Discovering High-Affinity Ligands for Proteins: SAR by NMR

Suzanne B. Shuker, Philip J. Hajduk, Robert P. Meadows, Stephen W. Feek*

A nuclear magnetic resonance (NMR)–based method is described in which small organic molecules that bind to proximal subsites of a protein are identified, optimized, and linked together to produce high-affinity ligands. The approach is called “SAR by NMR” because structure-activity relationships (SAR) are obtained from NMR. With this technique, compounds with nanomolar affinities for the FK506 binding protein were rapidly discovered by tethering two ligands with micromolar affinities. The method reduces the amount of chemical synthesis and time required for the discovery of high-affinity ligands and appears particularly useful in target-directed drug research.

Drugs are typically discovered by identifying active compounds from screening chemical libraries or natural products and optimizing their properties through the synthesis of structurally related analogs. This is a costly and time-consuming process. Suitable compounds with the requisite potency, compound availability, or desired chemical and physical properties cannot always be found. Furthermore, even when such compounds are found, optimization often requires the synthesis of many analogs.

We now describe a method for identifying high-affinity ligands that can aid in the drug discovery process. The technique, which is called “SAR by NMR,” is a linked-fragment approach wherein ligands are constructed from building blocks that have been optimized for binding to individual protein subsites (Fig. 1). In the first step of this process, a library of low molecular weight compounds (1) is screened to identify molecules that bind to the protein. Binding is determined by the observation of "N- or "H-amide chemical shift changes in two-dimensional "N-heteronuclear single-quantum correlation ("N-HSQC) spectra (2) upon the addition of a ligand to
an $^{15}$N-labeled protein. These spectra can be rapidly obtained, making it possible to screen a large number of compounds (3). Once a lead molecule is identified, analogs are screened to optimize binding to this site (Fig. 1, step 2). Next, a ligand is sought that interacts with a nearby site. Binding to a second pocket is determined by observing changes in a different set of amide chemical shifts in either the original screen or a screen conducted in the presence of the first fragment (Fig. 1, step 3). From an analysis of the chemical shift changes, the approximate location of the second ligand is identified. Optimization of the second ligand for binding to this site is then carried out by screening structurally related compounds (Fig. 1, step 4). When two “lead” fragments have been selected, their location and orientation in the ternary complex are determined experimentally either by NMR spectroscopy or by x-ray crystallography. Finally, on the basis of this structural information, compounds are synthesized in which the two fragments are linked together (Fig. 1, step 5) with the goal of producing a high-affinity ligand.

The SAR by NMR method is illustrated in the discovery of ligands that bind tightly to a protein that forms a complex with the potent immunosuppressant FK506 (1). This protein, which is called the FK506 binding protein (FKBP), inhibits calcineurin (a serine-threonine phosphatase) and blocks T cell activation (4) when it is complexed to FK506. Using NMR spectroscopy, we screened FKBP against our library of compounds (5). Many compounds were found that bound weakly to FKBP, including methyl-3-isouquinolinecarboxylate ($K_d = 1.0$ mM), 4-carboxethoxy-3-methyl-2-cyclohexen-1-one ($K_d = 0.5$ mM), and 2-phenylimidazole ($K_d = 0.2$ mM). The trimethoxyphenyl pipercolinic acid derivative, 2, showed the highest affinity for FKBP ($K_d = 2.0$ μM), which is consistent with earlier reports on the avidity of pipercolinic acid derivatives for FKBP (6). This compound was chosen for study without further optimization. The binding site for 2 is the same as that of the pipercolinic acid moiety of FK506 (7), as evidenced by the amide chemical shift changes of FKBP (Fig. 3).

To identify molecules that interact with FKBP at a nearby site (Fig. 1, step 3), we screened our library of compounds in the presence of saturating amounts of 2 (8). We found an initial candidate (3) that bound to FKBP with an affinity of 0.8 mM (Fig. 4) as measured from the changes in the $^{15}$N-HSQC spectrum of uniformly $^{15}$N-labeled FKBP upon binding (9). An analysis of these chemical shift changes indicated that the binding site for 3 is near the binding site for 2 (Fig. 3).

To probe the structural requirements for the binding of benzaldehyde derivatives to the second site (Fig. 1, step 4), we obtained analogs of 3 from our corporate sample collection and commercial sources and tested these compounds for binding to FKBP by NMR. From the structure-activity relationships (SAR) that are observed, it appears that both aromatic rings and the amide linkage are important for binding to FKBP, as evidenced by the decrease in affinity for analogs lacking these chemical moieties (9). The para-hydroxyl group (Fig. 4, R$_3$) on the aniline ring is also important for binding (4); whereas, the hydroxyl group on the benzyl ring at R$_4$ does not contribute to the binding affinity (5). However, when R$_3$(6) or R$_4$(7) is a hydroxyl group, an increase in binding affinity is observed relative to the parent compound (8). Thus, both substituted para-hydroxyl substituents (R$_3$ and R$_4$) appear to contribute to the binding to FKBP. The benzaldehyde containing two para hydroxyls, 9, was not commercially available.

**Fig. 1.** An outline of the SAR by NMR method.

**Fig. 2.** A superposition of $^{15}$N-HSQC spectra for FKBP in the absence (magenta contours) and presence (black contours) of compound 3. Both spectra were acquired in the presence of saturating amounts of 2 (2.0 mM). Significant chemical shifts changes are observed for labeled residues.

**Fig. 3.** A surface representation of FKBP showing the locations of 2 and 9, as determined from $^{15}$N-$^{13}$C-filtered NOE data (11). Residues that exhibited the largest chemical shift changes on the binding of 2, 9, or both 2 and 9 are colored in magenta, cyan, and yellow, respectively. Chemical shift changes for 9 (cyan and yellow) are those observed on the addition of 9 to FKBP in the presence of saturating amounts of 2 (2.0 mM). Weighted average chemical shifts changes were calculated ($\Delta\delta$) = $\frac{3\delta[\delta]}{3\delta(\delta)}$ + $\frac{3\delta[\delta]}{3\delta(\delta)}$, and colored residues are those for which $\Delta\delta$ [m] exceeded 0.15 and 0.05 ppm for 2 and 9, respectively.
but was readily synthesized (10) and served as our best ligand for the second site ($K_d = 100 \mu M$).

A model of the ternary complex of 2, 9, and FKBP was generated on the basis of isotope-filtered NMR studies (Fig. 3) (11). In this model, the methyl ester of 2 is close to the hydroxyl group on the benzoyl ring of 9. Thus, we designed linkers that would attach to these groups, span the distance between the two fragments, and have no steric clashes with the protein. Four compounds were synthesized (12) in which the pipolicin acid moiety of 2 is connected to the hydroxyl group on the benzoyl ring of 9 with linkers of different lengths (10 to 13). In addition, a compound, 14, in which the benzoyl ring of 9 is linked at a position ortho to the hydroxyl group was synthesized (13) to explore a different regiochemical linkage. The binding affinities of these five compounds for FKBP were measured in a fluorescence-based assay (14), and all of the compounds exhibit nanomolar affinities for binding to FKBP (Fig. 4).

To determine whether the linked compounds bind to the same site as the unentethered compounds, we compared the intermolecular nuclear Overhauser effects (NOEs) observed in the NMR structure of the FKBP-14 complex (15) (Fig. 5) to those observed in the ternary complex composed of 2, 9, and FKBP. The pipolicin acid moieties and trimethoxyphenyl groups of 2 and 14 bind to the same FKBP binding site and form hydrophobic interactions with the protein (Fig. 5). In addition, the benzanilide moieties of 9 and 14 bind in similar locations in the linked and unlinked compounds (Fig. 5). However, several NOEs were observed between the benzoyl ring of 9 and Gin53, Ile36, and Arg27 which were not detected between the linked compound (14) and FKBP. This suggests that although the benzanilide moieties of 9 and 14 are in very similar locations in the two complexes, the introduction of the linker causes a small shift (approximately 1 to 2 Å) in the position of this group.

As demonstrated here, SAR by NMR is useful for discovering high-affinity ligands with experimentally derived information. Although the small untethered molecules might only bind in the micromolar to millimolar range to the target protein, the binding affinity of a linked compound is, in principle, the product of the binding constants of the individual fragments plus a term that accounts for the changes in binding affinity that are due to linking (16). Thus, molecules with submicromolar affinity can be obtained by linking two compounds that bind in the millimolar range. In the application that we describe, a compound that binds to FKBP with a $K_d = 19 \text{nM}$ was synthesized by linking two molecules with binding affinities of 2 $\mu M$ and 100 $\mu M$. These two fragments were rapidly identified and optimized (less than 2 months), and only five linked molecules were synthesized, all of which exhibit nanomolar affinities for FKBP.

One of the advantages of SAR by NMR is the use of $^{15}N$-HMQC spectra to detect the binding of small, weakly bound ligands to an $^{15}N$-labeled target protein. Because of the $^{15}N$ spectral editing, no signal from the ligand is observed. Thus, binding can be detected even at high compound concentrations. This is an important advantage over conventional screening assays, such as fluorometric or colorimetric assays, where high compound concentrations can give rise to large background signals. Another advantage of the use of $^{15}N$-HMQC spectra is the ability to rapidly determine the different binding site locations of the fragments, which is critical for interpreting structure-activity relationships and for guiding the synthesis of linked compounds.

A number of linked-fragment based approaches that utilize molecular modeling have been described (17). However, computer-based methods have difficulties in reliably predicting which fragments will interact with the protein and the manner in which they bind. This kind of task is challenging because of (i) possible conformational changes of the protein and ligand when complexes are formed, (ii) the difficulties in treating bound water and counter ions, (iii) the neglect of entropy, and (iv) the use of relatively simple force fields (17). To circumvent these limitations, the SAR by NMR method uses experimentally derived information to guide fragment selection, optimization, and linking. An experimental approach for mapping the binding surface of proteins has been described (18), which in-

---

**Fig. 4.** Summary of the SAR by NMR method as applied to FKBP. Compounds with affinities from 19 to 228 nM were obtained from two fragment leads with affinities of 2 and 100 $\mu M$.

**Fig. 5.** Ribbon (23) depiction of the structure of FKBP (gray) when complexed to 14 (green carbon atoms). Shown in yellow are those residues that have NOEs to the ligand.
volves solving multiple x-ray crystal structures of a protein in different solvents. With this method, however, only organic solvents are examined as potential ligands, and the binding affinities are not measured.

The SAR by NMR method is conceptually similar to combinatorial chemistry (19) in that both techniques utilize a building block approach in the construction of molecules. However, in practice, the two are quite different. In combinatorial chemistry, ligand discovery requires the preparation and testing of a large number of linked compounds. Unfortunately, the development of synthetic protocols and assays to identify active compounds is frequently difficult and time-consuming. Furthermore, the need for relatively uniform and straightforward coupling conditions limits the range of useful molecular building blocks and compound diversity.

In contrast, when SAR by NMR is used, only a few compounds need to be synthesized because the tethered ligands are optimized prior to linking. In principle, the linked compounds are selected from a large virtual library composed of all combinations of the fragments. For example, for a protein with two independent binding sites, a compound collection containing 10^4 fragments represents a virtual library of more than 10^10 members (20).

Actual chemistry, however, is highly focused on linking only those molecules that have shown to bind to the protein and is guided by structural information on their relative locations and orientations. Furthermore, unlike combinatorial chemistry, the range of small molecules in the screening library in SAR by NMR is limited only by the requirement for aqueous solubility at millimolar concentrations, resulting in a diverse sample collection of compounds.

Although the SAR by NMR method is only applicable to small biomolecules (MW < 30 kD) that can be obtained in large quantities (≥200 mg) (21), many small proteins or protein domains fit these criteria and may serve as drug targets (22). When applied to these proteins, SAR by NMR can play an important role in drug discovery. Indeed, preliminary use of this technique with other proteins indicates that the method will be applicable to a variety of protein targets and an extremely valuable tool in drug research.

REFERENCES AND NOTES

1. Our library of compounds consists of a diverse set of 10,000 small molecules (average molecular weight, 213) dissolved in deuteriated dimethylsulfoxide (DMSO).
3. A two-dimensional (2D) HMQC spectrum of a uniform 15N-labeled H. N-acetylated and sidechain resonances of FKB in the complex were assigned by standard methods [G. M. Clore and A. M. Gronenborn, Methods Enzymol. 239, 349 (1995)]. Intra-molecular distance restraints were obtained from 13C- and 1H-unresolved NOESY spectra and 2D 13C-1H filtered NOESY spectra. Distance restraints were classified into three categories: non-recoverable, 5-10 Å cross-peak intensity and given a lower bound of 1.8 Å and upper bounds of 3.0 Å, 4.0 Å, and 5.0 Å. Hydrogen bonds, identified for slowly exchanging amides, were defined by two restraints: 1.8 to 2.3 Å for the H-O distance and 2.6 to 3.3 Å for the N-O distance. Structures were calculated with the X-PLOR 3.1 program [A. T. Brünger, X-PLOR 3.1 Manual (Yale Univ. Press, New Haven, CT, 1992)] on Silicon Graphics computers with a hybrid distance geometry-simulated annealing approach [M. Nilges et al., FEBS Lett. 229, 517 (1988)]. A total of 977 approximate interproton distance restraints were derived from the data, which included 34 hydrogen bonds, 63 intermolecular distance restraints, and 19 intra-ligand distance restraints. After simulated annealing, 21 structures were obtained with good covalent geometry, no distance violations greater than 0.2 Å, and an overall root-mean-square deviation (rmsd) to the averaged structure of 1.0 Å for 4010 restraints. All heavy atoms. Final energies (E) for the averaged, minimized structure were E = 129.2, E_{bond} = 15.6, E_{angle} = 124.7, E_{dihe} = 19.4, E_{aprot} = 22.8, and E_{non} = 3.0 kcal/mol.
4. In principle, the binding energy for the linked compound (\Delta G_{bind}) should be the sum of the intrinsic binding energies for the unlabeled compounds (\Delta G_{bind}) plus a term to account for the linking of the two fragments (\Delta G_{link}): \Delta G_{bind} = \Delta G_{bind} + \Delta G_{link} + \Delta G_{link}[W. P. Jencks, Proc. Natl. Acad. Sci. U.S.A. 78, 4046 (1981)]. In terms of binding affinities, this translates to K_{d}[\text{L}] = K_{d}[\text{L}][\text{L}][\text{L}] = \text{K}_{d}[\text{L}]^{2} \times 10^{-7} \text{M}^{-1} and \text{K}_{d}[\text{L}] = \text{K}_{d}[\text{L}]^{2} \times 10^{-7} \text{M}^{-1} and L less than 1.0, the expected affinity for the linked compound would be K_{d}[\text{L}]^{2} \times 10^{-7} \text{M}^{-1}.
8. In general, the number of potential linked ligands in a virtual library is n^p, where n is the number of fragment libraries contained in the database, p is the number of adjacent, independent binding sites in the protein, and n is the number of linkers that could be used to join the fragments.
9. To reduce the amount of protein required by the technique, the protein is recycled by removing the ligands with dialsis after the NMR experiment.
10. The SAR by NMR method is even applicable to proteases, provided that the first ligand inhibits the enzyme.
12. We thank G. Waring for the synthesis of 2 and tri-methoxybenzoylformic acid, T. Holzmann and J. Severin for the preparation of biotinylated FKB; A. Phillips, A. Peterson, and L. Tucker-Garcia for sample preparation and aid in building our library of compounds; K. Swift and E. Matyszewski for assistance in performing the fluorescence measurements; and D. Norbeck for helpful comments on the manuscript.
13. 15 July 1996; accepted 1 October 1996.