NMR structural studies of protein-small molecule interactions

Shah, D.M.

English Summary

NMR methods are used for a variety of applications in a drug discovery process to guide hit to lead development and ultimately to generate potential drug candidates. The availability of 3D structural information on protein-ligand complexes is crucial during various stages of drug development. NMR based methods are sensitive toward weak protein-ligand interactions and are applicable where X-ray crystallography is least effective i.e. for weakly bound ligands. However, traditional NMR approaches that involve uniform isotopic labeling are labor intensive and limited by the size of protein targets (e.g., < 30-40 kDa). Therefore, these have not been widely adopted in drug discovery. The main focus of the work described in this thesis was to develop and implement efficient NMR methods that are capable of providing 3D structural information on protein-ligand complexes in the early stages of the drug discovery projects. The work presented in this thesis covers a wide range of classical and newly developed NMR techniques applied to diverse molecular systems.

Chapter 1 provides general introduction to the drug discovery process, in particular, to the fragment-based drug discovery. The chapter also describes NMR methods that are often used for screening purposes and to obtain 3D structural information on protein target-small molecule complexes.

In Chapter 2, a fragment based- small molecule discovery approach was sought to discover inhibitors that disrupt the DNA binding capability of TEL, a target for oncology. TEL is a DNA binding protein involved in transcriptional regulation of many cellular proteins that are involved in the regulation of angiogenesis. TEL is considered to be a therapeutic target for pathological angiogenesis in tumor cells. Chapter 2 describes the techniques and approaches that were undertaken to discover small molecule ligands that specifically bind the DNA binding ETS-domain of TEL. In order to discover the primary hit matter, a TINS NMR fragment screen was

applied. To demonstrate the DNA binding capability of TEL_{ETS}, different array of techniques were used. Biochemical and structural analysis were performed using protein observed NMR, Surface Plasmon Resonance (SPR) and gel-shift assay to characterize DNA binding activity of TEL. The availability of the resonance assignments allowed determining the binding site of DNA and fragments hits. Based on the chemical shift perturbation analysis and CSP mapping, it was clearly observed that the binding site for the DNA and three fragments was common on the protein surface. To investigate whether these fragments were capable of disrupting the TEL_{ETS}-DNA complex, a gel retardation assay was performed. The assay results indicated that the TEL_{ETS}-DNA complex could be disrupted in the presence of high (millimolar) concentrations of 2 fragments. Clearly, the affinity for these fragments appears to be very low and the specificity, in relation to the other ETS-containing proteins remains to be tested. However, these fragments represent valuable starting points for further stages of drug discovery on TEL_{ETS}. This work demonstrates how the application of NMR screening can be efficiently used to discover fragment hits and protein-observed NMR studies can be used to obtain low resolution structural information on protein-ligand complexes.

In Chapter 3, a fragment based- small molecule discovery approach was initiated to discover small molecule inhibitors for Rit1 GTPase, a target for rheumatoid arthritis. Cellular studies carried out by Galapagos BV showed that inactivation of Rit1 reduced the levels of matrix metalloproteinases (MMPs) in bone cartilage tissue and that increased expression of Rit1 caused MMP levels to increase. MMPs are responsible for degradation of cartilage tissue and hence cause RA. Therefore, development of small molecule inhibitors that could inactivate Rit1 is important for a potential RA treatment. Chapter 3 describes the results obtained from

fragment based screening, crystallization, analoging (compounds with similar chemical structures), hit development and a structural study on the most potent compound. A diverse range of techniques and methods was employed. Fragment hits were discovered using TINS NMR. The fragment hits were further tested in an *in vitro* biochemical assay. These hits and analog compounds were able to inhibit the GDP-GTP nucleotide exchange. In other words, compounds were able to stabilize the GDP bound inactive state of Rit1. Despite their functional inhibition, the mechanism of action of compounds remained elusive. High resolution 3D structures of the protein-compound would give a direct insight into the mechanism. The crystallography efforts yielded a high resolution 3D structure of the Rit1 bound to GDP nucleotide but efforts to obtain 3D structures on protein-ligand complexes failed. Substantial efforts to crystallize the complex were not successful due to the low solvent content of the protein crystals and cracking of protein crystals in the presence of the most potent compound. To address this issue, a solution based NMR approach was necessary and a paramagnetic NMR based approach was sought, whereby a spin label was introduced on GDP to obtain GDP-spin label. The paramagnetic NMR approach was combined with CSP analysis to obtain low resolution information on the binding site of ligands. Paramagnetic studies using the GDP-spin label followed by docking calculations proposed a novel mechanism by which the compound inhibits GDP-GTP exchange of Rit1. In one of the binding sites, the ligand sterically blocks the GDP nucleotide by binding over the top of the nucleotide binding site of the protein. This explains that by physically trapping the GDP nucleotide, the ligand stabilizes the GDP bound form of the protein and the nucleotide exchange with GTP is inhibited. This work shows how NMR-based techniques can provide structural information when other biophysical approaches to

obtain binding site information are unsuccessful. Here, using the PRE based NMR method, we were able to understand the mechanism exhibited by the compounds to achieve the inhibition of nucleotide exchange. PRE based methods present an alternative to obtain binding site information when other high resolution techniques fail. In principle, GDP-SL can be used to investigate any other GTPase.

In Chapter 4, a solution NMR method was developed to obtain 3D structures of protein-small molecule complexes in rapid and efficient manner. The method was developed using the Hsp90-fragment complex as a model system. Hsp90 interacts with a large number of client proteins and it is thought that ATP hydrolysis is the main driver for these interactions and for this reason many drug discovery approaches have focused on targeting the ATP binding pocket in the N-terminal domain of Hsp90. The method described in Chapter 4 makes use of a small molecule fragment that binds in the ATP binding pocket of the N-terminal Hsp90 discovered by TINS NMR screening. The main goal was to use this protein-ligand system to develop an efficient way to obtain 3D structural information on protein bound to a ligand. In Chapter 4, an NMR method that is based on the production of a protein target that is selectively isotopically labeled with the methyl groups of isoleucine, leucine and valine residues is described. In this method standard NOESY experiments were used to obtain a set of intermolecular NOEs between the protein-ligand complex. The intermolecular NOEs were used as distance restraints to obtain 3D co-structures by molecular docking. The crystal structures were available for the identical ligand with Hsp90 in the PDB. The crystal structures show that two or more low energy conformations of the ligand are present in the complex. In contrast, the NMR data indicate that in solution, there is one predominant ligand conformation. Importantly, despite the moderate number of intermolecular restraints, the NMR 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. The method shows 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. This chapter demonstrates how a combination of selective methyl group labeling, standard NMR experiments and computational docking can be used to rapidly determine the 3D structure of a small molecule bound “weakly” to a protein target. The approach requires only a sparse set of intermolecular NOEs and is an alternative to time consuming traditional NMR approaches that involve uniform isotope labeling. This method is also amenable to large molecular weight targets.

Chapter 5 presents the general discussion of the work and provides future perspectives for NMR methods developed and implemented in this thesis.