Application of fragment-based drug discovery to membrane proteins

Früh, V.

Citation

Früh, V. (2009, October 7). *Application of fragment-based drug discovery to membrane proteins*. Retrieved from https://hdl.handle.net/1887/14047

Version: Not Applicable (or Unknown)

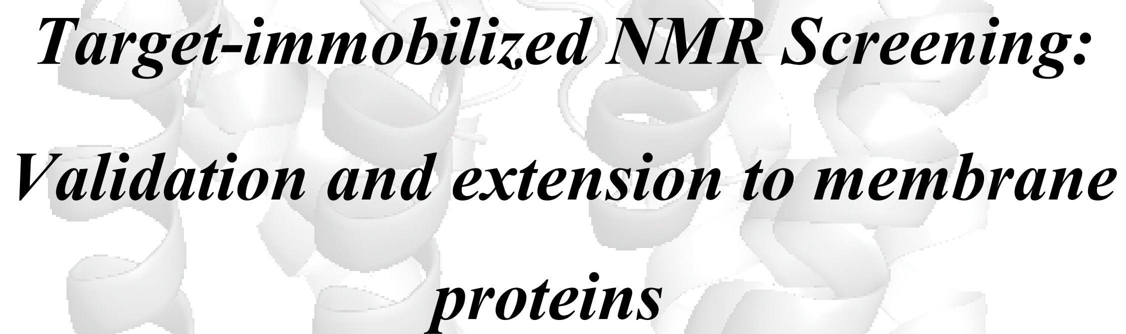
License: Leiden University Non-exclusive license

Downloaded from: https://hdl.handle.net/1887/14047

Note: To cite this publication please use the final published version (if applicable).

Chapter

4



Target-immobilized NMR Screening: Validation and extension to membrane proteins

Due to their involvement in a large number of pathologies and ease of access, membrane proteins are at the forefront of pharmacological interest. A number of high throughput screening technologies, including membrane and cell-based assays, have been used to develop lead compounds. Emerging biophysical techniques such as nuclear magnetic resonance (NMR) can potentially be used to find novel ligands, however to date they have only been applied to soluble proteins. We seek to apply our newly developed ligand screening technology, Target Immobilized NMR Screening (TINS), to membrane proteins. In TINS the target to be screened is immobilized on a porous support and flow-injection NMR spectroscopy is used to detect binders in mixes of compounds. Immobilization potentially allows TINS to be applied to membrane proteins. Here we address this issue by developing a simple and widely applicable immobilization protocol and applying it to model bacterial membrane proteins. Since TINS is a comparative method, we have developed a reference system to control for non-specific binding of hydrophobic compounds to lipids or detergents used to solubilize the membrane proteins. This control protein is of the same size as our target and refolded in the same lipid micelles. The proof of principle was tested with a limited screen small fragment (150 - 300 Da) library on a bacterial ion channel solubilized in dodecylphosphocholine micelles and a larger screen of approximately 1000 fragments was carried out on a membrane enzyme the Disulphide Bond Forming protein B. For both screens, 7 % of the fragments showed substantial changes in the NMR spectrum that were specific to the target and were considered binders. No loss of target binding capacity was detected after 1 month of storage of the samples. With the proof of principle validated, TINS is now ready to be applied to pharmaceutically[2] important, membrane bound drug targets.

[2] This chapter is a modified version of the published book chapter: Früh, V.; Heetebrij, R.; Siegal, G. Target Immobilized NMR Screening: Validation and Extension to Membrane Proteins. In *Fragment-Based Drug Discovery: A Practical Approach*; John Wiley & Sons: Chichester, 2008; Ch. 6.

Fragment based drug discovery (FBDD) methods have been widely embraced in the last few years. Nearly all of the major pharmaceutical firms have developed fragment screening and evolution programs and a number of biotech firms have sprung up that make exclusive use of the approach to develop small molecule therapeutics. Amongst the variety of fragment screening and evolution methods that have been implemented, there are two common themes. First, the collection of compounds to be screened consist of small (typically less than 300 Da), highly soluble molecules. As such, they typically interact with the target weakly, with binding constants in the range of 2 to 5,000 μM. Second, the low affinity hits discovered by screening such a collection must be developed into high affinity, high specificity ligands. This process is much more successful when 3D structures of target-compound complexes are available[21].

The promise of FBDD, that is compounds that through obeying Lipinski's rules[10] are more likely to make orally bioavailable, safe drugs, is starting to be put to the test as compounds begin to move into clinical trials. The number of such compounds is rising rapidly due to the successes of Plexxikon, Astex, Sunesis, SGX Pharma, and a host of other biotech companies that place FBDD at the core of their activities. However, a third common theme that applies to all FBDD to date is that is has been strictly applied to soluble targets. On the other hand, the attractiveness of membrane proteins as pharmaceutical targets has been well documented[204] with approximately 60% of all current targets being membrane proteins. Thus it would be a significant advantage to be able to apply FBDD to the class of targets that includes integral and membrane associated proteins.

We have developed a technology called Target Immobilized NMR Screening (TINS)[77,205] that in principle can be applied to screening of membrane proteins. In TINS, the target to be screened is immobilized on a commercially available chromatography resin in a simple and efficient process. The immobilized target, along with a second, reference sample, is placed in a flow-injection, dual cell sample holder in the magnet and the compounds to be screened are injected in mixes of about 5 compounds each[205]. Spatially selective spectroscopy[206] is then used to independently acquire a 1D ^{1}H spectrum of the compounds in the presence of the target or the reference. Comparison of the two spectra directly yields the identity of any compound that binds the target due to the simple reduction in peak amplitude of all resonances from the ligand. This

configuration yields a number of advantages for ligand screening. The combination of effective T_2 relaxation and chemical exchange endows the method with great sensitivity with specific binding as weak as 5 - 10 mM (K_d) being readily detected. On the other hand, the presence of a reference sample in routine use cancels the weak, non-specific interactions typically observed between many of the compounds to be screened and the target. Thus the presence of artefacts in TINS screens is greatly reduced as is the false positive rate. The sensitivity can also be used to reduce the concentration of immobilized target to as low as 5 μM solution equivalent, which combined with the fact that the entire compound collection is routinely screened with a single sample, means the screening can be carried out with as little as 5 nmols of the target.

TINS has been applied to a variety of soluble proteins and in this chapter we will present some of these results. In principle, immobilization should allow an extension of the range of targets to which TINS can be applied to include insoluble membrane proteins. This idea is not new and others have attempted to apply biophysical methods for detecting ligand binding to immobilized membrane proteins[151]. In particular, surface plasmon resonance (SPR) has been used for this application. Membrane proteins represent difficult targets for *in vitro* ligand screening studies however since they are insoluble, often require the presence of specific lipids for proper function, are highly challenging to purify, and rarely amenable to high resolution structural analysis. Furthermore, a general limitation that has always been encountered is the difficulty of functionally immobilizing membrane proteins in a form appropriate for the assay. SPR for instance requires a flat surface with an underlying metal layer (to provide the material with dielectric constant opposite that of water). While a few cases of successful immobilization of membrane proteins have been reported under these conditions, a widely applicable method is still lacking. Here we will report on our initial efforts in two areas, the ultimate goal of which is to allow routine *in vitro* fragment screening of a wide variety of membrane proteins.

General considerations for Fragment Screening

Fragments

We will focus on the principles and benefits of the TINS fragment library designed and tested as collaborative effort between ZoBio (www.zobio.com) and Pyxis Discovery (www.pyxis-discovery.com) of Delft, the Netherlands[207].

It is now a well accepted principle that the "rule of 3"[15] forms an approximate limit guiding the chemical nature of compounds that should be considered as a fragment for inclusion in a collection for ligand screening. At the other end of the spectrum, recent work from the Shoichet[208] lab suggests that including very simple fragments of less than approximately 150 Da could cause difficulties downstream during the lead evolution process. Clearly a number of *in silico* filters must also be employed to remove undesirable compounds such as known toxicophores or reactive groups. In our efforts we also placed great emphasis on water solubility of the compounds. In one of the first publications concerning fragment library design, only about 50 % of the selected fragments possessed sufficient solubility (1 mM) to be screened[18]. In more recent publications, better results for the water solubility of fragment libraries have been reported[31,209]. The prediction of water solubility however remains a challenge because one has to take into consideration both the crystal and solution state of the compound. Moreover, in our own analysis, we have not been able to find a simple correlation between the number of H-bond donors/acceptors and water solubility. Since computational methods for better prediction of water solubility are still under development, one must experimentally determine the solubility of a given fragment. However, by applying cut-off values based on experience, for properties that can be better predicted, such as cLogP and the number of hydrogen bond donors and acceptors, which have a profound influence on water solubility, the fraction of water soluble fragments can be increased considerably. In our own efforts, about 90 % of compounds that were selected were soluble as singletons at 500 μM in phosphate buffered saline and 5 % DMSO. Evotec has recently mentioned an in-house QSAR model to predict solubility which is claimed to be useful, but no data is presently available[210]. While originally our emphasis on water solubility stemmed

from practical aspects of making mixes of compounds at 500 μM each in aqueous buffer, this effort has been well served when screening membrane proteins since we feel that it is one of the important reasons that we have so far experienced a very low false positive rate.

Our library, which is intended to serve as a source of chemical diversity, is composed of compounds selected from four themes: (1) diversity using the scaffold-based classification approach (SCA)[61], (2) amino acid derivatives, (3) scaffolds found in natural products, and (4) shape diversity. All compounds were selected from a carefully prepared database representing 70,000 compounds that would make desirable starting points for drug discovery, including “rule of 3” compliance, and were commercially available from reliable suppliers. One of our explicit intentions in forming the library upon these design principles is to evaluate the performance of the various classes of compounds against different targets, both soluble and membrane bound. While it remains too early to make sensible conclusions from the roughly 10 targets that have been screened to date, in many cases there are up to two fold differences in hit rates between the different themes for a given target.

Immobilization and reference protein

The strength of TINS lies in the fact that it is a referential system. That is the signal acquired in the presence of the target protein is compared to the signal acquired in the presence of a reference sample consisting of a known protein immobilized at approximately the same density as the target. The requirement for a reference protein comes from the fact that TINS is highly sensitive to even very weak interactions between the compounds and the immobilized target. Therefore the choice of reference protein is important. Ideally one would like to have a reference protein which is convenient to produce in large quantities, can be readily immobilized, has the roughly “typical” amounts of exposed surface charge and hydrophobicity and has essentially no small molecule binding capacity. The PH domain of the cellular kinase AKT is a nearly ideal candidate which we use for screening of all soluble targets. Hajduk and colleagues showed that this protein was essentially refractory to small molecule binding using their well-known SAR by NMR assay[211].

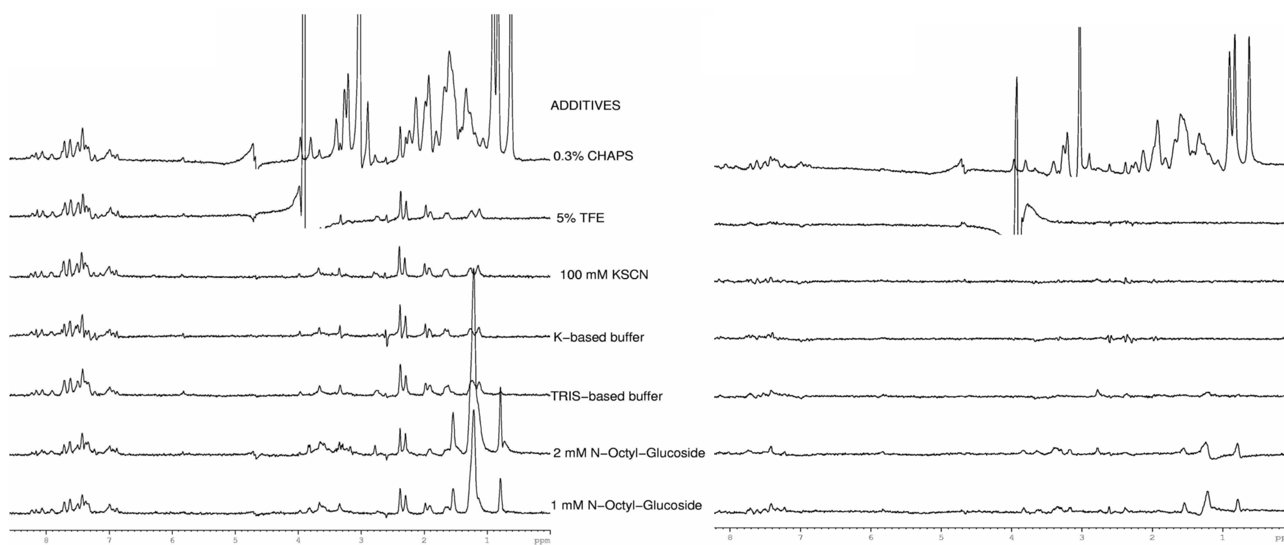


Figure 1. Cancellation of non-specific binding by the reference sample in TINS screening. The left hand panel shows difference ^{1}H NMR spectra of a mixture of non-binding compounds acquired in the presence of sepharose resin to which 6 mg/ml of an SH2 domain (111 amino acid residues) had been immobilized or just the resin itself. The indicated additive was included with each of the compound mixtures. The right hand panel shows the same difference spectra however the second spectra was acquired in the presence of a resin to which 6 mg/ml of FKBP had been immobilized. The improvement in cancellation when an immobilized protein is used as a reference is clear.

While we initially had concerns that this small protein would be unrepresentative of larger, potentially multi-domain targets, or that proper cancellation of non-specific binding would require accurate matching of total surface area, this turns out not to be the case as shown in Figures 1 and 2.

Immobilization is a constant source of questions with regards to TINS screening. In principle, one is free to choose any immobilization approach which is compatible with a) the biochemical function of the protein and b) the constraints of NMR. Specifically the major concern related to NMR is susceptibility mismatch between the solid support and the surrounding aqueous environment. The group of Meyer had originally demonstrated ligand binding to targets immobilized on glass beads[212]. However, the susceptibility mismatch was so severe in this case that magic angle spinning NMR was necessary to average out the inhomogeneity. Clearly this

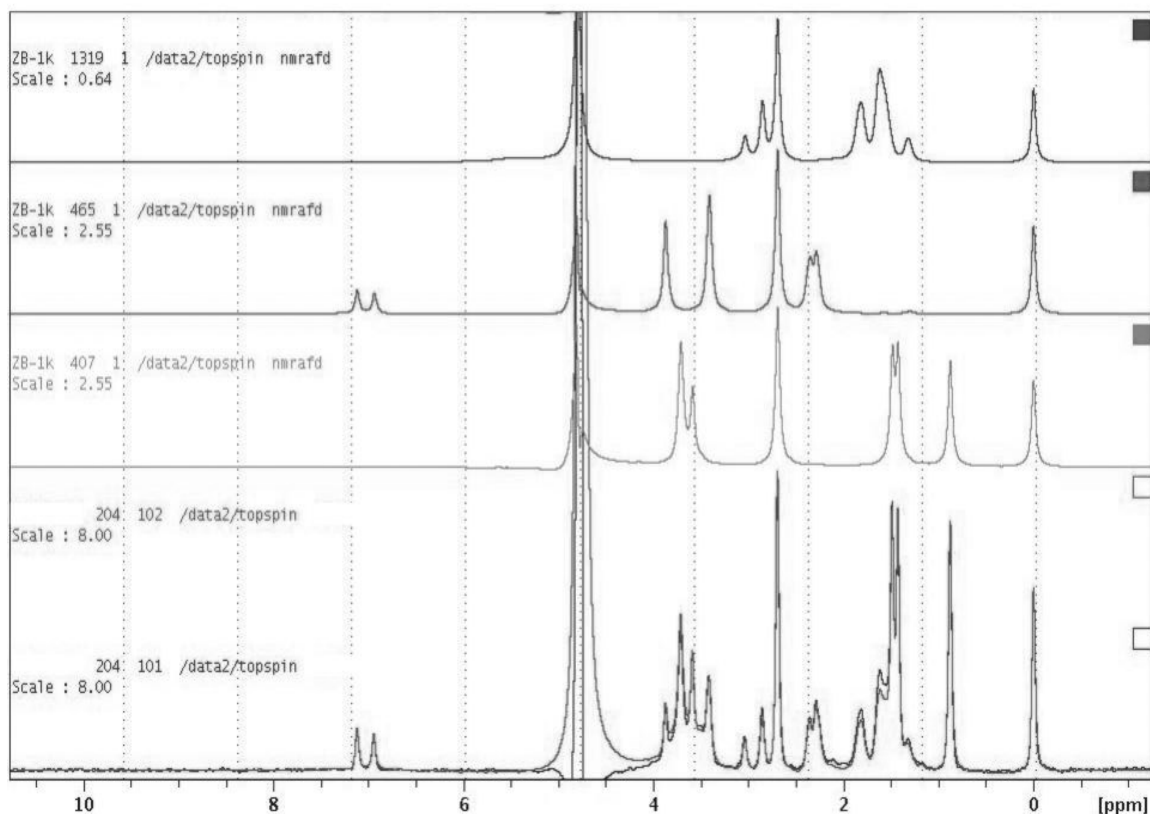


Figure 2. In this example taken from a screen of a soluble target, both the target and the reference protein (the PH domain of the kinase AKT) were immobilized on Actigel ALD (Sterogene, USA) at a solution equivalent of 100 μM. A mix consisting of three different compounds (upper three 1D 1H NMR spectra are of each compound in the mix separately) was applied simultaneously to the sample of immobilized target and reference protein in the dual cell sample holder. Spatially selective Hadamard spectroscopy was used to simultaneously acquire separate spectra of the compound mix in the presence of the immobilized target and reference. These spectra are overlaid at the bottom of the figure. The similarity of the two spectra indicates that none of the compounds specifically bind the target. The weak interactions with any immobilized protein that are observed for most compounds in the library are approximately the same for both the reference and target.

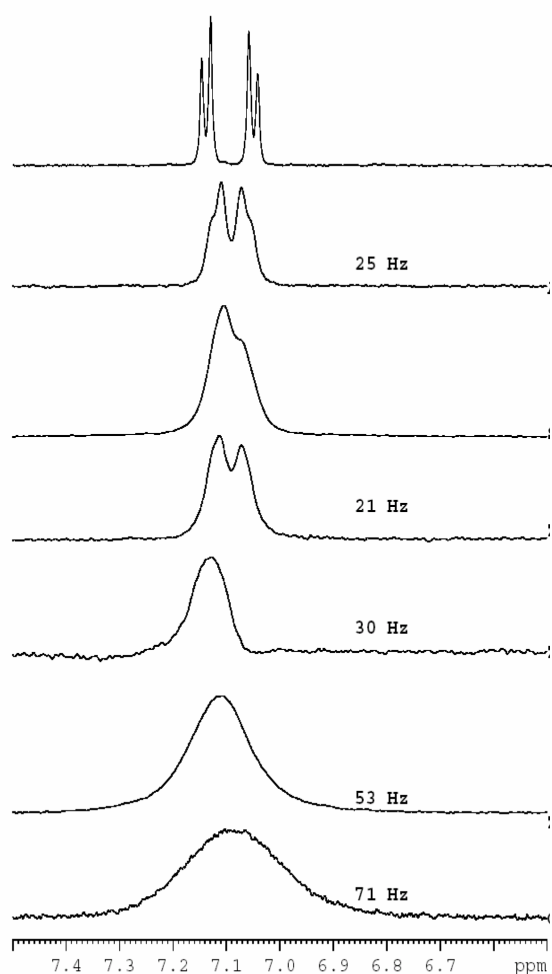


Figure 3. Effect of immobilization chemistry on the linewidth of compounds in solution. 1D 1H spectra of the aromatic protons of phosphotyrosine (pY) are shown with the fitted linewidth. From top to bottom, pY in solution, in the presence of Actigel ALD, streptavidin-sepharose, Zn-IDAA sepharose, Zn-NTA sepharose, Zn-NTA silica and controlled pore glass beads (for comparison).

arrangement would not be compatible with flow-injection NMR and so we sought a solid support which would not bind the compounds, would provide high capacity to immobilize proteins, and would minimize susceptibility differences. Sepharose based affinity resins turn out to be very useful in that they are very good matches for this list of requirements. In contrast to glass beads, sepharose beads can be more readily described as a three dimensional, bio-compatible mesh which is highly hydrated, yet sufficiently rigid to maintain good flow characteristics even after 300 applications of compound mixes. The susceptibility mismatch is minimal such that under our current screening setup, using the dual-cell sample holder made from KelF, we routinely obtain a linewidth of about 12 Hz. However, the nature of the immobilization chemistry of the sepharose bead also appears to play a role in the linewidth observed for the compounds, as can be seen in Figure 3.

A wide range of immobilization chemistries are commercially available in conjunction with sepharose beads. We have investigated a limited subset of these possibilities which include: direct, non-oriented immobilization via Schiff's base chemistry, oriented non-covalent immobilization *via* immobilized metal affinity chromatography

resins and oriented non-covalent immobilization via biotin-streptavidin binding. At present we favour direct, covalent attachment of proteins via primary amines since it is highly efficient (typically better than 85 % yield), minimizes leaching, and provides the best NMR results (Figure 3).

At the pH we typically carryout immobilization (pH 7.4), this reaction is rather specific for the amino terminus. In principle one could imagine that immobilization might interfere with the functionality of certain proteins, such as kinases that contain a lysine at an active site. Thus far we have not encountered this issue, but it is always possible to block access to this lysine by immobilizing in the presence of high levels of an ATP mimic such as AMPPNP. Kinases have been successfully immobilized for Biacore studies using related chemistry[213]. We have investigated the use of IMAC resins to immobilize proteins via a 6 his tag. While this method is convenient, it is not possible to use Ni^{2+} as the ion for chelating the tagged protein due to the potent paramagnetic relaxation. It is possible to immobilize his tagged protein using Zn^{2+} instead and leaching does not pose a problem. However, despite the fact that a sepharose resin is used in conjunction with a diamagnetic ion, there appears to be additional linebroadening effects (Figure 3). These may result from non-specific interactions with available NTA sites on the resin which turn out to be difficult to block. We have also used streptavidin sepharose to immobilize biotinylated ribonucleotides for ligand binding studies. This system is convenient and yields high quality NMR spectra (not shown). By blocking unoccupied binding sites with free biotin (and naturally using streptavidin sepharose as the reference sample) one should be able to limit small molecule binding to sites that are not on the target, however we have not carried out a full screen on such a system so it not possible to make a definitive statement at this time. Other affinity tags can also form the basis of successful, NMR compatible immobilization as well. For example, Haselhorst and colleagues have recently reported the use of Strep-tactin sepharose, a variant of streptavidin sepharose, to perform saturation transfer difference (STD) studies[214].

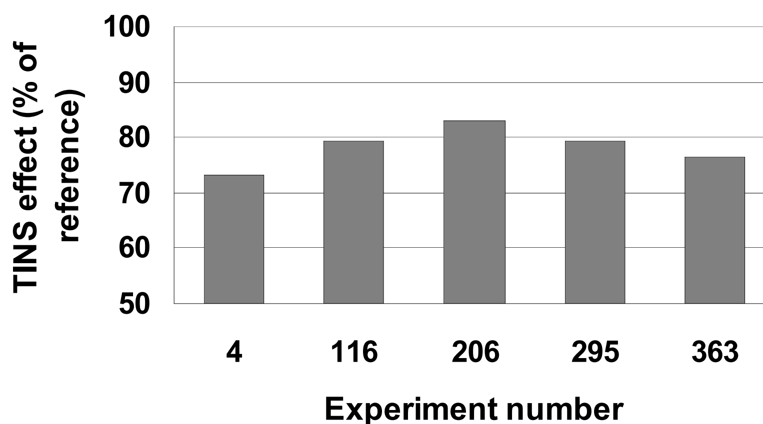
Ligand Screening

We decided to carry out our ligand screening studies using mixes of compounds at a very early stage in the process of developing TINS. This decision was made on the basis of throughput and robustness. Since our mixes consist of on average 5 compounds, obviously throughput is increased by a factor of 5 with respect to screening singletons. Also, since it is expected that only 1 (and occasionally 2) compounds per mix bind to the target, most peaks in the reference and target spectra should be of the same amplitude. If this is not the case it may be a sign that there is a problem with the screening sample. The use of mixes requires a strategy to design them properly. Given the constraint of increased linewidth generated by the heterogenous TINS system, the primary factor governing the selection of compounds for a mix is the number of well resolved peaks for each. We have therefore recorded a reference 1D ^{1}H spectrum of every compound in the ZoBio/Pyxis fragment collection at 500 μM in phosphate buffered saline (PBS) in the presence of a fixed amount of TSP. The reference spectra serve the dual role of quality control as well. The reference spectra are automatically peak picked and the peak positions stored in our database. We have developed an in-house algorithm to randomly select compounds from the collection and rapidly test them for TINS compatibility, that is at least 3 well resolved peaks for each compound (when available). This allows us to directly read out the ligand from the mix without further deconvolution (see below). The algorithm also places explicit limits on the number of aromatic compounds per mix and avoids mixing compounds with pKa extrema. Once designed, the mixes are then made at 500 μM for each compound in PBS. The mixes are stored at room temperature and subsequently visually inspected for signs of precipitation. About 1/3 of mixes are rejected at this point. Mixes that do not precipitate are subjected to ^{1}H NMR analysis where we expect to see that the NMR spectrum of the mix is a simple sum of the NMR spectra of the individual compounds using TSP as a reference. Changes to the NMR spectrum of the mix, which we rarely observe, are indicative of possible aggregation behaviour of the compounds.

In order to carry out a ligand screen, the resin bearing the target and reference proteins, which have been immobilized at a solution equivalent of about 100 μM, must be packed into the dual-cell sample holder. A homemade packing reservoir has been built to fit on top of the dual-cell

sample holder and double the volume of each cell. The resin (as a 50 % slurry) is pipetted in to each cell one at a time, allowed to settle by gravity and packed at a pressure of 0.5 bar. Once packed the cell can be connected to the sample delivery system via PEEK capillary tubes and inserted into the magnet using an aluminium arm. By attaching the cell to the aluminium arm we can readily orient it such that the plane that bisects each of the two cylindrical cells is parallel to one of the transverse gradients in our triple-gradient, flow injection probe[205]. In this way

Figure 4. Determination of target integrity during a TINS ligand screen. A known ligand was applied to both the target and reference cells and the reduction in peak amplitude was measured ('TINS effect'). This experiment was carried out serially after the indicated number of mixes had been applied to the immobilized target.



optimization of the NMR experiment for each screen is minimized. All that is necessary is to perform routine tuning and matching and shim, which we do using the FID of water. When known ligands are available, initial tests are performed to insure the integrity of the immobilized sample. This same experiment is repeated 4 - 5 times throughout and after the screen to detect possible target degradation (Figure 4). Once prepared, the mixes are placed in the Gilson autosampler in 96 positioned, deep-well plates and the Bruker HyStar software is programmed for each. We also use standard ICON NMR in Topspin to acquire the TINS data. A complete screen of about 1,500 unique compounds (including some replicates for quality assurance) requires about 7 days and runs without human intervention. Having evaluated a variety of different spatially selective NMR experiments, we have settled on the Hadamard sampling approach. The quality of the data using this experiment with carefully designed mixes is rather high, as can be seen in Figure 5.

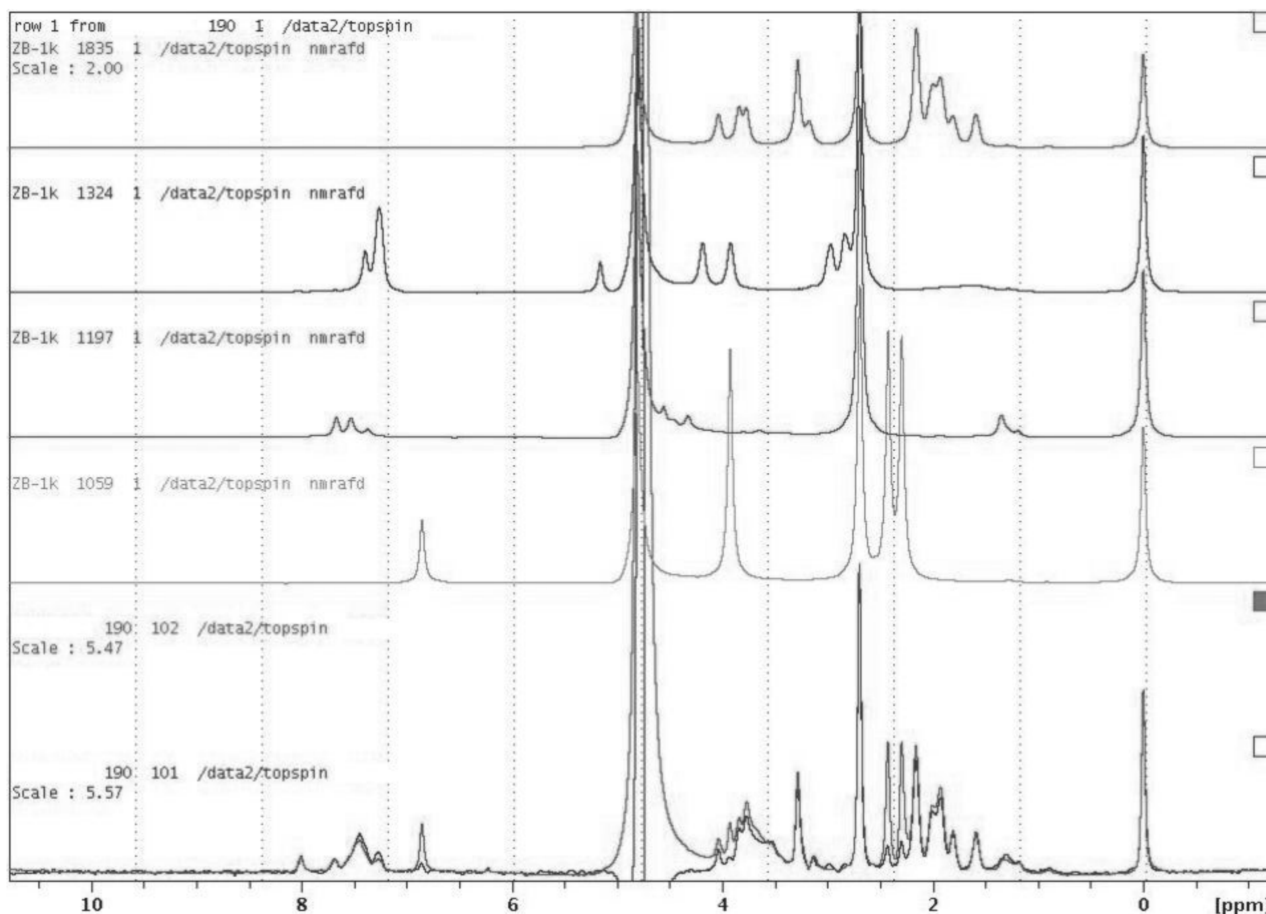


Figure 5. Direct determination of ligand identity using TINS. A mix of 5 compounds was applied to the dual sample holder containing immobilized target and the PH domain of AKT, both at 100 μM solution equivalent. The individual spectra of each cell, acquired with 30’ of measuring time, are overlaid at the bottom of the figure. The ^{1}H spectra of four of the five compounds are shown above for reference. The identity of the ligand (fourth spectrum identifier 1059) is readily obtained by simple inspection.

We have now screened a number of different targets, both soluble and membrane bound, using TINS. The hit rate for targets has varied from a low of 3 % to a high of about 10 %, where we define a hit as having at least a 30 % difference in amplitude between the reference and target spectra for all well resolved peaks. This cut-off was chosen for practical reasons based on the fact that the difference was sufficiently large to overcome artefacts related to spectral noise, minor lineshape differences between the two samples and spectral crowding and therefore enabled

reliable detection of a hit. This latter fact is particularly important since we wish to automate the data analysis process. Since screening on these targets has only been carried out using TINS, it is not possible to directly compare the observed hit rates with other methods including High Concentration Screening (i.e. screens based on inhibiting and enzymatic activity). Where Hajduk and colleagues reported essentially a 0 % hit rate for the PH domain of AKT[211], we in fact do detect some compounds binding, but our "hit rate" is about 0.2 %, some 10 fold lower than the lowest rate obtained for a target that is expected to be "drugable". In their work Hajduk and colleagues reported hit rates of up to 1 % for SAR by NMR. Interestingly, the 3 % hit rate for TINS was found when screening a soluble "NTPase" in the NDP bound form. The hit rate for the apo protein was about 9 %. The low hit rate found when the nucleotide binding pocket is occupied is expected and suggests that the high hit rates that we observe are not due to artefacts, but rather to reliable sensitivity to binding events. This idea is further supported by follow-up biochemical studies that we have now performed for two targets with enzymatic activity. Considering a soluble enzymatic target for which we found a hit rate of 9.5 %, approximately 50 % of the TINS hits showed significant inhibitory activity at 500 µM, while we would expect this number to go up even further if tested at the 1 - 2 mM typically used in high concentration screening. A similar pattern has been observed for membrane proteins (see below).

Membrane protein considerations

Quantity limitations

Although TINS removes limitations such as size and solubility of the target protein to be applied, there still remain quantity limitations with regards to membrane proteins. At present the practical lower limit for screening is roughly 25 µM solution equivalent (*e.g.* nmol/ml settled bed volume). Since we typically prepare 500 µl of immobilized resin to fill one cell of the sample holder, we require about 15 nmol of target. For a 50 kDa protein this works out to slightly under 1 mg and therefore it is safe to use 1 mg as a lower limit. For soluble proteins in which structure

guided hit optimization is the primary means for evolving fragments, this limit does not generally present a problem. However, for many membrane proteins formidable efforts are required to produce even this quantity. Accordingly, efforts are underway in our laboratory to enhance the sensitivity of TINS towards an eventual goal of being able to screen recombinantly expressed proteins in their native membrane environment, that is, without purification. Below we present data demonstrating the feasibility of immobilizing such native membrane fragments. Since this approach is beyond the present sensitivity limits of our TINS ligand screening station however, current efforts utilize highly expressed, purified, and functionally solubilized membrane proteins.

Given the current requirement for about 1 mg of functional protein to carry out ligand screening, it is clear that an appropriate system must be available to produce large quantities. Due to the interest in pharmacology and structure of membrane proteins, tremendous efforts have been made in recent years in developing new means to express, purify and solubilize them. It is not our intention to catalogue these approaches here, merely to mention some which show promise with respect to producing sufficient quantities for ligand screening and subsequent structural studies. Conceptually the simplest method for membrane protein production is *via* cell-free expression. Recently 6 different GPCRs have been produced in mg quantities using an *E. coli* based expression system that included Brij78 as a solubilizing detergent[215]. Studies were performed to show that at least one of the *in vitro* expressed GPCRs was functional. Interestingly, all appeared to be dimeric. Bacterial expression of membrane proteins typically results in the protein being unfolded and located in inclusion bodies. While purification of proteins from inclusion bodies is easy, the requirement for refolding can represent a considerable hurdle. Nonetheless, companies such as M-fold have successfully produced isotope labelled GPCR using this approach and showed that the protein was amenable to NMR studies[216].

Beyond bacterial expression systems, a number of eukaryotic expression systems have also been developed. One simple method of producing functional membrane proteins is to generate recombinant transient or stable cell lines based on CHO or HeLa cells. Such cell lines have the benefit of providing appropriate posttranslational modifications such as glycosylation which are not available in prokaryotic expression systems[217]. Often these modifications are required for protein function as shown for rhodopsin where folding is inefficient when the glycosylation site

at its N-terminus is suppressed[218]. Unfortunately the yield of proteins from stable cell lines is more often than not insufficient for ligand screening studies. Transient expression of membrane proteins can increase the yield by as much as a factor of ten but results in other inconveniences such as repeatability issues. Alternatives that have seen increasing success include recombinant expression in insect SF9 cells[219], use of Semliki Forest virus infected cells[220], and expression in the yeast *Pichia pastoris*[58]. All of these systems are capable of yielding sufficient quantities of folded, functional membrane proteins for ligand screening and structural studies. Unfortunately none is perfectly general and the rate limiting step remains finding the best system for a particular target of interest.

The membrane environment

Membranes are structured as stable phospholipids bilayers which delimit the boundaries of the organelle or the cell. The membrane provides an environment where chemical signals can be emitted and detected, where energy can be converted into inter- and intra-cellular functions, and through which materials can be transported. For all these activities, there are complex networks of interactions between the membrane-associated proteins, such as receptors, ion channels, and enzymes, and the ligands which stimulate or inactivate them. The membrane itself plays more than a passive role in these processes. Current understanding suggests that interaction between the membrane and embedded proteins is at least required for and may regulate protein function. Therefore the ultimate goal of research in our group is to be able to perform NMR based ligand screening studies on membrane proteins in their native environment. However, in light of the discussion in the preceding section it is clear that this is not yet possible and therefore membrane proteins must be recombinantly expressed and purified. Given the intimate interaction between protein and membrane, functional solubilization represents a major hurdle.

In order to retain functionality of a membrane protein, it is imperative to refold it or reconstitute it into a synthetic lipid environment which mimics the properties of its natural membrane as closely as possible[61]. Integral membrane proteins must be solubilized before being purified, and this often calls for addition of detergents after the initial centrifugation steps. For

example, the potassium channel KcsA was extracted from the cell membrane by addition of DPC prior to purification using IMAC and gel filtration chromatography[221]. Transmembrane proteins have large hydrophobic domains which can cause aggregation during purification. This can be avoided by using high concentrations of urea to prevent random folding before reconstitution in lipids[222]. These solubilization and purification steps are important because lipid reconstitution success depends on the state of the protein at this point. Organic solvents are the simplest approach to mimicking a membranous environment, but have only been possible to use with proteins with stable native folds such as ATP synthase[223] or colicin E1 immunity proteins[224]. The simplest true mimic of a membrane occurs when ionic or non ionic surfactants in organic solvents or water create micellar vesicles[61]. Micelles, which are 10 - 100 kDa in size when there is low ionic concentration, are very convenient since they are readily formed and can be used to solubilize membrane proteins in a monomeric form amenable to high resolution structural studies. To date all TINS screening of has been applied to micelle solubilized membrane proteins. However, due to, at least in part the monolayer and the extreme curvature of micelles, they are only rarely compatible with native functioning of membrane proteins. Surfactants used for such preparations include, but are certainly not limited to, sodium dodecyl sulfate (SDS), cetyltrimethylammonium chloride and bromide (CTAC and CTAB), lysophosphatidylcholine (LPC), Triton X-100, and dodecylphosphocholine (DPC)[61]. For NMR studies, deuterated surfactants are at least convenient and many times may be required. At present only DPC and SDS are commercially available in this form, although the latter tends to denature some proteins[225]. Micelles are formed when the surfactant is in a higher concentration than its critical micellar concentration (CMC), which can vary from 0.01 mM for non-ionic ones to 10 nM for short chain ionic ones, such as SDS[61]. The equilibrium shifts from micellar to monomeric forms of the surfactant when diluted with buffers that do not contain the detergent and therefore buffers must always contain a concentration of surfactant above the CMC to prevent micelle disruption and loss of protein conformation. In our hands, there is rapid exchange of surfactant molecules from the micellar to the monodispersed form resulting in rapid breakdown of micelle bound proteins when the surfactant is not included (see below). Bicelles are micelles which are composed of phospholipids rather than detergents and are slightly more complex than micelles.

Usually bicelles are composed of long chain phospholipids such as dimiristoyl-phosphatidylcholine (DMPC) forming bilayers and one shorter chain phospholipid such as dihexanoylphosphatidylcholine (DHPC) which lines the hydrophobic edges of the bilayer[226]. Bicelles, being mostly planar, represent a better membrane mimic than micelles and should be more compatible with protein function. The utility of bicelles for functionally solubilizing membrane proteins has recently been demonstrated by their use in crystallization of the GPCR, β_2-adrenergic receptor[43]. However, we have not yet tested bicelles for compatibility with TINS. In addition, there are more complex stable bilayer or multilayer vesicles of synthetic phospholipids which can be used to immobilize and orient membrane proteins on glass slides in solid-state NMR[227], but there usefulness for membrane protein immobilization on supports that are compatible with static NMR studies is not yet known.

Immobilization

The TINS methodology, by definition, requires immobilized protein to allow flow-through screening of ligands. Clearly, the choice of the surface upon which the protein will be immobilized and the choice of the immobilization chemistry have to be made within the limitations of the TINS equipment. The general requirements for immobilization compatible with high resolution NMR have been discussed so we focus on issues specifically related to membrane proteins here. We have taken a pragmatic approach when attempting to apply the TINS methodology to membrane proteins by beginning with what has worked for soluble proteins. To date we have immobilized three purified, micelle solubilized membrane proteins KcsA, OmpA and DsbB, all of which are from bacterial sources. All three membrane proteins were solubilized in dodecylphosphocholine micelles (DPC). In all three cases we have simply utilized the same immobilization scheme that has been successfully applied to soluble proteins i.e. Schiff's base chemistry to primary amines. We have found that the yield of immobilized micelle solubilized protein is nearly identical to that of soluble proteins. Further, immobilization has not had any detectable effect on the functionality of the immobilized, micelle solubilized proteins. This has been checked in two ways. For KcsA a panel of known ligands was available and we simply

assayed for binding using TINS. Since DsbB has an enzymatic activity, we adapted a spectrophotometric assay[1] for use with beads containing immobilized protein.

Enzyme inhibition studies were carried out by adding a reduced partner enzyme, and ubiquinone, whose reduction can be monitored by measuring the absorption decrease at 275nm over time. In order to reduce non-specific interactions to the resin and thus to compare enzymatic activity of the target prior to and post immobilization, there was an equivalent presence of resin in both cases. Results showed an efficient enzymatic activity post immobilization. Considering the imprecision in determining the amount of immobilized enzyme, the rate of the reaction of immobilized enzyme (3 M Ubiquinone-5/M DsbB s^{-1}) was close to that of the enzyme in presence of, but not immobilized to, the resin (4 M Ubiquinone-5/M DsbB s^{-1}) (Figure 6).

Naturally more complex strategies can be envisioned and may prove necessary for membrane proteins that are less robust than those used so far. One interesting strategy immobilizes protein first, followed by subsequent reconstitution into a synthetic lipid environment[108]. As with soluble proteins, active site blockers may be necessary in cases where illicit immobilization of lysine side chains in close proximity to the binding site may occur and thereby inhibit protein function. Various native or synthetic lipid assemblies have been extended to encompass the use of high affinity immobilization reagents such as biotin and streptavidin[66,163,165,228], antibodies[229,230,231], or nickel affinity[173,174] in order to immobilize the protein in more oriented manners. Thus as with soluble proteins, these approaches should also be compatible with TINS.

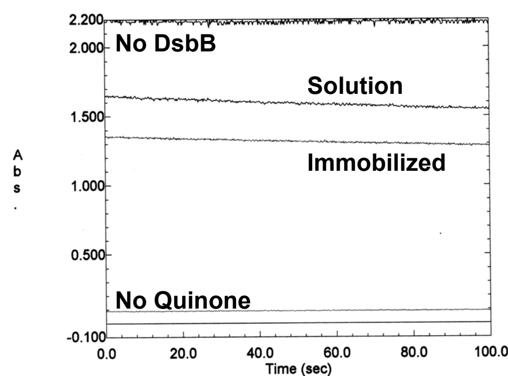


Figure 6. The target immobilized to the resin shows significantly similar enzymatic activity to the target in the presence of, but not immobilized to, the resin.

As a first step along the road to enabling TINS ligand screening for a truly broad range of membrane targets, we have begun to immobilize GPCRs in native membrane fragments (Chapter 3). In this experiment the idea was to use standard, stable animal cell expression systems such as CHO or HeLa cells as a source of material. In this way, all membrane proteins that can be

recombinantly expressed in these simple systems could potentially be used in fragment screening campaigns. Thus far we have succeeded in immobilizing membrane fragments produced by pottering (gentle disruption of animal cells) of post centrifugation membrane preparations. We have applied the procedure to both histamine receptors and adenosine receptors and in both cases, the pharmacology of immobilized receptors was similar to non-immobilized receptors. The efficiency of immobilization is quite reasonable with approximately 35 % of total receptors functionally immobilized and in comparison to non-immobilized ones; the immobilized receptors appear significantly more stable. At present the density of receptors is insufficient to perform NMR ligand screening but work is in progress to address this issue.

Screening

We have developed a diversity library for use in TINS and it is our intention to screen it against all targets. The design requirement for high solubility (to maximize oral bioavailability) pays dividends when used in membrane protein ligand screening since partitioning to the lipid phase is minimized. Nonetheless, as with soluble proteins, it remains important to use an appropriate reference system to cancel out non-specific binding events. We have used the *E. coli* protein OmpA as a successful reference protein in one partial screen of about 200 compounds and one complete screen of about 1,300 compounds. Its advantages include easy expression and purification, solubility in DPC and low small molecule binding. One potential way to avoid the use of a reference protein would be to screen using a known, competitive ligand. We are presently adapting the hardware of the TINS ligand screening station to enable competition ligand screening studies. In this arrangement the target is immobilized in both cells of the sample holder and the same mix applied to both cells while the competitor is added to only one of the cells. Competition ligand screening will eliminate the need for a separate reference protein but has the drawback that one can only find ligands to known binding pockets. When it becomes possible to screen proteins in native membrane vesicles, then a preparation of membrane vesicles of parental cell lines not expressing the target should serve as an ideal reference.

In order to further improve the robustness of TINS we include a reference compound in all mixtures that can be used to scale the two spectra post acquisition. With membrane proteins, even more than with soluble proteins, it is important to ascertain whether the reference compound interacts with the target or the surfactant used to solubilize it. The ideal reference compound has only one peak outside of the spectral range of all compounds and naturally, does not interact with the reference, target or surfactant. TSP fulfils most of these requirements but does bind to some targets. Alternatives we have used include glycine and tetramethylammonium chloride (TMA). A crude scaling factor for the two cells can be experimentally determined by integrating the water signal from each cell using a standard 1D imaging experiment with a single scan. Binding of potential reference compounds can readily be established by simply conducting TINS experiments on all, applying the scaling factor and analyzing the spectra for equal peak intensity in both cells. So far we have not encountered a case where more than one of the three potential reference compounds bound to the target.

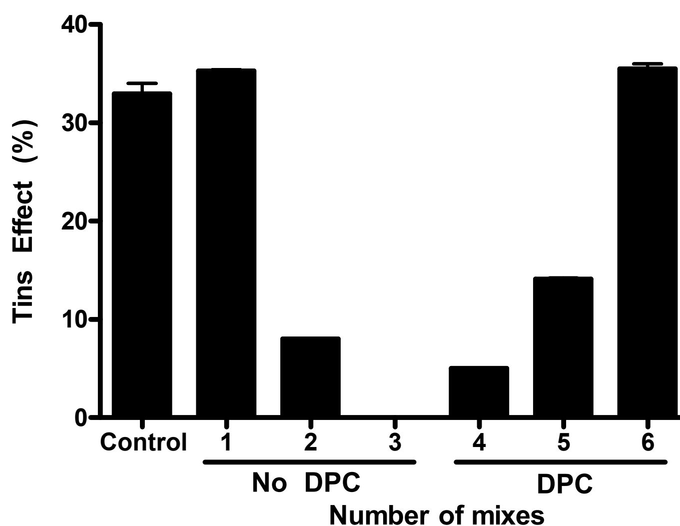


Figure 7. Requirement for the presence of detergent while screening micelle solubilized membrane proteins. In this series of experiments both the target (KcsA) and the reference (OmpA) were immobilized at a solution equivalent of 150 μM. The histogram represents the fractional difference in peak amplitude of a known ligand of KcsA in the presence of KcsA and OmpA. The bar labelled control represents the first application of the ligand. Subsequently 3 injections of the ligand were performed using buffers that contained no detergent. A further 3 injections were performed where the buffer used to wash the immobilized samples contained deuterated DPC.

As previously noted, individual detergent molecules rapidly exchange between the micellar and monomeric forms. Thus, washing of immobilized micelles in buffer without detergent leads

to rapid loss of protein functionality, as shown in Figure 7. At least for the case of KcsA, which consists of a single polypeptide, the loss of functionality (as measured by binding of a known ligand) appears to be perfectly reversible. Nonetheless, it is clear that DPC must be applied throughout the screening procedure. Since DPC is available in deuterated form its presence does not interfere with the acquisition of the NMR spectra of the compounds. For convenience we chose to include DPC only in the buffer used to wash the compounds out of the cells of the sample holder and not in the mixes themselves. Since this approach has lead to two successful screens of membrane proteins we are optimistic it will be general. In this way it may prove possible to acquire NMR spectra even in the presence of non-deuterated detergents since the concentration of the monomer is reduced by application of the compound mix in the absence of detergent. However, we have yet to test this hypothesis. Once the immobilized protein functionality has been verified, it is also important to create checkpoints at different time points of the screen with mixes containing a known binder as a positive control to check that protein functionality and thus conformation is maintained through the screen.

One final issue deserves special attention when considering carrying out ligand screening studies on a membrane protein, the kinetics of ligand binding. While low affinity ligands for soluble proteins nearly always exhibit rapid exchange kinetics on the NMR time-scale, this may not be the case for membrane proteins. For example, histamine binds the human H1 receptor with a K_d of 20 μM[232]. Such a small molecule (histamine fits well within the definition of a "fragment") binding with moderate affinity would normally imply a fast on rate. However, in this solid state NMR study, the on rate was found to be in the order of minutes! Likely mechanisms for such slow binding include access to the active site of the protein *via* the membrane or slow conformational exchange of the protein due to interaction with membrane (or membrane mimetic). Since the dynamic behaviour of detergents and phospholipids are strongly temperature dependent, it may be necessary to carryout screening at near physiological temperature where the long term stability of the target may be less than optimal. In such situations it may be necessary to prepare multiple samples in order to successfully carryout a screen of a complete fragment library.

Application of TINS to Ligand Discovery

Soluble Targets

To date TINS has been applied to five different soluble targets. We have immobilized the target at a range of concentrations for the various screens, from as high as 500 μM to as low as 100 μM solution equivalent. We now typically screen at 100 - 150 μM which represents an optimal balance between sensitivity, artefact suppression and protein consumption. In all cases we have used the PH domain of AKT as the reference. Typically we immobilize the target and reference on the activated sepharose, Actigel ALD (Sterogene, USA). The efficiency of immobilization is monitored by UV absorption of the supernatant and visual inspection to insure that no precipitation has occurred. If an enzymatic assay of the target is available, we use it at this stage to confirm that the immobilized protein remains functional. The derivatized supports are subsequently packed into the dual-cell sample holder under pressure (0.5 Bar/cell), connected to the solvent delivery lines from the sample handling system, and then placed into the magnet. In most cases a small number of known weak ligands (up to 6) are available to test whether the target has been functionally immobilized and to demonstrate that we can indeed detect ligand binding. One of the known ligands is then selected for use in monitoring the condition of the target during screening. We routinely monitor the condition of the immobilized target through repeated injection of the known ligand throughout the screen.

Once the immobilized target has been deemed functional we carry out the actual screen. The mixes are delivered in 1 ml volumes in deep 96 well plates to the Gilson autosampler. Sample handling is controlled by Bruker HyStar software which communicates with Bruker TopSpin to acquire the NMR data. Using the Hadamard sampling experiment described earlier we currently acquire data for 30 minutes with an additional 5 minutes for sample handling resulting in a cycle time of about 35 minutes. In a recent screen 324 experiments were run in total to assess binding of 1,393 compounds from our fragment collection. This number includes repeated assaying of the

positive control to assess target condition and some overlap of compounds (*e.g.* compounds appear in two different mixtures). This design allows us to assess the repeatability of the screening data. Such a screen was carried out without human intervention in under eight days. Finally, since the target and reference are immobilized, it is possible to change buffer conditions to closely match crystallography conditions without regard to protein stability. We routinely screen under solution conditions in which the reference protein would precipitate if not immobilized. Nonetheless it ligand binding characteristics vary only very moderately from one set of solution conditions to the next.

TINS proof of principle application to a bacterial membrane protein

TINS is a comparative method, where detection of ligand binding to the immobilized target is quantitated by comparison to an immobilized reference. With membrane proteins, partitioning of ligands can occur on the native or synthetic lipids surrounding the target present on the resin. An appropriate reference system had to be developed to control for non-specific binding of hydrophobic compounds to lipids or detergents used to solubilize the membrane proteins. An appropriate choice for such a reference protein would be one with few known binders, in order to minimize the chances of non-specific binding. The *E. coli* Outer membrane protein A (OmpA) was chosen for such qualities. This reference protein was of similar size as our intended target and also refolded in DPC micelles. To get an initial feel for whether we could detect specific binding to a membrane protein using TINS, we conducted a proof of principle study with a screen of a small subset (about 100 compounds) of our compound library using KcsA from Streptomyces as the target and OmpA as the reference.

Prior to screening it was necessary to establish an appropriate 1) level of DPC to include in the wash buffer to maintain the integrity of the immobilized, micelle solubilized target and 2) internal reference compound. If the DPC concentration in the environment of the target decreased to below its CMC, the micelles formed by DPC would start to slowly dissociate into monomers and be flushed away. Simple calculation suggested that it was necessary to use DPC at 5 mM in the wash buffer to in order to maintain the concentration above the CMC (1 mM) upon dilution with

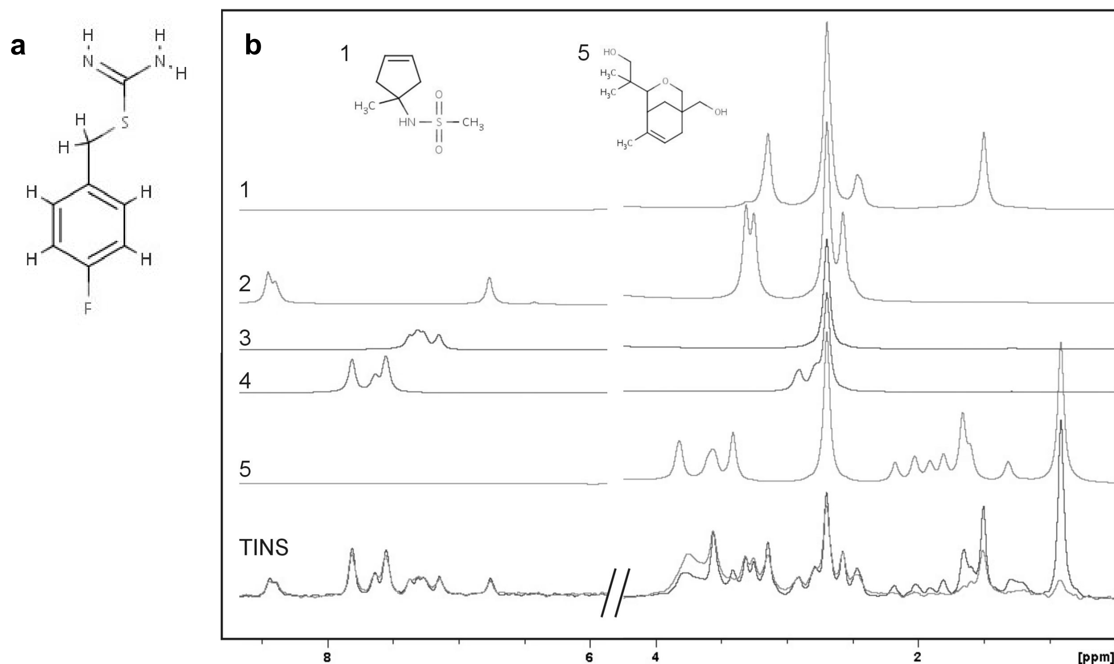


Figure 8. Proof of principle ligand screen against a bacterial membrane protein. **a**. Structure of the known ligand (4-fluorophenyl)methylsulfanylmethanimidamide used determine the integrity of the immobilized KcsA. **b**. Detection of ligand binding in one mix during the screen. A mix containing 5 different compounds was applied simultaneously to the cell containing immobilized KcsA and to the cell containing OmpA. The individual ^{1}H NMR spectra of each cell are overlaid (labeled TINS). The ^{1}H NMR spectrum of each individual compound, which has been intentionally linebroadened to approximately match the linewidth of the TINS spectra, is shown above (numbered). All peaks from compounds **1** & **5** were reduced in amplitude in the presence of the immobilized KcsA with respect to OmpA, indicating that these compounds bind to KcsA. The structures of compounds **1** & **5** are shown.

the compound mix absent DPC. We tested both TSP and TMA as possible internal standard by including both in a mixture with (4-fluorophenyl)methylsulfanylmethanimidamide (FPMSMA, Figure 8**a**), a known ligand for KcsA. These tests indicated that both TSP and the known ligand FPMSMA specifically bind KcsA and we therefore chose to use TMA as an internal standard.

Repeated application of TMA and FPMSMA, followed by washing with buffer plus 5 mM DPC demonstrated stability of the immobilized KcsA and so these conditions were used for the limited library screen. During the screen the immobilized target showed insignificant loss of binding capacity for the control compound and only 12 % loss after 3 months of storage. Out of

the 95 fragments that were screened, 7 % showed substantial changes in the NMR spectrum that were specific to the target and were considered binders after analysis of spectra intensities (Figure 8**b**). This is in line with target hit rates obtained for soluble proteins applied to TINS. Of the potential new hits, 2 structures had a similar scaffold to the known binder. The other hits had a variety of scaffolds with a variety of shapes and numbers of rings.

Development of a high affinity inhibitor of bacterial membrane protein DsbB using TINS

Very recently we have undertaken a program to develop high affinity inhibitors to the bacterial inner membrane protein DsbB in collaboration with Prof. John Bushweller's group at the University of Virginia (USA). DsbB is a redox enzyme involved in the production of toxin in gram negative bacteria[94] and as such is a potentially medically interesting target. The crystal structure of DsbB bound to its redox partner, DsbA has been solved[92] and the Bushweller group has solved the solution structure of a disulfide mutant of DPC solubilized DsbB[62]. For ligand screening we immobilized both the functional wild type DsbB (see above) and OmpA (as a reference) at a solution equivalent of 100 μM. We used the compound Ubiquinone-5 (Figure 9) which binds competitively with the native DsbB ligand. Similarly to KcsA, deuterated DPC was included only in the wash buffer.

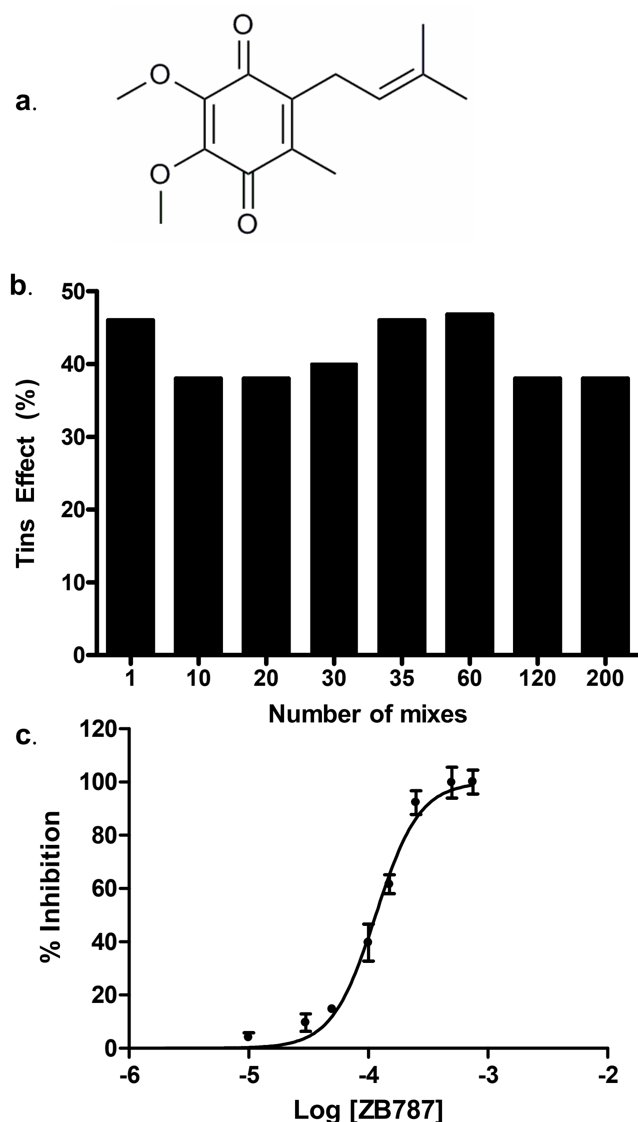


Figure 9. Ligand screening of a bacterial membrane protein. **a**. The structure of Ubiquinone-5 used to asses the integrity of immobilized DsbB during the screen. **b**. Ubiquinon-5 binding to immobilized DsbB during the screen. Binding is defined as in Figure 2. **c**. Enzyme inhibition curve of a hit from the screen.

Using this arrangement 1,270 fragments were screened in mixtures that averaged a little over five compounds each. Figure 9 demonstrates that the immobilized DsbB remains intact throughout the screen. In the screen we found 93 compounds that specifically bind DsbB for a hit rate of 7.3 %. The hits have been investigated for enzyme inhibition at 250 μM and the best 9 of these compounds had IC50s of 150 μM or better where a representative curve is shown in Figure 3. We have carried out both competition binding and competition enzyme inhibition analyses on a limited subset of the hits. Most of the hits are competitive with ubiquinone binding and this seems to represent the major small molecule binding pocket. However, one of the subsets of hits is not competitive with ubiquinone. Follow up biochemical and biophysical analyses are presented in Chapter 5 of this thesis.

Outlook

In the past decade an impressive repertoire of methods has been developed to enable drug development against soluble targets at the molecular level. In addition to fragment screening methods, structural biology has played a key role in this process. Although at present no drugs are marketed that are the exclusive result of the fragment approach the principles can clearly be seen in the remarkable specificity and potency of recently marketed kinase inhibitors such as Imatinib and Gefitinib and indeed, many fragment based drugs are in the late stages of clinical trials[21]. Membrane proteins represent a similar pharmacological challenge in that one would like to be able to specifically address individual targets from amongst large numbers of closely related members of a protein family. However, it is presently not possible to use the molecular methods developed for soluble proteins for drug discovery efforts on membrane proteins.

A major goal of the research in our laboratory is to adapt methods developed for soluble targets to membrane proteins or to develop alternative ones. While we are clearly only at the beginning stages of this process we have nonetheless made a promising start. We have been able to immobilize a variety of membrane proteins in functional form and have carried out ligand screening on two. Our current efforts are geared toward finding new ways to solubilize and immobilize membrane proteins that can be more widely applied. We are also looking towards a variety of methods to improve the sensitivity of TINS including experiments that are better optimized for the diffusion limited nature of the heterogeneous system we employ and possible implementation of a TINS cryoprobe.

Once one finds and validates hits, it is of course necessary to evolve these towards high affinity, high specificity ligands. The hit evolution process is greatly aided by the availability of three dimensional structural information of target-ligand complexes for soluble targets. Since crystallography of membrane proteins is not yet widely applicable it will be imperative to develop alternate approaches. We envision a number of such approaches that utilize the power of liquid or solid state NMR. In recent years both solid state NMR[233] and solution state NMR[234] have made significant progress in elucidating 3D structures of either the membrane protein itself or ligands bound to membrane proteins. While it is vital that these efforts continue, it is also

logical that NMR should be employed to take advantage of its unique ability to rapidly generate local, low-resolution structural information. For this we foresee new applications in chemical shift perturbation based modelling of protein-ligand complexes[235], sparse NOE based methods[236,237] and paramagnetic NMR[238]. With the foreseeable advancements in ligand screening and structural analysis, the era of molecular drug discovery on membrane protein targets should soon be upon us.