NMR structural studies of protein-small molecule interactions

Shah, D.M.

Chapter 4

Rapid protein ligand costructures from sparse NOE data

*Dipen M. Shah,[‡,†] Eiso AB,[†] Tammo Diercks,[φ] Mathias A. S. Hass,[‡] Nico A. J. van Nuland,[§] and Gregg Siegal[‡,†]**

[‡]Leiden Institute of Chemistry, Leiden University, Leiden 2300RA, the Netherlands.

[†] ZoBio BV, Leiden 2300RA, the Netherlands.

[φ] Centre for Co-operative Research in Biosciences, Derio 48160, Spain.

[§] Jean Jeener NMR Centre, Structural Biology, Vrije Universiteit Brussel and Molecular Recognition Unit, VIB Department of Structural Biology, Brussels 1050, Belgium.

This chapter was published as a research article:
Shah, D.M.; Ab, E.; Diercks, T.; Hass, M.A.S.; van Nuland, N.A.J.; Siegal, G.
Journal of Medicinal Chemistry, Volume 55, Number 23, pp. 10786-10790 (2012)

ABSTRACT

An efficient way to rapidly generate protein-ligand co-structures based on solution-NMR using sparse NOE data combined with selective isotope labeling is presented. A docked model of the 27 kDa N-terminal ATPase domain of Hsp90 bound to a small molecule ligand was generated using only 21 intermolecular NOEs which uniquely defined both the binding site and the orientation of the ligand. The approach can prove valuable for the early stages of fragment based drug discovery.

INTRODUCTION

The availability of 3D structural information on protein-ligand complexes has become an important driver to guide the pre-clinical stages of drug discovery. Structure based drug design (SBDD) has enabled the development of a wide array of drugs that are currently on the market such as inhibitors of the HIV protease and kinases.[1,2] In the past 15 years fragment-based drug discovery (FBDD), which uses very small, soluble drug fragments as starting points to develop new medicines, has become widespread. Due to their small size, fragments typically bind the target with low affinity (K_D > 10 μM). The development of weak fragment hits to more potent lead-like structures is 2-3 times more successful when 3D structural information is available.[3] Thus the success of both SBDD and FBDD is heavily dependent upon the availability of structural information. Presently both approaches are primarily driven by X-ray crystallography. Crystallography has the advantage that when successful, it is both rapid and high resolution. However, there are a number of cases in which crystallography is not successful such as when crystals are not available, the crystal packing precludes access to the active site or most commonly, weakly binding fragments simply do not give rise to electron density. In the latter case there can be many causes including insufficient occupancy of the binding site or multiple possible binding orientations. In principle, NMR-based solution methods can also generate atomic resolution structural information and NMR has indeed successfully supported FBDD campaigns.[4] Since NMR is extremely sensitive to weak protein-ligand interactions, it should be applicable exactly where crystallography is least effective, i.e. for complexes of weakly binding ligands. However, traditional NMR approaches involving uniform isotopic labeling are labor intensive and limited to proteins of moderate size (e.g. < 30 kDa) and have therefore not been widely adopted in drug

discovery.

In drug discovery it is often the case that the 3D structure of a target or a homologous protein is available. If the resonance assignment of such a protein is available, mapping of chemical shift perturbations (CSPs) onto the 3D structure is a simple, fast procedure that can provide low resolution information about a small molecule binding site.[5] CSP mapping is most commonly accomplished through analysis of ^{15}N- or ^{13}C-edited heteronuclear correlation spectra of the protein. However, the interpretation of ^{1}H-^{15}N or ^{1}H-^{13}C HSQC CSPs becomes ambiguous when the chemical shift perturbations are caused by changes in protein dynamics or a shift in equilibrium between two (or more) conformations. Various computational methods primarily based on chemical shift perturbation analysis are most successful at determining the location of the binding site to low resolution. The precise nature of the intermolecular contacts are well beyond the capabilities of structural analysis based on CSPs.[6,7] In principle a limited set of intermolecular NOEs could provide sufficient information to determine the orientation and binding mode of a small molecule bound to a protein. Indeed, a number of approaches that use a combination of amino acid selective labeling and intermolecular NOEs have been proposed.[8-10] While these are quite powerful, they suffer from one or more of the following limitations: (i) requirement for *a priori* knowledge of the ligand binding site in order to be able to efficiently select the right combination of residues to label, (ii) requirement for a large number of intermolecular NOE contacts, (iii) the resonance assignment uses the pattern of chemical shift perturbations induced by ligand binding and therefore may not be robust and (iv) requirement for extensive calculations to generate structures that match the experimental data. We sought an NMR based method capable of providing structures of sufficient resolution and reliability to

support elaboration of “weakly” binding fragments to target proteins that met the following criteria: widely applicable to small and large proteins, rapid (2-3 weeks per complex), standard solution conditions and NMR experiments, requires samples that can be easily made and the data interpretation is unambiguous. Here, we demonstrate such a method based on selective ILV methyl-labeling and a sparse set of intermolecular NOEs.

RESULTS AND DISCUSSION

The approach as outlined in Figure 1, requires an NMR sample of the protein in which the target protein is highly deuterated and where NMR visible, isotopic labels have been introduced along the backbone and selectively at ILV sidechains as described by Tugarinov and Kay.[11] Standard, through-bond double and triple-resonance NMR spectra are used to obtain the backbone and ILV sidechain resonance assignments of the protein. The ligand is then titrated into the protein in small increments such that the shift in both backbone and sidechain resonance positions can be readily followed using 2D heteronuclear correlation spectra. Once binding of the compound to the protein is saturated, ^{13}C- and ^{15}N-edited NOESY-HSQC spectra are recorded to generate a set of intermolecular NOEs between the ligand resonances and ILV methyl groups as well as the backbone amide protons of the protein. A known 3D structure or structures of the protein target (or a homology model) is then used for molecular docking using HADDOCK to obtain the protein-ligand structure based solely on the intermolecular NOE restraints.[12] Given that the isotopic labeling method has been used to obtain assignments for large proteins, it is likely that the proposed scheme can also be used to determine experimentally derived molecular models based on sparse NOE data for large protein-small

molecule complexes in an efficient yet reliable manner.[11,13-15] Here the approach is illustrated using the moderately sized 27 kDa N-terminal domain of the molecular chaperone heat shock protein 90 (Hsp90), a known cancer target.

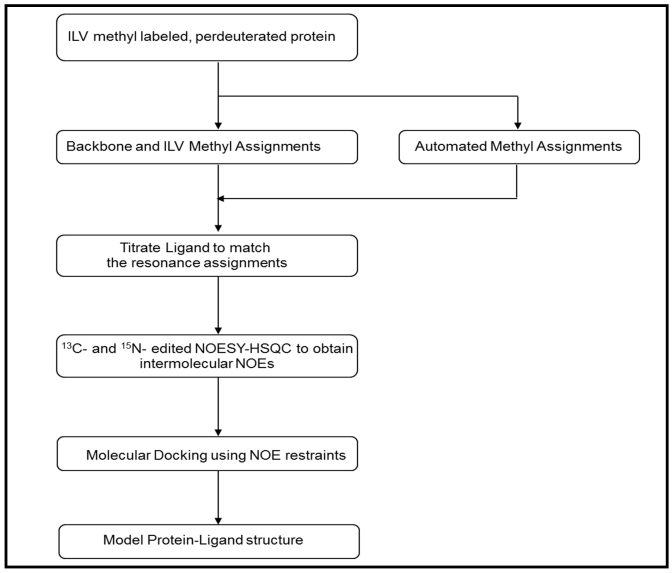

Figure 1. A flowchart scheme to illustrate the number of steps involved in the determination of the model protein-ligand complex by solution NMR based on the sparse NOE data using ILV methyl labeling approach.

The N-terminal domain of Hsp90 (9-233) was isotopically labeled in *E. coli* according to the published procedure.[11] Heteronuclear correlation experiments yielded highly resolved spectra with excellent sensitivity in which all expected methyl peaks can be observed (Figure 2). The triazine **1** had been discovered as a ligand of Hsp90 in a target immobilized NMR screen (TINS) of a fragment library constructed from commercially available compounds.[16,17] The binding of **1** to Hsp90 was confirmed using surface plasmon resonance technology and the equilibrium K_D was determined to be 58 μM *via* fitting to a 1:1 binding model (data not shown). We investigated the effect of **1** binding to the protein by recording a high-resolution CT-[^{1}H,^{13}C] HSQC spectrum in the presence and absence of the compound (Figure

2).[18,19] The methyl chemical shift perturbations are clear evidence of the binding of **1** to Hsp90. The peak pattern in the bound state is suggestive of a folded protein indicating that the binding is specific and does not result in protein denaturation or other undesirable affects. Surprisingly, the majority of the methyl resonances are affected by the binding of **1** and a similar pattern was observed in the [^{1}H, ^{15}N]-HSQC (not shown). Using the sequential assignment determined below, the methyl and backbone amide chemical shift perturbations (CSPs) have been mapped onto the crystal structure of Hsp90 (Figure 3). Although the CSPs surround the known ATP binding site, residues far from the site are also significantly affected. The widespread changes in the spectrum of Hsp90 could be indicative of conformational changes in the protein induced or stabilized by the binding of compound **1.** Indeed, conformational changes have been observed for a number of ligands binding to Hsp90.[20] Thus it can be difficult to determine the ligand binding site simply on the basis of CSPs (see below). Furthermore, the orientation of the ligand within the binding site, and therefore the nature of the protein-ligand interactions, is not defined by the CSP information.

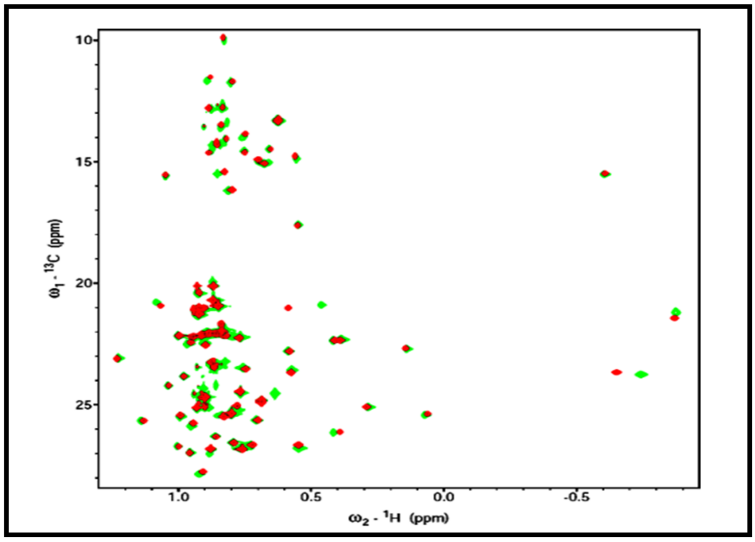

Figure 2. CT- [^{1}H,^{13}C] HSQC spectra of an ILV methyl protonated sample of deuterated N-terminal domain of Hsp90 recorded in the absence (in red) and in the presence of **1** (in green) at a protein to ligand ratio of 1:6. Significant chemical shift perturbations for various methyl groups are seen clearly in the presence of **1**, indicative of binding.

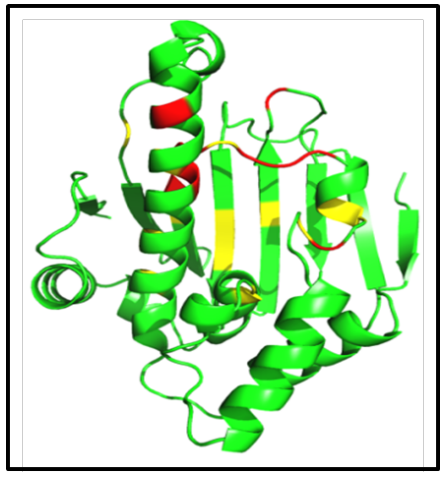

Figure 3. The methyl (yellow) and amide (red) chemical shift perturbations caused by the presence of **1** are mapped on the crystal structure of Hsp90 (PDB 1YER). The figure was created in PyMOL.[21]

The sequential backbone assignment of Hsp90 is available (e.g. BMRB 5355) but was confirmed by analysis of TROSY-based HNCACB and HNcoCACB NMR spectra. In total, 76% of the HN, N, Cα and Cβ chemical shifts were sequentially

assigned, including 87% of the ILV residues. The methyl resonances were subsequently correlated with the intraresidue Cα and Cβ assignments by a CCH-TOCSY experiment. Of the ILV residues that had been sequentially assigned, 92% of the ILV methyl assignments were obtained in a straightforward manner. It is not essential to obtain all ILV methyl assignments since only those within the binding site will give rise to intermolecular NOEs. Based on the crystal structure of a complex of **1** with Hsp90 (see below), all methyl groups that are within NOE distance (8 Å using a highly deuterated protein) to the ligand have been assigned using only data from these three experiments. Stereospecific methyl assignments for Leu and Val residues were obtained by producing a 10% ^{13}C labeled sample and CT-HSQC analysis as described previously.[22]

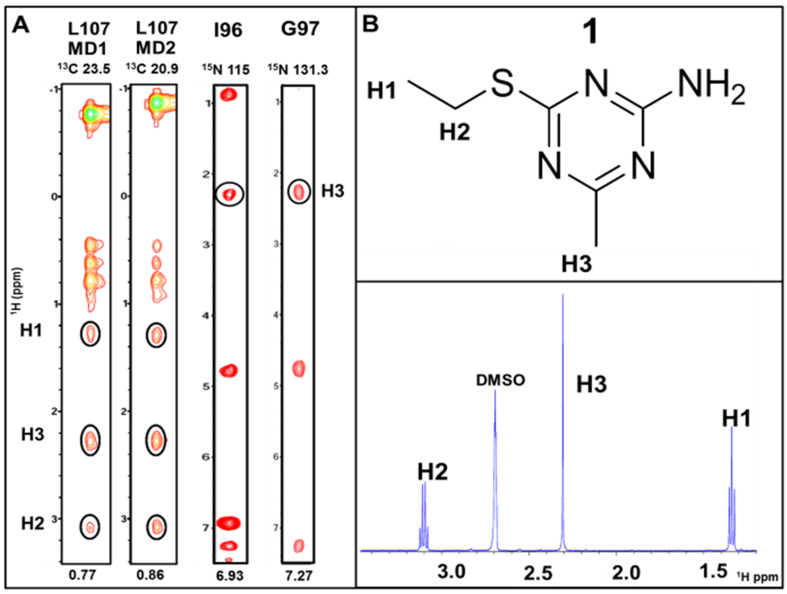

Figure 4. A] Representative strips from the 3D ^{13}C- and ^{15}N-edited NOESY-HSQC spectra of methyl protonated {I($^{13}CH_3$, δ1 only), L($^{13}CH_3$,$^{12}CD_3$), V($^{13}CH_3$,$^{12}CD_3$)} U-[^{15}N,^{13}C,^{2}H] Hsp90 in the presence of **1**. The intermolecular NOEs between the methyl groups of the protein (L107MD1, L107MD2) and the ligand H1, H2, H3 groups are circled. The strips for residues I96 and G97 from a ^{15}N-edited NOESY-HSQC spectrum are shown from which intermolecular NOEs to the H3 group of **1** are circled. The frequencies (ppm) of ^{1}H and ^{13}C or ^{15}N nuclei are shown at the bottom and the top of the strips respectively. B] The structure (top) and 1D ^{1}H spectrum (bottom) of **1** in D_2O with the resonance assignment.

3D ^{13}C-edited and ^{15}N-edited NOESY-HSQC spectra were recorded on the protein-ligand complex in order to detect intermolecular NOEs.[23,24] Due to the selective labeling scheme employed, there was minimal overlap amongst the protein resonances and therefore intermolecular NOEs were unambiguously assigned. Selected strips from the ^{13}C-edited NOESY-HSQC and ^{15}N-edited NOESY-HSQC spectra exhibiting intermolecular NOEs from the ligand are shown in Figure 4A. A total of 21 intermolecular NOEs were identified between resonances of **1** and protein residues including three to amide protons (I96, G97 and M98) and 18 NOEs to methyl groups (L107, L103, V150, V186 and I96). The peak intensities from the intermolecular NOEs were converted to distances and used as restraints to carry out the docking of the ligand using HADDOCK (See experimental section for more information).[12]

Typical of many pharmaceutical targets, a crystal structure of apo-Hsp90 (PDB 1YER) is available in the PDB. This structure was the starting point for the experimentally restrained docking procedure. We performed HADDOCK calculations using only the apo-Hsp90 structure and 21 intermolecular-NOE distances as unambiguous restraints (Table 1).[12] The lowest energy cluster exhibited a HADDOCK score and minimal restraint violation energy of -7.0 and 3.8 kcal/mol respectively, whereas the next ligand cluster exhibited a HADDOCK score of 3.4 and minimal restraint violation energy of 20.0 kcal/mol. The difference between the clusters indicates that the input data defines a single set of structures.

Table 1. Intermolecular NOE distance restraints included into HADDOCK calculations are shown. Note that a lower and upper distance margin of 1-2 Å were employed to above target distances.

Ligand atom	Protein residue	Protein residue atom	Target distance (Å)
H1	I96	HD1	3.8
H1	L103	HD+	4
H1	L107	HD1	4.3
H1	L107	HD2	4
H1	V150	HG1	3.8
H1	V150	HG2	3.8
H1	V186	HG2	4.3
H2	I96	HD1	4
H2	L103	HD+	5
H2	L107	HD1	4.6
H2	L107	HD2	4
H2	V150	HG1	5.3
H2	V186	HG2	5.3
H3	L103	HD+	4.9
H3	L107	HD1	3.9
H3	L107	HD2	3.6
H3	V150	HG1	4.8
H3	L186	HG2	5
H3	I96	HN	4
H3	G97	HN	4
H3	M98	HN	4

A crystal structure of the complex of **1** with Hsp90 was available in the PDB (3B24).[25] In 3B24 there are two protein molecules in the asymmetric unit and although **1** is bound in the same site in each, it is rotated by approximately 180^{o} in the different structures (Figure 5A). Superposition of the protein in the three structures (NMR model and two crystal structures) indicates that the sparse NOE method defines the binding site of the ligand uniquely and accurately. Most Hsp90 ligands form critical hydrogen bonds to residues D93 and T184 and these are present in both the X-ray structures and the NMR model. The orientation of **1** in the NMR based model is quite similar to one of the orientations in the asymmetric unit of the crystal structure (1.7 Å rmsd). In the NMR model, the orientation of **1** is defined by a network of 3 intermolecular NOEs from the ligand H3 group to the amide protons of residues I96,

G97 and M98 (Figure 5B) and is thus unambiguous. Interestingly, PDB entries 2WI2 and 2WI3, are structures of a similar triazine bound to Hsp90 in which two orientations differing by a 180^{o} axial rotation are found. In this case the different orientation was dependent on whether soaking or co-crystallization was used to form the complex.[26] Taken together the crystal structures suggest that there are two or more low energy conformations for the complex of these simple triazines with Hsp90. In contrast, the NMR data indicate that in solution, there is one predominant conformation and this is likely to be the most physiologically relevant. Importantly, despite the moderate number of intermolecular restraints, the NMR model defines precisely the same binding site as the crystal structure and the critical intermolecular hydrogen bonds suggesting that in addition to being fast, the method is robust.

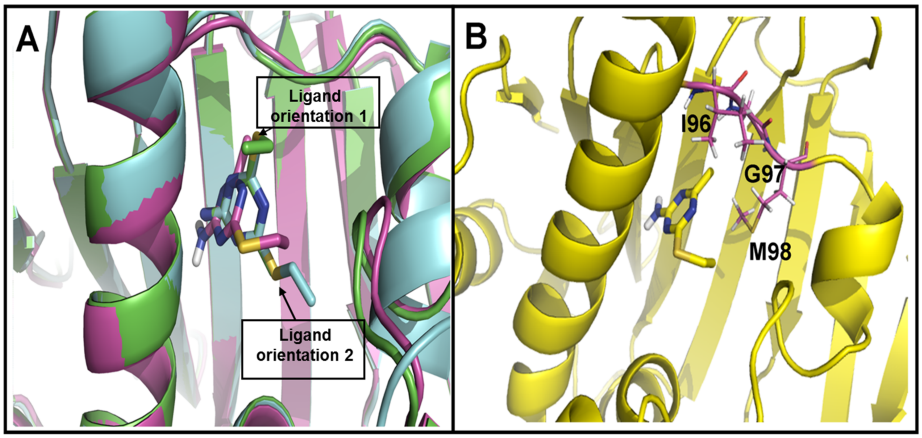

Figure 5. A] Overlay of the lowest energy HADDOCK model of **1** bound to Hsp90 with a crystal structure of the complex (PDB 3B24) in which **1** binds in two orientations (docked model- magenta, 3B24: green and sky blue). The orientation of **1** in the NOE-based model is similar to one of the binding modes (seen in sky blue color) of the ligand in PDB entry 3B24. B] The orientation of **1** in the docked model is largely determined by the three intermolecular NOEs observed between the H3 group of the ligand and HN of I96, G97 and M98 shown in pink within the protein. The figure was created in PyMOL.[21]

Numerous groups have attempted to use CSPs to determine the structure of protein-ligand complexes.[7,27] We used the ability of HADDOCK to include CSPs as ambiguous restraints for docking and modeled the structure of the Hsp90-**1** complex

using the most significant methyl and backbone amide CSPs (Figure 6). The HADDOCK calculations generated 2 clusters whose HADDOCK scores were similar and significantly lower than others. Interestingly, the clusters are both located at the binding site of **1** defined in both the NOE model and the crystal structures. However, despite the fact that electrostatics were used during the HADDOCK calculation, the hydrogen bonds to D93 and T184 are not present. In fact, the amine group of **1** is pointing away from these residues.

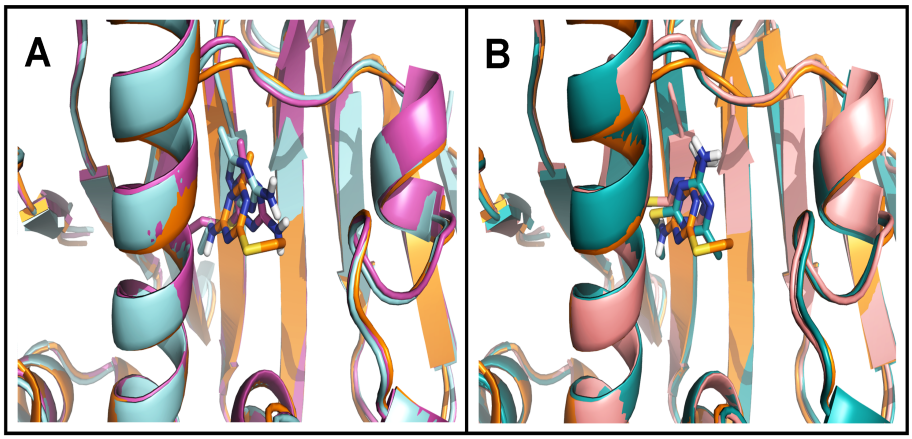

Figure 6.Overlay of the lowest energy HADDOCK model structures of **1** bound to Hsp90 obtained using the most significant methyl and backbone amide chemical shift perturbations ($\Delta\delta_i > \Delta\delta_{avg}$ + 2 x standard deviation). Two clusters of **1** are observed which have similar HADDOCK score. A] Overlay of the two structures from one of the clusters (magenta and sky blue) with NOE-derived docked model shown in orange. B] Overlay of the two structures from the other cluster (pink and dark green) with NOE-derived docked model shown in orange. The figure was created in PyMOL.[21]

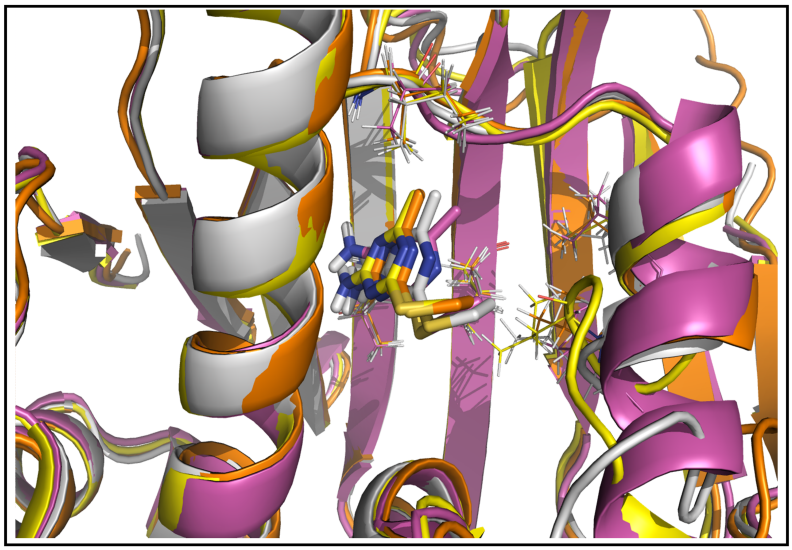

Figure 7. Overlays of the lowest energy HADDOCK model structures are shown when **1** is docked with various N-term Hsp90 structures with different sidechain conformations (shown as lines) in the PDB (1YER- yellow, 1OSF-orange, 2QF6-magenta and 3B26-grey). The ligand is found to bind in the same binding spot in comparison with the X-ray structure, irrespective of the conformational flexibility of the helix formed by residues 100-124 seen in the various PDB structures. The figure was created in PyMOL.[21]

Frequently ligand binding is accompanied by a conformational change in the protein. Conformational rearrangement of residues 100-124, which are in close proximity to the ligand binding site of Hsp90, has been observed for a number of ligands including in the two structures in the unit cell of 3B24 (Figure 5A). To investigate any possible influence of conformational rearrangement on the structures calculated from NMR data, we selected 3 different Hsp90 structures which differ in the conformation of residues 100-124 for docking using HADDOCK.[12] The binding site and orientation of the ligand cluster in each of the 3 different Hsp90 structures was very similar to that determined using the apo-protein (Figure 7). Importantly, the intermolecular hydrogen bonds to D93 and T184 are conserved in all 4 experimentally constrained docking structures. This data suggests that, at least for the case of Hsp90, binding induced conformational changes do not preclude

determining the essential features of a small molecule-protein complex even if those changes are unknown.

CONCLUSION

The method presented here appears to be capable of generating reliable protein-ligand structures in a quick and efficient manner. We have shown that a universal selective labeling scheme can be used to rapidly identify sufficient numbers of restraints for a small molecule ligand weakly bound to protein. Moreover, universal labeling precludes the necessity for *a priori* knowledge of the ligand binding site. However, the present method depends on the availability of selectively labeled, deuterated protein which can only be produced in *E. coli*. Note that the 3 NOEs involving backbone amide protons of the protein were critical for orienting **1** in the protein-ligand complex. This observation suggests that, at least in some cases, the use of exclusively methyl labeled protein could result in the loss of important constraints for the calculation procedure. Another obvious limitation is the requirement for methyl groups at the ligand binding site. Although ILV residues tend to be well located in proteins and clearly Hsp90 is a good example of this, it is not always the case.[13] However, it is also possible to selectively label all methyl containing residues providing even more complete coverage of protein structures and ligand binding sites.[28] While we have used standard, through-bond NMR techniques for resonance assignment here, recently steps towards automating the assignment of methyl resonances based either on intramolecular NOEs and/or through-space paramagnetic effects have been taken.[29] Implementation of selective methyl labeling in conjunction with automated resonance assignment should enable structure determination of complexes involving reasonably large proteins up to 75 kDa. We feel the present method could prove valuable for the early stages of FBDD by

providing 3D structure information on weakly binding fragment-protein complexes.

Experimental Section:

Sample Preparation

A methyl protonated {I($^{13}CH_3$, δ1 only), L($^{13}CH_3$,$^{12}CD_3$), V($^{13}CH_3$,$^{12}CD_3$)} U-[^{15}N,^{13}C,^{2}H] sample of N terminal domain (9-233) of Hsp90 was obtained by protein overexpression from a culture of *E. coli* BL21 (DE3) cells transformed with the plasmid pQTEV as described.[11] The protein was expressed in 1 L D_2O M9 medium supplemented with 2 g/L of U-[^{13}C, ^{2}H]-glucose (Sigma) as the main source of carbon and 0.3 g/L of $^{15}NH_4Cl$ (Sigma) as the nitrogen source. One hour prior to induction at ~0.5 OD, 50 mg of 2-keto-3,3-d^2-1,2,3,4-^{13}C-butyrate and 100 mg of 2-keto-3-methyl-d^3-3-d^1-1,2,3,4-^{13}C butyrate (^{13}C labeled α-ketoisovalerate deuterated at the β-position and with one of the methyl groups - $^{12}CD3$) were added to growth medium. The expression of the protein was induced by the addition of 1 mM IPTG for 16 hours at 18°C. The protein was purified on Ni-Hitrap FF (GE Lifescience) followed by purification on a superdex G75 gel filtration column (GE Lifescience). The seven-histidine tag was removed by proteolytic digestion with TEV protease (Invitrogen) according to the manufacturer recommended buffer condition. The cleaved tag and the TEV protease were removed by affinity purification on Ni-Hitrap FF, and subsequently the buffer of Hsp90 (9-233) was exchanged to 20 mM sodium phosphate pH 7.5, 50 mM NaCl. The NMR sample was 0.7 mM in protein, 95% H_20, 5% D_20, 20 mM sodium phosphate, pH 7.5, 50 mM NaCl. The NMR sample on the Hsp90-**1** complex was obtained by addition of 6 equivalents of **1** to the protein. The pH of the sample was adjusted carefully within +/- 0.05 units of 7.5 after addition of the ligand. Stereospecific methyl assignments were obtained by producing a 10% ^{13}C

labeled sample and CT-HSQC analysis as described. [22]

Compound **1** was purchased from Maybridge, catalog number RF 02707.

NMR backbone sequential assignments and ILV methyl assignments

In order to detect binding of **1** to Hsp90, a CT- [^{1}H,^{13}C] HSQC spectrum of 0.7 mM protein and protein plus ligand (1:6 ratio) was recorded on a Varian 800 MHz spectrometer equipped with a cryogenic probe. The chemical shifts of **1** in the bound state were also followed by acquiring 1D ^{1}H proton spectrum in the presence of protein at different protein plus ligand ratios (1:0, 1:2, 1:4 and 1:6).

In order to perform sequential assignment of the backbone resonances and correlate them to the intraresidue ILV methyl resonances of the Hsp90-**1** complex, the following experiments were acquired at 298K on 600MHz four-channel Bruker DMX NMR spectrometer equipped with a TXI cryo-probe; [^{1}H,^{15}N]-HSQC, trHNCACB recorded with (2048x80x110) in (^{1}HN,^{15}N,^{13}C) dimensions, trHNcoCACB recorded with (2048x80x180) in (^{1}HN,^{15}N,^{13}C) dimensions with a recycle delay of 1s and with 16 scans/FID.[30,31] To correlate the backbone Cα and Cβ frequencies to the intraresidue methyl resonances of ILV residues, a CCH-TOCSY experiment was recorded with (2048x80x152) in (^{1}H$_{m}$,^{13}C$_{m}$,^{13}C) dimensions with a recycle delay of 1s and with 8 scans/FID. The CCH-TOCSY pulse sequence was obtained by modification of a standard (H)CCH-TOCSY pulse sequence by removal of the initial INEPT pulses, addition of deuterium decoupling during the *t1* and *t2* period, starting with a direct carbon excitation pulse, followed by a DIPSI-3 TOCSY transfer sequence and final detection on the methyl protons. The optimal mixing time for the TOCSY sequence was found to be 23.3 ms.[32]

A series of 2D ^{1}H-^{1}H NOESY planes with mixing time between 100-600 ms were acquired in order to determine the optimal mixing time based on the decay of cross

peak intensity. The optimal mixing time was defined to be a balance of being able to detect long range NOEs and spin diffusion and was found to be 400 ms. In order to observe intermolecular NOE cross peaks between the ligand and the protein, a ^{13}C-edited NOESY-HSQC (1024x40x160 points in ^{1}H,^{13}C,^{1}H respectively) and a ^{15}N-edited NOESY-HSQC (1024x44x160 points in ^{1}H,^{15}N,^{1}H respectively) were recorded with , mixing time of 400 ms and a recycle delay of 1s and with 16 scans/FID respectively.[33-36] The NOESY experiments were acquired at 298K on a 900 MHz four-channel Bruker DMX spectrometer equipped with pulsed field gradient, conventional room-temperature triple resonance probe. The cross-peak volumes/peak intensities from the intermolecular NOEs were converted to distances based on the equation $V=C/b^{n}$ (V= cross-peak volume/peak intensity, C is a constant, b is the upper distance bound and *n* is an exponent 4).[37] The acquired NMR data was processed using the NMRpipe/NMRDraw software package and visualized in Sparky.[38,39] The backbone assignment process was guided by the predicted chemical shifts calculated by SHIFTX in the automatic assignment program MARS using PDB 1YER as input.[40,41] The use of ILV methyl labeling is attractive for a number of reasons including: (i) the favorable relaxation properties of methyl groups so that NMR spectra, even for larger systems, are of higher sensitivity and resolution,[14,42] (ii) the well-dispersed distribution of the methyl groups throughout the protein structure and particularly at ligand binding sites,[43] (iii) restriction of the NMR analysis to the backbone and sidechain resonances of ILV methyl groups of the protein simplifies the process and allows unambiguous identification of intermolecular NOEs and (iv) methods to produce ILV labeled samples in *E. coli* are robust and economical.[11,44,45]

Protein-Ligand Docking

Throughout the docking procedure using HADDOCK, 1000 protein structures were used for initial rigid body docking. Explicit solvent refinement was performed on 200 structures after initial docking iterations. The final docking solutions were selected by clustering of the structures and analysis of the HADDOCK scores. The Hsp90 structures used in HADDOCK calculations to investigate the conformational rearrangement of the helix formed by residues 100-124 are derived from a number of different complexes for which the ligand was removed prior to docking to compound **1**.[13]

Contributions

Tammo Diercks (CiCBiogune, Bilbao), Mathias A. S. Hass (Leiden University) and Nico A. J. van Nuland (Vrije Universiteit, Brussels) helped to set-up NMR experiments at the respective NMR facilities. Eiso AB (ZoBio BV, Leiden) assisted with the NMR data analysis and molecular docking. We also thank Dr. Hans Wienk for his assistance and the SONNMRLSF in Utrecht. We also acknowledge the contribution to Hsp90 crystallography from Thomas A. Ceska at UCB, UK.

REFERENCES

(1) Volarath, P.; Harrison, R.W.; Weber, I.T. Structure based drug design for HIV protease: from molecular modeling to cheminformatics. *Curr. Top. Med. Chem.* **2007**, 7(10), 1030-1038.

(2) Hardy, L. W.; Malikayil, A. The impact of structure-guided drug design on clinical agents. *Curr. Drug Discov.* **2003**, 3, 15–20.

(3) Hajduk, P.J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discov.* **2007**, 6, 211-219.

(4) Oltersdorf, T.; Elmore, S.W.; Shoemaker, A.R.; Armstrong, R.C.; Augeri, D.J.; Belli, B.A.; Bruncko, M.; Deckwerth, T.L.; Dinges, J.; Hajduk., P.J; Joseph, M.K.; Kitada, S.; Korsmeyer, S.J.; Kunzer, A.R.; Letai, A.; Li, C.; Mitten, M.J.; Nettesheim, D.G.; Ng, S.; Nimmer, P.M.; O'Connor, J.M.; Oleksijew, A.; Petros, A.M.; Reed, J.C.; Shen, W.; Tahir, S.K.; Thompson, C.B.; Tomaselli, K.J.; Wang, B.; Wendt, M.D.; Zhang, H.; Fesik, S.W.; Rosenberg, S.H. An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature*, **2005**. 435, 677-681.

(5) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science*. **1996**, 274, 1531-1534.

(6) Krishnamoorthy, J.; Yu, VCK.; Mok, Y-K. Auto-FACE: An NMR based binding site mapping program for fast chemical exchange protein-ligand systems. *PLoS ONE.* **2010,** 5(2), e8943.

(7) Schieborr, U.; Vogtherr, M.; Elshorst, B.; Betz, M.; Grimme, S.; Pescatore, B.; Langer, T.; Saxena, K.; Schwalbe, H. How much NMR data is required to determine a protein-ligand complex structure? *Chembiochem.* **2005**, 6 (10), 1891-1898.

(8) Pellecchia, M.; Meininger, D.; Dong, Q.; Chang, E.; Jack, R.; Sem, D.S. NMR-based structural characterization of large protein-ligand interactions. *J. Biomol. NMR.* **2002**, 22, 165-173.

(9) Constantine, K.L.; Davis M.E.; Metzler, W.; Mueller, L.; Claus, B.L. Protein-ligand NOE matching: a high-throughput method for binding pose evaluation that does not require protein NMR resonance assignments. *J. Am. Chem. Soc.* **2006**, 128, 7252-7263.

(10) Yu, L.; Sun, C.; Song, D.; Shen, J.; Xu, N.; Gunasekera, A.; Hajduk, P.J.; Olejniczak, E.T. Nuclear magnetic resonance structural studies of a potassium channel-charybdotoxin complex. *Biochemistry.* **2005**, 44**,** 15834–15841.

(11) Tugarinov, V.; Kay, L. E. Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J. Am. Chem. Soc*., **2003**, 125, 13868-13878.

(12) De Vries, S.J.; Van Dijk, M.; Bonvin, A.M.J.J. The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* **2010**, 5, 883-897.

(13) Sprangers, R.; Kay, L.E. Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature*. **2007**, 445, 618–622.

(14) Sibille, N.; Hanoulle, X.; Bonachera, F.; Verdegem, D.; Landrieu, I.; Wieruszeski, J.M.; Lippens, G. Selective backbone labelling of ILV methyl labelled proteins. *J. Bio. Mol.* **2009**, 43 (4), 219-227.

(15) Tugarinov, V.; Choy, W.Y.; Orekhov, V.Y; Kay, L.E. Solution NMR-derived global fold of a monomeric 82-kDa enzyme. *Proc. Natl. Acad. Sci.* **2005**,102, 622–627.
(16) Vanwetswinkel, S.; Heetebrij, R.J.; Van Duynhoven, J.; Hollander, J.G.; Filippov, D.V.; Hajduk, P.J; Siegal, G. TINS, target immobilized NMR screening: an efficient and sensitive method for ligand discovery. *Chem. Biol.,* **2005**, 12, 207-216.
(17) Siegal, G.; AB, E.; Schultz, J. Integration of fragment screening and library design. *Drug Discov. Today*, **2007**, 12, 1032-1039.
(18) Palmer, A.G. III; Cavanagh, J.; Wright, P.E.; Rance, M. Sensitivity improvement correlation NMR-spectroscopy. *J. Magn. Reson.* **1991**, 93, 151-170.
(19) Schleucher, J.; Schwendinger, M.; Sattler, M.; Schmidt, P.; Schedletzky, O.; Glaser, S.J.; Sorensen, O.W.; Griesinger, C. A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J. Biomol. NMR.* **1994**, 4, 301-306.
(20) Murray, C. W.; Carr, M.G.; Callaghan, O.; Chessari, G.; Congreve, M.; Cowan, S.; Coyle, J.E.; Downham, R.; Figueroa, E.; Frederickson, M.; Graham, B.; McMenamin, R.; O'Brien, A.M.; Patel, S.; Phillips, T.R.; Williams, G.; Woodhead, A.J.; Woolford, A.J.-A. Fragment-based drug discovery applied to Hsp90. Discovery of two lead series with high ligand efficiency. *J. Med. Chem*. **2010**, 53(16), 5942-5955.
(21) The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.
(22) Neri, D.; Szyperski, T.; Otting, G.; Senn, H.; Wuthrich, K. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ^{13}C labeling. *Biochemistry.* **1989**, 28, 7510-7516.
(23) Kay, L.E.; Keifer, P.; Saarinen, T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc***. 1992**, 114, 10663-10665.
(24) Davis, A.L.; Keeler, J.; Laue, E.D.; Moskau, D. Experiments for recording pureabsorption heteronuclear correlation spectra using pulsed field gradients. *J. Magn. Reson.* **1992**, 98, 207-216.
(25) Miura, T.; Fukami, T.; Hasegawa, K.; Ono, N.; Suda, A.; Shindo, H.; Yoon, D.O.; Kim, S.J.; Na, Y.J.; Aoki, Y.; Shimma, N.; Tsukuda, T.; Shiratori, Y. Lead generation of heat shock protein 90 inhibitors by a combination of fragment-based approach, virtual screening, and structure-based drug design. *Bioorg. Med. Chem. Lett.*, **2011**, 21, 5778-5783.
(26) Brough, P.A.; Barril, X.; Borgognoni, J.; Chene, P.; Nicholas G. M. Davies, Davis, B.; Drysdale, M.J.; Dymock, B.; Eccles, S.A.; Garcia-Echeverria, C..; Fromont, C.; Hayes, A.; Hubbard, R.E.; Jordan, A.M.; Jensen, R.M.; Massey, A.; Merrett, A.; Padfield, A.; Parsons, R.; Radimerski, T.; Raynaud, F.; Robertson, A.; Roughley, S.; Schoepfer, R.; Simmonite, H.; Sharp, S.; Surgenor, A.; Valenti, M.; Walls, S.; Webb, P.; Wood, M.; Workman, P.; Wright, L. Combining hit identification strategies: fragment-based and in silico approaches to orally active 2-aminothieno[2,3-d]pyrimidine inhibitors of the Hsp90 molecular chaperone. *J. Med. Chem*. **2009**, 52(15), 4794-4809.
(27) Stark, J.; Powers, R. Rapid protein-ligand costructures using chemical shift perturbations. *J. Am. Chem. Soc.* **2008**, 130, 535-545.

(28) Otten, R.; Chu, B.; Krewulak, K. D.; Vogel, H. J.; Mulder, F. A. A. A comprehensive and cost-effective NMR spectroscopy of methyl groups in large proteins. *J. Am. Chem. Soc*. **2010**, 132(9), 2952 – 2960.
(29) Venditti, V.; Fawzi, N.L.; Clore, G.M. Automated sequence- and stereo-specific assignment of methyl-labeled proteins by paramagnetic relaxation and methyl-methyl nuclear Overhauser enhancement spectroscopy. *J. Biomol. NMR.* **2011**, 51, 319-328.
(30) Salzmann, M.; Wider, G.; Pervushin, K.; Senn, H.; Wuethrich,K. TROSY-type triple- resonance experiments for sequential NMR assignments of large proteins. *J. Am. Chem. Soc.* **1999**, 121, 844-848.
(31) Eletsky, A.; Kienhoefer, A.; Pervushin,K. TROSY NMR with partially deuterated proteins. *J. Biomol. NMR.* **2001**, 20, 188-180.
(32) Kay, L.E.; Xu, G.Y.; Singer, A.U.; Muhandiram, D.R.; Forman-Kay, A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^{1}H and ^{13}C correlations in H2O samples of proteins J.D. *J. Magn. Reson.* **1993**, B 101, 333 – 337.
(33) Davis, A.L.; Keeler, J.; Laue, E.D.; Moskau, D. Experiments for recording pureabsorption heteronuclear correlation spectra using pulsed field gradients. *J. Magn. Reson.* **1992**, 98, 207-216.
(34) Palmer, A.G. III; Cavanagh, J.; Wright, P.E.; Rance, M. Sensitivity improvement correlation NMR-spectroscopy. *J. Magn.Reson.* 1991, 93, 151-170.
(35) Kay, L.E.; Keifer, P.; Saarinen, T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc*. **1992**, 114, 10663-10665.
(36) Güntert, P.; Braun, W., Wüthrich, K. Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **1991**, 217, 517-530.
(37) Delaglio, F.; Grzesiek, S.; Vuister, G.W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol NMR*. **1995**, 6(3), 277-93.
(38) Goddard, T.D.; Kneller, D.G. SPARKY 3, University of California, San Francisco.
(39) Neal, S.; Nip, A.; Zhang, H.; Wishart. D. Rapid and accurate calculation of protein ^{1}H, ^{13}C and ^{15}N chemical shifts. *J. Biomol. NMR*, **2003**, 26, 215-240.
(40) Jung, Y.S.; Zweckstetter, M. Mars -- robust automatic backbone assignment of proteins. *J. Biomol.NMR.* **2004**, 30(1), 11-23.
(41) Karagöz, G.E.; Duarte, A.M.; Ippel, H.; Uetrecht, C.; Sinnige, T.; van Rosmalen, M.; Hausmann, J.; Heck, A.J.; Boelens, R.; Rüdiger, S.G. N-terminal domain of human Hsp90 triggers binding to the cochaperone p23. *Proc. Natl. Acad. Sci*. **2011**, 108(2), 580-585.
(42) Hajduk, P.; Augeri, D.; Mack, J.; Mendoza, R.; Yang, J.; Betz, S.; Fesik, S. NMR-based screening of proteins containing 13C-labeled methyl groups. *J. Am. Chem. Soc.* **2000**, 122, 7898–7904.
(43) Sibille, N.; Hanoulle, X.; Bonachera, F.; Verdegem, D.; Landrieu, I.; Wieruszeski, J.M.; Lippens, G. Selective backbone labelling of ILV methyl labelled proteins. *J. Bio. Mol.* **2009**, 43 (4), 219-227.

(44) Tugarinov, V.; Kanelis, V.; Kay, L.E. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat. Protoc.* **2006**, 1, 749-754.
(45) De Vries, S.J.; Van Dijk, M.; Bonvin, A.M.J.J. The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc*. **2010**, 5, 883-897.