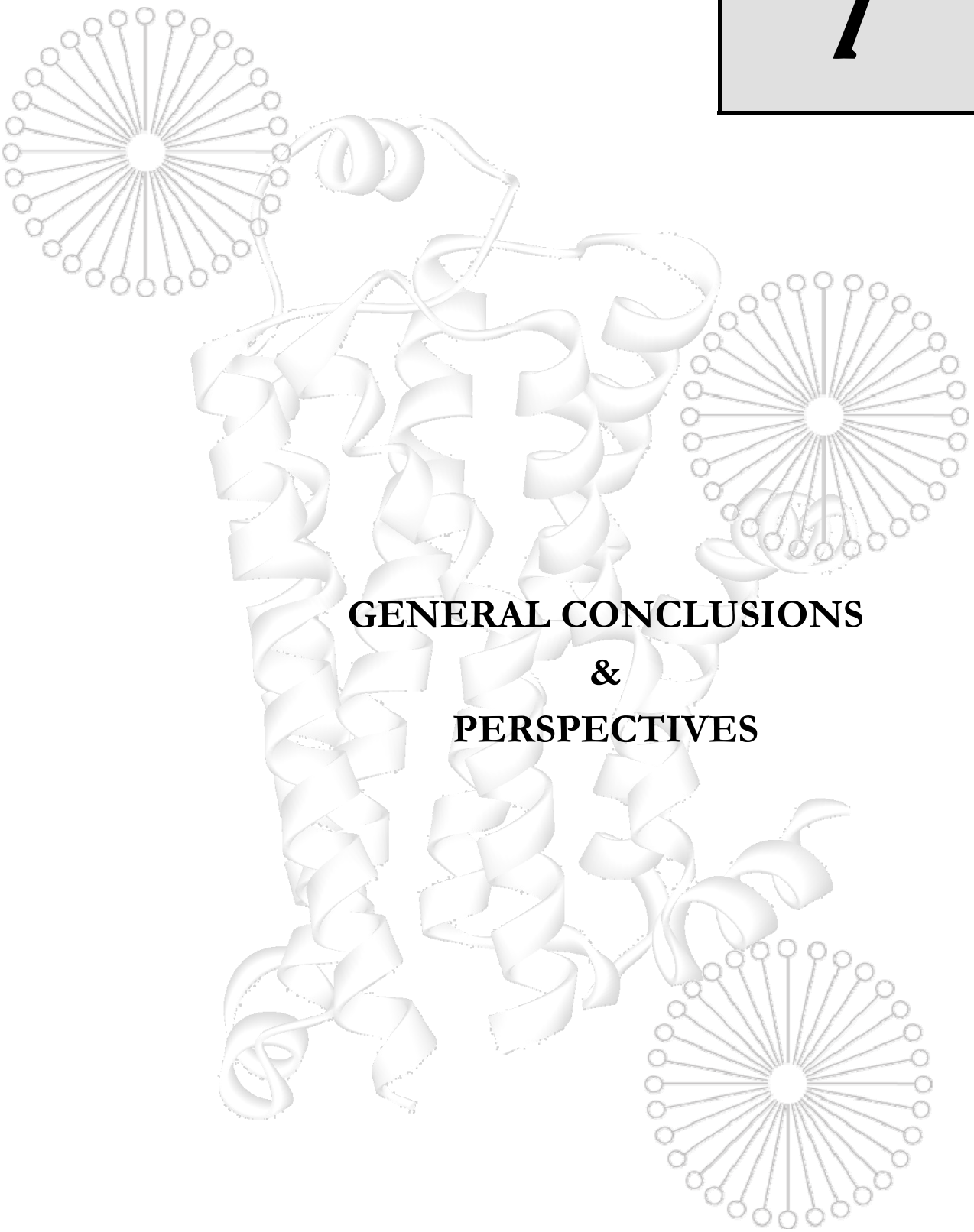
Application of fragment-based drug discovery to membrane proteins

Früh, V.

Chapter

7



GENERAL CONCLUSIONS & PERSPECTIVES

General Conclusions

As described in the review of membrane immobilization strategies in **Chapter 2**, there exists a variety of intricate protocols to express, solubilize, and immobilize membrane proteins for a range of applications ranging from 'macroscopic' functional cell-based assays to biophysical studies at the molecular level. Most of these methods, however, have not enabled fragment based drug discovery in a high throughput manner on membrane proteins due to the challenging inherent properties of this class of proteins and the lack of general methods for their solubilization and immobilization. With the simple and mild Schiff's base chemistry used to immobilize membrane proteins, combined with Target Immobilized NMR Screening, the research presented in this thesis has shown that fragment based drug discovery now has the potential to be applied to membrane proteins in general. As described below, the Schiff's base chemistry made it possible to immobilize a wide class of membrane proteins including the histamine H_1 and adenosine A_1 GPCRs, the potassium ion channel KcsA, the membrane enzyme DsbB, and the reference membrane protein OmpA. Although the GPCRs were immobilized in a functional manner, the intrinsic expression levels were too low to allow us to apply these proteins to the TINS methodology. On the other hand, the high levels of bacterial expression of KcsA, DsbB, and OmpA enabled us to apply these proteins to TINS. With available protocols for DsbB biochemical and biophysical characterization, the fragment hits identified for this enzyme were validated as specific binders, proving that fragment based drug discovery is now applicable to membrane proteins with exciting new perspectives, as described further.

Simple and functional immobilization chemistry generally applicable to membrane proteins

The GPCRs used in **Chapter 3** were an example of immobilizing membrane proteins within their native environment. The immobilization efficiency was low, with only 20 % of the initial receptors functionally immobilized on the resin, with a maximum population of 2 pmoles of

receptor per ml of resin. This is approximately 3 log units less than what is required for the current parameters of a TINS screen but nonetheless, the pharmacology profile of these immobilized receptors was very similar to the receptors which had not undergone the immobilization process. Immobilized receptors not only presented native pharmacology with regards to antagonist binding, but they also presented agonist two site binding profiles which confirmed the presence of G proteins as indicated by western blots of resin samples. Immobilizing GPCRs also had a positive effect on the relative stability of immobilized GPCRs compared to those in native membrane vesicles. This physiologically relevant pharmacology of the receptors on the resin suggests that there is potential to apply this class of proteins to molecular based screening methods such as TINS. The limiting step in obtaining high amounts of immobilized GPCRs for TINS however, mostly has to do with the current low expression levels rather than the immobilization procedure itself. For example, the increase in functionality which we report to occur upon immobilization with longer linkers between the protein and the surface has been reported for other GPCRs and is explained by the larger space available for the protein's extracellular domain movement[112]. Other suggestions for obtaining higher functionally immobilized GPCRs are described in the perspectives section.

The Schiff's base chemistry also enabled the functional immobilization of detergent solubilized membrane proteins from bacterial sources, KcsA and DsbB (**Chapters 4 & 5)** and DsbB in the alternative solubilization agent, the nanodisc (**Chapter 6**). These bacterial proteins have been extensively studied by molecular methods in the past, such as crystallography[92,253] and NMR[62,91], and a variety of protocols exists for their purification and functional solubilization in detergent micelles. These conditions enabled us to produce sufficient amounts in *E.coli* to optimize the immobilization conditions. The functional immobilization could be monitored either by injecting a known binder at different time points of the screen, or by an existing functionality assay. The only protein for which we had a functionality assay available, however, was DsbB[93]. We had no means to detect whether OmpA was properly folded upon immobilization, but our sole requirement for this protein was minimal small ligand binding properties, a quality which **Chapters 4 – 5** demonstrated to be true. These three detergent solubilized proteins were immobilized with an efficiency of 50 %, with a yield of 100 μM

(volume equivalent) of immobilized protein, as required for TINS. Upon the solubilization of DsbB and OmpA in the nanodisc complexes (DsbB/ND, OmpA/ND), the immobilization yield increased to 75 %. This suggests that in detergent micelles, the N-terminus of both DsbB and OmpA may have been slightly buried into the DPC micelles and therefore less accessible to the immobilization reaction with the aldehyde groups on the resin. In contrast, the N-terminus of the nanodisc membrane scaffold proteins may have participated in the immobilization reaction, thereby increasing the final yield of functionally immobilized DsbB/ND and OmpA/ND. In contrast to the GPCRs in this study, the higher immobilization yield of DsbB (in detergents or nanodiscs) may be because the enzymes were immobilized in a more stable conformation due to the presence of the endogenous ligand, UQ8. Both DPC-solubilized DsbB (DsbB/DPC) and DsbB/ND were 90 % active on the resin compared to in solution and remained stable for a month. DsbB/ND however showed approximately 16 % increase in turnover rate compared to that in detergent micelles. This may be explained by the increased stability and functionality of membrane proteins in a better mimic of the native membrane, such as in the lipid bilayer of the nanodisc formation, as opposed to the less stable detergent micelle formation[248]. The lower level of activity of detergent-solubilized DsbB may also be explained by the partial solubilization and consequent lowering of the effective concentration of the cofactor into the micelles during the enzymatic activity assay.

Target Immobilized NMR Screening in DPC detergent

The functional immobilization of membrane proteins in detergent micelles enabled us to carry out a TINS screen in the presence of detergent to identify fragments binding to KcsA (**Chapter 4**) and DsbB (**Chapter 4 and 5**) with OmpA as a reference. The immobilization of DsbB and OmpA enabled an entire screen on the same protein samples for one week at room temperature, which resulted in a 7 % hit rate with a final set of 93 fragments which were listed for validation as DsbB inhibitors. By comparing the intensities of the fragments injected from the same mix in the presence of the immobilized target or reference, one could immediately identify a target binder within a cocktail of fragments without any deconvolution. This was facilitated by the

reference protein which accounted for non-specific levels of binding to proteins and to detergents, thereby providing information on specific binders to the target. This proved the principle that TINS can be applied to membrane proteins, provided the appropriate distribution and concentration of detergent is used in the system. We tested the amount of detergent required in the wash buffer to enable continuous and stable protein functionality during a screen, and established that a concentration 5 x above CMC was sufficient. With KcsA, which is a tetramer protein, we could see an immediate drop in capacity to bind a control known binder upon an initial wash with an aqueous buffer void of detergents. Interestingly, this drop in intensity was reversible with a fresh injection of buffer containing DPC, suggesting that the immobilization does indeed somewhat stabilize the protein. We established that keeping the detergent in the wash buffer rather than in the mixtures containing the fragments would limit the unwanted interactions between them while still enabling a full screen without loss of protein functionality. DPC was used as a deuterated detergent, thereby rendering its signal invisible in the NMR spectra of the fragments.

At this point, there was no method for us to establish whether the detergents were partially solubilizing the fragments, and, in doing so, were causing loss of fragment ^{1}H 1D signal intensity. Fortunately, the signal intensities remaining were high enough that fragments binding to the targets could be detected thanks to the reference system composed of immobilized OmpA in the same detergent micelles.

DsbB hit validation by biochemical and biophysical analyses

Of the 93 DsbB/DPC hits identified from TINS, a number of observations suggested that the fragments were indeed specifically binding to DsbB. First, a well-distributed population of inhibitors was found by adding 250 μM of fragments to the DsbB assay and calculating the subsequent percentage of enzyme inhibition induced by the fragments. This resulted in 16 % showing less than 20 % inhibition or mild stimulation, 60 % of the fragments showing better than 30 % inhibition of DsbB, and 17 % showing more than 70 % inhibition. These were only single concentration point experiments which did not provide information on the mode of action, nor

whether this inhibition was due to protein precipitation or other assay artefacts. Thus, secondly, the 13 fragments showing highest inhibition were chosen to be validated further by competition assays to quantify the potency (IC50) of the fragments as well as their mode of action (competition for the cofactor UQ1 or the DsbA binding site of DsbB). These assays revealed artefactual inhibition by 3 fragments, but stoichiometric inhibition by the remaining 10 fragments. The IC50 values ranged between 10 and 200 μM, and further competition analyses revealed two distinct binding modes. The most 8 potent fragments were thus grouped into two categories. 3 fragments caused the expected competition with the ubiquinone cofactor in the endogenous binding pocket of DsbB and the remaining 5 presented a mixed model inhibition, as defined by lowering the affinities of both DsbA and UQ1 for DsbB.

Finally, because the biochemical assays were not providing structural insight into the mechanisms of DsbB inhibition by the fragments, an orthogonal biophysical method was used and confirmed the mechanism of the fragment modes of action in a structurally relevant manner. Unfortunately, as is the limiting factor for many membrane proteins, DsbB in DPC micelles was dynamically unstable and although part of the crystal structure was obtained in complex with DsbA[92], there was no full structure nor sequential assignments available from NMR studies. We therefore used ^{15}N labelled DsbB[CSSC] for which the sequential assignments were available[62]. In this mutant, cysteines 44 and 104 have been mutated to serine, which had the consequence of creating a stable disulfide bridge between Cys41 and Cys130[62], resulting in a stabilized and inactive form of DsbB. Clearly, using the native protein would have been more ideal, but the mutant was a physiologically relevant intermediate of the disulfide oxidation pathway[62] which, nonetheless, turned out to be useful in providing us with information regarding the structural interactions between the fragments and the enzyme. HSQC experiments revealed detectable changes in the NMR spectra of DsbB[CSSC], which contained the endogenously bound UQ8 cofactor, upon titration of UQ1 (positive control) and all 8 fragments. This shows that the fragments identified by TINS and validated by biochemical assays also bound the double cysteine mutant, with a somewhat lower affinity as expected from this conformational variant of DsbB. The titrations revealed interesting information which confirmed the biochemical assay validation results. Titrations of the positive control UQ1 yielded patterns which suggested slow

exchange of UQ1 into the DsbB ubiquinone binding pocket (already mapped by structural studies by crystallography on wildtype DsbB in complex with DsbA[92]) but also some elements of fast exchange. The 3 fragments showed similar slow and fast exchange patterns on the same residues as affected upon UQ1 titration: these residues were located close to the first periplasmic loop of DsbB[CSSC] which contains the active cysteine pair in the wildtype DsbB, involved in reducing the ubiquinone in its native binding pocket. Although the second group of fragments showed some similar patterns of chemical shift perturbations as the UQ1 and competitive fragment titrations close to the ubiquinone binding site, they also presented a distinctive pattern in the area of the second periplasmic loop. The second periplasmic loop contains residues involved in DsbA binding (Phe100 - Phe106) and oxidation (Cys104) but also the hydrophobic residues (Leu116 and Val120) involved in associating DsbB to the membrane upon charge transfer from DsbA to DsbB[92]. The effects of mixed model fragment titrations also differed from the competitive fragment titrations in that they only presented patterns of fast exchange. These differences can explain the mechanism of DsbB inhibition by these fragments. Upon competitive inhibition, such as with competitive fragments or UQ1, the dissociation of the quinone within the binding pocket may have happened in two phases. The initial fast exchange displaced the quinone moiety which was bound towards the hydrophilic exposed part of the enzyme, but the dissociation of the isoprenyl tail, which extended deep into the hydrophobic groove, making extensive contacts with the hydrophobic residues between TM1 and TM4, happened at a slower rate. Upon mixed model inhibition however, the fragments did not displace the ubiquinone but rather caused a conformational change which affected electron transfer to the ubiquinone and binding of DsbA to DsbB, as predicted by biochemical assays. In conclusion, the fragments identified by TINS were validated as specific binders of DsbB and turned out to be either specific inhibitors of native ligand binding, or inhibitors of protein-protein interactions.

Target Immobilized NMR Screening in alternative solubilization medium: the nanodisc

An attractive alternative to detergents is the use of amphiphilic polymers or complexes which on the one hand satisfy hydrophobic needs of membrane proteins while on the other maintain surface polarity contacts with the surrounding aqueous buffer. In **Chapter 6** we have shown that the nanodisc system has been a useful alternative to detergent in a TINS screen. Although there were quite some conditions to test before obtaining a good population of well formed complexes, the procedure was simple and easy to characterize by gel filtration and SDS-PAGE gel band volume analysis. Not only was the immobilization yield more efficient in nanodiscs than in detergent due to the participation of the scaffold protein in the immobilization reaction, but the activity of DsbB was higher as well. As previously mentioned, this suggests that either the protein was more active in lipid bilayers, as can be expected, or suggests that the detergent may indeed have partially solubilized the cofactor UQ1 in the assays, thereby resulting in lower final effective turnover rates. This last possibility has been further supported by the better quality of NMR spectra in the presence of nanodiscs as opposed to DPC. The signal to noise ratio was double to that in detergents, which, by NMR standards, suggests that there may have been a loss of 30 – 40 % of the effective amount of fragments into detergent micelles. We had obtained particles of appropriate size upon gel filtration of nanodisc-embedded OmpA (OmpA/ND), but using these complexes as references for nanodisc-embedded DsbB (DsbB/ND) enabled us to identify only 5 known binders as opposed to 17 with the empty nanodiscs as a reference, and 18 with the DPC screen of DsbB with the reference protein OmpA. We used the same cut-off for all three screens, suggesting that the OmpA reference was now showing higher non-specific binding properties when embedded in a nanodisc complex. Due to the highly stable nature of OmpA previously solubilized in a variety of detergents, however, it is unlikely that it was unfolded upon nanodisc entrapment, but it may more likely be a matter of mismatch in heights of the protein transmembrane domain of 30 Å and the corresponding lipid bilayer of 50 Å in the nanodisc[254]. This mismatch may have exposed areas with high fragment binding properties, either from OmpA hydrophobic residues or from the lipids in the nanodisc. When using empty nanodiscs as

a reference, DsbB/ND proved to be stable throughout the screen without any addition of detergent. Naturally, in the event that a target protein can not be solubilized in nanodiscs, however, screening the protein in detergent micelles still remains an option due to the sufficient level of fragment ^{1}H 1D peak intensities remaining even after the fragments have been partially solubilized in the detergent micelles.

Perspectives

The possibility of using molecular methods such as NMR for fragment based drug discovery on membrane proteins opens an exciting new venture for drug discovery. Clearly, we are still limited from screening membrane proteins within their native membrane, and we could only obtain samples in the quantities and stabilities required for TINS by screening bacterial membrane proteins in detergent micelles. Work is underway to improve the conditions which would facilitate screening of membrane proteins in their native membrane. This includes enrichment of native membranes *in situ*, the use of more powerful NMR probes such as cryoprobes, improved NMR parameters, and downscaling of sample size by immobilizing proteins on the glass surfaces of microfluidic chambers compatible with the system. In the meantime however, in combination with new solubilization and stabilization strategies[70,71,255], there are a variety of perspectives for TINS on membrane proteins *in vitro*.

Membrane protein immobilization

In future studies, it would be interesting to see how we can maximize the population of functionally immobilized GPCRs and tailor the chemistry so as to immobilize these and other membrane proteins in a more oriented manner rather than the random Schiff's base chemistry. There exists a variety of immobilization chemistries which can be used with the TINS compatible aldehyde resin used in this thesis. In our case, although there were no structures available for the H_1 and A_1 receptors, mutagenesis studies on the adenosine receptor class has revealed the close homology between the recently solved A_{2A} receptor structure[45] and the A_1 receptor[256]. This type of information can be used to pinpoint structural elements of the receptors that can be involved in immobilization without affecting functionality, such as biotinylation of cysteines[110]. For instance, the C-terminus has been found to regulate G protein coupling to GPCRs[257], and should therefore be avoided in future immobilization reactions. GPCRs can also be embedded in biotinylated forms of nanodiscs, which can then be immobilized to avidin-covered surfaces with

the benefit of leaving the target GPCR unaffected by the immobilization reaction. In addition, GPCRs which are in the ligand bound state tend to maintain a higher conformational stability, as can be seen by successful crystallization of the β_2-adrenergic receptor in the presence of the partial inverse agonist carazolol[34]. This points to the potential of obtaining higher populations of functional GPCRs when they are immobilized in the presence of their ligands. The immobilization of membrane proteins will not only enable the elaboration of new drugs, but will also be interesting to use in studies looking at the physiologically relevant mechanisms. With regards to GPCRs for example, this includes allosteric modulation, dimerization, and coupling to G proteins and other GPCR-interacting proteins (GIPs) such as the receptor-activity-modifying-proteins (RAMPs)[257,258]. It may for example be useful to immobilize a GPCR in the presence of the ligand to establish whether it influences G protein coupling, as is the case for the C5a receptor, where C5a binding activates G protein coupling[259]. G protein coupling can therefore be potentially targeted in drug discovery by screening different mutants of a GPCR target with TINS to identify fragment scaffolds which interact with specific residues relevant for G protein coupling. The current immobilization procedure also has the potential to be used with other forms of solubilized membrane proteins, such as those in cubic lipid phases[260], and is simple and generally adaptable to use with other chemical immobilization strategies, and with other proteins, surfaces, and assays.

Target Immobilized NMR Screening

Perspectives for TINS on membrane proteins are numerous because the technique can be adapted to various solubilization and immobilization strategies for primary screening of fragment binding. Screening a focused library based on known important scaffolds for the targeted membrane proteins can minimize the time of a screen down to a few days rather than an entire week, limiting the problems faced by low protein stability or quantities. There may soon be improved NMR parameters with increased sensitivity which can enable one to downscale the amount of protein required per screen, thus enabling screens on proteins in their native membrane vesicles. In the meantime however, detergent or nanodisc solubilized membrane proteins may be

the only preparation available for current day fragment based drug discovery with TINS. This is possible with the new stabilization mutations which have found to enable one to produce, for example, high quantities of GPCRs in *E.coli*, stabilized in either the agonist or the antagonist state, as has been previously reported for the A_{2A} receptor[261] and the β_1-adrenergic receptor[70]. The advantage of this technique is that *E.coli* can be grown in fermentors to provide large quantities as opposed to the adherent mammalian cell cultures, and carrying out a TINS screen on a specific state of the receptor minimises the need to carry out future assays to determine whether ligands are agonists or antagonists. Furthermore, synergetic effects or allosteric modulation can be studied by designing appropriate libraries.

The application of TINS to membrane proteins in nanodisc formations also enabled us to carry out a screen in complete absence of detergents. The lower signal to noise ratio obtained in screens containing detergents was still high enough to identify potent fragment hits. This leads to the advantageous possibility that screening nanodisc-embedded proteins in the absence of detergent can be carried out with 30 – 40 % lower protein and fragment concentrations and still obtain detectable fragment ^{1}H 1D peak intensities. GPCRs and GPCR-like proteins have been previously successfully embedded in nanodiscs, both as monomers (b2-adrenergic receptor)[74] and as trimers (bacteriorhodopsin)[262], suggesting that there are exciting possibilities of screening nanodisc embedded Class A and Class C GPCRs, the latter of which are obligate dimers[263], with TINS. The use of empty nanodiscs as an appropriate and generalized reference, has now also been established for SPR technology[264], and we believe future TINS applications should, when possible, be applied to nanodisc embedded targets. Furthermore the use of the empty nanodisc as a reference would enable one to minimize the energy spent in finding an appropriate reference protein and developing a protocol for the appropriate self assembly into nanodiscs.

Research is currently replacing old ideas about physiological processes, as can be seen by the increase in reports on GPCR multimerization[50] and the multiprotein networks into what is now being coined 'receptosomes'[257]. Therefore, there are a variety of other topics aside from fragment based drug discovery to explore with TINS, in the event of obtaining better information for modulating a biological system in the body. Without much information available on orphan receptors, TINS can be used as a tool to compare hits between different classes of protein,

leading to deorphanizing studies. As previous studies have shown, classification of GPCRs can be predicted by which class of G proteins they bind[265,266], and therefore, deorphanizing studies in principle could also be carried out by immobilizing an orphan GPCR and injecting different G proteins instead of fragments. Furthermore, the pharmaceutical industry is finding new and exciting ways of targeting drugs, such as pharmacogenomic profiling[267,268] and the use of multitarget drugs[269,270]. Thus, why not screen several targets within a specific pathway, or screen several physiologically relevant target mutants and identify a common hit between them, as has been suggested previously for drugs targeting multiple mutated version of kinases as an efficient treatment for cancer[271].

Fragment hit elaboration

As we mentioned, structurally relevant information can be used to facilitate the elaboration of weak fragment hits into stronger, more potent and more specific ligands. For the hits which were found for KcsA and DsbB, clearly, obtaining structural key information of fragment binding to the wildtype protein would be an ideal next step. The intrinsic movement of the periplasmic loops upon DsbB enzymatic activity or upon KcsA dynamics may however hinder these kinds of experiments no matter how stabilizing the solubilization medium is. As with the stabilized GPCR mutants in literature[70,261], and the DsbB[CSSC] mutant we used in Chapter 5, combining mutagenesis and molecular methods has the potential to provide worthwhile information on ligand-protein interactions. However, some interactions are inherently weak, such as protein-protein interactions involved in kinase domains[272,273] and between G proteins and GPCRs[258]. Perhaps weak interactions in biological processes are currently undermined and targeting such interactions may enable to provide the medicinal chemistry realm with more specific inhibitors with lower amounts of side effects. The advantage of using fragment based drug discovery in such a context is that fragment hits may suffice as weak inhibitors of protein-protein interactions, for example. There are a variety of NMR based applications which can be applied to obtain valuable information on molecular interactions between ligands and proteins, as described in the outlook section of Chapter 4, such as modelling chemical shift perturbations[235], sparse NOE

based methods[236,237] and paramagnetic NMR[238], but it is important to integrate other applications of medicinal chemistry, biology, and computational methodologies in order to be fully effective[237]. The emerging era of new internet based software should be used to link data[274] across different databases, such as ones concerning diseases, targets, drugs, and clinical trials, in the global aim to make medicinal chemistry research more efficient by enabling the discovery of new connections between diverse data sets[275]. Currently, work is underway to better understand how we can modulate the NMR parameters in such a way that TINS can be used for fragment binding quantification and hence, the ranking of fragment hits, and also as a competition binding assay which can provide information on fragment kinetics in relation to a given target. TINS can be applied as a primary screening step, but once elaborated fragments have been made, it can also be worthwhile to use TINS as a secondary screening step.

In conclusion, TINS can now pave the way to applying fragment based drug discovery to membrane proteins in general because it enables one to identify weak fragment binding to membrane proteins, whether solubilized in detergents or nanodiscs, in smaller quantities than those required for biophysical methods in general. The interesting perspectives which follow such applications are numerous because the immobilized target and references can be modulated to include, for example, different states or mutants of the same protein, and libraries can be modulated to include studies on allosteric modulation, and synergetic or competitive effects on weakly interacting biomolecules.