Nuclear magnetic resonance as a breakthrough frontier in fragment-based drug design

Morkos A. Henen1,2, Laila Abou-Zeid1, Alaa El-Din Barghash1 and Hassan M. Eisa1

1 Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Mansoura University, Mansoura 35516, Egypt. 2 Max F. Perutz laboratories, Vienna University, Vienna, Austria

morkos_henen@yahoo.com

Abstract: The fragment-based drug design (FBDD) is one of the main two approaches that are used in drug design (in addition to the high-throughput screening). The main obstacle that faces this approach is the weak binding affinity of the tested fragments with the biological targets (the proteins for example). From here, the need for a sensitive technique to elucidate this binding had emerged. Among the most sensitive techniques, nuclear magnetic resonance is one of leading approaches. NMR not only provides the sensitivity but also flexibility due to the richness of the approach with different techniques, which can be used. It’s possible to monitor the binding depending on the resonances of the binding ligand (which means that the protein spins will be only in the background not recorded). This can offer a quick way to measure the binding of moderate to weak binding ligands without the need for labeling the protein or running long multidimensional experiments. Alternatively, measuring the binding by observing the protein spins offers a unique prospect not only to test the binding but also to map the binding site on atomic scale. Herein, we describe and review different techniques and examples of the use of NMR to test binding of small fragments with different biological targets.

Keywords: NMR, Fragment-based drug design, HSQC, STD-NMR, AFP-NOESY

1. Targets for the drug-design

1.1. Protein-Ligand interaction

Protein-ligand interaction is a key process required to perform diverse of biological functions, hence, interfering with this interaction is a therapeutic concept. Ligand is a small molecule that binds a biomolecule as protein to perform a specific biological function. The binding occurs by intermolecular forces, for example, hydrogen bonds, ionic bonds and van der Waals forces, this association is usually reversible (dissociation). Association between ligand and the biomolecule leads sometimes to conformational changes and this leads to modification of the function of the biomolecule. The interaction between a protein and a ligand is usually specific, so the protein can discriminate between many different molecules and bind only one particular molecule or one of a number of very closely related (i.e. chemically and structurally similar) molecules. Protein–ligand interactions range from weak and transient to strong and persistent, depending on the strength of the non-covalent bonds.

To describe the affinity between a protein and ligand, the term dissociation constant ($K_D$) is used[1]. $K_D$ is the molar concentration of the ligand at which half of a specific site on the protein is occupied. The lower the $K_D$ the higher the affinity between the protein and the ligand. Dissociation constant can be determined by several methods, one of the most important techniques is isothermal titration calorimetry[2].

![Figure 1. Ligand binding site, which is often formed in a cleft or a pocket on the protein surface. Protein is shown in green while ligand is shown in magenta.](image)

1.2. Inhibitors vs Stabilizers

For designing ligands to interact with biological target, there are two main options:

- Compounds that work as inhibitors where they bind to the targeted protein, thus inhibiting binding of other binding partners. This binding could be in the protein active site or could be allosteric binding which changes the conformation of the protein making it unfit for the other binding partners.
- Compounds that work as stabilizers where they bind at the interface of the complex between two proteins. This approach is more difficult to design
due to the ambiguity of finding a compound that binds two or more proteins simultaneously.

Figure 2. Schematic representation of 2 main kinds of ligands that interacts with biological targets.

Surveying literature, many compounds have been reported as ligands for biological targets as inhibitors and recently as stabilizers. The pharmaceutical company Astex has reported about a small fragment (compound 1) as inhibitor for beta-secretase 1 (BACE1) which is a potential target for treatment of Alzheimer's disease. NMR techniques have been used to test the fragment[3].

Richter et al.[4] reported about a synthetic pyrazole derivative (compound 2) as stabilizer for the protein complex between 14-3-3 E and PMA2 [plant H(+)-ATPase].

Figure 3. Shows the structure of 14-3-3 E form (green) in complex with PMA2 (blue). The crystal structure revealed that a newly synthesized pyrazole derivative (comp. 2) stabilizes this complex. (the compound is shown in red within the complex)

2. Fragment based drug design approach

The indispensable starting point of FBDD is the identification of small molecule that are weak binders in the size range of 100–300 Da and characterized by being drug-like[5]. In comparison, High-throughput screening (HTS) in which diversity of compounds are distributed well-plates where the targeted protein is added and the interaction is then measured using different methods like measuring the reflectivity[6], the HTS screen can contain up to 10^5 compounds and this covers only small fraction of possible small molecules. On top of that, HTS sometimes fails to provide compounds of biological relevance in drug industry because the compounds tested are of non-drug nature and this makes them not-applicable in the pharmaceutical industry[5].

Figure 4. The main principles for fragment based drug design, which include library design, a method to test the binding of the fragments, and then the development into larger compounds using growing or linking.

The fragment library usually based on selection of a list of available chemical compounds with exclusion of toxic or reactive compounds[7]. More focused libraries also can be designed taking in account substructure of know pharmacologically active compounds[8]. The challenging point in FBDD is that the fragments tested are usually of weak binding to the tested biomolecule, hence a sensitive method are needed to test the binding. These fragments are then developed into larger compounds either via merging/linking fragments together or via growing the fragments to enhance the binding. In linking, two or more fragments that are known to be binding to different sites on the target are linked together with a linker to have the product of the K_D of the linked fragments. On the other hand, growing means to augment the binding via addition of functional groups.

There are wide variety of techniques that can be used in order to measure binding between the targeted protein and the selected fragments like NMR and surface plasmon resonance[9]. Indeed, detection of the binding between these fragments and the proteins is challenging due to the weak binding (K_D usually in the mM range). Therefore, NMR provides a unique technique due to its high sensitivity and reliability in this area. The first published fragment based lead discovery by NMR was described by...
Abbott group[10] and the binding was detected by the perturbation of the HSQC spectrum of labeled protein upon addition of the ligands. Different NMR techniques can be used to elucidate the binding events; some of them have been used in this investigation and will be discussed in details.

2.1. Detection of possible fragments for a biological target:

Most libraries are constructed based on cheminformatics pipelines. These pipelines are subject to refinements that exclude reactive or toxic molecules, assess solubility and chemical diversity, and considering similarity to other compounds that are known for similar targets[11]. Finding ligands for the targeted biomolecule can be achieved via different methods.

Examples of known methods are:

- Protein-structure-similarity clustering (PSSC): it’s used to investigate potential ligands for specific protein[12]. In this technique, proteins that share the similar structure rather than sequence identity are the ones of interest, and hence, it serves as a guiding mechanism to select natural products to target structurally similar proteins. The main disadvantage of this technique is that the 3D structure of the binding motif should be conserved for comparison, but the lack of 3D information about a lot of proteins that have pharmacological interest stands as a limitation for using this method.

- Recently, meta-structure approach can be used to design a library of different fragments based on the similarity between the targeted protein and other proteins in the database[13]. The meta-structure parameters are used as a powerful comparison tool between proteins based on sequence. Meta-structure concept was developed by Prof. Robert Konrat[13] in 2009. In this technique, 3D structure of different protein were taken from PDB, and each 3D is converted into a topological map in which each node represent an amino acid and each edge indicates a neighborhood between the two residues (neighborhood means that the distances between the Ca-Cα are less than 8 Å). From this topological map, the shortest path length connecting two residues in the network quantifies a topological parameter called θ. This parameter depends on the primary sequence distance between the two amino acids and their nature. Parameter θ is statistically evaluated using the PDB[14], and stored as a pairwise distribution functions.

Only by knowing the sequence of the protein, it’s possible to detect the meta-structure parameters, which are then aligned with other proteins in the database to predict structure and possible ligands by doing protein-Meta-Structure similarity clustering (PMSSC). As it depends only on the sequence, so it can be used even for proteins that have unknown 3D structure and of biological relevance.

2.2. NMR as a sensitive technique in detection of fragment binding:

In fragment based drug design (FBDD), different fragments are tested for binding to the target biomolecule. As mentioned before, the problem with this method is the weak affinity of these fragments, which limits other conventional methods from detecting this. Therefore, NMR serves here as a sensitive technique, that can be used to demonstrate binding of those low affinities fragments.

Two main methods are used for protein-ligand binding investigation using NMR; either observing target resonance or ligand resonance.

2.2.1. Ligand resonance observing method

Different sets of experiments have been designed in order to see the protein-ligand interaction like saturation transfer difference NMR, 13C-NMR, water-LOGSY and NOE-ROE experiments. Usually these methods provide the advantage of having short time measurement experiments in addition to the absence of limitation of the macromolecule size. The main limitation of this method is the lack of ability to detect tight binding ligands as the slow dissociation rate inhibits the transfer of the magnetization on the small bound fraction to the bulk unbound one which is the one that is recorded in this kind of experiments[15].

In order to overcome the drawback of high affinity limitation, different approaches have been adopted such as reporter screening[16] and spin labeling[17].

*Ligand resonance NMR methods in drug discovery*

Different NMR experiments have been used in different literatures, here we present summary and example of some of these methods

2.2.1.1. 1D $^1$H-Saturation Transfer Difference NMR: (STD)[18]

Saturation transfer difference approach can provide a detailed picture on the interaction between ligands and protein targets, and also to detect possible epitopes. The STD experiment can also be applied as screening technique in the dynamic combinatorial libraries to identify high-affinity ligands. The method requires only small amount of the unlabeled target protein and a 50-100-fold excess of ligand(s).

In this technique, RF irradiation of the protein spins at a resonance where no ligand signals are present is done; this is achievable because of the unique chemical environment of the protein and because of the huge line width of the protein signal in comparison to sharp peaks of the small ligands, so the protein spins have significant signal intensity.
even in segments outside the spectral window of low-molecular-weight ligands as in the negative ppm region (less than zero ppm) or above 10 ppm. This selective protein irradiation leads to saturation of the entire protein via spin diffusion. If the ligand binds the protein, saturation will spread onto the ligand. So the ligand spins that are closer to the protein receive more intermolecular NOE. As a result, intensity of the ligand signal is attenuated. Subtraction of resulting spectrum from the reference 1D $^1$H-spectrum (where the RF irradiation is done so far from the area of interest) yields the positive 1D $^1$H-STD spectrum, which contain only signals from the binding ligand (s). If the ligand is not binding, the resulting subtraction spectrum will be just a zero $^1$H-STD spectrum. The experiment is repeated in the absence of the protein to be sure that a zero $^1$H-STD spectrum is produced (negative control experiment).

$$\text{Figure 5. Schematic representation of the saturation transfer difference NMR experiment. The degree of saturation of the ligand protons is represented via the size of the proton in the figure. The larger the proton size in the figure means the closer it's to the protein surface and the more amount of magnetization it gets. (Figure is adapted from Mayer and Meyer[18] and Bruning et al.[19])}$$

### 2.2.1.2. $^{19}$F-NMR screening

$^{19}$F atom has unique chemical properties that attracted great attention in the pharmaceutical industry. It can affect the physico-chemical properties of the molecule, which is translated into better penetration into the membrane for example. In addition, C-F bond is 7 Kcal stronger than the C-H bond and that makes it more resistant to enzymatic attack. $^{19}$F-NMR ligand screening offers unique advantages; a) high sensitivity and the natural abundance of the $^{19}$F nucleus, b) no interference from protonated buffers or reagents, c) absence of overlap because of the wide dispersion of the peaks, d) the $^{19}$F transverse relaxation rate $R_2$ in bound state is highly sensitive to binding due to high exchange contribution[20]. This $R_2$ sensitivity to binding is directly reflected on the peak width. Hence, when a fluorinated ligand binds to a protein, a broadening of the $^{19}$F-peaks involved in the binding site is observed. In addition, competitive-based fluorine screening has been described as a tool for detection of binding constant. In this technique, the fluorinated ligand acts as a (spy) molecule that could be displaced by non-fluorinated ligand, which gives the advantage of avoiding the overlap between the spy, and the non-fluorinated ligand.

$$\text{Figure 6. $^{19}$F-NMR of fluorinated ligand (s) in presence and in absence of the binding protein showing broadening of the peak upon binding in comparison to the control sample (c). (The figure is adapted from Dalvit[20]).}$$

$$\text{Figure 7. Structure of compounds TFBC (3), TFMCPP (4) and FMBC (5). $^{19}$F spectra of the three compounds with trifluoroacetate (TFA) as internal standard were recorded in absence of serine protease bovine trypsin (spectrum a) and in presence of 50µM of the protein (spectrum b). (Figure is adapted from Lee et al[22].) }$$

With conventional NMR spectroscopy, it is possible to screen libraries of compounds containing CF3 and CF groups at concentrations of 18 and 35µM, respectively[21]. Due to the spectrum
simplicity of the 19F spins, it’s possible to study the binding of mixture of fluorinated ligands at the same time to one target.

Lee et al.[22] studied the binding of serine protease bovine trypsin with three different fluorinated ligands; 4-(trifluoromethyl)benzencarboximidamide hydrochloride (TFBC), 4-(trifluoromethyl)-1,5,6,7-tetrahydro-2H-cyclopenta[b]pyridin-2-one (TFMCPP), and 3-fluoro-4-methylbenzencarboximidamide hydrochloride (FMBC). (compounds 3, 4 and 5 respectively). Trifluoroacetate has been used as internal standard, which has no interaction with the protein. The comparison of the 19F-spectra in absence and in presence of the protein, revealed the binding of the 3 compounds to the protein based on reduction of peak height, increase in line broadening and chemical shift change.

2.2.1.3. 2D-homonuclear $^1$H-$^1$H NOESY

NOESY is very illuminating NMR experiment in terms of structure and it gives a clear idea about the neighborhood between protons can be extracted. The main building block of NOE experiment consists of 90˚ pulse to label the protons with their chemical shifts, followed by evolution time. Afterwards, another 90˚ pulse is applied which takes the magnetization to the z-axis, followed by a mixing time delay ($\tau_m$) during which the perturbed spins return to equilibrium via reshuffling of the spins population and hence transfer the magnetization to other spins in the system via cross relaxation ($\sigma$) in distance range 5 Å. Finally, another 90˚ pulse is applied to bring the magnetization back to the transverse plane, and followed by detection[23].

![Figure 8. The building block of NOE experiment pulse sequence.](http://www.jofamericanscience.org)

In the observed spectrum, we see mainly two types of peaks; the diagonal peaks which appear at offset ($\Omega_1, \Omega_2$) of spin 1, and cross peaks, which appear at ($\Omega_1, \Omega_2$) for spin 1 and 2 respectively. The cross peaks are the ones that carry the information about possible interaction between spin 1 and 2. The intensity of these peaks at short mixing time is ($\sigma \tau_m$). In case of fast tumbling molecules, the cross relaxation rate is positive so we see negative peak, while for large slowly tumbling molecules, $\sigma$ is negative and the cross peaks appear with positive sign (same as the diagonal peaks). The change of sign of the NOE cross relaxation rate arises from the fact that $\sigma=W_2-W_0$ where $W_2$ is the probability to have double quantum transition, while $W_0$ is the probability to have zero quantum transition. For small molecules, $W_2$ is larger than $W_0$ which makes the sign of the cross relaxation rate positive, while for large molecules, $W_2$ starts to decline while $W_0$ continues to linearly increase resulting in negative sign for $\sigma$.

![Figure 9. 2D homonuclear $^1$H-$^1$H NOE showing diagonal and cross peaks signs.](http://www.jofamericanscience.org)

The cross relaxation rate is proportional to the reciprocal of distance between the two spin ($r^6$), hence usually it’s observed if the distance between the 2 spins is in the range of 5 Å. In addition, the cross peak intensity is dependent on the mixing time, where the magnetization starts to build up, followed by a maximum and then the magnetization starts to decay. Therefore, one has to be careful with choosing the right mixing time depending on what kind of NOE wanted to be observed; in case of spin diffusion, longer mixing time should be chosen, as the intensity is now dependent on quadratic term ($\sigma^2 \tau_m^2/2$).

Measuring 2D $^1$H-$^1$H NOE at different mixing time for a small ligand in presence of protein can provide an efficient tool to observe change of the intensity of the cross peaks and to map the ligand protons that are imbedded in the binding site. From the NOE build up curves, it’s possible to see spin diffusion effect on some protons of the ligand which can be translated into binding between these protons and the protein.

2.2.1.4. Cross relaxation during Adiabatic Fast Passage[24].

The nuclear overhauser effect (NOESY) is an information-rich NMR technique for structure determination studies and it refers to the longitudinal cross relaxation. In addition, it’s also proved that Rotating Frame overhauser spectroscopy (ROESY) can provide unique information about dynamics of proteins[25]. ROESY refers to the transverse cross relaxation.
In this method, the NMR experiment measures homonuclear (1H-1H) cross relaxation rates (NOEs and ROEs) during adiabatic past passage (AFP) which is convenient method for the examination of protein ligand binding. Unlike conventional AFP methods, the RF field is not small but of comparable strength to the sweep frequency (γB1=Δω) and so leading to significant contribution of transverse relaxation to the effective spin lock relaxation rate. The pulse scheme for this experiment is a conventional NOESY in which the AFP that has parabolic phase shift in order to create a time dependent offset, replaces the original longitudinal mixing period. Therefore, the effective relaxation rate for two spins i and j is

\[ \sigma_{ij}^{\text{eff}} = \sigma_{\text{NOE}} \cos \theta \cos \theta_i + \sigma_{\text{ROE}} \sin \theta \sin \theta_j \]

Where \( \theta \) is the angle between the offset and the effective fields, \( \sigma_{\text{NOE}} \) is the NOE cross relaxation and \( \sigma_{\text{ROE}} \) is the ROE cross relaxation. For small molecules like ligands, both NOE and ROE enhancements are the same (+50%) and the spin lock relaxation rate is independent on the strength of the applied AFP strength during mixing time. For large molecules, NOE and ROE enhancements aren’t the same as they have different value and different sign (NOE becomes negative -100%, while ROE remains positive and has higher value +200%). As a result, for large molecules, the enhancement will be dependent on the AFP strength. As the strength of the AFP increases, the relaxation shifts from NOE to ROE regime. During this moving from the NOE to ROE, there will be a zero crossing. For rigid macromolecules with no internal mobility, the zero crossing happens at angle \( \theta = 35.5^\circ \).

![Figure 10. NOE and ROE enhancement (\( \eta \) is the enhancement; \( \omega^2 T_2^* \) is a parameter of size based on correlation time and the larmor frequency). (Figure is adapted from Auer et al.[24]).](http://www.jofamericanscience.org)

It’s noteworthy that this zero crossing angle can be less than 35.5° because of high internal mobility (order parameter \( S^2 > 0.5 \)), or it can be larger in case of spin diffusion.

![Figure 11. Adenosine monophosphate (compound 6) tested for binding with alcohol dehydrogenase. The ribose proton H1’ was inverted and used as a source of transferring magnetization to H2 and H8. (The figure is adapted from Auer et al.[24])](http://www.jofamericanscience.org)

2.2.2. Target resonance observing method

Herein, the changes of the macromolecule chemical shift (e.g. protein) are observed upon the binding to the ligand. One of the main advantages of this method is the possibility to detect the binding site on the protein if the assignment of the target is available. On the other hand, the biggest obstacle in this process is the size limit of the macromolecules which is usually limited to 40-50 KDa[26], as at larger size, the spectrum usually suffer from faster relaxation properties which lead in turn to line broadening. Moreover, this method is limited by the need to prepare expensive labeled sampled (e.g. \(^{15}\)N and/or \(^{13}\)C labeled).

One of the most widely used ways here is following the peak shift in assigned 2D \(^{15}\)N-\(^{1}\)H Heteronuclear single quantum coherence transfer “HSQC” spectrum of the \(^{15}\)N labeled protein (or the transverse relaxation optimized spectroscopy “TROSY” for bigger protein). Every peak in the spectrum represents the amide NH of one residue.
Upon addition of ligand, the residues that are in the binding site or affected by binding event, they show new chemical environment, which is translated in turn into chemical shift deviation. Hence, the binding site can be mapped by naming all the residues that showed change in their chemical shift upon their binding. In addition, doing titration experiments by using increasing amount of the ligand and following the peak shift can allow the estimation of binding constant.

In a new study done Boehringer Ingelheim (Canada)[27], they were able by using target-NMR to find novel inhibitor binding site on HIV-capsid N-terminal domain. The interaction between compound 7 and the Capsid N-terminal domain was recorded using the assigned 2D-TROSY spectrum of the protein. In addition, the ligand was titrated into the protein in different ratios, the peak shift deviation was recorded as a function of ligand concentration and used to extract the value of $K_D$ which was around 40µM.

Figure 12. The TROSY spectrum of the HIV capsid N-terminal domain and compound 7 in different ratios. The peaks that were shifted upon binding are highlighted with blue circle. The spectrum in black is the reference where there’s no ligand. The ligand was added to the protein in the following ratios 1.5:1 (red) 4:1 (cyan) and 8:1 (blue). (The figure is adapted from Goudreau et al.[27])

Conclusions
NMR can provide a very sensitive approach in order to investigate the binding of weakly binding fragments in FBDD using diverse techniques. In addition, the binding affinity can be estimated and evaluated using the same approach. Eventually, NMR is considered the only biophysical method that gives atomic-scale information about the drug’s binding site in solution.

Acknowledgements:
The 1st author is so thankful for Professor.Robert konrat, MFPL, Vienna University, Austria, for supervising his thesis. Our gratitude also goes to professor Dr. Magda Nasr, Vice Rector for graduate studies, Mansoura University, Egypt, for her endless support.

Corresponding Author:
Morkos Adel Henen
Department of Pharmaceutical organic chemistry
Faculty of pharmacy, Mansoura University
35516, Egypt
E-mail: morkos_henen@yahoo.com

References