NMR spectroscopy is an advantageous tool for fragment-based lead discovery because it combines high sensitivity and robustness to loose fragment binders with unique structural resolution and throughput. Favourable applications include screening against shallow binding pockets (such as protein-protein interactions) and side-pockets outside the highly conserved active centre of the major target classes (such as protein kinases and proteases) to gain selectivity. A set of protein- and ligand-detected NMR assay techniques allows for the identification of novel chemotypes and the determination of protein-ligand structures for soluble target proteins of up to 100kD.

AN ATTRACTIVE STARTING POINT FOR LEAD DISCOVERY

High-throughput screening (HTS) using compound collections with average molecular weights (MW) of around 400Da or higher has often resulted in lead compounds with inappropriate ADME properties due to violations of the ‘Rule of five’ MW limit (1). This finding prompted substantial compound collection enhancement initiatives and re-profiling towards a lower MW in many pharmaceutical enterprises. Fragment-based lead discovery is a logical extension of this trend towards ‘small is beautiful’, using collections of unparalleled low MW compounds between 100 and 300 Da – designated ‘fragments’, ‘needles’ or ‘shapes’ (2). Fragments are attractive starting points for lead discovery because they: (1) show better binding efficiencies for a given protein pocket; (2) give more room for MW increases during hit-to-lead optimisation; (3) yield new unexpected chemotypes for the generation of novel chemical matter; and (4) allow for simpler synthesis protocols. The improved binding efficiency of fragments – defined as the free energy of binding divided by the number of heavy atoms or MW – results from their additional degrees of rotational and translational freedom, allowing for an optimal positioning within the protein pocket (Figure 1A). By contrast, identical sub-structures in higher MW compound collections are constrained in their freedom of positioning by the presence of bulky substituents (Figure 1B).

ADVANTAGES OF NMR FOR FRAGMENT SCREENING

Fragments exhibit significantly weaker binding affinities, typically in the micromolar to millimolar range, than compounds of higher MW. Hence, exceptionally sensitive and robust detection technologies, such as nuclear magnetic resonance (NMR) spectroscopy, are required for assaying fragment binding to target proteins. In the absence of structural information, fragment-to-lead optimisation is more challenging for loose fragment hits than for tightly binding hits because derivatives of fragments are more flexible to change their binding geometry, thereby showing potentially misleading structure activity relationships (SAR). Therefore 3D structural information about the binding sites of fragments, which are obtainable from NMR assays for tens to hundreds of derivatives, strongly assists SAR studies. This unique combination of high sensitivity to loosely binding fragments and medium- to high-throughput for binding site determination renders NMR an advantageous fragment screening technique, as evidenced by numerous successful applications of NMR-screening in the literature (3).
Innovations in Pharmaceutical Technology

PROTEIN-DETECTED NMR-SCREENING
The best known and structurally most informative NMR-screening technique is the SAR-by-NMR method pioneered by Abbott Laboratories. It senses local chemical shift perturbations (CSPs) for the NMR signals of the protein when a ligand binds, which led to the name ‘Protein-Detected NMR-Screening’. The method uniquely allows direct determination of the binding site of the ligand from the local cluster of typically 5-10 CSPs surrounding a fragment binder (Figure 2). Numerous publications show that this method is amenable to ligand molecular weights ranging from minute Mg²⁺ cations to macromolecular proteins or nucleic acids. Likewise, ligand affinities ranging from two digit millimolar down to subnanomolar Kd values are robustly detected.

By using powerful isotope labelling strategies and relaxation-optimised NMR pulse programmes, soluble proteins of up to 50kD can be studied. This limit includes the active domains of many therapeutically relevant target proteins. More serious limitations arise from the necessity for stable isotope labelling with ‘N, ‘H or ‘C nuclei, which is generally cost-efficient only if the protein can be produced from E. coli or yeast expression systems with yields of larger than approximately 2-5mg/L. Alternatively, moderate quantities of ‘N-labelled protein can be obtained from SP9 expression systems provided that expression yields are sufficient.

LIGAND-DETECTED NMR-SCREENING
Ligand-detected NMR-screening constitutes the ideal back-up technique for target proteins that are not amenable to protein-detected NMR-screening due to MW >50kD, inaccessibility to isotope labelling, low expression yields or low solubility of <30µM. According to their name, ligand-detected methods measure the ‘H NMR signals of the small molecule ligand, which differ in their intensity, relaxation properties and sign between bound and unbound ligands. Usually ligands are assayed in 10- to 50-fold excess over protein, which yields improved signal-to-noise but requires that ligands bind in the fast exchange regime on the NMR time-scale. This fast exchange requirement limits ligand-detected NMR assays to an affinity window between approximately 0.1µM and 10mM, which perfectly covers the typical affinity range of fragment binders, but precludes ligands with high potency and slow off-rates.

WATERLOGSY AND STD NMR-SCREENING
WaterLOGSY (Water ligand optimised gradient spectroscopy) is one of the most sensitive ligand-detected NMR-screening techniques, especially for drug-like synthetic compounds with limited solubility. In the WaterLOGSY NMR experiment, magnetisation is transferred to the abundant solvent H₂O, from where it diffuses with opposite sign onto bound and unbound ligands. When the ligand magnetisation is finally detected in a 1D ‘H NMR experiment, a binding ligand shows positive ‘H signals, whereas a non-binding ligand shows negative ‘H signals (Figure 3B).

In the ligand-detected Saturation-Transfer-Difference (STD) experiment, the hydrogen atoms of the target protein (usually the methyl protons) are selectively excited. This magnetisation diffuses across the protein and onto bound ligands, whereas non-binding ligands are not affected. By recording 1D ‘H NMR difference spectra between on- and off-resonance excitation of the protein, only interacting ligands show up as positive ‘H signals, while non-interacting ligands are differenced to zero intensity.

STRUCTURAL INFORMATION FROM LIGAND-DETECTED NMR ASSAYS
Structural information about the binding site of fragments can be obtained by running WaterLOGSY or STD NMR assays in the presence of competitor or reporter compounds. For example, to discriminate between binding to the catalytic and allosteric sites of a given protein, a competitor – whose binding site is known – can be used to displace a fragment binder. Displacement is monitored by

<Figure 1: Fragments yield better binding efficiencies>
(A) A fragment hit is free to position its donor (D), acceptor (A) and hydrophobic (Ph) pharmacophore points within optimal distances to the complementary points of the protein pocket. (B) A higher molecular weight binder is constrained in its freedom of positioning, resulting in a sub-optimal interaction of pharmacophore points and a reduced binding efficiency of the identical substructure.

<Figure 2: The concept of protein-detected NMR-screening>
Compounds of the fragment library (A) are mixed with the Drug Target (B). When the fragment binds, local protein chemical shift perturbations (red H atoms) are detected by NMR, resulting in an output of novel fragment hits (C).
the concentration-dependent weakening of the positive WaterLOGSY or STD \(^1\)H signals of the fragment binder. Such competition NMR assays are commonly applied as a secondary assay on the hits of a direct primary assay.

Alternatively, a reporter compound of a known binding site can immediately be used in the primary NMR assay. Novel fragments binding to the site of the reporter are detected through the weakening of the reporter’s WaterLOGSY or STD signals (Figure 3). Fragments binding to other sites show up as additional positive signals elsewhere in the \(^1\)H spectrum. A suitable reporter compound must bind to a single site of the protein in the fast exchange regime, and should possess an affinity between approximately 10 and 500\(\mu\)M. Substructures of potent lead compounds are good candidates for matching these criteria.

Moreover, atomic resolution structures of protein-ligand complexes can be determined by using the STD-NMR assay, provided that the 3D structure of the protein is known (4). Distance restraints between protein and ligand are obtained by recording STD-NMR spectra on four to six amino acid type labelled and perdeuterated protein samples in the presence of the ligand of interest. Combining this set of distance restraints during structure calculation results in high resolution protein-ligand complexes. Thus, STD-NMR can yield complex structures for protein-ligand pairs not amenable to co-crystallisation or soaking.

**LEAD DISCOVERY FOR PROTEIN-PROTEIN INTERACTION TARGETS**

Target proteins with shallow binding pockets, such as protein-protein interactions, require screening technologies with enhanced sensitivity for loose binders and excellent binding efficiency because conventional screening assays usually fail for such targets. By contrast, NMR-fragment-screening routinely identifies such weak fragment binders which better fit into such pockets. In addition, the structural information guides the hit-to-lead process in which the chemical substructures of fragments binding next to each other are merged into a larger, more potent and more specific compound. In the absence of structural data, a much larger and often prohibitive number of derivatives has to be synthesised for each fragment hit.

In-house research in collaboration with the academic group of Professor Hartmut Oschkinat (Forschungsinstitut für Molekulare Pharmakologie, Berlin) supports this concept (5). The PDZ domain of the human protein AF6 mediates protein-protein interactions through a shallow groove that binds to the C-terminal 4 to 7 residues of membrane receptors and ion channels (Figure 4, page 50). The affinity to natural 6-residue peptide mimetics is in the weak three digit micromolar range indicating low drugability. However, by using protein-detected NMR-screening, three chemically distinct classes of fragment hits can be identified. One chemical class was synthetically optimised with support from the structural binding site data. The high resolution 3D structure of the complex between the PDZ domain and the best derivative confirms binding to the active site, and guides further hit-to-lead chemistry.
FRAGMENT-SCREENING FOR SELECTIVE LEADS

Lead structures for the major target classes – such as protein kinases, protein phosphatases and proteases – frequently suffer from insufficient selectivity for the single targeted protein, because enzymatic pockets are similar within a class. This selectivity problem can be addressed by using NMR-fragment-screening to find fragment binders to protein sub-pockets surrounding the enzymatic site. Such sub-pockets are usually not conserved across the target class, thus allowing for interactions selective to the targeted protein.

A successful application of this concept is the generation of highly selective leads for the challenging anti-diabetes target, protein-tyrosine-phosphatase 1b (PTP-1b) (6). By protein-detected NMR-screening, novel fragment hits to a neighbouring sub-pocket were identified and chemically linked to the known core lead structure in the enzymatic pocket. The resulting new lead structure shows better selectivity over side effect-relevant phosphatases than numerous other PTP-1b inhibitors disclosed at that time.

Dependent on the molecular weight and accessibility of the target protein to isotope labelling, any of the previously described NMR-fragment-screening methods can be applied to search for sub-pocket binders. The common advantage of all NMR methods stems from the fact that loose binders to sub-pockets outside the enzymatic cavity can be detected, which is impossible by biochemical assays.

NOVEL CHEMOTYPES FOR LIBRARY DESIGN

The chemical space for novel inhibitor scaffolds to major target classes, such as protein kinases and proteases, has been narrowed down through extensive screening and lead optimisation projects. Thus, novel chemotypes are of paramount interest for the design of new focussed libraries and lead generation. The 20 different binding geometries of protein kinase inhibitor scaffolds show that fragment-like scaffolds can adopt distinct and unexpected binding geometries in which donor and acceptor atoms, as well as substituents, are positioned in novel ways (7). NMR-fragment-screening is ideally suited to measure the binding efficiency and binding geometry of loosely interacting combinatorial chemistry scaffolds. First applications have been explored but are not yet published.

CONCLUSION

Originating from the statistical analysis of successful low MW drugs of the past, fragment-based lead discovery has become an attractive tool because it yields novel chemotypes with better binding efficiencies and ample room for MW increases during hit-to-lead optimisation. The technological challenges of fragment-based discovery are ideally met by the high sensitivity and the unique binding site information and throughput of NMR fragment assays. These advantages open the door for the screening of protein-protein interactions and non-conserved sub-sites next to the catalytic centres of validated targets. Soluble proteins up to 50 and 100kD are accessible to protein- and ligand-detected NMR-screening respectively, but membrane proteins and proteins with low expression yields remain intractable.

The author can be contacted at Schade@Combinature.com

References