Journal of Biomolecular NMR, **22:** 165–173, 2002.
KLUWER/ESCOM

NMR-based structural characterization of large protein-ligand interactions

Maurizio Pellecchia*, David Meininger, Qing Dong, Edcon Chang, Rick Jack & Daniel S. Sem
Triad Therapeutics, Inc. San Diego, 92121, U.S.A.

Received 26 October 2001; accepted 10 December 2001

Key words: drug design, drug discovery, ligand binding, ligand docking, NMR, protein structure

Abstract

Genomic research on target identification and validation has created a great need for methods that rapidly provide detailed structural information on protein-ligand interactions. We developed a suite of NMR experiments as rapid and efficient tools to provide descriptive structural information on protein-ligand complexes. The methods work with large proteins and in particular cases also without the need for a complete three-dimensional structure. We will show applications with two tetrameric enzymes of 120 and 170 kDa.

Introduction

As a result of genomics efforts, the number of protein drug targets is expected to increase 10-fold (Drews, 1998). However, there is a significant gap in going from the sequence of a newly identified drug target, to its three-dimensional structure, and then to a high affinity (< 100 nM) inhibitor. With the overwhelming amount of information arising from functional and structural genomics efforts (Stevens et al., 2001) there is a great need for novel techniques capable of exploiting this information rapidly. While structure-based drug design techniques are certainly powerful tools in the drug discovery and design process, in general these approaches are limited by the uncertainty of success and the time required for the structure determination process. Here we present data demonstrating that a novel structural determination method based on nuclear magnetic resonance spectroscopy, NMR-DOC (Nuclear Magnetic Resonance DOcking of Compounds), has the unique ability to provide in a relatively short amount of time an experimentally derived molecular model describing ligand-protein interactions once the three-dimensional structure of the target protein is known. The structural information obtained allows for better synergistic application of structural genomics with structure-based drug design approaches and will lead to the acceleration of inhibitor design and discovery for a given protein target. In a particularly useful application called NMR-SOLVE (Structurally Oriented Library Valency Engineering), we show that NMR data can be gathered which will have utility in the design of bi-ligand compounds with high-affinity, even without knowledge of the three-dimensional structure of the target protein. Both NMR-DOC and NMR-SOLVE significantly lift the molecular weight limitation of protein NMR, by avoiding the need for complete resonance assignments.

*To whom correspondence should be addressed. E-mail: mpellecchia@triadt.com

Materials and methods

Protein expression, purification and isotope labeling

The gene encoding *E. coli* DHPR in pET11a (Novagen) was obtained from Dr J.S. Blanchard (Albert Einstein College of Medicine, NY). The PCR product was subcloned into pET21a+ (Novagen) using the Nde 1 and Bam H1 restriction sites. BL21 (DE3) Gold *E. coli* (Stratagene) were transduced with the expression constructs and either grown on 2YT media in shaker flasks or D_2O growth adapted and grown on

modified M9 minimal media (Meininger et al., 2000; Metzler et al., 1996) in 1 L fermentations primarily as described (Meininger et al., 2000) using a BioFlo 3000 fermentor (New Brunswick Scientific). Natural-abundance and selectively-labeled protein was purified via Q Sepharose Fast Flow anion-exchange chromotagraphy and Blue Sepharose-affinity chromatography (Amersham Pharmacia Biotech). Maximum purified yields for DHPR were >650 mg l^{-1} for unlabeled protein and >450 mg l^{-1} for selectively-labeled protein. *E. coli* DHPR mutant-containing plasmids were produced with the QuickChange^{TM} Site-Directed Mutagenesis Kit (Stratagene). Mutant protein was purified by the same protocol used for wild-type protein.

The DOXPR gene was cloned from *E. coli* gDNA by PCR utilizing the following primers: 5′-gccactgcatatgaagcaactcaccattctgg and 3′-gccactgggatcc tcagcttgcgagacgcatc. We expressed and purified DOXPR as reported (Meininger et al., 2000). Maximum purified yields obtained for DOXPR were >500 mg l^{-1} for unlabeled protein and >400 mg l^{-1} for [MIT] selectively-labeled protein as determined by amino acid analysis.

NMR spectroscopy

NMR experiments were performed on a Bruker DRX700 spectrometer operating at 700 MHz ^{1}H frequency and equipped with a triple resonance probe and a triple axis gradient coil. DHPR concentration was $\sim$75 μM (300 μM monomer) in 25 mM Tris-D_{11} in D_2O buffer, pH = 7.8 and T = 303, the sample volume was 150 μl in *shigemi* tubes. Protein-ligand complexes were prepared by slowly adding 2.5 μl of DMSO-D_6 solution of compounds (30 mM to 100 mM) to the solution containing the target protein. Based on the large chemical shift difference of Thr $^{13}C^{\gamma}$ ($\sim$18 ppm) and $^{13}C^{\beta}$ ($\sim$70 ppm), selective WURST adiabatic decoupling (Kupce and Freeman, 1995) of $^{13}C^{\gamma}$ from $^{13}C^{\beta}$, during the ^{13}C evolution resulted in much narrower lines in the Thr $^{13}C^{\gamma}$ dimension. This line narrowing dramatically reduced the overlap among the 14 $^{13}C/^{1}H^{\gamma}$ resonances in DHPR wherein Thr $^{13}C^{\gamma}/^{1}H^{\gamma}$ cross-peaks are much narrower than those corresponding to Ile $^{13}C/^{1}H^{\delta}$ (broadened by the 35 Hz $^{13}C^{\delta}$-$^{13}C^{\gamma}$ coupling constant). Instead of the HSQC (heteronuclear single quantum correlation) scheme, we adopted a HMQC (heteronuclear multiple quantum correlation) magnetization transfer since, from theoretical principles, ^{1}H-^{13}C dipole-dipole relaxation mechanism, responsible for the fast ^{13}C transverse relaxation rates, is largely attenuated (Cavanagh et al., 1996). In uniformly labeled protein samples, HSQC sequences exhibit better relaxation properties than HMQC due to strong dipole-dipole relaxation between protons introduced during the heteronuclear evolution time (Bax et al., 1990). The selectively labeled samples, however, were mostly deuterated and proton-proton dipole-dipole interactions can occur (in this particular case) only between Met, Ile and Thr residues.

As Thr and Met residues of interest are not in close proximity to each other ($>$ 6 Å), these dipole-dipole interactions are very small, hence HMQC is superior. Typical 2D [^{13}C,^{1}H] spectrum was recorded in 30 minutes. Resonance assignments for key residues M17 and T104 were assigned as described in the text and these assignments were subsequently confirmed with single point mutants, M17I and T104S, respectively. The assignments were obtained with 2D ^{13}C,^{1}H correlation spectra of [MIT]-M17I-DHPR and [MIT] –T104S-DHPR, respectively (data not shown).

Typical 2D [^{1}H,^{1}H] NOESY (Anil-Kumar et al., 1980) spectra were acquired with 256 $\times$ 2048 complex points and with mixing times between 50 ms and 500 ms. Thr $^{13}C^{\gamma}$ decoupling during *t1* evolution was achieved with a ^{13}C 180 degree refocusing pulse. ^{13}C decoupling during the acquisition was achieved with a GARP composite decoupling sequence (Shaka et al., 1985). The measuring time for a 2D [^{1}H,^{1}H] NOESY varied from $\sim$ 12 h to 48 h, depending on the ligand concentration (500 μM to 2 mM). Ambiguities due to proton overlap among Thr and Met methyl proton chemical shifts were removed by recording a 3D [^{13}C,^{1}H] resolved [^{1}H,^{1}H] NOESY experiment (Fesik and Zuiderweg, 1988) (data not shown). QUIET-NOESY (Quenching Undesirable Indirect External Trouble in NOESY) experiments (Neuhaus and Williamson, 2000; Vincent et al., 1996) were also performed to avoid indirect NOE cross-peaks arising from spin diffusion. The magnetization transfer via cross-relaxation (Neuhaus and Williamson, 2000) is particularly efficient in slowly tumbling proteins. Because of the particular labeling scheme adopted, every proton in the binding site in our samples is at a distance $>$ 6 Å to any other proton in the protein, with the exception of intra-residue protons. Therefore, the magnetization is very efficiently transferred from the binding site protons to the protons of the ligand, which explains the high sensitivity of the NOESY experiments with such large proteins.

NMR experiments with DOXPR were measured with [MIT]-DOXPR at a concentration of 75 μM (300 μM monomer), pH = 7.5 and T = 303 K. 2D [^{1}H,^{1}H] NOESY and QUIET-NOESY experiments were measured as described above. ^{13}C, ^{1}H correlation spectra were obtained with a 2D HMQC sequence as described above with the exception that the selective WURST ^{13}C homonuclear decoupling was applied at 27 ppm to decouple Ile $^{13}C^{\delta}$ (resonating at ∼ 10 ppm) from Ile $^{13}C^{\gamma}$ (resonating at ∼ 27 ppm). Typically each 2D [^{13}C,^{1}H] spectrum was recorded in about 30 minutes.

Selective cross-saturation experiments were performed with a train of IBURP pulses (Geen and Freeman, 1991) each of 8 ms duration spaced by 10 ms and shifted at 1 ppm (cross-saturated spectrum) and at −2.5 ppm (reference spectrum).

Molecular modeling

Docking of the inhibitor TTM2000.029.A85 into the binding site of the target enzyme was calculated based on the X-ray coordinates of DHPR when complexed with NADH and PDC (Scopin et al., 1997) and the NMR-derived constraints with torsion angle dynamics as implemented in DYANA (Güntert et al., 1996). The position of the substrate analog and the coordinates of the enzyme were fixed. The energy minimized coordinates of the inhibitor were obtained using InsightII (Molecular Simulation Inc.). This model was subsequently linked by a dummy linker of ∼ 100 Å encompassing 56 dummy torsion angles, to the model of DHPR bound to PDC. Random torsion angles were assigned to the linker to generate a model of the complex with random initial positioning of the ligand, while the coordinates of DHPR and PDC were kept fixed. Subsequently, a variable target function was minimized in the linker torsion angle space versus the NOE distance constraints between the ligand and both protein and substrate analog (PDC). 20 structures were calculated with 5000 iterations per structure. The best 7 structures with the lower target function converged into the final structure shown.

Results and discussion

NMR spectroscopy has long been known for its ability to provide atomic resolution structural information on protein structure and dynamics, and is increasingly being used in pharmaceutical drug discovery and development (Sem and Pellecchia, 2001). However, its use has been limited due to the difficulty in obtaining structural information on protein-ligand complexes that exceed 30 kDa, as well as by the lengthy and potentially ambiguous process of resonance assignments (Cavanagh et al., 1996). In theory, a structural description of protein-ligand binding interactions requires a description of only the key interactions occurring in the binding site of the target protein. The NMR-DOC method employs several protocols to rapidly obtain structural information by selectively detecting these key interactions, thus enabling a structural characterization of a given protein-ligand complex, even for complexes involving very large multimeric proteins (MW >100 kDa).

We illustrate the various implementations of NMR-DOC by describing experiments involving the enzyme dihydrodipicolinate reductase (DHPR), a homotetramer of 120 kDa, which is involved in the biosynthesis of lysine and bacterial cell wall components (Fakas and Gilvarg, 1965). Since this enzyme is well beyond the current molecular weight limit for full structure determination with NMR spectroscopy, we first employed selective labeling schemes and deuteration for the acquisition of high-resolution data. The three-dimensional X-ray structure of the enzyme (Scopin et al., 1997) revealed that several threonine residues (T80, T103, T104, T170) occur in both the cofactor NADH and the substrate binding sites (Figure 1A). A methionine residue (M17) is also present at the interface between the cofactor NADH and a substrate analog, pyridine 2,6 dicarboxylate (PDC) (Figure 1A). We therefore prepared a sample of DHPR that was selectively labeled in these amino acid residues, as follows: $^{13}C^{\varepsilon}/^1$H Met, $^{13}C^{\delta}/^1$H Ile$^{\delta}$, ^{13}C/^{1}H Thr and U-^{2}H. We will call this labeled sample [MIT]-DHPR (see methods). In order to observe all the expected cross-peaks, one for each Met, Ile and Thr methyl in DHPR in a ^{13}C,^{1}H correlation spectrum, we optimized the NMR parameters for maximum sensitivity and resolution (see methods). Indeed, all expected cross-peaks are clearly observed in the 2D ^{13}C,^{1}H correlation spectrum of [MIT]-DHPR, even at a relatively low enzyme concentration (Figure 1B). Such sensitivity and resolution are remarkable for molecules of this size as previous studies with such large systems were limited to 1D ^{13}C spectroscopy (Kato et al., 1991).

Although the [MIT] labeling scheme largely eliminates the problems of spectral complexity and line broadening that characterize the NMR spectra of large proteins, the traditional methods for the resonance assignments (Cavanagh et al., 1996) are no longer

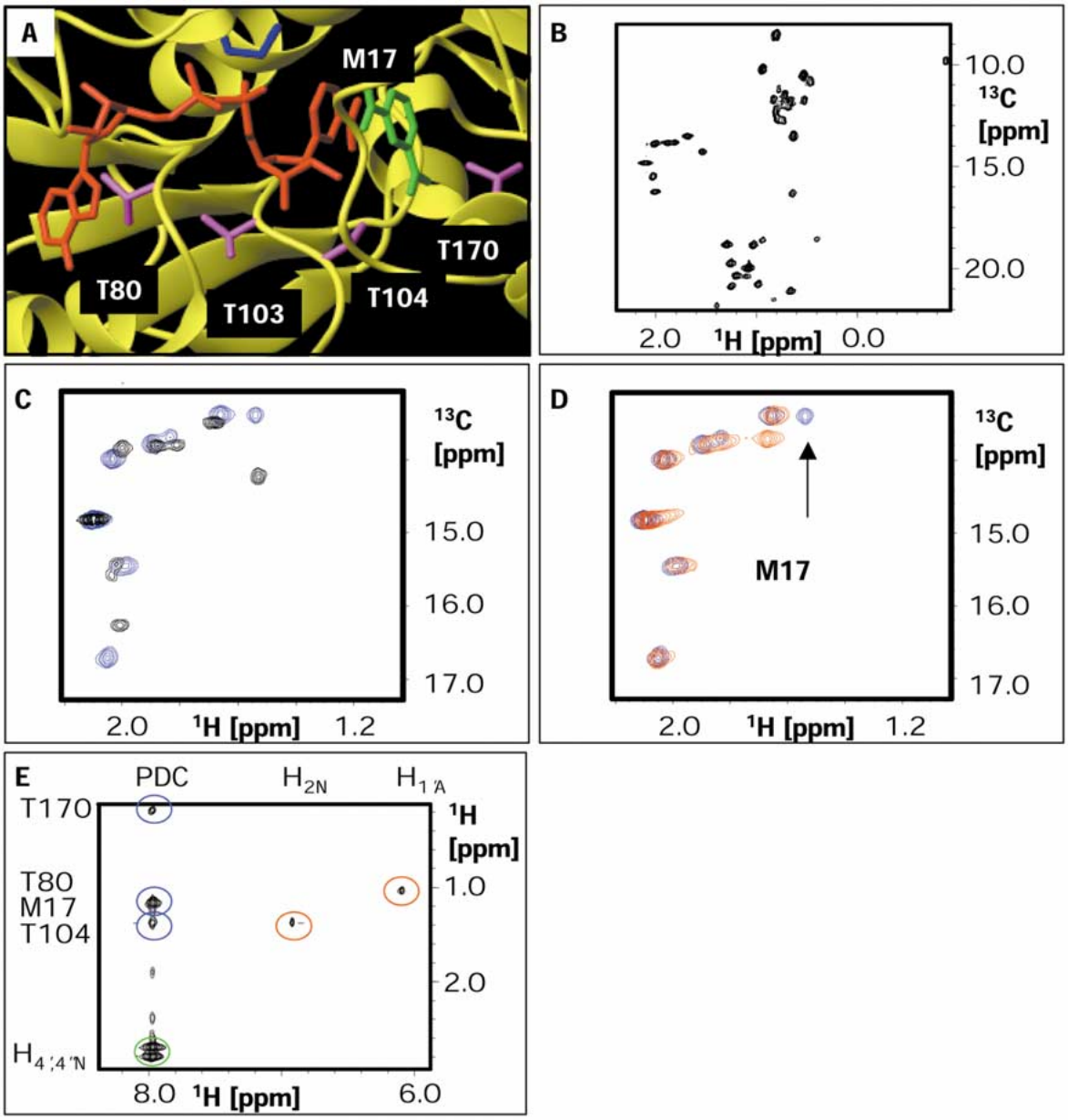

Figure 1. Selective labeling and binding site resonance assignments in DHPR. (A) Ribbon drawing representing the binding site region of DHPR (PDB code 1arz) in complex with NADH and PDC (Scopin et al., 1997). The cofactor NADH is in red. The substrate analog PDC is in green. Heavy atom side chains of Thr and Met residues close to either ligands are displayed and numbered. All figures displaying three-dimensional structures were generated with MOLMOL (Koradi et al., 1996). (B) 2D [^{13}C,^{1}H] HMQC spectra recorded with a 150 µl of a 50 µM (200 µM monomer) sample of U-^{2}H, ^{13}C$^{\varepsilon}$/^{1}H$^{\varepsilon}$ Met, ^{13}C/^{1}H Thr, ^{13}C$^{\delta}$/^{1}H$^{\delta}$ Ile labeled DHPR ([MIT]-DHPR). (C) and (D): Met ^{13}C$^{\varepsilon}$/^{1}H$^{\varepsilon}$ sub-spectra. (C) Black, unbound [MIT]-DHPR; blue, [MIT]-DHPR bound to PDC. (D) Red, [MIT]-DHPR bound to 4-Cl PDC; blue, [MIT]-DHPR bound to PDC. The assignment of active site M17 is indicated. (E) Portion of a 2D [^{1}H,^{1}H] NOESY spectrum recorded with a ternary complex between NADH, PDC and [MIT]-DHPR. Inter-molecular NOE cross-peaks are marked with circles: Red, protein-NADH NOEs; blue, protein-PDC NOEs; green, inter-ligand NADH–PDC NOEs. The three protons of PDC are degenerate in the complex. The chemical shift of proton H_{2N} of the nicotinamide ring and the $H_{1'A}$ of the ribose of the adenosine of NADH are indicated. The spectra were acquired and processed with XwinNMR (Bruker AG) and analyzed with XEASY (Bartels et al., 1996).

applicable to these samples. However, assignment of binding site Met and Ile methyl ^{13}C,^{1}H resonances may be accomplished in several alternative ways based on the introduction of a perturbation in the resonances residing in the binding site of the enzyme. The simplest but crudest is the detection of chemical shift perturbation introduced by ligand binding. For instance, the $^{13}C^{\varepsilon}/^{1}H^{\varepsilon}$ resonances of M17 may be assigned via the change in chemical shift upon binding of PDC to DHPR (Figure 1C). However, since ligand binding often induces several changes in the spectra due to protein conformational changes (Figure 1C), we developed a differential chemical shift perturbation method in which the spectra of the target protein bound to two slightly different ligands were compared (Figure 1D), related to the work of Medek et al. (2000). The data show that when [MIT]-DHPR binds PDC by contrast to the 'chemically perturbed' variant inhibitor, 4-Cl PDC, distinct changes in chemical shift for only one of the methionine $^{13}C^{\varepsilon}/^{1}H^{\varepsilon}$ resonances are detected, which therefore identifies the signals associated with M17 (Figure 1D). Both PDC and 4-Cl-PDC bind to DHPR with micromolar dissociation constants, so that, at the concentrations used, the protein is saturated in both samples. Therefore, the resultant chemical shift differences originate solely from the small perturbation introduced by binding slightly different ligands.

Similarly, we obtained resonance assignments for residues T104 and T103 with differential chemical shifts with slightly modified cofactors (NADH vs. 3-acetyl pyridine NADH and thio-NADH, data not shown). Thus, we mapped the binding site of DHPR using a very fast approach based on strategically directed chemical shift perturbations of structurally characterized reference compounds.

A more straightforward way to obtain binding-site resonance assignments is via protein–ligand NOEs. In this case, we perturbed the ligand through either a selective inversion (transient Nuclear Overhauser Effect [NOE]) or complete saturation (steady-state NOE) of its resonances using radio-frequency pulses, and we detected the effect on the protein spectra in the form of an NOE (Neuhaus et al., 2000). A portion of a 2D [^{1}H,^{1}H] NOESY spectrum (Anil-Kumar et al., 1980) of [MIT]-DHPR in complex with the cofactor NADH and the substrate analog PDC is shown (Figure 1E). Due to the selective labeling scheme, there is little overlap between the protein methyl-proton resonances and the ligand-proton resonances (Figure 1B). Given the resonance assignments of the ligands, which are easily obtained with conventional 1D and 2D NMR experiments, assignments of the structurally neighboring Met and Thr methyl protons in the binding site are straightforward. Binding site residues are structurally mapped relative to protons on the structurally characterized NADH reference ligand. Thus, binding site residues are mapped according to their proximity to the different protons on a reference ligand (NADH or PDC in this case).

The detection of inter-ligand (PDC to NADH in this example) NOEs (Li et al., 1999) provides extremely useful information on the orientation and the distance of two binders relative to each other. These NOEs also appear in the 2D [^{1}H,^{1}H] NOESY of the ternary complex (Figure 1E). Protein deuteration limits spin diffusion (Neuhaus et al., 2000) and, most importantly, transverse nuclear spin relaxation and therefore permits the use of more concentrated protein samples without inducing excessive broadening of the resonance lines of the bound ligands. Therefore, protein deuteration provides increased signal to noise ratio for both protein-ligand and inter-ligand NOEs. Observation of bound-state inter-ligand NOEs provides an enormous advantage over the transferred inter-ligand NOE effect, which is limited to fast dissociating molecules, since this effect becomes rather small when one or both ligands dissociates very slowly, as for tight binders (Li et al., 1999).

Once the binding site residues are assigned, compound screening can be performed with simple experiments that provide information in a relatively short time both on binding affinity (via titration), and on inhibitor binding mode. We demonstrate this with the results of NMR-DOC on [MIT]-DHPR using nicotinamide mononucleotide (NMNH, Figure 2A) as a test and validation cofactor mimic, since its binding mode is easily extrapolated from the X-ray structure of NADH bound DHPR (Scopin et al., 1997). These experiments were performed with a low ($\sim$ 10 μM) concentration of labeled protein and 1 mM of the test ligand. The saturation of binding site nuclei was achieved through saturation of the aliphatic region of the spectrum that contains the resonances of the labeled Thr and Met residues. The saturation is transferred to the ligand only if it is bound to the protein, as evidenced by a difference spectrum (Figure 2B). While cross saturation has been used (Dalvit et al., 2000; Klein et al., 1999; Neuhaus et al., 2000) with unlabeled protein, using selectively labeled samples makes these experiments more informative as most of the protonation localized in its binding site. Therefore,

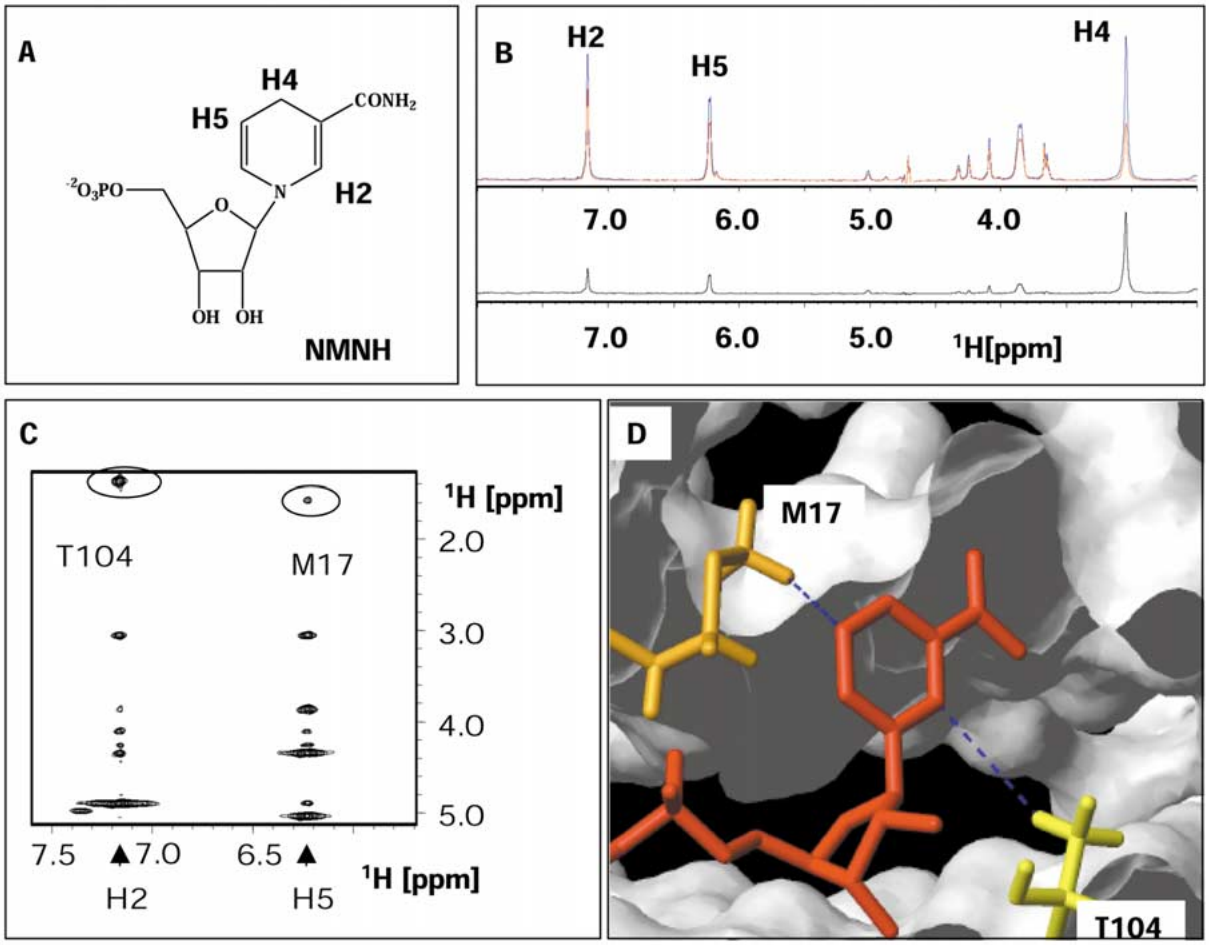

Figure 2. NMR-DOC with [MIT]-DHPR and NMNH as validation ligand. (A) Structure of NMNH. (B) Reference spectrum and selective binding site-saturated spectrum of NMNH (500 μM) in the presence of [MIT]-DHPR at 2.5 μM (10 μM monomer). (C) Portion of a 2D [^{1}H,^{1}H] NOESY spectrum of NMNH (500 μM) in complex with [MIT]-DHPR (75 μM). Protein-NMNH NOEs are labeled. Other cross-peaks represent intra-molecular NMNH NOEs. (D) The NOEs labeled in (C) are displayed as dotted blue lines on the three-dimensional structure of NADH in complex with DHPR.

observation of a cross-saturation effect is a likely indicator of specific binding in the binding pocket of the enzyme and not the result of aspecific binding. Thus, once such a structurally screened ligand is discovered, its orientation in the binding site can be more accurately determined by measuring inter-molecular NOEs with previously assigned resonances (using a reference ligand) in the binding site of the enzyme.

In addition to intra-molecular NOEs, that can be used to determine the bioactive conformation of the ligand, very strong inter-molecular NOEs between the nicotinamide ring protons and the methyl groups of the binding site residues M17 and T104 were observed (Figure 2C). These NOE cross-peaks are in agreement with the X-ray structure of the DHPR-NADH-PDC ternary complex (Figure 2D) and can therefore be used as constraints for docking the ligand in the binding site of the enzyme. For example, we produced a model of the complex of DHPR with a weak-binding inhibitor ($K_D \sim 500$ μM) (Figure 3A) in complex with PDC based on DHPR-ligand and PDC-ligand NOE-derived constraints (Figure 3B) within a few days (Figure 3C).

An attractive application of NMR-DOC experiments is in the design of bi-ligand compound inhibitors of bi-ligand enzymes by linking compounds such as those shown in Figure 3C, to obtain higher affinity ligands. In fact, we are in the process of characterizing the binding properties of several bi-ligand inhibitors obtained by properly linking NADH mim-

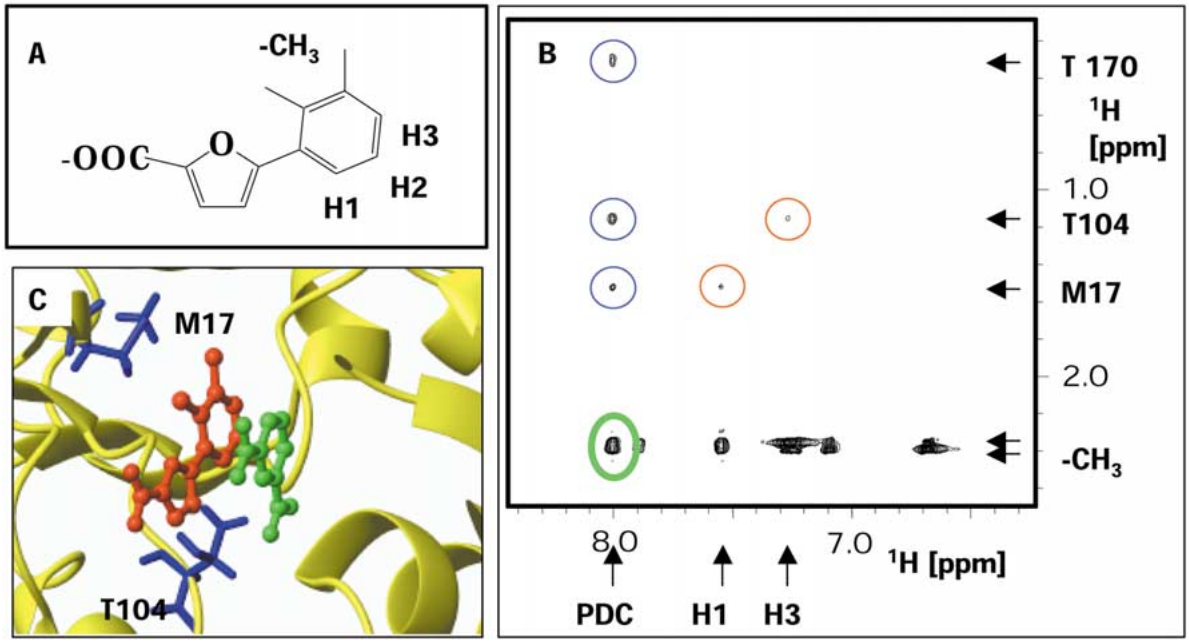

Figure 3. NMR-DOC with the novel inhibitor TTM2000.029.A85. (A) Structure of the in house discovered NADH mimic TTM2000.029.A85. (B) Portion of a 2D [^{1}H,^{1}H] NOESY spectrum of TTM2000.029.A85 (500 μM) in complex with [MIT]-DHPR (75 μM). The NOEs are labeled and the color code is the same as reported in Figure 1E. (C) Docked structure of TTM2000.029.A85 (red) into the three-dimensional structure of DHPR (yellow) bound to PDC (green). The structure was calculated with DYANA based on the NOE data obtained with the ternary complex TTM2000.029.A85, [MIT]-DHPR and PDC, and the X-ray structure of DHPR bound to PDC.

ics, such as the one shown in Figure 3A, to PDC (data not shown).

For enzymes that posses two adjacent binding pockets, as dehydrogenases, the structural characterization of a ligand in the first binding site can be used to design a linker reaching the second binding site, and a combinatorial library aimed to target the second binding pocket can be placed at the end of the linker. This strategy, called NMR-SOLVE (Structurally Oriented Library Valency Engineering), is very attractive as it works also for protein targets for which no three-dimensional structure is available. It differs from NMR-DOC in that there is a strategic focus on the region of the binding site at the interface of the cofactor and substrate binding sites, proximal to where the enzymatic chemical reaction occurs.

This is illustrated with the enzyme 1-deoxy D-xylulose 5-phosphate reductoisomerase (DOXPR) (Kuzuyama et al., 2000), a homotetrameric enzyme of 174 kDa for which there is no protein with > 20% homology in the Protein Data Bank. We uniformly ^{2}H/^{12}C labeled DOXPR, except for the methyl groups of Met, Thr and Ile$^\delta$ that were ^{13}C and ^{1}H labeled to obtain [MIT]-DOXPR. The MIT amino acid labeling was chosen based on our survey of oxidoreductase three-dimensional structures that revealed an average of four to five of these residues in the NAD-binding sites (data not shown). Selective side-chain ^{13}C/^{1}H labeling for the amino acids Val, Tyr, Phe, Trp and His could also be obtained (Goto and Kay, 2000).

The [^{13}C,^{1}H] HMQC spectrum for [MIT]-DOXPR and the 2D [^{1}H,^{1}H] NOESY in complex with its cofactor NADPH are shown (Figure 4). From their unique ^{1}H chemical shifts, the observed NOEs clearly indicated that there is an Ile as well as a Met proximal to the nicotinamide ring portion of NADPH (Figure 4B). Proximity of another Ile residue to the adenosine ring was also observed (Figure 4B). Assignments of an 'interface residue' was made based on NOEs from the H_{2N} proton of NADPH to a Met in DOXPR (Figure 4B) and to a Thr in DHPR (Figure 1E). Although the knowledge of the crystal structure of DHPR allowed the assignment of this interface residue to Thr104, for the purpose of NMR-SOLVE all that matters is the knowledge of the chemical shifts of an interface residue. Indeed, in the case of DOXPR, for which there is no three-dimensional structure available, we cannot identify an interface residue with its location in the primary sequence, but we can identify the chemical shift of an interface Met residue (Fig-

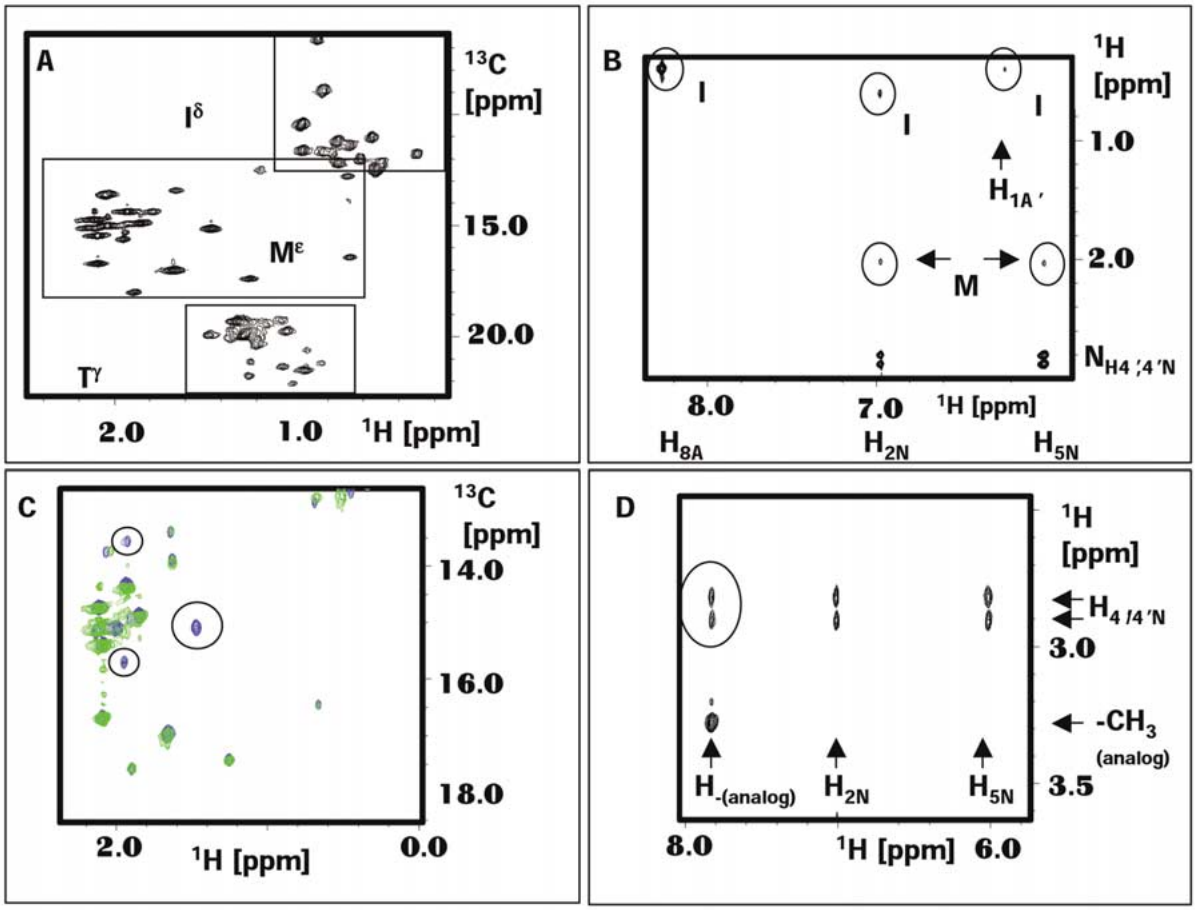

Figure 4. NMR-SOLVE with DOXPR. (A) 2D [^{13}C,^{1}H] HMQC spectrum recorded with a 75 μM (300 μM monomer) sample of U-^{2}H, ^{13}C/^{1}H Thr, ^{13}C$^{\delta}$/^{1}H$^{\delta}$ Ile, ^{13}C$^{\varepsilon}$/^{1}H$^{\varepsilon}$ Met labeled DOXPR ([MIT]-DOXPR). Met, Ile and Thr regions are enclosed in dashed rectangles. (B) Portion of a 2D [^{1}H,^{1}H] NOESY spectrum of the complex between NADP^{+} and [MIT]-DOXPR with intermolecular NOEs circled. (C) Met region of the 2D [^{13}C,^{1}H] HMQC spectrum for [MIT]-DOXPR. Blue, without catalytic Mn^{2+}; green, with 10 μM catalytic Mn^{2+}. The Met residues that are broadened beyond detection due to proximity to the active site Mn^{2+} are labeled with a circle. (D) Inter-ligand NOEs between a reactive intermediate analog and NADPH in the ternary complex of NADPH – [MIT]-DOXPR – reactive intermediate analog.

ure 4B). Subsequently, the detection of NOEs between a novel inhibitor and this residue would provide useful information on the orientation of the inhibitor with respect to the reference cofactor. As such, it is possible to identify ligands, and portions of these ligands, that reside in the catalytic part of the binding site close to the nicotinamide ring of NADH. This provides the structural information needed to optimally place a linker on the cofactor mimic for creating a bi-ligand combinatorial library that targets the substrate binding site. Thus, bi-ligand combinatorial library construction is guided without knowledge of the three-dimensional structure of the enzyme target, based solely on the NMR-SOLVE data.

For metal-binding enzymes, identification of interface and active-site residues can also be achieved through detection of line broadening using a paramagnetic metal ion probe. It has recently been proposed that DOXPR binds a Mn^{2+} ion with a catalytic role (Kuzuyama et al., 2000). We performed a comparison of protein in the presence and absence of Mn^{2+} and identified three binding site Met residues (Figure 4C). One Met residue also exhibited NOEs with the cofactor NADPH, therefore further defining its position at the interface between the cofactor and substrate binding sites. Furthermore, inter-ligand NOEs in DOXPR between a stable version of an enolate intermediate analog that binds to DOXPR with a K_i of 470 μM (D.S. Sem, unpublished results), and the

cofactor NADPH were observed (Figure 4D). Again, these inter-ligand NOEs can be used to identify molecules that bind in the catalytic portion of the cofactor binding site of the enzyme, and to determine their orientation relative to the substrate binding pocket.

Conclusion

In conclusion, we have demonstrated that the NMR-DOC method is applicable to very large proteins and can rapidly provide cogent structural information once the three-dimensional structure of a given enzyme is known. For enzymes with two adjacent binding sites, as in the dehydrogenase gene family, the NMR-SOLVE method also works in the absence of knowledge of a given target's three-dimensional structure. This is based on the fact that unique structural information on ligand-binding mode can be gathered and used to guide medicinal and/or combinatorial chemistry in the design of bi-ligand libraries, by focusing on the catalytic region of the binding site.

The resonance assignments obtained for binding site residues can be used both to detect ligand binding and determine ligand orientation. In addition, such assignments should also enable further studies on binding site dynamics, hydration, and on enzyme mechanism, therefore extending the use of NMR spectroscopy to proteins of molecular weight well above 100 kDa. Identification of the amino acids present in a binding site and their position relative to a reference ligand will also provide information critical to guiding, refining and validating the construction of homology-modeled protein structures.

Acknowledgements

We thank Prof Kurt Wüthrich for helpful suggestions, and Drs Stephen Coutts and Lin Yu for helpful discussions and review of this manuscript, and Dr Victor Hong for enzyme kinetic support.

References

Anil-Kumar, Ernst, R.R. and Wüthrich, K. (1980) *Biochem. Biophys. Res. Comm.,* **95**, 1–6.

Bartels, C., Xia, T., Billeter, M., Güntert, P. and Wüthrich, K. (1995) *J. Biol. NMR,* **6**, 1–10.

Bax, A., Ikura, M., Kay, L.E., Torchia, D.A. and Tschudin, R. (1990) *J. Magn Reson*., **86**, 304–318.

Cavanagh, J., Fairbrother, W.J., Palmer III, A.G. and Skelton, N.J. (1996) In *Protein NMR Spectroscopy, Principles and Practice*, Academic Press, New York, pp. 244–552.

Dalvit, C., Pevarello, P., Tato', M., Veronesi, M., Vulpetti, A. and Sundström, M. (2000) *J. Biomol. NMR,* **18**, 65–68.

Dragovich, P.S., Barker, J.E., French, J., Imbacuan, M., Kalis, V.J., Kissinger C.R., Knighton, D.R., Lewis, C.T., Moomaw, E.W., Parge, H.E., Pelletier, L.A., Prins, T.J., Showalter, R.E., Tatlock, J.H., Tucker, K.D. and Villafranca J.E. (1996) *J. Med. Chem*., **39**, 1872–1884.

Drews, J. (1998) *In Search of Tomorrow's Medicines*, Springer, New York.

Geysen, H.M., Meloen, R.H. and Barteling, S.J. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 3998–4002 .

Farkas, W. and Gilvarg, C. (1965) *J. Biol. Chem.*, **240**, 4717–4722.

Fesik, S.W. and Zuiderweg, E.R.P. (1988) *J. Magn. Reson.*, **78**, 588–593.

Geen, H. and Freeman, R. (1991) *J. Magn. Reson*., **93**, 93–141.

Goto, N.K. and Kay L.E. (2000) *Curr. Opin. Struct. Biol.*, **10**, 585–592.

Güntert, P., Mumenthaler, C. and Wüthrich, K. (1997) *J. Mol. Biol.*, **273**, 283–298.

Kato, K., Matsunaga, C., Odaka, A., Yamato, S., Takaha, W., Shimada, I. and Arata, Y. (1991) *Biochemistry*, **30,** 6604–6610.

Klein, J., Meinecke, R., Mayer, M. and Meyer, B. (1999) *J. Am. Chem. Soc.*, **121**, 5336–5337.

Koradi, R., Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graph.*, **14**, 51–59.

Kupće, E. and Freeman, R. (1995) *J. Magn. Reson.*, **B109**, 329–333.

Kuzuyama, T., Takahashi, S., Takagi, M. and Seto, H. (2000) *J. Biol. Chem.*, **275**, 19928–19932.

Li, D., DeRose, E.F. and London, R.E. (1999) *J. Biomol. NMR*, **15**, 71–76.

Medek A., Hajduk, P.J., Mack, J. and Fesik, S.W. (2000) *J. Am. Chem. Soc.*, **122**, 1241–1242.

Meininger, D.P., Rance, M.,. Starovasnik, M.A., Fairbrother, W.J. and Skelton N.J. (2000) *Biochemistry*, **39**, 26–36.

Metzler, W.J., Wittekind, M., Goldfarb, V., Mueller, L. and Farmer II, B.T. (1996) *J. Am. Chem. Soc.*, **118**, 6800–6801.

Neuhaus, D. and Williamson, M.P. (2000) *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, Wiley-VCH, New York, pp. 129–279.

Scopin, G., Reddy, S.G., Zheng, R. and Blanchard, J.S. (1997) *Biochemistry*, **36**, 15081–15088.

Sem, D.S. and Pellecchia, M. (2001) *Curr. Opin. Drug Disc.*, **4**, 479–492.

Shaka, A.J., Barker, P.B. and Freeman, R. (1985) *J. Magn. Reson.*, **64**, 547–552.

Stevens, R.C., Yokoyama, S. and Wilson, I.A. (2001) *Science*, **294**, 89–92.

Vincent, S.J.F., Zwahlen, C., Bolton, P.H., Logan, T.M. and Bodenhausen, G. (1996) *J. Am. Chem. Soc.*, **118**, 3531–3532.