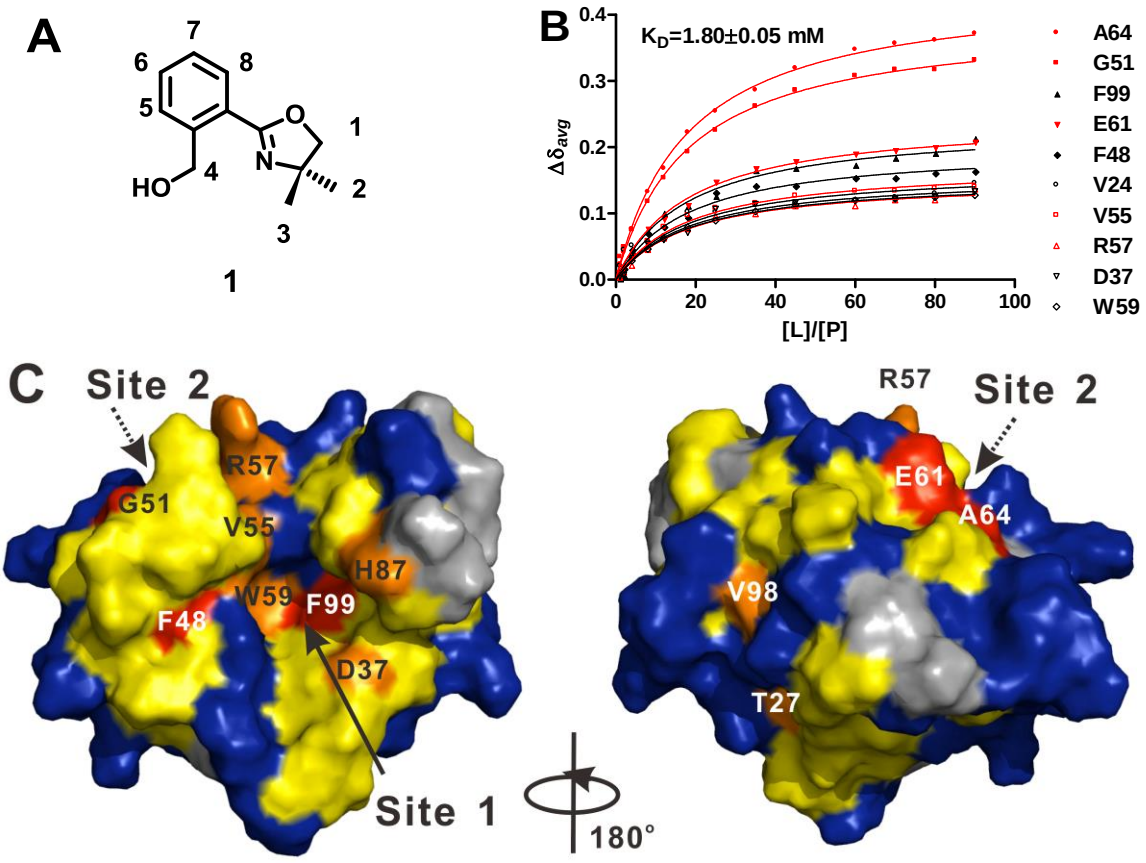

Figure 3.1: (A) Chemical structure of ligand 1. (B) Chemical shift changes of FKBP12 resonances as a function of increasing [L]/[P]. The top 10 residues which showed largest perturbations are shown. Residues in the site-1 and site-2 regions are shown in black and red, respectively. The dissociation constant of 1 was obtained by fitting simultaneously to a 1:1 binding model (equation 3.2, solid lines). (C) Mapping of CSPs from the binding of 1 on the structure of FKBP12 (PDB accession number

2PPN[29]). The positions of site 1 and site 2 are indicated. Color representation: red, $\Delta\delta_{avg} > 0.15$ ppm; orange, $0.15 > \Delta\delta_{avg} > 0.10$ ppm; yellow, $0.10 > \Delta\delta_{avg} > 0.03$ ppm; blue, $\Delta\delta_{avg} \leq 0.03$ ppm; dark grey, no data. Figures showing protein structures were prepared with PyMOL.[169]

NMR assignments

Assignment of the chemical shifts of free FKBP12[7] and FKBP12 bound to ascomycin[12] have been reported and these were used as the starting points. The complete assignments were verified as following: Backbone assignments (HN, CO, C_α and C_β) of ligand-free protein were obtained from standard triple resonance spectra and transferred to the bound state by following changes in [^{1}H,^{15}N]-HSQC and [^{1}H,^{13}C]-HSQC spectra upon ligand addition. Subsequently, the aliphatic side chains were assigned from a (H)CCH-TOCSY spectrum and the aromatic ones from 2D ^{1}H,^{1}H-NOESY, DQF-COSY and ^{1}H,^{1}H-TOCSY spectra.

NOESY measurements

NOEs develop due to through-space dipole-dipole interactions, thus containing information on the distances between atoms separated by 5Å or less in space. The intensity of an NOE is proportional to the inverse of the sixth power of the distance between the two nuclei, but not always a precise reflection of the distance. The structure of the FKBP12-**1** complex was initially determined using standard intermolecular NOE based methods. Intermolecular NOEs were derived from a combination of 3D ^{15}N-separated ω1-^{13}C/^{15}N-filtered NOESY-TROSY, 3D NOESY-[^{1}H,^{15}N]-TROSY, 3D ^{13}C-separated, ω1-^{13}C/^{15}N-filtered NOESY-[^{1}H,^{13}C]-HSQC, 3D NOESY-[^{1}H,^{13}C]-HSQC, 2D ^{13}C-edited,^{13}C/^{15}N-filtered NOESY and 2D NOESY spectra. Figure 3.2 and 3.3 show regions of NOESY spectra containing intermolecular NOEs between the ligand and backbone amides (Figure 3.2) and side chains (Figure 3.3) of the protein. In total, 66 intermolecular NOE crosspeaks were identified in these NMR spectra.

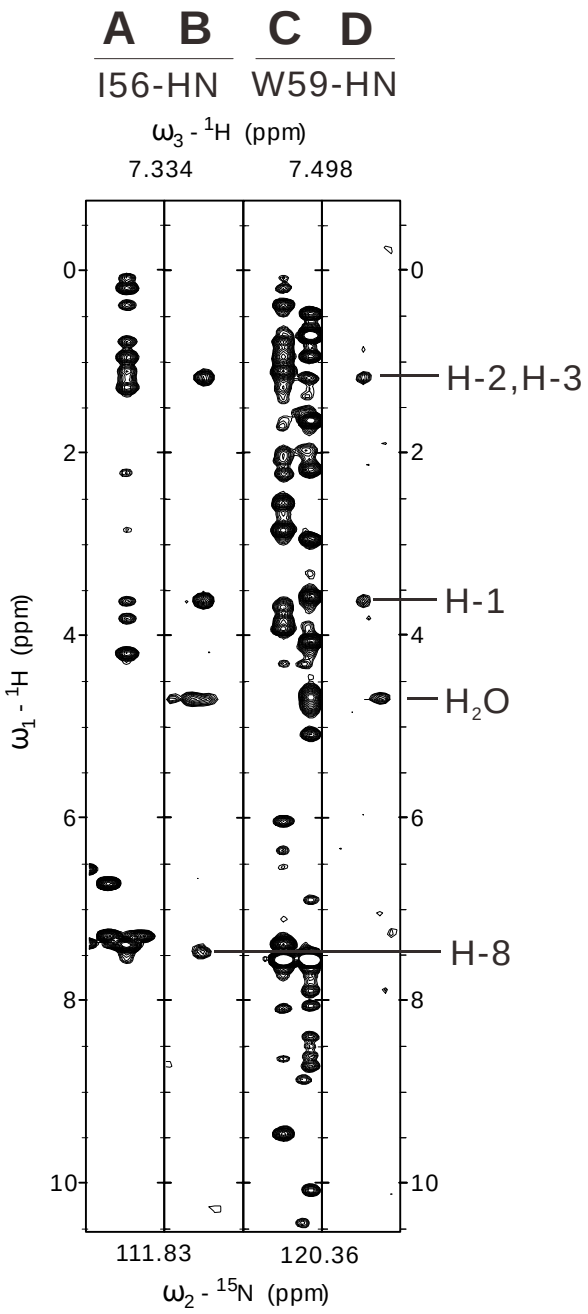

Figure 3.2: ^{1}H-^{1}H cross-sections from NOESY-^{15}N TROSY (A, C) and ^{13}C,^{15}N- filtered NOESY-^{15}N-TROSY (B, D) of FKBP12-1 complex, displaying intermolecular NOE cross-peaks between selected backbone amide groups of ^{13}C/^{15}N labeled FKBP12 and ^{1}H nuclei of the unlabeled ligand 1. (A, B) I56 at (δ_N,δ_H) = (111.83, 7.334) ppm and (C, D) W59 at (δ_N,δ_H) = (120.36, 7.498) ppm. Chemical shifts of ligand protons and H_2O are indicated.

Ligand conformation

There are two possible orientations of the five-membered oxazole ring relative to the six-membered aromatic ring. Between these two orientations, the five-membered ring flips by 180^{o}. The difference in intensities of intramolecular NOEs between protons 4/1, 4/2 and 4/3 (proton numbers as indicated on the structure in Figure 3.1A) in the presence of FKBP12 suggests that the presented conformation is the prefered bound conformation in solution. A possible explanation is that the hydroxyl group from the aromatic ring may form a hydrogen bond with the nitrogen atom, creating an extra 7-membered ring.[170]

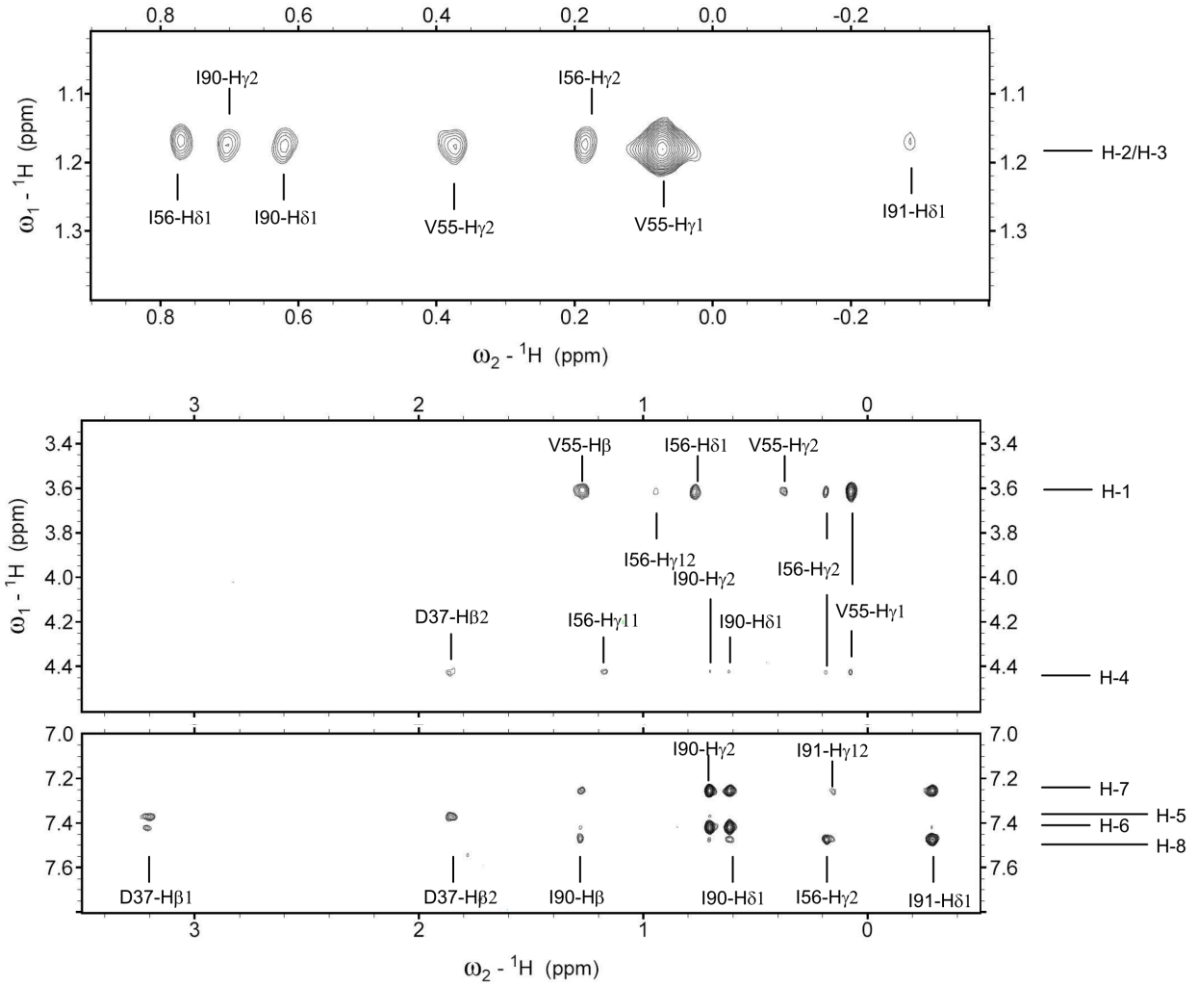

Figure 3.3: Regions showing intermolecular NOE evidences in 2D ^{13}C-edited,^{13}C/^{15}N-filtered NOESY. Chemical shifts of the ligand ^{1}H nuclei are indicated on the right.

NOE-based structure calculations

Among the 66 intermolecular NOE crosspeaks, 43 restraints were used for structural calculations (Table 3.1). The NOEs that were weak (20 peaks) or with uncertain assignments (3 peaks) were excluded. Superposition of various FKBP12 structures from the PDB (1FKR, 1FKS, 1FKT, 1D6O, 2PPN) indicated that the loops surrounding the hydrophobic pocket are variable and hence might undergo significant dynamic behavior in solution. Therefore, to cover the range of conformations of the loops, multiple structures were used as input for structure calculations.

Table 3.1: List of intermolecular NOE restraints and their corresponding distances (in Å) calculated from the final NOE structure. (A) Restraints from flexible residues. (B) Restraints from rigid residues. (C) Restraints that were excluded due to weak intensities or uncertain assignments.

(A) NOE restraints from flexible residues

no.	residue #	protein	ligand	r(exp)	r(calc)
1	37	Hβ1	H-5	3.4	4.1
2	37	Hβ2	H-5	3.4	3.3
3	55	Hα	H-1	3.4	2.6
4	55	Hα	H-2,H-3	3	3.6
5	55	Hα	H-8	3.4	6.5
6	55	Hδ	H-1	3.3	4.5
7	55	Hγ1	H-1	2.8	3.7
8	55	Hγ1	H-2,H-3	2.4	3
9	55	Hγ2	H-2,H-3	3.3	5.3
10	56	Hδ1	H-1	3.3	3.8
11	56	Hδ1	H-2,H-3	3.3	3.5
12	56	Hγ2	H-1	3.5	2.3
13	56	Hγ2	H-2,H-3	3.3	4.7
14	56	Hγ2	H-8	3.2	5
15	56	HN	H-1	2.8	2.4
16	56	HN	H-2,H-3	3	3.5
17	56	HN	H-8	3.4	6.4
18	82	Hδ	H-8	3.5	4.3
19	82	Hε	H-8	2.5	2.7
20	87	Hδ2	H-7	3.4	5.3
21	90	Hδ1	H-2,H-3	3.4	9.7
22	90	Hδ1	H-7	2.9	4.4
23	90	Hδ1	H-6	2.7	2.5
24	90	Hγ2	H-2,H-3	3.5	9.8
25	90	Hγ2	H-7	2.6	3.7
26	90	Hγ2	H-6	2.7	2.4
27	91	Hδ1	H-7	2.9	3.4
28	91	Hδ1	H-8	2.8	4.5

(B) NOE restraints from rigid residues

no.	residue #	protein	ligand	r(exp)	r(calc)
29	26	Hε	H-4	2.7	5.1
30	36	Hδ	H-5	2.9	3.3
31	36	Hε	H-5	3.5	3.9
32	46	Hδ	H-2	3.5	4.9
33	46	Hε	H-4	3.1	3.2
34	46	Hε	H-2	2.6	3.4
35	59	Hη2	H-3	3.4	3.7
36	59	Hη2	H-2	3.7	6.4
37	59	HN	H-1	3.5	6.4
38	59	HN	H-2,H-3	3.5	6.3
39	59	Hζ2	H-3	3.7	3.2
40	59	Hζ3	H-3	3.4	4
41	59	Hζ3	H-2	3.7	6.4
42	99	Hε	H-4	3.3	5.8
43	99	Hε	H-2	3.4	6.2

(C) Excluded NOE restraints

no.	residue #	protein	ligand	r(exp)	remark
44	37	Hβ1	H-4	3.8	weak
45	55	Hγ1	H-4	3.9	weak
46	55	Hγ2	H-1a	4.2	weak
47	55	HN	H-2,H-3	4.1	weak
48	56	Hγ11	H-4	3.5	weak
49	56	Hγ12	H-1	3.6	weak
50	56	Hγ2	H-1a	3.9	weak
51	56	Hγ2	H-1b	3.9	weak
52	56	Hγ2	H-4	3.9	weak
53	57	HN	H-2,H-3	3.8	weak
54	59	Hζ2	H-2	3.9	weak
55	60	HN	H-2,H-3	3.8	weak
56	90	Hβ	H-7	3.7	weak
57	90	Hβ	H-8	4.1	weak
58	90	Hδ1	H-4	3.6	weak
59	90	Hγ11	H-7	4.2	weak
60	90	Hγ2	H-4	3.8	weak
61	91	Hδ1	H-2,H-3	3.4	weak
62	91	Hγ11	H-7	3.9	weak
63	91	Hγ12	H-7	3.7	weak
64	82	Hε	H-1a	3.4	uncertain
65	99	Hδ	H-4	3.7	uncertain
66	99	Hζ	H-1a	3.4	uncertain

The position and orientation of the ligand is similar in all structures (overlay and individuals in Figure 3.4), suggesting that the conformation of the loops does not strongly influence the protein-ligand interaction. 33 out of 43 input NOEs were satisfied, resulting in a well-defined ligand orientation.

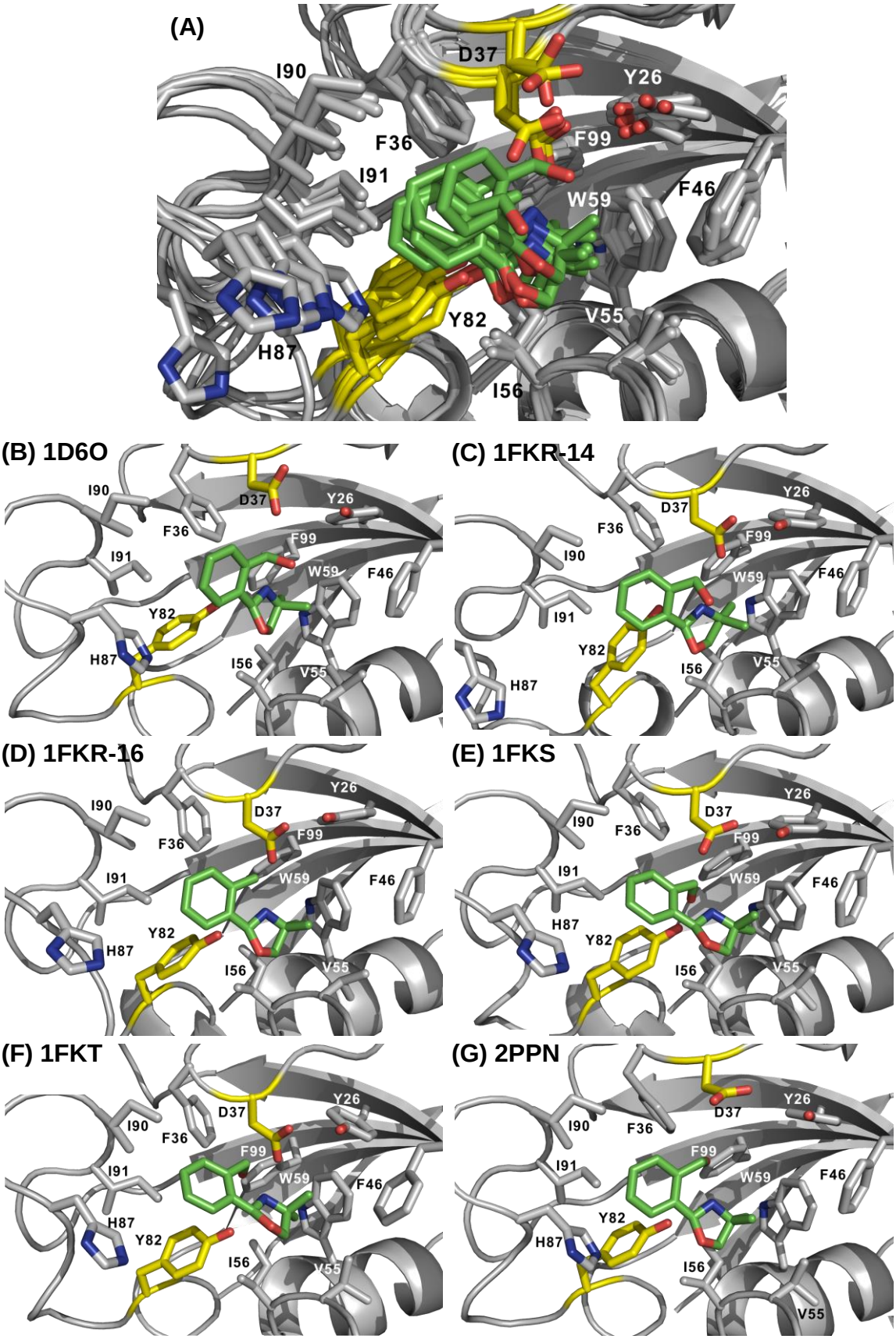

Figure 3.4: NOE-based structures calculated with selected PDB files. The side chains showing possible hydrogen bonding to ligand 1 are colored in yellow. The side chains of other residues showing intermolecular NOEs to the ligand are colored grey. (A) Overlay of the lowest energy

structures of the complex of **1** (in green sticks) with 6 previously determined structures of FKBP12 (1D6O, 1FKR models 14 and 16, 1FKS, 1FKT and 2PPN) as determined by intermolecular NOE restraints. The RMSD of all ligand atoms relative to the mean is 1.1 ± 0.4 Å. Each individual structure is presented in (B)-(G): (B) 1D6O (C) 1FKR model 14 (D) 1FKR model 16 (E) 1FKS (F) 1FKT (G) 2PPN.

However, some large violations of NOEs remain (Table 3.1 and Figure 3.5). The NOEs indicate **1** has contacts with residues located in site 1. The ligand interacts with the hydrophobic pocket formed by F36, F46, V55, I56, W59, I90, I91 and F99. The violations could be due to the motion of the flexible loop at residues 50-56 and 78-95 which are on opposite sides of the hydrophobic pocket.[163,168,171–173] As no single ligand orientation can satisfy all the restraints, it is possible that the dynamic behavior of the protein gives rise to time-averaged NOEs. Consistent with this idea, the backbone[163] and side chains[173] of FKBP12 were also shown previously to undergo chemical exchange on the μs-ms timescale. Alternatively, it is also possible that multiple ligand orientations are present. Therefore, the NOE structure presented here is only an approximation of the actual ligand binding mode. Analysis of the structures indicates possible hydrogen bonds present between (1) the ligand hydroxyl and Y82-hydroxyl, (2) the ligand hydroxyl and D37-Oδ and (3) the ligand nitrogen and Y82-hydroxyl. Previous studies have suggested that the side chains of D37[20,25] and Y82[15] are involved in hydrogen bonding with rapamycin and FK-506, two high affinity ligands of FKBP12. Our NOE-based structure model shares a similar hydrogen bonding framework with the previous studies.

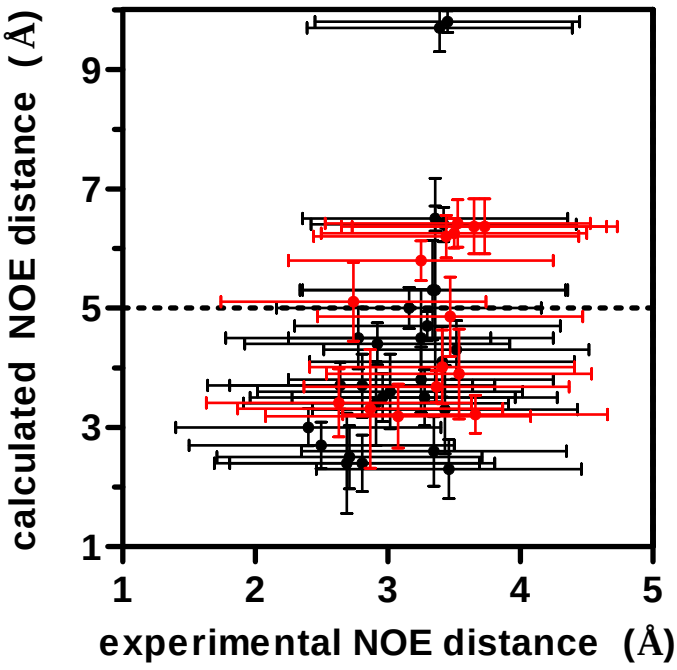

Figure 3.5: Correlation of experimental and back-calculated distances from NOE. Red symbols, restraints from rigid residues; black symbols, restraints from flexible residues. Horizontal error bars represent ±1 Å. Vertical error bars represent 1× standard deviation for the calculated structures. Horizontal dashed line represents a distance of 5 Å.

3.4 Conclusions

We have successfully determined the binding pose of ligand **1** bound to FKBP12 by intermolecular NOE restraints. Not all NOEs were satisfied in the final structure, indicating a certain extent of dynamics. Nevertheless, the final structure is convergent among different PDB structures and all the NOE crosspeaks indicate that ligand **1** binds at site 1. This final structure is used in Chapter 4 as a reference for an orthogonal comparison for the PCS-derived structure.