Universiteit
Leiden
The Netherlands

Biophysical characterization of membrane protein-small molecule interactions

Chen, D.

Citation
Chen, D. (2015, March 4). *Biophysical characterization of membrane protein-small molecule interactions*. Retrieved from https://hdl.handle.net/1887/32075

Version:	Corrected Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/32075

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

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/32075 holds various files of this Leiden University dissertation

Author: Dan Chen
Title: Biophysical characterization of membrane protein-small molecule interactions
Issue Date: 2015-03-04

Biophysical characterization of membrane protein-small molecule interactions

Dan Chen

Biophysical characterization of membrane protein-small molecule interactions

Proefschrift

Ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus Prof mr. C. J. J. M. Stolker,

volgens besluit van het College voor Promoties

te verdedigen op woensdag 4 maart 2015

klokke 15:00 uur

door

Dan Chen

Geboren te Dalian, P. R. China

in 1980

Promotiecommissie

Promotor:	Prof. Dr. M. Ubbink
Co-promotor:	Dr. G. Siegal
Overige leden:	Prof. Dr. J. Brouwer Prof. Dr. A. P. IJzerman Prof. Dr. I. de Esch (VU University Amsterdam) Prof. Dr. A. Milon (Université Toulouse III - Paul Sabatier)

Biophysical characterization of membrane protein – small molecule interactions

Dan Chen

Doctoral Thesis, Leiden University, 2015

The research described in this thesis was funded by an EU Seventh Framework Program [FP7/2007-2013] under grant agreement no. [211800].

ISBN: 978-94-6259-544-6

This thesis was printed by Ipaskamp Drukkers (Enschede, the Netherlands)

To my parents

献给我亲爱的父母

Table of Contents

List of abbreviations

ADMET	Absorption,distribution,metabolism, excretion, toxicology
APols	Amphipols
BSA	Bovine serum albumin
cAMP	Cyclic adenosine-5'-monophosphate
CHO	Chinese hamster ovary
CPMG	Carr-Purcell-Meiboom-Gill
CSP	Chemical shift perturbation
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethylsulfoxide
DTT	Dithiothreitol
EC_{50}	Half-maximal effective concentration (potency)
E_{max}	Maximal effect (efficacy)
FBDD	Fragment-based drug discovery
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HSQC	Heteronuclear single quantum coherence
HTS	High-throughput screening
IB	Inclusion body
ICL	Intracellular loop
ILOE	Inter-ligand Overhauser effect
IPTG	Isopropyl β-D-1-thiogalactopyranoside
GPCR(s)	G protein-coupled receptor(s)
LE	Ligand efficiency
MP(s)	Membrane protein(s)
MSP	Membrane scafford protein
NAM	Negative allosteric modulator
NCE	New chemical entity
ND(s)	Nanodisc(s)
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser effect spectroscopy
IC_{50}	Half-maximal inhibitory concentration (affinity)
K_d	Equilibrium dissociation constant
K_i	Equilibrium inhibition constant (absolute affinity)
k_{off}	Dissociation rate constant
k_{on}	Association rate constant
PAM	Positive allosteric modulator
PBS	Phosphate-buffered saline
PCS	Pseudocontact shift
PD	Parkinson's disease
PRE	Paramagnetic relaxation enhancement
SAR	Structure-activity relationship
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPR	Surface plasmon resonance
STD	Saturation transfer difference
TCEP	Tris(2-carboxyethyl)phosphine
TINS	Target immobilized NMR screening
TMD	Transmembrane helical domain
Tris-HCl	Trisaminomethane hydrochloride
VFTM	Venus flytrap module

Chapter 1

General Introduction

1. Drug Discovery and Development

Modern drug discovery consists of sequential steps (Figure 1), which usually start with identification and validation of a target in a disease process, using recent advances in human biology, genomics, proteomics and computational power. The majority of targets selected for drug discovery efforts are proteins. Two classes predominate: G protein-coupled receptors (GPCRs) and kinases. Next step involves discovery of lead compounds that interact with the target chosen. High-throughput screening (HTS) process is the most common way that leads are usually found. Hundreds of thousands of compounds can be tested for their ability to affect the activity of the target protein, and then identify any that might be potential drugs. Based on the results, several lead compounds are selected for further optimization with suitable pharmacodynamics (affinity, selectivity, efficacy, physiochemical properties) and pharmacokinetics (metabolic stability and oral bioavailability). Once a compound that fulfills all of these requirements has been identified, it will be submitted to pre-clinical trials (on animals) and eventually clinical trials (on humans).

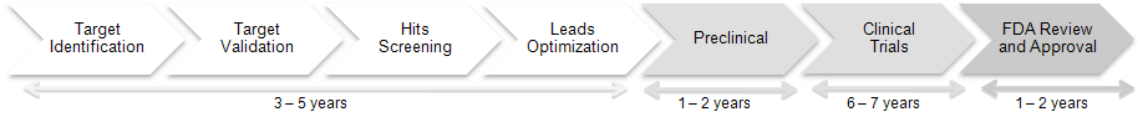


Figure 1. Modern drug discovery and development is a 10–17 year process from idea to marketed drug. For every 5,000 - 10,000 compounds that enter the R&D process, ultimately only one receives approval.

2. GPCRs: the Most Important Drug Targets

G protein-coupled receptors (GPCRs) are critical eukaryotic signal transduction gatekeepers and form a large and important protein family in the human genome, over 800 members.[1] They are located in the plasma membrane of all cell types. GPCRs respond to a variety of extracellular signals, from photons, protons, ions and small organic molecules, to peptides and glycoproteins. GPCRs mediate the actions of extracellular signals across the

plasma membrane over a distance of ~30 Å, and then convert them into intracellular responses via coupling to cytoplasmic heterotrimeric G-proteins, β-arrestins and kinases, which then activate downstream effectors and trigger cascades of cellular and physiological responses.[2] GPCRs play a crucial role in regulating the pathophysiology of numerous diseases, so they are one of the most important target families for drug discovery.

2.1 GPCRs Classification and Pharmacology

All GPCRs share a common structural topology of seven-transmembrane helical domains (TMD), an extracellular N-terminus, and an intracellular C-terminus. They can be classified into five major groups: Rhodopsin-like; Secretin-like; Adhesion; Glutamate; and Frizzled/Taste 2.[1] These families can be further divided into numerous subfamilies based on structural and sequence similarities. The largest subfamily is the rhodopsin family (family A), having a short extracellular N-terminus and having several highly conserved amino acids in the TMD bundle. In addition, an eighth helix is present at the C-terminus (Figure 2). The family A receptors can be activated by a variety of small molecules including amines, purines, lipids, peptides, as well as large glycoproteins. Almost a quarter of marketed small molecule drugs target this subfamily of GPCRs.[3] In this family, the endogenous small molecule (orthosteric) binding site is in the top of the TMD bundle (Figure 2, green hexagon).[4] Many molecules also bind to this site and are competitive with the natural ligands. As shown in Figure 3, the binding of agonists results in an increase of receptor activity. The binding of inverse agonists reduces receptor activity. Antagonists act to block GPCR activation by preventing the binding of agonists to the receptor. Molecules that bind to the receptor at distinct sites from orthosteric binding site, are called allosteric modulators. Binding of allosteric ligands can either enhance or inhibit the binding of the orthosteric agonist, and ligands are termed positive or negative allosteric modulators (PAM and NAM), respectively. Allosteric ligands can be promising drug candidates since they are more specific to the subtype receptors and have

less side effects.[5] For family A receptors, it is thought that the binding of allosteric modulators often specifically takes place on the top of the helices and the extracellular loops of receptors (Figure 2, red triangle).[6]

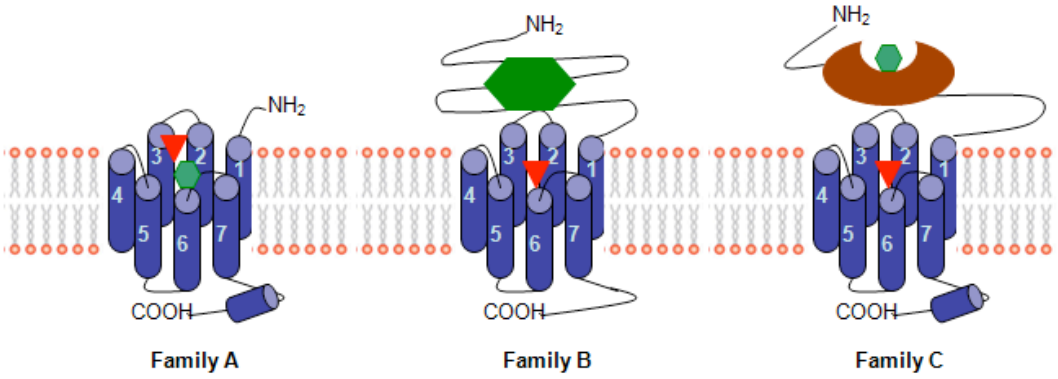


Figure 2. Schematic representations of overall structures of family A, family B, and family C GPCRs. The binding pockets of orthosteric and allosteric ligands are present as green hexagon and red triangle, respectively. The Venus flytrap module of family C GPCRs is present as brown cavity.

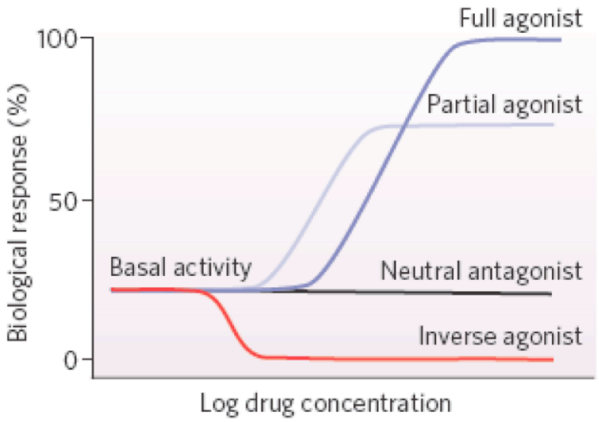


Figure 3. Classification of ligand efficacy for GPCRs. Permission and copyright 2009 Nature.

The secretin and adhesion families (family B) have a large extracellular N-terminal domain containing several conserved cysteine residues that form a network of disulphide bridges. The secretin family is activated by hormone peptides including glucagon, secretin, calcitonin and parathyroid hormone. The binding occurs at the N-terminal domain and the upper part of the TMD, which results in an essentially unoccupied TMD bundle. Moreover, allosteric binding sites for family B GPCRs have been identified within the TMD[7] (Figure 2).

Several drugs (peptides) were launched on the market, targeting this family of receptors, such as Byetta[8] (exenatide, glucagon-like peptide 1 agonist) and Miacalcin[9] (calcitonin). However, to date, this family has proven to be very difficult for small molecule drug discovery. For the adhesion family, the extracellular domain usually contains a wide variety of protein modules that are known to be involved in protein-protein interactions, such as lectin-like, immunoglobulin-like and cadherin-like motifs. The majority of this family consists of orphan receptors, and few attempts have been made to de-orphanizate to yield new drug targets.[10]

The final major class is the glutamate family (family C), which contains receptors for the metabotropic glutamate and γ-aminobutyric acid (GABA) as well as the calcium sensing receptors.[11] These receptors are characterized by a long N-terminus (500-600 residues) that forms a separate ligand-binding domain called the Venus flytrap module (VFTM, composing of two domains that form a cavity), distinct from the TMD.[4] Drugs acting on this class of receptors either bind at N-terminal domain (e.g., baclofen,[12] agonist of the $GABA_B$ receptors) or are allosteric modulators that bind within the TMD (e.g., cinacalcet,[13] allosteric modulator of the calcium-sensing receptors).

2.2 High-Resolution GPCR Structures

Until recently, the field of membrane protein structural biology was 20-30 years behind the study of soluble proteins. However, this field has quickly caught up by developing better ways to produce, purify and crystallize membrane proteins. The first crystal structure of a membrane protein, a bacterial photosynthetic reaction centre, was solved at 3 Å resolution in 1985.[14] It supported the fact that membrane proteins could be crystallized. Fifteen years after, the first high-resolution three-dimensional structure of a GPCR, rhodopsin, was solved by a collaboration of the groups of Okada and Palczewski.[15] This structure revealed the architecture of the seven transmembrane helical domain, which has been used as the basis for most GPCR homology models. Okada

developed a simple method for purifying rhodopsin from rod outer segment membranes using detergent extraction, which can improve the protein stability and facilitate the crystallogenesis. However, rhodopsin is an unusual GPCR because of its relatively high biochemical stability and its high expression level from natural tissue, bovine retina.[16] The major obstacles to obtain structures of other GPCRs include protein production, purification (inherent hydrophobicity), protein stability and conformational flexibility. In terms of production, it is now possible to generate sufficient quantities (tens of milligrams) of GPCRs using insect cell (sf9 and Hi5 cells) expression systems.[17] To overcome the thermodynamic and protein instability, several innovative protein engineering techniques have been developed (described in further detail below), which resulted in an increasing number of solved structures since 2007, when researchers solved the structure of a second GPCR: the human β_2-adrenergic receptor.[18,19] In the past thirteen years, using these newly developed techniques, more than 75 structures of 21 different family A GPCRs have been determined in complex with ligands, peptides, antibodies and a G-protein (heterotrimeric Gs), as shown in Table 1.

Table 1. GPCR crystal structures.

Receptors subfamilies	Receptors
rhodopsin	bovine rhodopsin [20], squid rhodopsin [21]
aminergic GPCR family	β-adrenergic receptors (avian β_1AR [22] and human β_2AR [19], muscarinic acetylcholine receptors (human M_2R [23] and rat M_3R [24]), human H_1 histamine receptor [25], human D_3 dopamine receptor [26], human serotonin receptors (5-HT_{1B} [27] and 5-HT_{2B} [28])
nucleoside-binding GPCR	human adenosine A_{2A} receptor (A_{2A}R) [29]
peptide-binding GPCRs	human CXCR1 [30] and CXCR4 [31] chemokine receptor, opioid receptors (human nociceptin receptor [32] and κ-OR [33] and mouse μ-OR [34] and δ-OR [35]), rat neurotensin receptor (NTSR1) [36], human protease-activated receptor (PAR1) [37]

lipid-binding GPCRs	human sphingosine-1 phosphate ($S1P_1$) [38] receptor, human GPR40 free fatty-acid receptor 1 (FFAR1)[39]
glycoprotein-binding GPCR	smoothened receptor (SMO) [40]
Class B GPCRs	corticotropin releasing factor receptor 1[41], glucagon receptor[42]
Class C GPCRs	human metabotropic Glutamate Receptor 1 (mGlu1)[43], human metabotropic Glutamate Receptor 5 (mGlu5)[44]
glutamate GPCR	GABA(B) receptor[45]

2.3 Human Adenosine A_{2A} Receptors – Therapeutic Potential

The adenosine A_{2A} receptor ($A_{2A}R$) is one of four adenosine receptor subtypes (A_1, A_{2A}, A_{2B} and A_3) belonging to the family A GPCRs.[46] It is widely distributed throughout the body, and can be activated by extracellular adenosine, thereby stimulating the cyclic AMP-protein kinase A pathway by coupling to G_s protein. $A_{2A}R$ activation occurs in response to stress, cell damage or ischemia,[47] and protects tissues by controlling inflammation. For this reason, $A_{2A}R$ agonists are candidates for development of anti-inflammatory drugs. Nearly all of the known $A_{2A}R$ agonists are derived from purine nucleosides;[48] a representative selection of $A_{2A}R$ ligands is shown in Figure 4. Several of them are promising cardiovascular clinical candidates.[49–51] The $A_{2A}R$ agonist BVT.115959 is currently in phase II clinical trials for the treatment of diabetic neuropathic pain.[52] Regadenoson is the only clinically approved $A_{2A}R$ agonist and is used as a coronary vasodilator in myocardial perfusion imaging.[53]

$A_{2A}R$ antagonists have generally been derived directly from xanthine, or contain a heterocyclic ring in place of the xanthine core.[54] The natural products caffeine and theophylline act as weak $A_{2A}R$ antagonists. The purine-like derivative ZM241385, a potent and selective $A_{2A}R$ inverse agonist, has been widely used in research. $A_{2A}R$ has been shown to be a novel, non-dopaminergic therapeutic target for Parkinson's disease (PD). $A_{2A}R$ is co-expressed with dopamine D_2 receptor in the striatum and functionally opposes the actions of the

D_2 receptor.[55] The pathologic hallmark of PD is the gradual loss of dopamine and hence reduced dopamine D_2 receptor activity. Blockade of $A_{2A}R$ might reduce the striatal dopamine depletion and dopaminegic cell loss.[56] Several $A_{2A}R$ antagonists/ inverse agonists have been evaluated in clinical trials as alternative therapeutic agents for the treatment of PD, because of their efficacy and advantages in reducing motoric dysfunction and suppressing L-DOPA-induced dyskinesias, which experienced by patients following long-term levodopa therapy.[52] Preladenant (SCH420814; Merck) is currently being evaluated in phase III clinical trials for the treatment of PD.[57] The $A_{2A}R$ antagonists also have potential indications for other neurodegenerative diseases, such as Alzheimer's disease, depression, and addiction.[58]

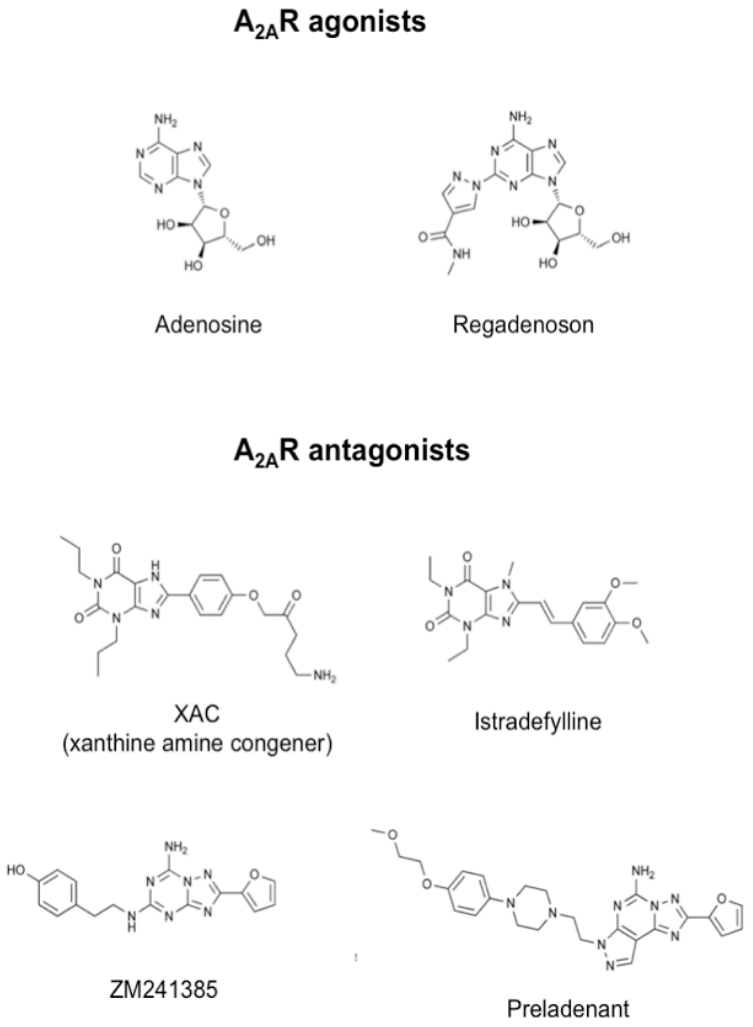


Figure 4. A representative selection of $A_{2A}R$ ligands.

To date, twelve high-resolution crystal structures of $A_{2A}R$, with agonist or antagonist/inverse agonist bound, have been solved. The binding site position and protein conformation have been determined in both active and inactive states. This structural information could be used to aid in designing specific compounds with drug-like properties for $A_{2A}R$.

2.4 GPCRs Drug Discovery – Successful and Challenging

GPCR drug discovery has been very successful, a large number of small molecule drugs have been developed in multiple therapeutic areas. Sixty-three new GPCR drugs have been launched in the past 10 years and in total they account for an estimated 30-40% of all drugs currently on the market.[59] GPCR drug discovery presents an important area for the pharmaceutical industry. More than a quarter of the 100 top-selling drugs target GPCRs, with annual sales of over 65 billion dollars.[60] However, research in this area is still very challenging, and only one novel GPCR has been drugged per year by a new chemical entity (NCE). In addition, over 100 clinically relevant GPCRs still remain undrugged. They are orphan GPCRs with unknown endogenous ligands and in most cases unknown function.[61] Traditional GPCR drug discovery efforts have relied on cell-based assays combined with HTS of large collections of compounds (tens to hundreds of thousands of compounds), with typical molecular weights ranging from 350 to 500 Da. While this approach has been successful for many GPCRs (like adrenergic, muscarinic, and histamine receptors)[60], recent clinically validated high-value targets such as neuropeptide receptors, chemokine receptors, peptide-hormone receptors and metabotropic glutamate receptors have represented a big challenge and been intractable to small molecule drug discovery. Fragment-based drug discovery (FBDD) is now established as an alternative strategy to HTS for ligand discovery, where low molecular weight compounds, known as fragments, are screened as new chemistry starting points. The screening of fragment libraries usually employs biophysical screening methods, several techniques have been validated for GPCRs drug discovery.[62–64]

In addition, recent impressive progress in obtaining GPCR structural information has enabled the use of structure-based drug design methods for hit identification and optimization. Until recently, application of both fragment- and structure-based drug discovery to GPCRs is still challenging, primarily because of the instability of the proteins when isolated and purified from the cell membrane. New advances in protein stabilization by using different protein engineering methods have facilitated GPCR structural and biophysical studies.[65] These approaches will be discussed in next section.

3. Fragment-Based Drug Discovery

3.1 Small is Beautiful: Why Fragments?

Over the past two decades, pharmaceutical industry invested significantly into HTS for identifying potent compound "hits" against a drug target, typically proteins.[66] HTS involves screening of large compounds libraries, typically containing a million complex "drug-like" compounds which have been selected to comply with the Rule of Five as formulated by Lipinski *et al.* [67] : molecular mass < 500 Da, cLogP (the calculated logarithm of the partition coefficient between water and 1-octanol) < 5, number of hydrogen-bond donors ≤ 5 and number of hydrogen-bond acceptors ≤ 10. The Rule of Five is used to maximize an oral drug molecule's probability of surviving development, *i.e.* has good pharmacokinetic and ideal ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) properties. Despite many success cases with high-value hits, there are many obvious limitations to drug-sized compounds screening. Hit rates are often low, particularly when screening against challenging targets; and many hits fail in the process of lead optimization.[68,69] In the recent decade, HTS has not delivered on its promise of increasing the numbers and quality of new drugs entering clinical trials. This lack of success is in part due to the complexity and the relatively large size of the compounds routinely being screened. Large HTS compound libraries represent only a small fraction of chemical diversity

space (i.e. all possible compounds that could exist) and poor sampling by HTS libraries is a limitation for finding a good starting point for further lead optimization. Moreover, optimization into more potent compounds often tends to reduce their initial drug likeness properties,[70–72] such as bioavailability, absorption and solubility. Finally, even if the three dimensional (3D) structure of an HTS hit binding to the target protein can be obtained, it may still not be clear which parts of the molecule contribute most to the binding energy, leading to ambiguity about how to increase potency.

In 1981, William Jencks first proposed the theory behind FBDD.[73] Over the past 15 years, FBDD has developed as an alternative approach to HTS for hit identification. Screening starts with low-molecular weight drug fragments, which typically conform to the "Rule of Three" [74]: molecular mass of <300 Da, up to three hydrogen bond donors, up to three hydrogen bond acceptors, and cLogP of ≤3, in addition, up to three rotatable bonds and polar surface area ≤ 60 $Å^2$. Due to the simplicity of the fragments it is possible to screen a large proportion of chemical space with a small collection of fragments.[71] The basic promise is that, instead of searching a huge collection (million) of potential drug-like molecules (up to 30 heavy atoms, 10^{60} possible molecules),[75,76] one could screen smaller (few thousands) collections of fragments (up to 12 heavy atoms, 10^7 possible molecules).[77] Lower complexity of fragments gives a higher probability at probing the key binding areas of a protein binding site, and results in a higher and more reliable hit rate than HTS.[70] Because of the small size of fragment molecules, they typically bind to the target proteins with lower affinity [78] (micromolar to millimolar range) compared to drug-like compounds, which can form more interactions (nanomolar to micromolar range). Although fragments are weak binders, they can adopt optimum orientations in the active site, demonstrated by high ligand efficiency (LE, that is, binding energies per atom of a ligand to its target protein).[79] LE has emerged as a useful guide to optimize fragment and lead selection in the discovery process. All the factors mentioned above are crucial in successful drug development. However, the low K_D requires a more elaborate optimization to obtain a lead compound. Traditional bioassays used in

high-throughput screening are generally unable to detect such small fragments because of their low potency binding (K_D > 1 μM) [80] to the protein target. In order to detect these weak binders, FBDD employs highly sensitive biophysical techniques, such as nuclear magnetic resonance (NMR) spectroscopy, surface plasmon resonance (SPR) and protein X-ray crystallography for screening. Today, these biophysical techniques are extensively used in fragments screening, because they are highly sensitive to detect low affinity fragments binding, and more interestingly, provide the structural information on target-fragment binding.

The first practical implementation of fragment-based screening was carried out in the pioneering studies at Abbott using structure−activity relationship (SAR) by NMR [81] in the late 1990s, and at Astex using protein X-ray crystallography since the early 2000s.[82,83] This sparked a revolution in FBDD and there is a substantial number of published works to support this approach, from both drug discovery companies and academic institutes. Initially, FBDD was employed against soluble targets, principally on kinases that had shown the ability to bind small molecules with high affinity. Then, FBDD approach was successfully applied to other soluble protein families for which structural information was readily available. There are examples of success with fragment-based methods against targets that have proven to be very difficult with traditional HTS approaches.[84] For example, the aspartic protease β-secretase (BACE) was considered an intractable target fifteen years ago. But now several companies have used FBDD approach to find high affinity inhibitors and have advanced compounds into clinical trials.[85,86] Fragment-based methods are now being used against targets as diverse as protein−protein interactions,[87] transcription factors,[88] protein chaperones, and RNA.[89] The approach has yielded one market drug in 2011: vemurafenib.[90,91] It took only 6 years since Plexxikons' researchers started to look for the low-molecular-weight hits against kinases until the FDA approval of vemurafenib for metastatic melanoma. Implementation of FBDD in the GPCR field has lagged behind the soluble protein families, primarily because GPCRs are highly unstable when extracted from cell membranes,

making them very difficult to isolate, purify and obtain in sufficient quantities for biophysical screening. However, in recent research, biophysical fragment screening methods such as NMR, SPR have been validated for GPCR target families.[62,63]

3.2 Methods for Finding Fragments

Reliably detecting ligand binding in the affinity range of fragments is a technical challenge, as traditional bioassay methods are often unsuitable to screen weak protein-ligand binding. Over the past decade, this issue has been successfully overcome by developing a variety of sensitive biophysical methods for detection, of which NMR, SPR and X-ray crystallography (which will be discussed in next section) have been the most widely used. A significant advantage of biophysical screening methods is that we can directly detect the binding of hits with target in a reversible manner.

Nuclear Magnetic Resonance

NMR-based fragment screening was firstly introduced in 1996 as “SAR by NMR” by Fesik [92] and co-workers at Abbott laboratories. The protein is ^{15}N isotopically labelled and [^{1}H,^{15}N] Heteronuclear Single Quantum Coherence (HSQC) NMR spectra are recorded in the presence and absence of fragments that are added in a mixture of up to 30 at a time. Chemical shift changes in protein spectra in the presence of a fragment indicate binding, and if the resonance assignment is available, the location of ligand-binding site can be determined. This approach is called “protein-observed” NMR, which relies on changes in the NMR spectrum of the protein. Such a powerful approach with very low false positive rates, has led to a large number of high-affinity inhibitors, and some of them have moved to clinical trials,[93] for example: ABT-518 [94] and ABT-263 [95] for the targets matrix metalloproteinase 3 and Bcl-2, respectively. However, it is limited to relatively small proteins (around 30 - 40 kDa), due to relaxation time and signal broadening. Moreover, the approach requires large

quantities of labelled protein, generally at least 200 mg.[92] Furthermore, both the protein and the fragments need to be soluble in the experimental buffer. In addition, a low hit rate requires a rather large fragment library, typically more than 15,000 compounds.

For all the above reasons, a number of NMR-based methods have been developed to detect binding via changes in the NMR spectrum of the ligand. This approach presents large advantages in that it doesn't suffer the limitation of protein size, and the protein concentration can be 10-100 fold lower compared to protein observed methods. Moreover, the target proteins need not be isotopically labelled. In addition, ligand-based approaches get a higher hit rate (in general 5-10 fold), therefore, the library size can be reduced. There are a number of methods in use: the most popular in practice are the saturation transfer difference (STD) [96] method and the WATERLOGSY [97] method. Both methods rely on the selective transfer of magnetization from the protein to the ligand upon target binding. An interesting ligand-detected approach called Interligand nuclear Overhauser Effects (SAR by ILOE) identifies pairs of weak binders for a target protein, facilitating linking,[98] although false positives need to be avoided due to aggregation of compounds.[99] Another interesting ligand-detected method is called target-immobilized NMR screening (TINS) (Figure 5).[100] TINS uses a single sample of immobilized protein to screen the entire library with very low false positive rates, due to the use of a reference protein.[101] Utility of TINS methodology for identification of fragment ligands for challenging targets such as integral membrane proteins will be discussed in this thesis. TINS is compatible with competition mode screening in which displacement of a known ligand is used as the readout.[62] It should be noted that the tight binding ligand are not suitable for detection by ligand observed NMR methods, which are dependent on the "fast exchange regime" between the ligand and the target protein.

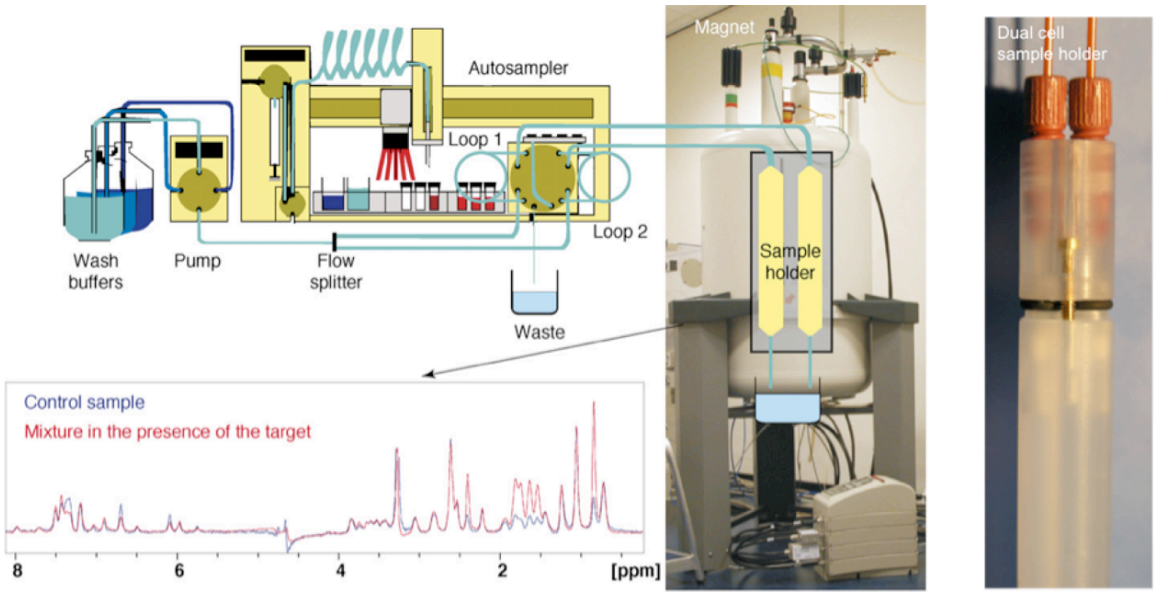


Figure 5. A patent protected TINS screening station has been developed by ZoBio B.V. in collaboration with Bruker. The system is built around a 500 MHz NMR spectrometer. The target and reference proteins are immobilized on a sepharose resin and packed into separate cells of the dual cell sample holder (right). The sample holder is inserted into the magnet and the mixtures of fragments to be screened are injected into each cell. A spatially selective NMR experiment is used to acquire a 1D ^{1}H spectrum of the soluble material in each cell (blue & red overlaid spectra) (left). Binding of a fragment is detected by a reduction in the height of all its peaks, conveniently expressed as a ratio of peak height in the target to that in the presence of the reference (T/R ratio). Figure reproduced from Bartoli *et al*.[103] Permission and copyright 2006 Drug Discovery Today: Technologies.

Surface Plasmon Resonance

Today, the use of surface plasmon resonance (SPR) to characterize fragment binding becomes a primary screening technique, because of recent improvements in SPR biosensors and instrumentation. In direct SPR fragment screening, solutions of fragments flow past a target biomolecule that has been immobilized onto a coated chip and the binding can be detected by the increase of mass at the sensor surface. In some cases, the binding kinetics can be measured and the binding affinity calculated from the fragment association−dissociation response.[104] Now, SPR allows detection of binding of small molecules with molecular weights as low as 100 Da.[105] The use of suitable control proteins could significantly reduce the number of false positives from non-specific binders. For example, we can choose unrelated proteins that are

solubilized and immobilized in the same way as target protein, or proteins with binding sites being blocked with known high affinity ligands.

Another interesting approach has been developed by Graffinity Pharmaceuticals.[106] All compounds are linked to the same proprietary spacer molecule (ChemTag®), which serves as an attachment point for the covalent immobilization on the array surface. Assessment of the protein binding is then carried out by measuring the change in refractive index. The experiments could also be performed in a competition mode, in which a known ligand is immobilized on the surface and the target is allowed to bind. When a competing fragment is introduced, the dissociation of the target from the surface is detected as a large signal change. SPR experiments are relatively rapid and straightforward to set up, so screening of entire fragment libraries by SPR is now possible within a few weeks, requiring as little as 25−50 μg of protein. SPR can be a very useful tool: not only can it provide the affinity and dissociation constants, it can also provide the stoichiometry of binding.

SPR is emerging as a powerful technique for fragment screening against difficult targets such as GPCRs.[107] As with TINS, SPR is also a “label-free” method that measures the direct binding of ligand to the receptor. Moreover, SPR measures in real time and kinetics and affinity can be determined. In addition, SPR can be applied to purified, detergent solubilized thermostable receptors[108] and also native receptors directly captured from crude cell lysates. Nevertheless, a known ligand must be available to assess the validity of the assay.

X-Ray Crystallography

X-ray Crystallography and protein-observed NMR are unique in providing detailed structural binding information. Unlike NMR, crystallography has no limitation on the protein size and can provide validated hits and structural binding information in one step. The use of high-throughput X-ray crystallography for FBDD was pioneered by Astex Therapeutics [109] and rapidly increased over the past 15 years. In a successful case in Astex, approximately 500 fragments were

screened by X-ray crystallography against cyclin-dependent kinase 2 by soaking apo CDK2 crystals. A fragment progressed into a drug candidate, AT7519, that is currently in phase II clinical trials.[110]

In fragment screening, crystals of the target protein are soaked with high-concentrations of cocktails containing 3-10 compounds (in general 10 mM per compound). In some case, multiple hits can bind to the target with partial occupancies. To avoid the ambiguous electron density assignment, the compounds in each cocktail should contain distinctive shapes and the average hit rate should be less than one compound per cocktail.[111] The protein crystal must be robust, exhibit a good diffraction, generally better than 2.5 Å resolution. Otherwise, the low-resolution can yield a wrong position or conformation of the ligand. The ligand binding site should be accessible in the crystal, and the fragment binding should not induce a large conformational change that cracks the crystal,[112] and the number of fragments in each cocktail should be kept low. X-ray crystallography provides very limited information on ligand binding affinity, and thus it must be correlated with other experimental techniques, such as NMR and SPR.[85,113]

Mass Spectrometry

NMR, X-ray crystallography, and SPR are the best-known biophysical methods for FBDD today, but several other approaches can also be used, such as Mass Spectrometry (MS). MS can be used to detect fragments that bind to a protein either covalently or non-covalently. Sunesis Pharmaceuticals developed an approach called “Tethering”,[114] which relies on formation of a disulfide bond between a cysteine residue in the target protein and sulfur-containing fragments. Fragments that bind near the cysteine form disulfide bonds with the protein, increasing the mass of the protein and allowing the detection of the fragment binding by mass spectrometry. Because of the equilibrium of the disulfide bond formation, Tethering can discriminate weakly bound fragments from those that have no intrinsic binding affinity to the protein. A potent selective inhibitor of

aurora kinases SNS-314 was optimized after identification of a precursor using Tethering.[115] It is also possible to detect fragment binding to proteins via native MS, successfully developed by NovAlix.[116]

In silico Fragment Screening

In recent years, virtual screening has become widely used and is maturing as a hit identification strategy. In most experimental FBDD studies, a few thousand fragments can be screened. In contrast, at least 250,000 fragments are commercially available, leaving a large portion of potential fragment libraries untested. Because commercially available fragments are too numerous to be screened experimentally, virtual fragment screening by molecular docking is a complementary approach. Moreover, *in silico* screening offers a way to identify novel scaffolds as starting points for further structure-based lead optimization.

There are many ways to successfully find and characterize fragments *in silico*, each with its own set of strengths and weaknesses. Several orthogonal methods should be used in combination. The general fragment-based process starts by fragment library design and target validation; it is followed by fragment screening, using biophysical techniques to detect non-covalent binding to a target; then fragment is optimized by a combination of structural information, *in silico* design, and chemical synthesis.

4. New Advanced Strategies to Enable GPCRs for FBDD and SBDD Campaigns

4.1 Methods for Stabilizing GPCRs

GPCRs are inherently dynamic proteins, and have numerous thermodynamic conformations between active and inactive states. In addition, low expression and poor stability once extracted from cell membrane environment represent major obstacles for biophysical studies and crystallogenesis.[17] In order to overcome these limitations with GPCRs, several

innovative protein engineering methods have been developed. The disordered extra-membranous loop regions and flexible N- and/or C-terminal domains of the receptor have been deleted. The receptor was predominately stabilized in one conformation by adding high affinity ligands. A recent method introduces a number of points mutations into the receptor construct in order to increase the thermostability of the protein.[117] This latter approach, called stabilized receptors or "StaR", locks the receptor in one specific conformation that greatly facilitates protein purification, structure determination, and biophysical screening for drug discovery. The StaR approach has been used to obtain the structure of avian β_1AR,[118] human adenosine A_{2A}R,[119,120] and neurotensin receptors [36]. Use of A_{2A} StaR for NMR based fragment screening is demonstrated in this thesis. An alternative strategy, as the T4 lysozyme (T4L) engineering method, the third intracellular loop (ICL3) was replaced with lysozyme from bacteriophage T4. Usually when the third intracellular loop (ICL3) of the receptor is present in the construct, it leads to conformational heterogeneity and is found disordered in crystallized GPCRs. Here, the fusion proteins were introduced to minimize conformational flexibility and improve the crystal contacts, rather than increase thermal stability. So far, this method enabled the crystallization of 13 GPCRs.[121] More recently, replacement of ICL3 by the thermostabilized apocytochrome b562RIL showed great utility in the crystallization of multiple GPCRs.[27,28,40] A final strategy introduces a fusion protein and creates a complex of the receptor with monoclonal antibody fragments (Fab5,[28] Fab2838,[122] and Nanobody Nb80[123]), which were generated to recognize conformational epitopes of GPCRs. This strategy has been applied to a growing number of GPCRs.

4.2 New Solubilization Strategies for GPCRs in Aqueous Solution

Membrane proteins (MPs) are naturally embedded in a mosaic lipid bilayer, which is a heterogeneous and dynamic environment. Biophysical studies require an isolated pure stable protein in a functional folded form. Detergents play a critical role in the MP extraction and purification, because they present an

amphipathic environment that can mimic lipid bilayers, thus maintaining the function and structure of the protein.[124] In fact, detergents do not stabilize membrane domains as efficiently as natural lipidic membranes, and the excess detergent micelles pose a hazard to MP stability as well as significant levels of non-specific partitioning of compounds into micelles during biophysical fragment screening.[102] Detergents often co-concentrate with protein and lead to a loss of activity/stability during protein extraction and purification.[125] This limits the subsequent biophysical studies, such as NMR and X-ray crystallography.

Recently, non-detergent surfactants and novel detergent alternatives have been developed in order to better solubilize and stabilize MPs for structure determination and function studies.[126] Amphipols (APols) are amphipathic polymers which have been designed to keep individual MPs stable in aqueous solutions.[127] APols are relatively short polymers (8-20 kDa) that carry a hydrophilic backbone grafted with hydrophobic side chains. APols present high-affinity/low off-rate for the transmembrane domain and form a protein-APol complex stable for long periods of time.[128] APols can be used to trap detergent solubilized MPs into less destabilizing conditions, keep their native structure and preserve their functionality. Several membrane proteins have been studied in complexes with APols including the acetylcholine receptor,[129] OmpA, FomA, and bacteriorhodopsin.[130] APols have proven to be efficient for GPCR refolding. This procedure can start from the protein in a denaturing environment, typically urea for β-barrel MPs and SDS for α-helical ones, and to (re)fold the protein to their native state by adding the APol and removing the detergent by dialysis.[131]

One novel model membrane system called “Nanodiscs” (NDs) has also been used to study integral membrane proteins. Nanodiscs are self-assembling complexes that consist of a phospholipid bilayer core surrounded by two molecules of an amphipathic α-helical membrane scaffold protein (MSP). The MSP is a fragment of human apolipoprotein A-1 engineered by Sligar and colleagues.[132] Target protein can be incorporated into the self-assembly process in which detergent is removed from initial detergent-phospholipid micelles, either by dialysis or addition of detergent-removal Bio-beads. This procedure

reconstitutes the target protein into a native-like environment and the final complex can be handled in aqueous conditions. Many MPs (bacteriorhodopsin, human β_2AR, and cytochrome P450)[132,133] have been successfully embedded into NDs, with preserved functionality, and subsequently used for biochemical and biophysical studies. In the thesis, the first complete biophysical screen of a fragment library against an integrate membrane protein DsbB which was solubilized in both detergent micelles and NDs is reported.[102] It has been shown that NDs are suitable for biophysical studies of fragment binding to membrane proteins, with reduced non-specific partitioning of compounds into solubilization medium, specifically for hydrophobic compounds.

Another model membrane system, bicelles, micelles formed of lipid bilayers, make an excellent membrane mimic environment for membrane proteins. Bicelles have found successful applications in solution and solid-state NMR,[134,135] and more recently in membrane protein crystallography.[136]

4.3 GPCR Overexpression in Bacterial Inclusion Bodies

Currently, there are three major approaches to obtain structural information on the ligand-target interaction: X-ray crystallography, solution-state NMR (solNMR) and solid-state NMR (ssNMR). X-ray crystallography and ssNMR can be performed with large proteins, in conditions similar to those of native cell membrane. However, they present some significant difficulties, for example, producing crystals and obtaining good diffraction, and lack of dynamic information. SolNMR methods have therefore become more commonly used in the study of protein-ligand interactions. SolNMR can be used to investigate weak protein-ligand interactions and obtain 3D structural information of target-ligand complexes, where X-ray crystallography is not successful.

Jean-Louis Banères developed a new strategy to produce large amounts of functional GPCRs from *Escherichia coli*, which enable structural characterization of GPCRs in solution.[131] Once the target can be expressed in *E. coli*, then stable isotope labels can be incorporated, enabling the full range of

NMR tools to be applied. Addition of an α_5 integrin fragment (α_5I, 31 kDa) as a fusion partner can facilitate the recombinant expression of many rhodopsin-like GPCRs from *E. coli* inclusion bodies (IBs). Expression of GPCRs from IBs presents many advantages. IBs are stable and resistant to proteolytic degradation, and they are not toxic to the cell. GPCR production in IBs can be achieved at very high level. This strategy has been used to successfully over-express rat olfactory OR_5 receptor, human leukotriene BLT1 and BLT2 receptors, and human serotonin 5-HT_{4A} receptor.[137–140] However, the expressed GPCRs must be subsequently folded into their native state. Before refolding, the recombinant GPCRs accumulated in IBs need to be solubilized in denaturing conditions (SDS and/or urea). After removing the fusion partner, transfer of the receptor from SDS to APol, with lipid supplementation, can be used to achieve folding. Using this APol folding strategy, the leukotriene receptors BLT1 and BLT2 and the serotonin 5-HT_{4A} can be folded as functional proteins with an efficiency of 60% as shown in radioligand binding assays.[141] Most biochemical and biophysical techniques can be applied to APol-trapped MPs.[126] In the past few years, several MPs in complex with APols have been successfully used for solNMR for studying the structure of ligand binding and conformational transition.[142,143] Banères and colleagues used transferred nuclear Overhauser effects (trNOEs) to determine the structure of ligand binding on an APol-trapped GPCR.[144] A deuterated BLT2 receptor was folded and stabilized using a partially deuterated APol A8-35, the ligand leukotriene LTB4, was then added and the receptor-ligand complex structure was determined from the trNOE signals. These studies provide a novel approach for GPCR ligand binding determination.

Aim and Outline of this Thesis:

"To demonstrate an NMR-based methodology for characterization of fragment binding on solubilized integral membrane proteins, such as GPCRs"

TINS has been successfully applied to soluble proteins for investigating small molecule binding. In order to validate the proof of concept on membrane proteins, which are at the forefront of pharmacological interest, in **Chapter 2**, TINS screening has been performed against a bacterial membrane protein, the Disulphide Bond Forming protein B (DsbB). DsbB is chosen as a TINS candidate since it can be produced in large quantities from *E. coli*, has a robust enzymatic activity, and its 3D structure is available. A bacterial Outer Membrane Protein (OmpA), which has minimal small molecule binding, is selected as reference protein. The reference protein serves to cancel non-specific binding of hydrophobic compounds to lipids and/or detergents used to solubilize the membrane proteins. The reference protein was solubilized in the same detergent micelle conditions as the target protein. The suitability of TINS for fragment screening of membrane proteins, which was functionally solubilized in a non-detergent medium, has also been tested. The ND system has been explored as an alternative solubilization medium, which consists of a lipid bilayer surrounded by two amphipathic α-helical membrane scaffold proteins (MSP). The small molecule binders from both detergent- and nanodisc-solubilized DsbB screens have been validated and characterized using an enzymatic inhibition assay. The binding modes of validated fragments were confirmed by Heteronuclear Single Quantum Coherence (HSQC) using ^{15}N enriched protein. These results demonstrated the feasibility of using a fragment-based strategy to discover new chemistry starting points for subsequent development of compounds targeting membrane proteins. **Chapter 3** addressed the feasibility of TINS for fragment screening on an inverse agonist stabilized, StaR GPCR, the human adenosine A_{2A} receptor. A moderately sized fragment library was screened for binding to the A_{2A} StaR with OmpA as a reference protein, both solubilized in DDM micelles.

Multiple hits were validated by pharmacologic assays with wild-type receptor. Potential $A_{2A}R$ allosteric modulators discovered by TINS are described in this chapter. In parallel, the same fragment library was screened using molecular docking against an $A_{2A}R$ crystal structure, as described in **Chapter 4**. Further, the virtual screening was extended to 328,000 commercially available fragments. Novel $A_{2A}R$ ligands were then explored by structure-activity relationship, guided by molecular docking and molecular dynamics simulations. Elaboration of weak binding hits to lead-like compounds is still challenging for FBDD. Obtaining high-resolution 3D structures of protein-fragment complexes is a key to efficient hit optimization. In order to enable future NMR structural studies, in **Chapter 5**, a novel methodology to over-express A_{2A} StaR from *Escherichia coli* inclusion bodies has been employed. The receptor is subsequently refolded in its native state using Amphipols. The preliminary protein production results are presented. Finally, **Chapter 6** presents concluding remarks about the research work in this thesis and perspectives regarding the future of GPCR research when applying fragment- and structure-based approaches to GPCR drug discovery.

References

1. Fredriksson, R., Lagerström, M. C., Lundin, L.-G. & Schiöth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol. Pharmacol. 63, 1256–72 (2003).
2. Klabunde, T. & Hessler, G. Drug design strategies for targeting G-protein-coupled receptors. ChemBioChem. 3, 928–944 (2002).
3. Hopkins, A. L. & Groom, C. R. The druggable genome. Nat. Rev. Drug Discov. 1, 727–730 (2002).
4. Jacoby, E., Bouhelal, R., Gerspacher, M. & Seuwen, K. The 7 TM G-protein-coupled receptor target family. ChemMedChem 1, 761–782 (2006).
5. Müller, C. E., Schiedel, A. C. & Baqi, Y. Allosteric modulators of rhodopsin-like G protein-coupled receptors: Opportunities in drug development. Pharmacol. Ther. 135, 292–315 (2012).
6. Katritch, V., Cherezov, V. & Stevens, R. C. Diversity and modularity of G protein-coupled receptor structures. Trends Pharmacol. Sci. 33, 17–27 (2012).
7. Conn, P. J., Christopoulos, A. & Lindsley, C. W. Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. Nat. Rev. Drug Discov. 8, 41–54 (2009).
8. Briones, M. & Bajaj, M. Exenatide: a GLP-1 receptor agonist as novel therapy for Type 2 diabetes mellitus. Expert Opin. Emerg. Drugs 7, 1055–1064 (2006).
9. Lewiecki, E. M. Emerging drugs for postmenopausal osteoporosis. Expert Opin. Emerg. Drugs 14, 129–144 (2009).
10. Yona, S., Lin, H.-H., Siu, W. O., Gordon, S. & Stacey, M. Adhesion-GPCRs: emerging roles for novel receptors. Trends Biochem. Sci. 33, 491–500 (2008).
11. Pin, J.-P., Galvez, T. & Prézeau, L. Evolution, structure, and activation mechanism of family 3/C G-protein-coupled receptors. Pharmacol. Ther. 98, 325–354 (2003).
12. McLean, B. N. Intrathecal baclofen in severe spasticity. Br. J. Hosp. Med. 49, 262–267 (1993).
13. Iqbal, J., Zaidi, M., Schneider, A. E. Cinacalcet hydrochloride (Amgen). IDrugs. 6, 587–592 (2003).
14. Deisenhofer, J., Epp, O., Huber, M.R., Michel, H. Structure of the protein subunits in the photosynthetic reaction centre of Rhodopseudomonas viridis at 3 A resolution. Nature 318, 618–624 (1985).
15. Okada, T. et al. X-Ray diffraction analysis of three-dimensional crystals of bovine rhodopsin obtained from mixed micelles. J. Struct. Biol. 130, 73–80 (2000).
16. Kobilka, B. K. The gatekeepers revealed. Nature 465, 823–826 (2010).
17. Rosenbaum, D. M., Rasmussen, S. G. F. & Kobilka, B. K. The structure and function of G-protein-coupled receptors. Nature 459, 356–363 (2009).
18. Cherezov, V. et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science 318, 1258–65 (2007).
19. Rasmussen, S. G. F. et al. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. Nature 450, 383–7 (2007).
20. Palczewski, K. Crystal structure of rhodopsin: A G protein-coupled receptor. Science. 289, 739–745 (2000).
21. Murakami, M., Kouyama, T. Crystal structure of squid rhodopsin. Nature 453, 363–367 (2008).
22. Warne, T. et al. Structure of a beta1-adrenergic G-protein-coupled receptor. Nature 454, 486–491 (2008).
23. Haga, K. et al. Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. Nature 482, 547–551 (2012).
24. Kruse, A. C. et al. Structure and dynamics of the M3 muscarinic acetylcholine receptor. Nature 482, 552–556 (2012).
25. Shimamura, T. et al. Structure of the human histamine H1 receptor complex with doxepin. Nature 475, 65–70 (2011).
26. Chien, E. Y. T. et al. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. Science 330, 1091–1095 (2010).
27. Wang, C. et al. Structural basis for molecular recognition at serotonin receptors. Science 340, 610–614 (2013).
28. Wacker, D. et al. Structural features for functional selectivity at serotonin receptors. Science 340, 615–619 (2013).
29. Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, S. R. The 2.6 angstrom crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. Science 322, 1211–1217 (2008).
30. Park, S. H. et al. Structure of the chemokine receptor CXCR1 in phospholipid bilayers. Nature 491, 779–783 (2012).
31. Wu, B. et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. Science 330, 1066–1071 (2010).
32. Thompson, A. A et al. Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. Nature 485, 395–399 (2012).
33. Wu, H. et al. Structure of the human κ-opioid receptor in complex with JDTic. Nature 485, 327–332 (2012).
34. Manglik, A. et al. Crystal structure of the μ-opioid receptor bound to a morphinan antagonist. Nature 485, 321–326 (2012).
35. Granier, S. et al. Structure of the δ-opioid receptor bound to naltrindole. Nature 485, 400–404 (2012).
36. White, J. F. et al. Structure of the agonist-bound neurotensin receptor. Nature 490, 508–513 (2012).
37. Zhang, C. et al. High-resolution crystal structure of human protease-activated receptor 1. Nature 492, 387–392 (2012).

38. Hanson, M. A et al. Crystal structure of a lipid G protein-coupled receptor. Science 335, 851–855 (2012).
39. Srivastava, A. et al. High-resolution structure of the human GPR40 receptor bound to allosteric agonist TAK-875. Nature 513, 124–127 (2014).
40. Wang, C. et al. Structure of the human smoothened receptor bound to an antitumour agent. Nature 497, 338–343 (2013).
41. Hollenstein, K. et al. Structure of class B GPCR corticotropin-releasing factor receptor 1. Nature 499, 438–443 (2013).
42. Siu, F. Y. et al. Structure of the human glucagon class B G-protein-coupled receptor. Nature 499, 444–449 (2013).
43. Wu, H. et al. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. Science 344, 58–64 (2014).
44. Doré, A. S. et al. Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. Nature 511, 557–562 (2014).
45. Geng, Y., Bush, M., Mosyak, L., Wang, F. & Fan, Q. R. Structural mechanism of ligand activation in human GABA(B) receptor. Nature 504, 254–259 (2013).
46. Foord, S. M. et al. International union of pharmacology. XLVI . G protein-coupled receptor list. Pharmacol. Rev. 57, 279–288 (2005).
47. Fredholm, B. B., Johansson, S. & Wang, Y.-Q. Adenosine and the regulation of metabolism and body temperature. Adv. Pharmacol. 61, 77–94 (2011).
48. Bennett, K. a et al. Pharmacology and structure of isolated conformations of the adenosine A_{2A} receptor define ligand efficacy. Mol. Pharmacol. 83, 949–958 (2013).
49. Awad, A. S. et al. Adenosine A_{2A} receptor activation attenuates inflammation and injury in diabetic nephropathy. Am. J. Physiol. Renal. Physiol. 290, 828–837 (2006).
50. Desai, A. et al. Adenosine A_{2A} receptor stimulation increases angiogenesis by down-regulating production of the antiangiogenic matrix protein thrombospondin 1. Mol. Pharmacol. 67, 1406–1413 (2005).
51. Udelson, J. E. et al. Randomized, controlled dose-ranging study of the selective adenosine A_{2A} receptor agonist binodenoson for pharmacological stress as an adjunct to myocardial perfusion imaging. Circulation 109, 457–464 (2004).
52. Müller, C. E. & Jacobson, K. a. Recent developments in adenosine receptor ligands and their potential as novel drugs. BBA-Biomembranes 1808, 1290–1308 (2011).
53. Al Jaroudi, W. & Iskandrian, A. E. Regadenoson: a new myocardial stress agent. J. Am. Coll. Cardiol. 54, 1123–1130 (2009).
54. Müller, C. E. & Jacobson, K. A. Xanthines as adenosine receptor antagonists. In Methylxanthines. Handb. Exp. Pharmacol. B.B. Fredholm, ed., Springer 200, 151–199 (2011).
55. Gomes, C. V, Kaster, M. P., Tomé, A. R., Agostinho, P. M. & Cunha, R. A. Adenosine receptors and brain diseases: neuroprotection and neurodegeneration. Biochim. Biophys. Acta 1808, 1380–1399 (2011).
56. Andrews, S. P. & Tehan, B. Stabilised G protein-coupled receptors in structure-based drug design: a case study with adenosine A2A receptor. Med. Chem. Commun 4, 52–67 (2013).
57. Müller, C. E. & Ferré, S. Blocking striatal adenosine A_{2A} receptors: a new strategy for basal ganglia disorders. Recent Pat. CNS Drug Discov. 2, 1–21 (2007).
58. Clementina, M. & Giuseppe, S. A_{2A} receptor ligands: past, present and future trends. Curr. Top. Med. Chem. 10, 902–922 (2010).
59. Overington, J., Al-Lazikani, B. and Hopkins, A. L. How many drug targets are there? Nat. Rev. Drug Discov. 5, 993–996 (2006).
60. Congreve, M., Langmead, C. J., Mason, J. S. & Marshall, F. H. Progress in structure based drug design for G protein-coupled receptors. J. Med. Chem. 54, 4283–4311 (2011).
61. Lagerström, M. C. & Schiöth, H. B. Structural diversity of G protein-coupled receptors and significance for drug discovery. Nat. Rev. Drug Discov. 7, 339–357 (2008).
62. Congreve, M. et al. Fragment screening of stabilized G protein-coupled receptors using biophysical methods. Method. Enzym. 493, 115–136 (2011).
63. Chen, D., Errey, J. C., Heitman, L. H., Marshall, F. H., IJzerman, A. P., and Siegal, G. Fragment screening of GPCRs using biophysical methods: Identification of ligands of the adenosine A_{2A} receptor with novel biological activity. ACS Chem. Biol. 7, 2064–2073 (2012).
64. Zhukov, A. et al. Biophysical mapping of the adenosine A_{2A} receptor. J. Med. Chem. 54, 4312–4323 (2011).
65. Andrews, S. P., Brown, G. A & Christopher, J. A. Structure-based and fragment-based GPCR drug discovery. ChemMedChem 9, 256–75 (2014).
66. Hertzberg, R. P. & Pope, A. J. High-throughput screening: new technology for the 21st century. Curr. Opin. Chem. Biol. 4, 445–451 (2000).
67. Lipinski, C. a, Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug. Deliv. Rev. 46, 3–26 (2012).
68. Mestres, J. & Veeneman, G. H. Identification of “Latent Hits” in compound screening collection. J. Med. Chem. 46, 3441–3444 (2003).
69. Gribbon, P. & Sewing, A. High-throughput drug discovery: what can we expect from HTS? Drug Discov. Today 10, 17–22 (2005).
70. Hann, M. M., Leach, a R. & Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. J. Chem. Inf. Comput. Sci. 41, 856–864 (2001).

71. Oprea, T. I., Davis, a M., Teague, S. J. & Leeson, P. D. Is there a difference between leads and drugs? A historical perspective. J. Chem. Inf. Comput. Sci. 41, 1308–1315 (2001).
72. Teague, S., Davis, A., Leeson, P. & Oprea, T. The design of leadlike combinatorial lLibraries. Angew. Chem. Int. Ed. Engl. 38, 3743–3748 (1999).
73. Jencks, W. P. On the attribution and additivity of binding energies. Proc. Natl. Acad. Sci. U. S. A. 78, 4046–4050 (1981).
74. Congreve, M., Carr, R., Murray, C. & Jhoti, H. A " Rule of Three " for fragment-based lead discovery? Drug Discov. Today 8, 876–877 (2003).
75. Bohacek, R. S., McMartin, C., and Guida, W. C. The art and practice of structure-based drug design: A molecular modeling perspective. Med. Res. Rev. 16, 3–50 (1996).
76. Ertl, P. Cheminformatics analysis of organic substituents: Identification of the most common substituents, calculation of substituent properties, and automatic identification of drug-like bioisosteric groups. J. Chem. Inf. Comput. Sci. 43, 374–380 (2003).
77. Fink, T., and remond, J. L. No Virtual exploration of the chemical universe up to 11 atoms of C, N, O, F: Assembly of 26.4 million structures (110.9 million stereoisomers) and analysis for new ring systems, stereochemistry, physicochemical properties, compound classes, and drug disc. J. Chem. Inf. Model. 47, 342–353 (2007).
78. Murray, C. W. & Rees, D. C. The rise of fragment-based drug discovery. Nat. Chem. 1, 187–192 (2009).
79. Carr, R. A. E., Congreve, M., Murray, C. W. & Rees, D. C. Fragment-based lead discovery: leads by design. Drug Discov. Today 10, 987–992 (2005).
80. Card, G. L. et al. A family of phosphodiesterase inhibitors discovered by cocrystallography and scaffold-based drug design. Nat. Biotechnol. 23, 201–207 (2005).
81. Hajduk, P. J. et al. Discovery of potent nonpeptide inhibitors of stromelysin using SAR by NMR. J. Am. Chem. Soc. 119, 5818–5827 (1997).
82. Murray, C. W. & Blundell, T. L. Structural biology in fragment-based drug design. Curr. Opin. Struct. Biol. 20, 497–507 (2010).
83. Congreve, M., Murray, C. W. & Blundell, T. L. Structural biology and drug discovery. Drug Discov. Today 10, 895–907 (2005).
84. Coyne, A. G., Scott, D. E. & Abell, C. Drugging challenging targets using fragment-based approaches. Curr. Opin. Chem. Biol. 14, 299–307 (2010).
85. Wang, Y.-S. et al. Application of fragment-based NMR screening, X-ray crystallography, structure-based design, and focused chemical library design to identify novel microM leads for the development of nM BACE-1 (beta-site APP cleaving enzyme 1) inhibitors. J. Med. Chem. 53, 942–950 (2010).
86. May, P. C. et al. Robust central reduction of amyloid-β in humans with an orally available, non-peptidic β-secretase inhibitor. J. Neurosci. 31, 16507–16516 (2011).
87. Abdel-Rahman N, Martinez-Arias A, B. T. Probing the druggability of protein-protein interactions: targeting the Notch1 receptor ankyrin domain using a fragment-based approach. Biochem. Soc. Trans. 39, 1327–1333 (2011).
88. Basse, N. et al. Toward the rational design of p53-stabilizing drugs: probing the surface of the oncogenic Y220C mutant. Chem. Biol. 17, 46–56 (2010).
89. Chen, L., Cressina, E., Leeper, F. J., Smith, A. G. & Abell, C. A fragment-based approach to identifying ligands for riboswitches. ACS Chem. Biol. 5, 355–358 (2010).
90. Yang, H. et al. RG7204 (PLX4032), a selective BRAFV600E inhibitor, displays potent antitumor activity in preclinical melanoma models. Cancer Res. 70, 5518–5527 (2010).
91. Bollag, G. et al. Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. Nature 467, 596–599 (2010).
92. Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. Science 274, 1531–1534 (1996).
93. Hajduk, P. J. & Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. Nat. Rev. Drug Discov. 6, 211–219 (2007).
94. Wada, C. K. The evolution of the matrix metalloproteinase inhibitor drug discovery program at abbott laboratories. Curr. Top. Med. Chem. 4, 1255–1267 (2004).
95. Park, C.-M. et al. Discovery of an orally bioavailable small molecule inhibitor of prosurvival B-cell lymphoma 2 proteins. J. Med. Chem. 51, 6902–6915 (2008).
96. Mayer, M. and Meyer, B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. Angew. Chem. Int. Ed. 38, 1784–1788 (1999).
97. Dalvit, C. et al. Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. J. Biomol. NMR 18, 65–68 (2000).
98. Becattini, B. and Pellecchia, M. SAR by ILOEs: An NMR-based approach to reverse chemical genetics. Chemistry 12, 2658–2662 (2006).
99. Sledz, P., Silvestre, H. L., Hung, A. W., Ciulli, A., Blundell, T. L., and Abell, C. Optimization of the Interligand Overhauser Effect for Fragment Linking: Application to Inhibitor Discovery against Mycobacterium tuberculosis Pantothenate Synthetase. J. Am. Chem. Soc. 132, 4544–4545 (2010).
100. Vanwetswinkel, S. et al. TINS, target immobilized NMR screening: an efficient and sensitive method for ligand discovery. Chem. Biol. 12, 207–216 (2005).
101. Marquardsen, T. et al. Development of a dual cell, flow-injection sample holder, and NMR probe for comparative ligand-binding studies. J. Magn. Reson. 182, 55–65 (2006).

102. Früh, V. et al. Application of fragment-based drug discovery to membrane proteins: identification of ligands of the integral membrane enzyme DsbB. Chem. Biol. 17, 881–891 (2010).
103. Bartoli, S., Fincham, C. I. & Fattori, D. The fragment-approach: An update. Drug Discov. Today Technol. 3, 425–431 (2006).
104. Perspicace, S. et al. Fragment-based screening using surface plasmon resonance technology. J. Biomol. Screen. 14, 337–349 (2009).
105. Navratilova, I. & Hopkins, A. L. Fragment screening by surface plasmon resonance. ACS Med. Chem. Lett. 1, 44–48 (2010).
106. Neumann, T., Junker, H.-D., Schmidt, K. & Sekul, R. SPR-based fragment screening: advantages and applications. Curr. Top. Med. Chem. 7, 1630–1642 (2007).
107. Rich, R. L., Errey, J., Marshall, F. & Myszka, D. G. Biacore analysis with stabilized G-protein-coupled receptors. Anal. Biochem. 409, 267–272 (2011).
108. Christopher, J. a et al. Biophysical fragment screening of the β1-adrenergic receptor: Identification of high affinity arylpiperazine leads using structure-based drug design. J. Med. Chem. 56, 3446–3455 (2013).
109. http://www.astex-therapeutics.com.
110. Davies, T. G. & Tickle, I. J. Fragment screening using X-ray crystallography. Top Curr. Chem. 317, 33–59 (2012).
111. Hennig, M., Ruf, A. & Huber, W. Combining biophysical screening and X-ray crystallography for fragment-based drug discovery. Top Curr. Chem. 317, 115–143 (2012).
112. Davis, A. M., St-Gallay, S. A & Kleywegt, G. J. Limitations and lessons in the use of X-ray structural information in drug design. Drug Discov. Today 13, 831–841 (2008).
113. Zhu, Z. et al. Discovery of cyclic acylguanidines as highly potent and selective beta-site amyloid cleaving enzyme (BACE) inhibitors: Part I--inhibitor design and validation. J. Med. Chem. 53, 951–965 (2010).
114. Erlanson, D. A., Wells, J. A. & Braisted, A. C. Tethering: fragment-based drug discovery. Annu. Rev. Biophys. Biomol. Struct. 33, 199–223 (2004).
115. Kollareddy, M. et al. Aurora kinase inhibitors: progress towards the clinic. Invest. New Drugs 30, 2411–2432 (2012).
116. Hannah, V., Atmanene, C., Zeyer, D., Van Dorsselaer, A., Sanglier-Cianférani, S. Native MS: an ESI, way to support structure- and fragment-based drug discovery. Futur. Med. Chem. 2, 35–50 (2010).
117. Robertson, N. et al. The properties of thermostabilised G protein-coupled receptors (StaRs) and their use in drug discovery. Neuropharmacology 60, 36–44 (2011).
118. Warne, T. et al. The structural basis for agonist and partial agonist action on a β(1)-adrenergic receptor. Nature 469, 241–244 (2011).
119. Lebon, G. et al. Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. Nature 474, 521–525 (2011).
120. Doré, A. S. et al. Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthines XAC and caffeine. Structure 19, 1283–1293 (2011).
121. Chun, E. et al. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. Structure 20, 967–976 (2012).
122. Hino, T. et al. G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody. Nature 482, 237–240 (2012).
123. Rasmussen, S. G. F. et al. Structure of a nanobody-stabilized active state of the β(2) adrenoceptor. Nature 469, 175–180 (2011).
124. Bowie, J. U. Stabilizing membrane proteins. Curr. Opin. Struct. Biol. 11, 397–402 (2001).
125. Gohon, Y. & Popot, J. Membrane protein – surfactant complexes. Curr. Opin. Colloid Interface Sci. 8, 15–22 (2003).
126. Popot, J.-L. Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. Annu. Rev. Biochem. 79, 737–775 (2010).
127. Popot, J.-L. et al. Amphipols from A to Z. Annu. Rev. Biophys. 40, 379–408 (2011).
128. Popot, J.-L. et al. Amphipols: polymeric surfactants for membrane biology research. Cell. Mol. Life Sci. 60, 1559–1574 (2003).
129. Martinez, K. L. et al. Allosteric transitions of Torpedo acetylcholine receptor in lipids, detergent and amphipols: molecular interactions vs. physical constraints. FEBS lett. 528, 251–256 (2002).
130. Pocanschi, C. L. et al. Amphipathic polymers: tools to fold integral membrane proteins to their active form. Biochemistry 45, 13954–13961 (2006).
131. Banères, J.-L., Popot, J.-L. & Mouillac, B. New advances in production and functional folding of G-protein-coupled receptors. Trends in Biotechnol. 29, 314–322 (2011).
132. Bayburt, T. H. & Sligar, S. G. Membrane protein assembly into Nanodiscs. FEBS lett. 584, 1721–1727 (2010).
133. Leitz, A., Bayburt, T., Barnakov, A., Springer, B. & Sligar, S. Functional reconstitution of β2-adrenergic receptors utilizing self-assembling Nanodisc technology. Biotechniques 40, 601–612 (2006).
134. Poget, S. F. & Girvin, M. E. Solution NMR of membrane proteins in bilayer mimics: small is beautiful, but sometimes bigger is better. BBA-Biomembranes 1768, 3098–3106 (2007).
135. De Angelis, A.A., Opella, S. J. Bicelle samples for solid-state NMR of membrane proteins. Nat. Protoc. 2, 2332–2338 (2007).
136. Faham, S. & Bowie, J. U. Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. J. Mol. Biol. 316, 1–6 (2002).
137. Kiefer, H. et al. Expression of an olfactory receptor in Escherichia coli: purification, reconstitution, and ligand binding. Biochemistry 35, 16077–16084 (1996).

138. Baneres, J.-L. et al. Structure-based analysis of GPCR function: Conformational adaptation of both agonist and receptor upon leukotriene B4 binding to recombinant BLT1. J. Mol. Biol. 329, 801–814 (2003).
139. Arcemisbéhère, L. et al. Leukotriene BLT2 receptor monomers activate the G(i2) GTP-binding protein more efficiently than dimers. J. Biol. Chem. 285, 6337–6347 (2010).
140. Banères, J.-L. et al. Molecular characterization of a purified 5-HT4 receptor: a structural basis for drug efficacy. J. Biol. Chem. 280, 20253–20260 (2005).
141. Dahmane, T., Damian, M., Mary, S., Popot, J.-L. & Banères, J.-L. Amphipol-assisted in vitro folding of G protein-coupled receptors. Biochemistry 48, 6516–6521 (2009).
142. Catoire, L. J. et al. Solution NMR mapping of water-accessible residues in the transmembrane beta-barrel of OmpX. Eur. Biophys. J. 39, 623–630 (2010).
143. Raschle, T., Hiller, S., Etzkorn, M. & Wagner, G. Nonmicellar systems for solution NMR spectroscopy of membrane proteins. Curr. Opin. Struct. Biol. 20, 471–479 (2010).
144. Catoire, L. J. et al. Structure of a GPCR ligand in its receptor-bound state: leukotriene B4 adopts a highly constrained conformation when associated to human BLT2. J. Am. Chem. Soc. 132, 9049–9057 (2010).

Chapter 2

Application of Fragment-Based Drug Discovery to Membrane Proteins: Identification of Ligands of the Integral Membrane Enzyme DsbB

This work has been published as part of:

V. Früh, Y. Zhou, D. Chen, C. Loch, E. Ab, Y. N. Grinkova, H. Verheij, S. G. Sligar, J. H. Bushweller, G. Siegal. Application of fragment-based drug discovery to membrane proteins: identification of ligands of the integral membrane enzyme DsbB. *Chem. Biol.* 2010, 17, 881−891.

ABSTRACT

Membrane proteins are important pharmaceutical targets, but they pose significant challenges for fragment-based drug discovery approaches. Here, we present the first successful use of biophysical methods to screen for fragment ligands to an integral membrane protein. The *Escherichia coli* inner membrane protein DsbB was solubilized in detergent micelles and lipid bilayer nanodiscs. The solubilized protein was immobilized with retention of functionality and used to screen 1071 drug fragments for binding using target immobilized NMR Screening. Biochemical and biophysical validation of the eight most potent hits revealed an IC_{50} range of 7–200 μM. The ability to insert a broad array of membrane proteins into nanodiscs, combined with the efficiency of TINS, demonstrates the feasibility of finding fragments targeting membrane proteins.

1. Introduction

With 60% of currently marketed drugs targeting membrane proteins[1], it is clear that finding small molecules to modulate the function of such proteins is essential. High throughput screening (HTS) methods have been successful in identifying such compounds, but because the methods of detection rely on functional assays, they are generally only sensitive to submicromolar interactions. Such relatively tight interactions are generally only observed for larger compounds (300-500 Da). However, it has proved challenging to simultaneously optimize potency and absorption, distribution, metabolism, and excretion (ADME) properties of these "lead-like" or "drug-like" compounds. Furthermore, such large compounds inefficiently explore the binding sites of proteins[2]. Fragment-based drug discovery (FBDD) has become a powerful complementary approach to HTS for generating novel chemical modulators of pharmaceutical targets. FBDD screens small libraries (1,000-20,000 compounds) of so-called drug "fragments" that are often described by a "rule of threes"[3] (Ro3, M_r < 300 Da, cLogP < 3, H-bond donors < 3, H-bond acceptors < 3, number of rotatable bonds < 3 and TPSA (total polar surface area) < 60 $Å^2$) for binding to the target. Ro3 compliant compounds typically bind the target with K_D greater than 10 μM. In order to detect such weak binding, sensitive biophysical techniques are typically required, particularly when the target s not an enzyme. Commonly used techniques for detecting fragment binding include NMR, X-ray crystallography and surface plasmon resonance (SPR)[4].

Although biophysical methods have been successfully applied to an array of soluble protein targets[5], they have failed in one way or another when applied to membrane proteins. There are two primary reasons for this failure: insufficient quantity of the target and problems related to the solubilization media. Many biophysical methods require tens or even hundreds of mg of purified, functional protein and most membrane proteins are difficult to produce in these quantities. However, recent advances have enabled the production of low mg quantities of a variety of MPs[6-8]. Membrane proteins that can be produced in sufficient quantity

must then be solubilized in a surfactant while maintaining their functional state, which is also often challenging. Finally, non-specific partitioning of fragments into the surfactant has been a severe problem leading to high levels of false positives.

The use of detergent micelles to solubilise MPs has only met limited success in retaining the native function of the protein while at the same time the micelles often interfere with biophysical assays. A possible solution to this bottleneck would be to employ non-detergent media to functionally solubilize MPs. The Nanodisc (ND) has been developed as an alternative, surfactant free approach to solubilize MPs. NDs consist of a lipid bilayer that is surrounded by an amphiphilic α-helical membrane scaffold protein (MSP). A variety of proteins have been functionally solubilized in NDs[9-11], which are much better mimics of the native membrane than detergent micelles. However, the suitability of NDs for biophysical assays of ligand binding to MPs has yet to be determined.

An NMR-based fragment screening approach has been developed and has proven capable of overcoming many of the challenges posed by membrane proteins[12]. The approach, called Target Immobilized NMR Screening (TINS)[13], involves immobilizing a target and a reference in two compartments of a dual-cell sample holder[14] and simultaneously injecting mixtures of fragments in an automated process. For each mixture a 1D ^{1}H NMR spectrum is recorded while fragment binding to the target protein results in a decrease in peak amplitude. The reference, which is selected for minimal specific small molecule binding, serves to cancel out non-specific binding of fragments to protein surfaces. Hits can therefore be detected by comparing spectra of the compounds recorded in the presence of the target to those recorded in the presence of the reference. By repeatedly using the same sample to screen the entire fragment collection (>1,000 compounds), typically only ~25 nmol of protein is required, thus bringing many MPs within the requirements for TINS. Furthermore, the reference system is expected to account for non-specific binding of fragments to the media in which the membrane protein is solubilized.

It was sought to apply TINS to a *bona fide*, integral membrane pharmaceutical target that could be functionally solubilized in detergent micelles

and NDs. The inner membrane protein of *E. coli* Disulphide bond forming protein B (DsbB), and its homologs in other gram-negative bacteria, is an oxidoreductase that is essential for protein disulfide bond formation in the periplasm[15]. Periplasmic DsbA functions as the catalyst for protein disulfide bond formation and is reoxidized by DsbB with concomitant reduction of bound ubiquinone or menaquinone. Since many bacterial virulence factors are secreted proteins that require disulfide bonds for proper folding and function, the DsbA/DsbB system is a potential antimicrobial drug target[16-18]. DsbB is an ideal candidate to test the TINS methodology since it can be produced in large quantities and solubilized in detergent micelles where it retains a robust enzymatic activity which is easily assayed. In addition, a wealth of biochemical data is available that describes the enzymatic activity of the wildtype as well as numerous relevant mutants[15,19-21]. Finally, the 3D structures of wildtype DsbB bound to its redox partner DsbA[22] and of a mutant representing an enzymatic intermediate are available[23]. Selection of an appropriate reference system is critical to insure the robust performance of TINS. Our previous experience using the *E. coli* Outer membrane protein A (OmpA) transmembrane domain, which has native structure under the same detergent micelle conditions as DsbB suggested that it had minimal small molecule binding[24] and would therefore serve as a good reference.

In this chapter, the joint work between the author and previous PhD candidate Dr. Virginie Früh is presented. The first complete screen of a fragment library against an integral membrane protein is reported. The applicability of TINS for fragment screening using both micelle and ND solubilized protein has been tested. The initial work of fragment screening using micelle solubilized DsbB has been performed by Dr. Früh, including validation and hit characterization with respect to the mode of action using an enzyme inhibition assay and investigation of the binding mode of two classes of inhibitors by analysis of chemical shift perturbations in ^{15}N-labeled mutant DsbB. These results are included in the appendix. The author's work is focused on characterization of small molecule - ND solubilized protein interactions and the results are compared with those for micelle solubilized DsbB.

2. Results

DsbB Functional Immobilization and Enzymatic Activity

Wildtype DsbB (containing endogenous quinone) has previously been solubilized in DPC micelles, which we refer to as DsbB/DPC, with retention of enzymatic function[23]. Next DPC solubilized DsbB and OmpA were trapped in NDs. Gel filtration analysis of our preparations revealed Stokes diameters of 9.63, 9.68, and 9.52 nm for empty NDs (-/ND), NDs with embedded DsbB (DsbB/ND), and NDs with embedded OmpA (OmpA/ND), in accordance with literature values[25]. The DsbB/ND was immobilized on a Sepharose resin *via* a Schiff's base intermediate. At the pH selected (7.4), this reaction is relatively specific for the free N-terminus. A final concentration of approximately 100 μM DsbB/ND (nmol protein per ml settled bed volume) was achieved with an overall yield of 75%. Non-immobilized and immobilized DsbB/ND were assayed for enzymatic activity for comparison to DsbB/DPC. As shown in Table 1, both preparations had a k_{cat} that was somewhat greater than the micelle solubilized protein, indicating that they remained completely functional. The increased k_{cat} for DsbB in NDs could possibly result from a more native functionality of the enzyme in the lipid bilayer environment of the ND.

Table 1. Enzymatic activity of solubilized vs immobilized DsbB.

Enzyme	k_{cat} (M UQ1/M DsbB-min^{-1})
DsbB wt/DPC	267 ± 14
DsbB wt/DPC (immobilized)	238 ± 27
DsbB wt/ND	346 ± 13
DsbB wt/ND (immobilized)	329 ± 26

Stability of the Immobilized Protein to Repeated Sample Application Cycles

In a method such as TINS where a single sample of the target is used to screen an entire compound collection, the integrity of the immobilized protein is clearly critical. Soluble proteins are routinely stable over more than 200 cycles of

sample application and washing[13]. Solubilized MPs however, include two components, the surfactant and the protein itself, both of which must remain stable in order to ensure proper ligand screening. Preliminary studies of DsbB/DPC and OmpA/DPC clearly demonstrated that repeated cycles of compound application and washing in the absence of added detergent resulted in rapid degradation of DsbB activity[12]. Therefore deuterated DPC was included at a minimum concentration of 5 mM (approximately 3 x critical micellar concentration) in the buffer used to wash the compounds from the sample holder. The library, which consisted of 1,071 fragments at the time, was then screened in mixtures that averaged approximately 5 compounds each. Including control experiments designed to monitor the physical integrity of the target and reference samples, approximately 200 sample application/washing cycles were performed. To monitor the integrity of the DsbB sample during the screen, the binding of synthetic UQ1 was observed (Figure 1a). In TINS, binding of a fragment is best described by the T/R (target/reference) ratio, defined as the average ratio of the amplitude of peaks in the presence of the target, DsbB, to that in the presence of the reference, OmpA. It is clear from Figure 1a that binding of UQ1 to DsbB/DPC remained relatively constant throughout the screen which required 5.5 days to complete.

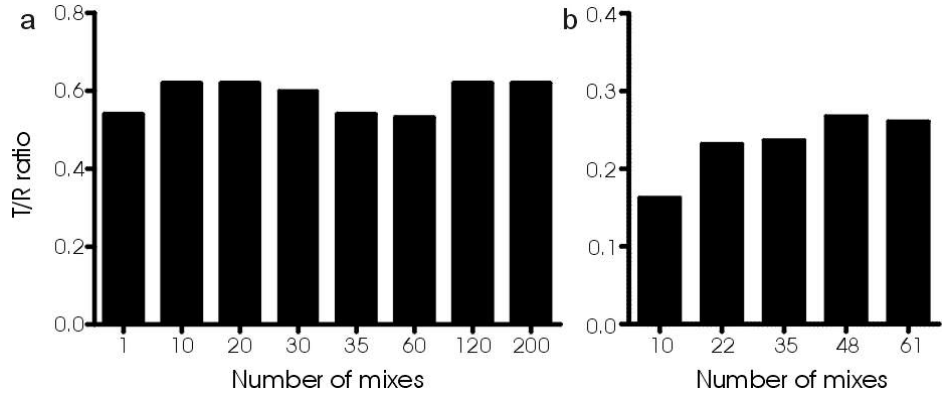


Figure 1. Stability of the DsbB in Micelles and NDs. The stability of DPC solubilized (a) and ND solubilized (b) DsbB after multiple cycles of compound application and washing was assessed by binding of a known ligand. Binding is displayed as the average ratio of peak heights for the compound in the presence of DsbB over that in the presence of the reference (T/R ratio). The reference in (a) was DPC solubilized OmpA and in (b) -/ND.

Since it was not practical to rescreen the entire fragment collection multiple times, we selected a subset of 20 compound mixtures containing positive hits from the DsbB/DPC screen and 20 mixtures containing no positive hits from the DsbB/DPC screen (a total of 183 compounds) to assess the suitability of the ND system for ligand screening. The 40 mixes, along with the control experiments used for DsbB/DPC, were applied sequentially to the immobilized DsbB/ND using -/ND as the reference. We wanted to assess whether DsbB/ND and -/ND were stable in the absence of added phospholipid. As with DsbB/DPC, we monitored the integrity of DsbB/ND, as determined by binding of a known ligand (UQ1), during multiple cycles of compound application and washing in lipid free buffers. The T/R ratio for ligand binding to DsbB/ND vs -/ND is shown in Figure 1b**.** This data suggests a possible, initial small degradation in binding behaviour (although the variation is similar to that seen in Figure 1a), after which the ligand binding capacity of DsbB/ND remained constant. The constant T/R ratio during cycles 22 through 61 suggests that both DsbB/ND and -/ND remained intact.

Comparison of Micelle Solubilized vs ND Solubilized Protein for Ligand Binding Studies

The influence of detergent or ND on the quality of the NMR spectra of the fragments is shown in Figure 2(d and e). In both cases the compound whose spectrum is shown in 3c can be identified as specifically binding to DsbB. However, the signal-to-noise ratio of the aromatic resonances of the compounds (Figure 2a and b) in 2e is nearly double that in 2d**.** The improved quality of the spectra allows more reliable analysis of the peaks at 7.3 and 7.4 ppm, which are now clearly seen to indicate specific binding of this compound to DsbB/ND, consistent with behaviour of the peaks at 2.8 and 2.2 ppm. The reduced signal in the presence of detergent solubilized protein is likely due to non-specific partioning of 30-40% of the compounds into the micelle.

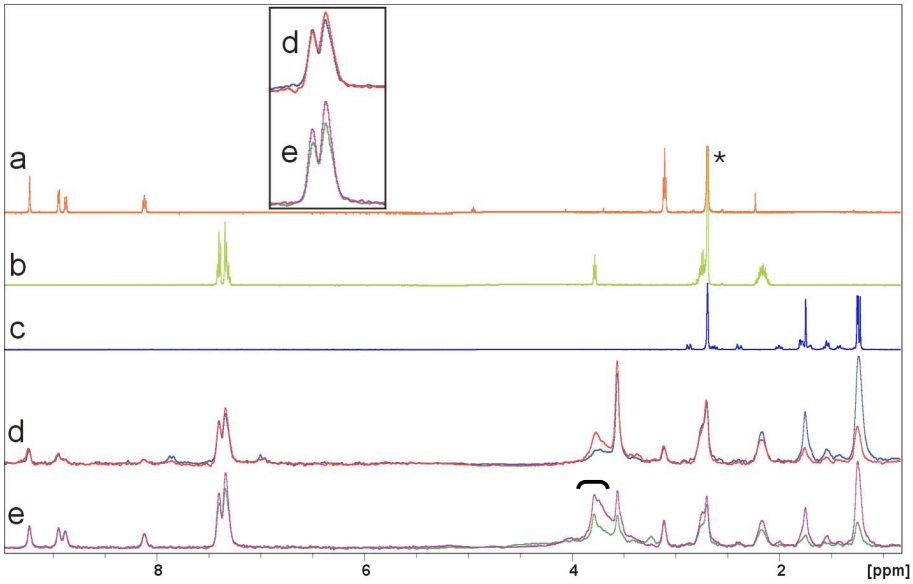


Figure 2. Detection of ligand binding to immobilized DsbB using TINS. The 1D ^{1}H NMR spectrum of 3 different fragments in solution (**a - c**) is shown for reference. The ^{1}H NMR spectrum of a mix of the 3 fragments in the presence of DsbB/DPC (red spectrum) or OmpA/DPC (blue spectrum) that have been immobilized on the sepharose support is shown in d. The spectra of the same mix recorded in the presence of DsbB/ND (green) or -/ND (magenta) is shown in e. The asterisk indicates the resonance from residual ^{1}H DMSO and the bracket shows residual sugar ^{1}H resonances from the sepharose media. The residual H_2O resonance at 4.7 ppm has been filtered out.

The stability of the empty ND (-/ND) as shown in Figure 1b, affords the possibility to use NDs directly as a generic reference to account for non-specific ligand binding to the phospholipid bilayer and the scaffolding protein. To investigate this, we screened all 183 compounds for binding to DsbB/ND using either OmpA/ND or -/ND as a reference. By plotting the T/R for each compound from the screen using -/ND versus that using OmpA/ND we derive a two-dimensional plot that gives an overview of the performance of the screen (Figure 3a). Overall there was a reasonable correlation in ligand binding to DsbB/ND using either empty NDs or OmpA/ND as the reference (R^2 =0.78). In general however, the T/R ratio of fragments is lower with -/ND as a reference, indicating that specific binding to DsbB/ND is more pronounced. Since the NMR spectra of the fragments in the presence of DsbB/ND in the screen vs -/ND or OmpA/ND

are similar, this suggests a higher level of non-specific binding of the fragments to OmpA/ND. We conclude therefore that -/ND is the preferred reference.

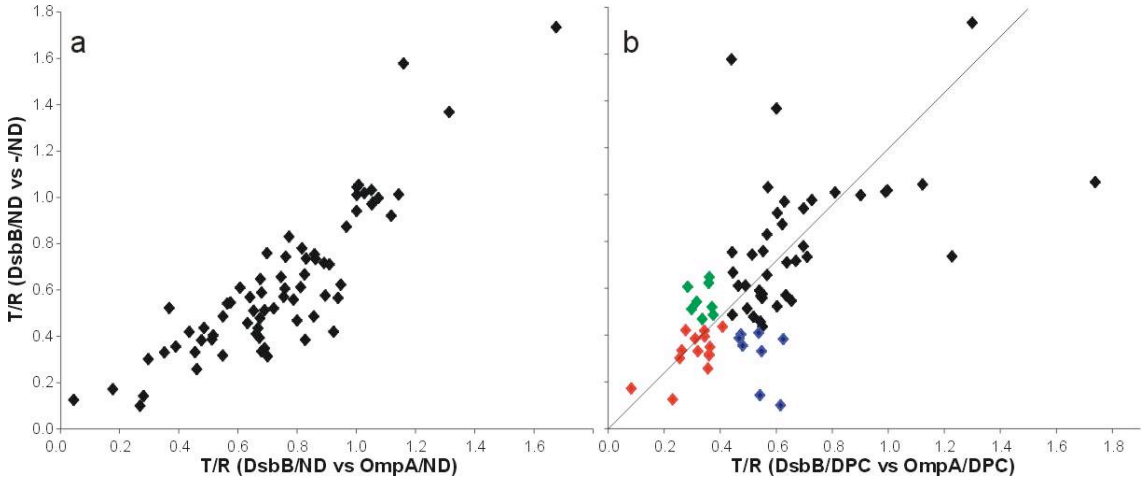


Figure 3. Comparison of TINS screening in micelles vs NDs. A total of 70 fragments were assayed for binding to DsbB solubilized in either detergent micelle or ND. a) The 70 fragments were screened for binding to DsbB/ND using either empty ND (-/ND) or OmpA/ND as a reference. The T/R (see text) for each compound is plotted for one screen versus the other. R2 = 0.78. b) The T/R for each compound in the DsbB/ND vs -/ND screen is plotted against the value from the DsbB/DPC vs OmpA/DPC screen. Hits common to both screens are show in red. Hits found only in the ND screen are shown in blue while those found only in the DPC screen are in green.

Table 2. Fragment hits from the screen of DsbB in micelles and NDs.

	Hits	cLogP	BioAssay in ND	BioAssay in DPC
Micelle	8	1.34	-	+
ND	7	2.21	++	++
Both	14	2.13	++	++

Note: - poor correlation between enzyme inhibition and binding assay, + reasonable correlation (approximately 50% hits bioactive), ++ good correlation (80-90% bioactive)

We then compared the ligand screening results from DsbB/DPC (OmpA/DPC as reference) to those from DsbB/ND (-/ND as reference). Upon inspection of the raw NMR data from the DPC screen we observed that although 183 compounds were present in the 40 mixes selected, only 127, about 2/3, gave observable NMR spectra. Presumably, those compounds missing from the NMR data had non-specifically adsorbed to the micelle. In contrast, 164 of 183 compounds gave observable spectra in the ND screen. Of the 127 compounds

with observable spectra in the DPC screen, 70 were of sufficiently high quality to allow a reliable comparison with the ND screen and we therefore focused our efforts on these. Inspection of Figure 3b clearly shows that the correlation between the micelles and NDs is much less pronounced than between the two ND references. Using the same criteria for hit selection for both, 22 hits were identified for DsbB/ND and 22 were identified for DsbB/DPC. Of these biophysically detected hits, 14 were common to both the micelle and ND (red) screen while 7 were unique to the ND screen (blue) and 8 unique to the micelle screen (green, see also Table 2). We analyzed the solubility of each of the fragment hits using the calculated Log of the octonal/water partition coefficient (cLogP). Interestingly, the hits specific for the ND screen are on average slightly less soluble in water than the hits found in both screens, but the hits specific to the micelle screen are considerably more soluble (Table 2). A possible explanation for this observation is that the less soluble fragments exhibited greater non-specific binding to the micelle, thus masking specific binding to DsbB. This observation is consistent with the NMR data in Figure 2.

Hit Validation using Enzymatic Assays

The TINS assay simply identifies compounds that bind to DsbB, but not necessarily in a biologically relevant manner. Therefore we felt it was critical to validate the hits in terms of biological activity. An enzymatic assay was used to assess the ability of the compounds to inhibit electron transfer mediated by DsbB. Each of the 93 fragments identified as TINS hits in the micelle screen was assayed for inhibition of DsbB-dependent reoxidation of DsbA at 250 μM, as shown in appendix. The bioassay was also used to compare hits selected in the micelle screen to those selected in the ND screen (see Table 2). As expected, fragments common to both the micelle and ND screens yielded a strong correlation with biological activity with 12/14 exhibiting medium (30-70%) or high (>70%) inhibition of DsbB in both micelles and NDs. We observed a good correlation between ligands detected in the ND screen and biochemical activity

against both micelle and ND solubilized DsbB where 6/7 compounds had medium inhibitory activity and the seventh was a mild stimulator. In contrast, while the micelle specific ligands correlated reasonably well with the bioassay using detergent solubilized DsbB where 5/8 were medium or strong inhibitors, none inhibited DsbB/ND.

3. Discussion

The use of Ro3 compliant, “drug fragments” as a starting point for drug discovery has delivered a number of innovative compounds against soluble targets which are currently in clinical trials[26]. Membrane proteins have not made good targets for FBDD due to their challenging physicochemical properties. In particular, the difficulty of generating sufficient quantities of purified, functional protein and of detecting specific binding to the target, as opposed to non-specific partitioning into hydrophobic phases, have limited the applicability of biophysical ligand screening approaches. Here we have addressed these two issues by a) immobilizing the target and reusing a single sample to screen an entire fragment collection and b) using a reference sample to cancel out non-specific interaction of the fragments with the hydrophobic phase. Using TINS we have screened a collection of nearly 1,100 fragments with a single sample of less than 2 mg of protein and demonstrated that the protein was stable throughout the procedure. The stability of DsbB to repeated cycles of fragment application and washing depends on detergent micelles and the quinone cofactor. The detergent requirement could be overcome by including it in the buffer or using NDs to solubilize the protein. Endogenous UQ-8 binds DsbB very tightly and is quite resistant to repeated detergent washing[27].

Screening of the fragment library resulted in 93 ligands that were specific for DsbB. A number of observations suggest that most of these ligands are directly binding to DsbB and not indirectly *via* the micelle (see appendix). First, the DsbB binding detected using TINS was relative to OmpA solubilized in identical conditions. Second, there is a range of potencies in the enzyme

inhibition studies that includes a small number of non-inhibitors and activators. Third, and perhaps more critically, inhibition is saturable and occurs over 2 log orders, strongly suggesting a stoichiometric interaction. Fourth, titration of 8 different fragments into ^{15}N labelled DsbB resulted in chemical shift perturbations at well defined sites in both solvent exposed and micelle buried portions of the protein. In particular, the similarity of the chemical shift perturbations induced by the synthetic quinone UQ1 and fragment **2** indicate the compounds are binding to the same or overlapping sites. An additional, likely important, factor contributing to the low false positive rate is that the fragments that make up the collection are highly soluble, with each having been tested at 500 μM in an aqueous buffer alone and in a mixture. Nonetheless, an appreciable fraction of these fragments exhibit sufficient nonspecific binding to the micelle that they were only poorly or even not detected in the NMR spectra. These data suggest that ligand screening in the presence of micelle-solubilized membrane proteins may bias the chemical nature of the fragment library. The median cLogP for observable compounds in the presence of ND-solubilized proteins is slightly greater (0.9 versus 1.1) suggesting a reduced tendency of hydrophobic compounds to partition into the nanodisc.

As described in appendix, the eight fragments with greatest potency in the single concentration enzyme inhibition assay were fully characterized for potency, mode of action, and binding site on DsbB. A simplistic analysis suggests that these fragments can be divided into two groups, one that competes only with quinone for DsbB binding and a second that perturbs the apparent affinity of DsbB for both quinone and DsbA. The different mode of action is best exhibited by the differing effect on the apparent K_m for UQ1 or DsbA for each. Addition of fragment **2** reduced K_m for UQ1 more than 8-fold, while it had only a marginal effect on the K_m for DsbA (only 5% greater than experimental error). In contrast, addition of **8** reduced K_m for UQ1 more than 4.4-fold and K_m for DsbA more than 2-fold.

Functional solubilization of membrane proteins in detergent micelles is a challenging process that must be individually optimized for each protein. NDs

offer the potential to enable a more generic approach to handling membrane proteins since they can be used to functionally solubilise a variety of membrane proteins and can obviate the requirement for an intermediate micelle solubilized step[10,25]. Furthermore, immobilization can be made generic for all MPs with minimal effect on functionality by appropriate engineering of the MSP portion of the ND. Use of NDs as a generic solubilization/immobilization system for ligand screening is further enabled by the stability of the empty ND and the fact that it represents a high quality reference system to remove false positives. This conclusion is strongly supported by the observation that the 8 DsbB/DPC specific hits failed to inhibit DsbB/ND while 7 of 8 DsbB/ND hits also inhibited DsbB/DPC. Apparently, despite the reference sample, some compounds interact with DsbB in a micelle specific manner. This problem would be eliminated by using NDs.

4. Significance

Integral membrane proteins make up a significant portion of the human genome, carrying out numerous important normal as well as disease-related functions. More than 50% of drugs currently on the market target a membrane protein, demonstrating the utility of targeting this class of proteins. Approaches for the development of drugs targeting this class of proteins have focused principally on the use of high-throughput screening methods. Recently, FBDD has emerged as a powerful additional drug discovery approach. Because of the typically modest binding affinities of the small, functional group rich, i.e., high ligand efficiency, compounds utilized in FBDD, various biophysical techniques, including NMR, are typically used to detect binding. To date, several compounds developed using FBDD have advanced to clinical trials.[26] However, thus far FBDD has been demonstrated only for soluble proteins, not membrane proteins.

Herein, we describe the first complete screen of a fragment collection against an integral membrane protein. The screen was performed using a detergent micelle-solubilized protein using a simple and rapid ^{1}D NMR method we described previously (TINS). The 93 hits were subsequently validated in an

enzyme inhibition study. The use of a reference sample in the TINS experiment eliminated the well-documented problem of nonspecific binding of compounds to the detergent. As membrane protein activity is enhanced in more bilayer-like environments, we have also solubilized the protein in NDs and shown that the screening approach is effective with this preparation as well. The use of NDs further ameliorates issues with nonspecific binding and should also extend the method to proteins which typically do not behave well in detergents such as GPCRs.

Our results clearly establish the feasibility of using a fragment-based approach for finding starting matter for subsequent development of compounds targeting membrane proteins, including the all-important GPCR class of proteins. In addition, increasing success in the preparation of membrane proteins in reasonable quantities should make many such proteins amenable to the use of TINS for fragment screening, thereby increasing its general utility.

5. Methods

Protein Purification

DsbA, DsbB, and OmpA were expressed and purified as previously reported[28-30]. All proteins have a 6x-His tag at the N-terminus or C-terminus for affinity purification.

ND Self-Assembly

The ND self assembly procedure was repeated the same way for both OmpA and DsbB with slight adaptations from the previously reported procedures[25]. The reconstitution mixture contained Membrane Scaffold Protein MSP1D1(-) which lacked the His-tag, with mixed micelles of POPC and cholate at a ratio of 1:65:130. This reconstitution mixture was added to the OmpA or DsbB in detergent micelles (each with 10x the detergent CMC) in a volumetric ratio of 1:1 and incubated on ice for 4 hours. We always ensured a stoicheometry of MSP1D1(-) to OmpA or DsbB of 2:1. Upon addition of 0.7 g/ml of the

hydrophobic adsorbent Bio-Beads SM-2 (Biorad, Hercules, CA) and gently mixing for 4 hours at 4 ºC, the NDs underwent self-assembly, incorporating DsbB or OmpA in the lipid bilayer. This step was limiting, whereby detergent removal before 4 hours resulted in incomplete ND formation, but caused ND complex malformation if carried out for longer (i.e. 16h, data not shown). The His-tag of the embedded OmpA and DsbB was used to separate the empty non-tagged MSP1D1(-) complexes from the mixture by IMAC chromatography using Ni-NTA resin in a buffer containing 100 mM Tris, 300 mM NaCl, and imidazole at 0 mM, 10 mM, and 100 mM for loading, washing, and elution respectively. The eluted fractions were applied to a gel filtration column (Superdex 200 10/300 from GE Healthcare) in order to remove the remaining aggregated, non-embedded OmpA and DsbB, and to exchange the ND-embedded proteins into phosphate buffered saline (PBS, pH 7.4) for compatibility with the immobilization step required for TINS. A set of standard proteins was used to calibrate the Stokes' diameter *vs*. the retention time of the column.

Protein Immobilization

Actigel ALD resin (Sterogene, Carlsbad, CA, USA) was used as a 50% slurry and all experiments were carried out at 4 ºC when possible. The resin was washed with cold phosphate buffer (50 mM Na_2HPO_4, 50 mM KH_2HPO_4, 100 mM NaCl) at pH 7.5. 200 nmol of DPC solubilized DsbB was added to 1 ml bed volume of resin. The reductant sodium cyanoborohydride ($NaCNBH_3$) was added to a final concentration of 0.1 M. After an overnight incubation at 4 °C, residual unreacted aldehydes were blocked by addition of 50 mM Tris buffer and $NaCNBH_3$ for another 2 hours. The same procedure was repeated for DPC solubilized OmpA. Quantification of immobilized protein was monitored by absorption of the supernatant at 280 nm before and after immobilization, and by SDS-PAGE gel with a known standard curve and band volume analysis. These data indicated that a final concentration of 100 μM of both immobilized DsbB and OmpA was achieved, equating to a 50% yield. The procedure was repeated identically to immobilize DsbB and OmpA solubilized in ND (after pooling

fractions containing particles ranging between 9.2 to 9.7 nm) and empty ND at a similar final concentration as the micelle solubilized protein. ND preparations could not be quantified by UV absorption; therefore, they were loaded on SDS-page gels with a BSA standard curve for band volume quantification by Quantity One (Biorad), providing information on the concentration and ratio of ND molecules and incorporated proteins. The yield of immobilized, ND-solubilized protein was 75%.

DsbB Activity Assays

DsbB activity was quantified by measuring the capacity of the enzyme to reoxidize the protein DsbA or reduce its cofactor Ubiquinone-5, also called Coenzyme Q1 (UQ1) at pH 6.2. DsbA was reduced with 10 mM DTT for 10 minutes on ice. DTT was subsequently removed by gel filtration on a PD-10 column pre-equilibrated with degassed, distilled water containing 0.1 mM EDTA. DsbA fluorescence (excitation at 295 nm and emission at 330 nm) was measured in the presence of DsbB and UQ1 in 50 mM sodium phosphate, 100 mM NaCl, 0.1 % detergent [DPC or DDM depending on which was used to solubilize DsbB] and 0.1mM EDTA) at 30 ºC. Both UQ1 and DsbA were added at final concentrations of 30 μM. DsbB was added at a final concentration of 20 nM. The activity of DsbB in terms of moles ubiquinone reduced/moles DsbB min^{-1} could be calculated by using the initial slope of the fluorescence decrease upon DsbA oxidation, or by using the slope of absorption decrease at 275 nm upon reduction of UQ1[28].

To measure activity of immobilized DPC or ND solubilized DsbB, resin was aliquoted and diluted with degassed activity assay buffer to a final protein concentration of approximately 20 nM. For an appropriate baseline, an equivalent amount of resin without protein (blank resin) was prepared in the same manner. Quinone reduction was monitored in DPC samples after addition of 20 μM coenzyme Q1 and 20 μM DsbA and DsbA oxidation was measured for ND solubilized DsbB.

Target Immobilized NMR Screening

Immobilized, DPC solubilized DsbB or ND solubilized DsbB and OmpA were each packed into a separate cell of a dual-cell sample holder[14]. Mixes of the 1,071 fragments were made by 200-fold dilution of a 100 mM stock of each compound in d_6-DMSO such that the final DMSO concentration was never greater than 5%. Upon injection of each mix into the dual-cell sample holder, flow was stopped and spatially selective Hadamard spectroscopy[31] was used to acquire a 1D ^{1}H spectrum of each sample separately. A CPMG T2 filter of 80 ms was used to remove residual broad resonances from the sepharose resin. The cycle time was about 35 min, with 30 min required for the NMR experiment and 5 min for sample handling, resulting in a total time of about 5.5 days to complete the screen. To maintain the proper fold of each protein, 5 mM deuterated DPC was included in the buffer (20 mM phosphate buffer in D_2O, 100 mM NaCl, pH 7.6) used to wash the fragment mixes from the sample holder.

Contributions

Yelena N. Grinkova (University of Virginia) provided the empty NDs. Dr. Virginie Früh performed TINS against micelle solubilized DsbB. Francis Figaroa (ZoBio BV) assisted with compounds preparations for TINS screening. Eiso AB (ZoBio BV) helped with ^{15}N HSQC measurement.

Reference

1. Zheng, C.J., Han, L., Yap, C.W., Xie, B., and Chen, Y.Z. (2006). Progress and problems in the exploration of therapeutic targets. *Drug Discov. Today* 11, 412–420.
2. Carr, R.A.E., Congreve, M., Murray, C.W., and Rees, D.C. (2005). Fragment-based lead discovery: leads by design. *Drug Discov. Today* 10, 987–992.
3. Congreve, M., Carr, R., Murray, C., and Jhoti, H. (2003). A rule of three for fragment-based lead discovery? *Drug Discov. Today* 8, 876–877.
4. Siegal, G., AB, E., and Schultz, J. (2007). Integration of fragment screening and library design. *Drug Discov. Today* 12, 1032–1039.
5. Hajduk, P.J., Meadows, R.P., and Fesik, S.W. (1997). Drug design discovering high-affinity ligands for proteins. *Science* 278, 497–499.
6. Rasmussen, S.G.F., Choi, H.K., Rosenbaum, D.M., Kobilka, T.S., Thian, F.S., Edwards, P.C., Burghammer, M., Ratnala, V.R.P., Sanishvili, R., Fischetti, R.F., et al. (2007). Crystal structure of the human beta 2 adrenergic G-protein-coupled receptor. *Nature* 450, 383–387.
7. Serrano-Vega, M.J., Magnani, F., Shibata, Y., and Tate, C.G. (2008). Conformational thermostabilization of the beta 1-adrenergic receptor in a detergent resistant form. *Proc. Natl. Acad. Sci. USA* 105, 877–882.
8. Dahmane, T., Damian, M., Mary, S., Popot, J.L., and Baneres, J.L. (2009). Amphipol-Assisted in Vitro Folding of G Protein-Coupled Receptors. *Biochemistry* 48, 6516–6521.
9. Nath, A., Atkins, W.M., and Sligar, S.G. (2007). Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins. *Biochemistry* 46, 2059–2069.
10. Katzen, F., Fletcher, J.E., Yang, J.P., Kang, D., Peterson, T.C., Cappuccio, J.A., Blanchette, C.D., Sulchek, T., Chromy, B.A., Hoeprich, P.D., et al. (2008). Insertion of membrane proteins into discoidal membranes using a cell-free protein expression approach. *J. Proteome Res*. 7, 3535–3542.
11. Leitz, A.J., Bayburt, T.H., Barnakov, A.N., Springer, B.A., and Sligar, S.G. (2006). Functional reconstitution of beta(2)-adrenergic receptors utilizing self-assembling Nanodisc technology. *Biotechniques* 40, 601.
12. Fruh, V., Heetebrij, R.J., and Siegal, G. (2008). Target immobilized NMR screening: Validation and extension to membrane proteins. In Fragment- Based Drug Discovery, A. Practical Approach, E.R. Zartler and M. Shapiro, eds. (Chichester, UK: Wiley), pp. 135–158.
13. Vanwetswinkel, S., Heetebrij, R.J., van Duynhoven, J., Hollander, J.G., Filippov, D.V., Hajduk, P.J., and Siegal, G. (2005). TINS, target immobilized NMR screening: an efficient and sensitive method for ligand discovery. *Chem. Biol*. 12, 207–216.
14. Marquardsen, T., Hofmann, M., Hollander, J.G., Loch, C.M.P., Kiihne, S.R., Engelke, F., and Siegal, G. (2006). Development of a dual cell, flow-injection sample holder, and NMR probe for comparative ligand-binding studies. *J. Magn. Reson*. 182, 55–65.
15. Bardwell, J.C.A., Lee, J.O., Jander, G., Martin, N., Belin, D., and Beckwith, J. (1993). A pathway for disulfide bond formation in vivo. *Proc. Natl. Acad. Sci. USA* 90, 1038–1042.
16. Inaba, K., and Ito, K. (2002). Paradoxical redox properties of DsbB and DsbA in the protein disulfide-introducing reaction cascade. *EMBO J*. 21, 2646–2654.
17. Stenson, T.H., and Weiss, A.A. (2002). DsbA and DsbC are required for secretion of pertussis toxin by Bordetella pertussis. *Infect. Immun.* 70, 2297–2303.
18. Jagusztyn-Krynicka, E.K., Rybacki, J., and Lasica, A.M. (2009). Novel strategies for antibacterial drug discovery - antitoxin drugs. *Postepy Mikrobiologii* 48, 93–104.
19. Jander, G., Martin, N.L., and Beckwith, J. (1994). Two cysteines in each periplasmic domain of the membrane-protein DsbB are required for its function in protein disulfide bond formation. *EMBO J*. 13, 5121–5127.
20. Regeimbal, J., and Bardwell, J.C.A. (2002). DsbB catalyzes disulfide bond formation de novo. *J. Biol. Chem*. 277, 32706–32713.
21. Kadokura, H., Bader, M., Tian, H.P., Bardwell, J.C.A., and Beckwith, J. (2000). Roles of a conserved arginine residue of DsbB in linking protein disulfide-bond-formation pathway to the respiratory chain of *Escherichia coli. Proc. Natl. Acad. Sci. USA* 97, 10884–10889.
22. Inaba, K., Murakami, S., Suzuki, M., Nakagawa, A., Yamashita, E., Okada, K., and Ito, K. (2006). Crystal structure of the DsbB-DsbA complex reveals a mechanism of disulfide bond generation. *Cell* 127, 789–801.
23. Zhou, Y., Cierpicki, T., Jimenez, R.H., Lukasik, S.M., Ellena, J.F., Cafiso, D.S., Kadokura, H., Beckwith, J., and Bushweller, J.H. (2008). NMR solution structure of the integral membrane enzyme DsbB: Functional insights into DsbB-catalyzed disulfide bond formation. *Mol. Cell* 31, 896–908.
24. Arora, A., Abildgaard, F., Bushweller, J.H., and Tamm, L.K. (2001). Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. *Nat. Struct. Biol*. 8, 334–338.
25. Civjan, N.R., Bayburt, T.H., Schuler, M.A., and Sligar, S.G. (2003). Direct solubilization of heterologously expressed membrane proteins by incorporation into nanoscale lipid bilayers. *Biotechniques* 35, 556–560, 562–563.
26. Hajduk, P.J., and Greer, J. (2007). A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discov.* 6, 211–219.
27. Inaba, K., Takahashi, Y.H., and Ito, K. (2004). DsbB elicits a red-shift of bound ubiquinone during the catalysis of DsbA oxidation. *J. Biol. Chem*. 279, 6761–6768.
28. Bader, M., Muse, W., Ballou, D.P., Gassner, C., and Bardwell, J.C.A. (1999). Oxidative protein folding is driven by the electron transport system. *Cell* 98, 217–227.
29. Bader, M.W., Xie, T., Yu, C.A., and Bardwell, J.C.A. (2000). Disulfide bonds are generated by quinone reduction. *J. Biol. Chem*. 275, 26082–26088.

30. Arora, A., Rinehart, D., Szabo, G., and Tamm, L.K. (2000). Refolded outer membrane protein A of Escherichia coli forms ion channels with two conductance states in planar lipid bilayers. *J. Biol. Chem.* 275, 1594–1600.
31. Murali, N., Miller, W.M., John, B.K., Avizonis, D.A., and Smallcombe, S.H. (2006). Spectral unraveling by space-selective Hadamard spectroscopy. *J. Magn. Reson.* 179, 182–189.

Appendix

Target Immobilized NMR Screening (TINS) of DsbB/DPC

The fragment collection was screened for binding to DsbB at 500 μM each, in 182 mixtures. A spatially selective Hadamard NMR experiment[1] was used to simultaneously acquire a 1D ^{1}H spectrum of compounds in the presence of DsbB/DPC or OmpA/DPC. The data resulting from the screen could be analysed directly without deconvolution because fragments could be directly identified by comparing peaks from TINS spectra with reference spectra of the individual fragment. The screen resulted in 93 hits for DsbB, defined as fragments which had a T/R ratio less than 0.3. This particular cut-off was chosen by virtue of a step-like relationship between the observed TINS effect and the number of "hits" whereby even slightly raising the cut-off gave a large increase (> 2-fold) in the number of compounds that were selected as hits (not shown). The resulting hit rate for DsbB was 8.7% which is well within the range we typically observe with soluble proteins using TINS (3-9.5%). Application of the same criteria to OmpA/DPC binding identified seven compounds as hits for a hit rate of 0.6%, validating the earlier data suggesting that OmpA/DPC has minimal small molecule binding capacity.

Hit Validation using Enzymatic Assays

All fragments from the screen that were designated as positive for binding were assayed for DsbB inhibition at 250 μM. The amount of DMSO in all biochemical assay controls was adjusted to match the amount present when fragments were tested. Those compounds that showed more than 70% inhibition at 250 μM were further characterized by titration from 0.0001 mM to 10 mM to generate IC_{50} curves. The mode of action for the 8 most potent fragments was determined from competition enzyme assays. For this analysis either DsbA or UQ1 was titrated in from 0.2 to 40 μM, while the other was kept constant at 40 μM. For each titration point, slopes were measured in the presence of 5, 10, and 75 μM of the fragment. DsbB activity data was analyzed using the non-linear regression curve fitting routines in Graph Pad Prism v. 5.01 (Graph Pad, San Diego, CA, U.S.A.). Statistical significance was evaluated with the student's T-test. Depending on the light absorbing properties of the fragments, they were used in either the fluorescence or UV-absorbance assay. Compounds which were not compatible with the assays due to

high intrinsic fluorescence, high UV absorbance or irregular baselines were not included in the analysis.

Each of the 93 fragments identified as TINS hits in the micelle screen was assayed for inhibition of DsbB-dependent reoxidation of DsbA at 250 μM. Eight compounds interfered with the assay when run in either fluorescence or absorbance mode and therefore were left out of the analysis. The remaining 85 hits exhibited a distribution of potencies against DsbB, including 60% with better than 30% enzymatic inhibition and 16% with either less than 20% inhibition or stimulation. This data confirms that a very high percentage of the hits found in the biophysical assay also modulate the enzymatic activity of DsbB and are functionally relevant.

To avoid the possibility of artifacts in the biochemical assay we selected the 13 fragments from the micelle screen showing strong inhibition in the single concentration point assay for further analysis. We first assayed these 13 fragments for potency (IC_{50}) by dose-response experiments (Figure A1). Dose response experiments were carried out with increasing fragment concentrations, from 0.0001 to 10 mM, while both DsbA and UQ1 were kept in excess. Three of the 13 fragments indeed showed artefacts including signs of protein precipitation at higher compound concentration and/or steeper than expected Hill coefficients. The remaining 10 fragments titrated over 2 Log orders and exhibited a Hill coefficient close to unity. By these criteria, the ten fragments are reversible inhibitors with a 1:1 stoichiometry and are therefore well-behaved. The eight most potent compounds had IC_{50} values between 7 and 170 μM (Table A1) and consisted of a variety of scaffolds (Figure A1). The calculated binding efficiency index[2] (Table A1) indicates that these fragments are all very good or excellent starting points for hit elaboration projects.

Table A1. Biological characterization of fragment hits.

Fragment	IC_{50} (μM)	Hill coefficient	BEI[2]
1	7 ± 1	0.80 ± 0.10	21.8
2	100 ± 10	1.40 ± 0.15	14.7
3	193 ± 11	1.20 ± 0.11	15.3
4	13 ± 1	0.80 ± 0.10	21.2
5	46 ± 12	0.80 ± 0.10	22.2
6	70 ± 10	1.00 ± 0.10	14.0
7	115 ± 11	1.15 ± 0.05	19.1
8	168 ± 10	1.40 ± 0.10	14.2

Note: BEI – Binding Efficiency Index = pIC_{50}/MW

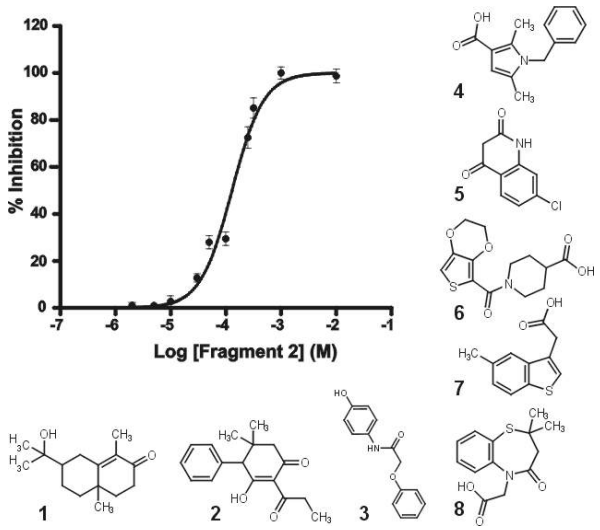


Figure A1. Potency determination of selected hits from the TINS screen. An example of an inhibition curve used to determine the IC_{50} for compound **2**. The curve represents the mean ± S.E.M. of three independent experiments performed in triplicate. The structures of the 8 most potent compounds are shown and the IC_{50} values are provided in Table A1.

As a second validation step, we carried out a more detailed kinetic analysis of the mode of action of the 8 most potent fragments. Substrate-velocity experiments were performed in which either DsbA or UQ1 were titrated in the presence of saturating amounts of the other. The titrations were then repeated in the presence of increasing amounts of the inhibitory fragment (Figure A2, Table A2). In this analysis, fragments **1-3** behaved similarly. This group is exemplified by fragment **2** where increasing concentrations result in moderate perturbation of the maximum enzymatic turn over rate (k_{cat}) and apparent affinity of DsbA but a dramatic reduction (> 6-fold) in the apparent affinity of UQ1. This result suggests that fragments **1-3** compete for the same binding site as UQ1. On the other hand, addition of fragments **4-8** simultaneously decreased both the apparent affinity and the k_{cat} for UQ1 and DsbA as best exemplified by fragment **8** (Figure A2). This data suggests a mixed model of inhibition of DsbB by these fragments. We next sought biophysical confirmation of these two different modes of fragment interaction with DsbB.

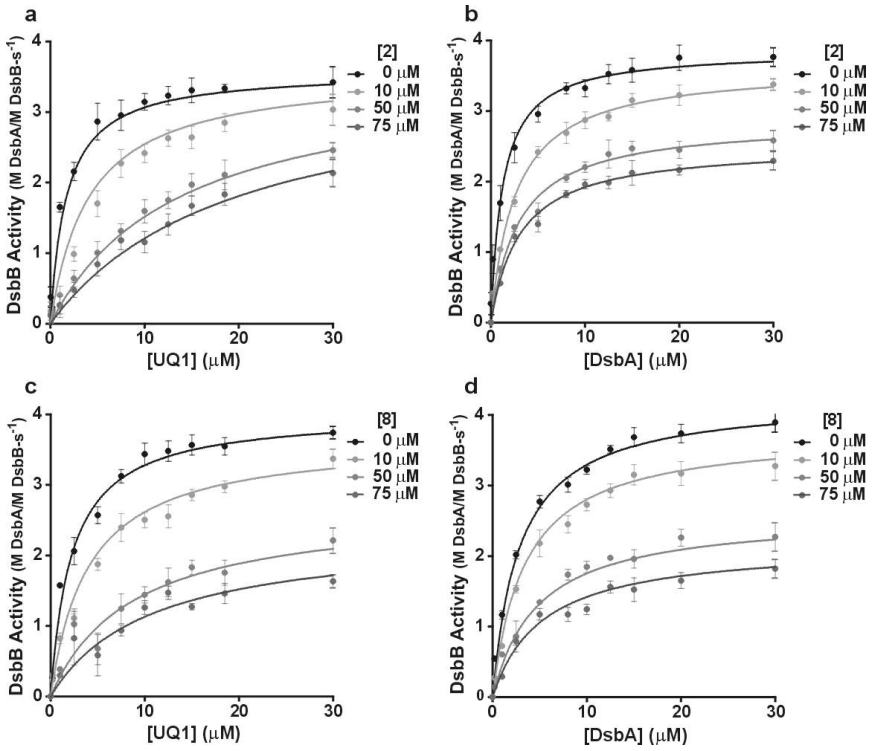


Figure A2. Mode of action determination for the most potent DsbB inhibitors. Fragment **2** was assayed in competition with synthetic UQ1 (**a**), the electron acceptor, or DsbA (**b**) the electron source. Fragment **8** was assayed in the same manner (panels **c** and **d** respectively). The k_{cat} and K_m apparent determined from the data are shown in Table 2 in the absence and presence of the indicated amount of each inhibitor.

Table A2. Kinetic analysis of two fragment hits.

Fragment		Substrate	k_{cat} (M DsbA/ M DsbB-s^{-1})	K_m (μM)
2	0 μM	UQ1	3.7 ± 0.1	1.6 ± 0.1
	75 μM		3.1 ± 0.2	13.2 ± 1.7
	0 μM	DsbA	4.3 ± 0.1	2.4 ± 0.2
	75 μM		2.5 ± 0.1	2.8 ± 0.1
8	0 uM	UQ1	4.0 ± 0.1	2.2 ± 0.0
	75 μM		2.3 ± 0.1	9.7 ± 1.5
	0 μM	DsbA	4.3 ± 0.1	1.9 ± 0.1
	75 μM		2.4 ± 0.2	4.1 ± 0.5

Confirmation of Different Modes of Interaction with DsbB by NMR

Due to the poor quality of the NMR spectra of the wild-type DsbB, it was necessary to use a mutant that represents an intermediate in the disulfide oxidation pathway[3]. Validated hits from the screen were titrated at 1, 5, and 10 mM into ^{15}N-

labelled DsbB[CSSC] mutant (C44S, C104S). [^{15}N,^{1}H] HSQC experiments were acquired at 40 ºC in a Bruker DRX 600 MHz spectrometer equipped with a cryoprobe. A reference titration of DMSO and a non-binding fragment from the screen were used to subtract chemical shift perturbations not related to fragment binding.

Chemical shift perturbations of the protein NMR spectrum (typically 2D ^{15}N-^{1}H HSQC or ^{13}C-^{1}H HSQC) in the presence of compounds can both confirm binding to the target and localize the binding site on a protein when resonance assignments are available[4]. While the sequential assignment of wildtype DsbB is not available due to the poor quality of the NMR spectra, spectra of the DsbB[CSSC] double cysteine mutant are of high quality, resulting in a complete backbone resonance assignment for this form of the protein[3]. When purified from *E. coli,* DsbB[CSSC] contains the endogenous ubiquinone-8[5], thus compounds specific for this site must compete with UQ-8 for binding. We first titrated the synthetic quinone UQ1 into a sample of ^{15}N DsbB[CSSC]. Addition of UQ1 to ^{15}N DsbB[CSSC] resulted in numerous chemical shift perturbations but two in particular afford a detailed analysis of the binding and allow a reliable comparison with the fragments found in TINS screening. As shown in Figure A3, the sidechain indole of Trp135 (in the vicinity of the quinone binding site) and the backbone amide of Arg109 (close to the DsbA binding site) respond very differently to addition of UQ1. Titration of UQ1 resulted in the simultaneous disappearance of the Trp135ε-NH peak from DsbB[CSSC] bound to endogenous quinone and the appearance of a new peak close by in the spectrum. Due to its proximity and the unique chemical shift of the Trpε NH proton, in combination with the high level of conservation of this residue, the new peak is likely from the Trpε NH proton of the UQ1 bound DsbB[CSSC]. This pattern of peak changes is indicative of slow exchange on the NMR time scale (e.g. k_{off} < 30 Hz Δδ in Figure A3a). In contrast, the backbone amide of Arg109 is essentially unchanged by the addition of UQ1.

Addition of all 8 fragments to ^{15}N labelled DsbB[CSSC] resulted in detectable changes in chemical shifts, suggesting that the fragments selected by TINS screening and biochemical assays on wild type protein also bind the cysteine mutated form. The presence of chemical shift perturbations both in solvent exposed loops and in portions of the protein buried within the micelle (data not shown) suggests that the fragments are specifically binding to the protein and not non-specifically partitioning into the micelle.

Fragment **2**, which competitively inhibited ubiquinone binding, induced chemical shift perturbations in a variety of amino acids, including W135 and R109. The pattern of chemical shift perturbations induced by fragment **2** closely resembles that induced by UQ1. First, titration of **2** into ^{15}N DsbB[CSSC] resulted in chemical shift changes in the W135ε-HN peak that were similar to those induced by UQ1 (i.e. slow exchange). Moreover, the resonance frequency of the new peak tentatively assigned to the DsbB[CSSC]-**2** complex is similar to that of the DsbB[CSSC]-UQ1 complex. Similarly, R109HN, which is minimally affected by UQ1, undergoes only minor chemical shift perturbations in the presence of **2**.

In contrast, the chemical shift changes induced by fragments **4-8** differ in both the overall pattern and the details from fragment **2** and UQ1 (Figure A3c). Addition of **8**, for example, induced concentration dependent shifts in the Trp135ε-NH resonance to an entirely different chemical shift than did fragment **2** or UQ1. This concentration dependent shift is indicative of rapid exchange on the NMR time scale. There was no evidence for slow exchange for any of the fragments **4-8**, although **4** and **7** show signs of line broadening of the backbone resonance of Q33 that may indicate intermediate exchange (not shown). In contrast, the backbone amide of R109, which is only mildly perturbed by UQ1 or **2**, is very dramatically perturbed by the addition of fragment **8**. This data suggests that fragments **4-8**, which exhibit mixed mode DsbB inhibition, bind in either a different mode or different site to fragment **2** which is competitive with ubiquinone.

The different behaviour of the resonances of the backbone amide of R109 and the sidechain indole of W135 upon titration with the fragments provides further support for two modes of action. Titration of UQ1 DsbB[CSSC] results in slow exchange between the endogenous quinone bound form and a newly arising peak at a nearby position which we assign to the UQ1 complex. Similarly, addition of **2** resulted slow exchange between the endogenous quinone bound form of W135ε-NH and the appearance of a new peak with similar chemical shift as the UQ1 complex. An additional chemical shift perturbation indicating fast exchange with the endogenous complex was also observed. The fast exchange is likely due to competition between **2** and the quinone moiety of the bound UQ8, consistent with the competitive kinetics observed for this inhibitor. However,

we have shown that the isoprenyl tail of UQ8 extends down the groove between TM1 and TM4, making extensive interactions with the protein[3]. Therefore, displacement of the

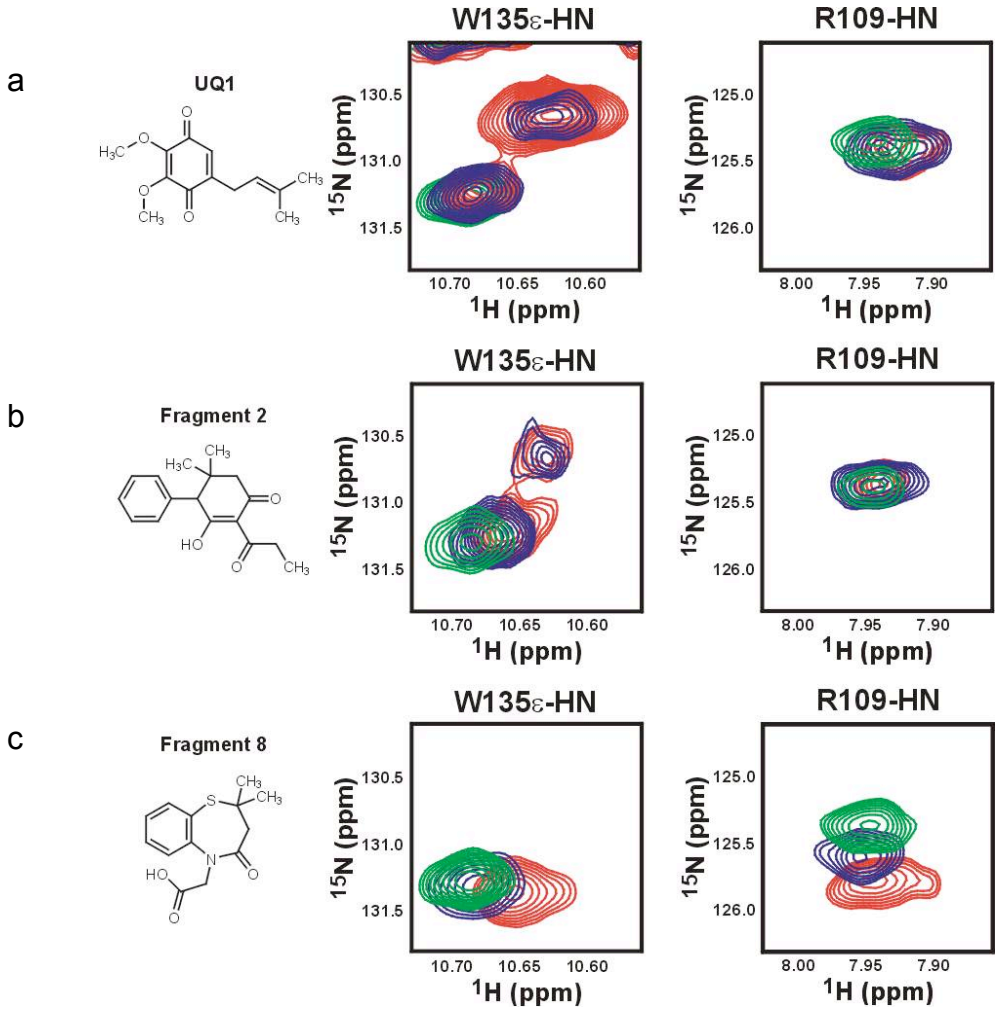


Figure A3: NMR analysis of fragment binding to DsbB. The eight most potent fragments were titrated into ^{15}N DsbB[CSSC]. Data for the synthetic quinone UQ1 (A), competitive fragment **2** (B), and the mixed model fragment **8** (C) are shown. For each of these three compounds, the structure of the compound is shown in the left column and the characteristic peak perturbations in the [^{15}N,^{1}H] HSQC spectrum (green 0 mM fragment, blue 5 mM fragment, and red 10 mM fragment) are shown in the middle (Tryptophan 135 side-chain indole) and right columns (Arginine 109 backbone amide).

quinone moiety apparently occurs on a slower timescale. Addition of **8** to DsbB[CSSC] also causes chemical shift perturbation of the W135ε-NH. But these suggest rapid exchange between a typical 2-state equilibrium rather than the more complex effects seen with **2**. In addition, the bound state has a different resonance frequency than the bound state of UQ1 or **2**. Additional large downfield chemicals shifts of the resonance of R109N indicative of rapid exchange are observed while UQ1 and **2** had no or only minor

effects on this peak. From a drug discovery perspective, this data is exciting because it suggests, as with soluble enzymes, it is possible to find small molecule inhibitors with different modes of action and possibly non-overlapping or even different binding sites on membrane proteins.

We note that the concentration of the fragments required to induce chemical shift perturbations in DsbB[CSSC] is significantly higher than the IC_{50} values measured for the wild type protein but the same as required for UQ1. A likely explanation is that the conformation of the mutant differs slightly from the wildtype protein, against which the fragments were selected. In addition, either the affinity for the quinone is higher for the DsbB[CSSC] mutant or more likely, the quinone binding site may be partially occluded. This latter possibility is clearly consistent with the reduced dynamic behaviour of DsbB[CSSC] with respect to the wild type protein, which results in the substantial improvement in the quality of the NMR spectra. This reduced dynamic behaviour of the disulfide mutant may be responsible for the slow exchange kinetics observed for UQ1 and **2** if release of the UQ8 from this binding site can only occur from a sparsely populated conformation.

Reference

1. Murali, N., Miller, W.M., John, B.K., Avizonis, D.A., and Smallcombe, S.H. (2006). Spectral unraveling by space-selective Hadamard spectroscopy. *J. Magn. Reson.* 179, 182–189.
2. Abad-Zapatero, C., and Metz, J.T. (2005). Ligand efficiency indices as guide-posts for drug discovery. *Drug Discov. Today* 10, 464–469.
3. Zhou, Y., Cierpicki, T., Jimenez, R.H., Lukasik, S.M., Ellena, J.F., Cafiso, D.S., Kadokura, H., Beckwith, J., and Bushweller, J.H. (2008). NMR solution structure of the integral membrane enzyme DsbB: Functional insights into DsbB-catalyzed disulfide bond formation. *Mol. Cell* 31, 896–908.
4. Shuker, S.B., Hajduk, P.J., Meadows, R.P., and Fesik, S.W. (1996). Discovering high-affinity ligands for proteins: SAR by NMR. *Science* 274, 1531–1534.
5. Bader, M., Muse, W., Ballou, D.P., Gassner, C., and Bardwell, J.C.A. (1999). Oxidative protein folding is driven by the electron transport system. *Cell* 98, 217–227.

Chapter 3

Fragment Screening of GPCRs Using Biophysical Methods: Identification of Ligands of the Adenosine A_{2A} Receptor with Novel Biological Activity

D. Chen, J. C. Errey, L. H. Heitman, F. H. Marshall, A. P. IJzerman, G. Siegal. Fragment screening of GPCRs using biophysical methods: identification of ligands of the adenosine A_{2A} receptor with novel biological activity. *ACS Chem. Biol.*, 2012, 7, 2064-2073.

ABSTRACT

Fragment-based drug discovery (FBDD) has proven a powerful method to develop novel drugs with excellent oral bioavailability against challenging pharmaceutical targets such as protein-protein interaction targets. Very recently the underlying biophysical techniques have begun to be successfully applied to membrane proteins. Here we show that novel, ligand efficient small molecules with a variety of biological activities can be found by screening a small fragment library using thermostabilized (StaR) G protein-coupled receptors (GPCRs) and Target Immobilized NMR Screening (TINS). Detergent solubilized, StaR adenosine A_{2A} receptor was immobilized with retention of functionality and a screen of 531 fragments was performed. Hits from the screen were thoroughly characterized for biochemical activity using the wild-type receptor. Both orthosteric and allosteric modulatory activity has been demonstrated in biochemical validation assays. Allosteric activity was confirmed in cell-based functional assays. The validated fragment hits make excellent starting points for a subsequent hit-to-lead elaboration program.

1. Introduction

G protein-coupled receptors (GPCRs) are one of the most popular classes of investigational drug targets due to their central role in a variety of biological and pathological processes, such as: neuromodulation, metabolic disorders, inflammation, cancer and viral infection.[1,2] Thus, it is not surprising that about 30% of all prescribed drugs on the market act on this class of receptors.[3] In the last ten years, GPCR drug discovery has relied on cell-based assays combined with high-throughput screening (HTS) of large compound libraries[4] for lead discovery as well as optimization. This paradigm has delivered compounds that modulate the behavior of GPCRs as desired, many of which have made it to patients. Nonetheless, only one small molecule drug targeting a novel GPCR has been approved in each of the last 10 years. Moreover, the majority of drugged GPCRs are members of families that are activated by small molecules such as the adrenergic, muscarinic and histamine receptors.[5] In contrast, peptide and protein hormone receptors have represented a much greater challenge. In addition, achieving subtype selectivity has been difficult and where successful, has required quite substantial medicinal chemistry efforts. Furthermore, hits from HTS screens frequently must be deconstructed to remove liabilities that result in non-ideal absorption, distribution, metabolism, and excretion (ADME) properties.[6]

Fragment-based drug discovery (FBDD) uses low-molecular-weight (< 300 Da), moderately lipophilic (clogP < 3) and highly soluble[7] molecules or “drug fragments” as starting points for developing novel drugs. FBDD is particularly advantageous for its ability to more completely assess “compound space” for molecules that interact with the target of interest.[8] Since the typical interaction of such small molecules with proteins is weaker than 1 μM KD, the types of assays employed in HTS screening are not easily capable of discriminating between positives and negatives. In order to detect such weak binding, sensitive biophysical techniques such as NMR, SPR or X-ray crystallography are most commonly used. FBDD has been successful in developing inhibitors of soluble

targets such as kinases,[9] proteases[10,11] and most recently protein-protein interactions.[12] In principle, it could be advantageous to apply FBDD to GPCRs to identify molecules with novel modes of action.[13] However, the application of biophysical techniques to membrane proteins is extremely challenging due to their instability when extracted from the cell membrane and the difficulty of large scale recombinant expression. Even when successful, significant levels of non-specific partitioning of compounds into the media used to solubilize membrane proteins remains a significant source of false positives with biophysical studies.

We have developed an NMR-based, fragment screening approach which has been used to find new chemical starting points for further development of compounds targeting membrane proteins, including GPCRs.[14,15] The approach, called target immobilized NMR screening (TINS),[16] uses a target and a reference protein immobilized on a compatible solid support. Binding of fragments to the immobilized protein is monitored by changes in the 1D ^{1}H NMR spectrum of a ligand. A powerful method, called StaRs (stabilized receptors), has been developed to overcome the instability and purification problems of GPCRs. StaR GPCRs, which contain a small number of point mutations for both enhanced thermostability and conformational homogeneity,[17-20] can be produced in significant quantities and purified in correctly folded form in detergent micelles for structural and biophysical studies.[21-23]

We have previously demonstrated the feasibility of the combination of TINS and StaRs to screen fragment libraries for target specific binding.[15] Here we use TINS to find fragment ligands for the human adenosine A_{2A} receptor ($A_{2A}R$) with novel modes of action. Screening of a moderate sized fragment library using an antagonist stabilized StaR yielded multiple hits that were pharmacologically validated with the wild-type (WT) receptor. Amongst these hits, we find multiple allosteric modulators of the $A_{2A}R$ with both positive and negative effects on binding and function of the orthosteric ligand.

2. Results

Functional Immobilization and Stability

In order to assess the feasibility of ligand screening studies, we first determined whether immobilized, micelle solubilized A_{2A}-StaR2 retained functionality and if so, whether the preparation was sufficiently stable over time (Figure 1). We used an immobilization method that was successfully applied to the StaR β_1 adrenergic receptor (β_1AR),[15] in which primary amines of the protein form a Schiff's base with an aldehyde group on the solid support. Using this procedure, the DDM-solubilized A_{2A}-StaR2 was efficiently immobilized on a sepharose resin at pH 7.4 with a yield of 70% and subsequently, unreacted aldehydes were blocked using deuterated Tris buffer. To determine whether the immobilized A_{2A}-StaR2 remained competent for ligand binding, the binding of [^{3}H]ZM241385, an inverse agonist that binds the orthosteric site of StaR and wild-type receptors with a K_D of approximately 1 nM and 1:1 stoichiometry, was assessed as previously described.[24] Initial measurement of [^{3}H]ZM241385 binding after immobilization indicated that nearly 95% of the immobilized A_{2A}-StaR2 was competent. Next, we compared the stability of immobilized A_{2A}-StaR2 with A_{2A}R-WT in HEK cell membranes. After 5 days stored at 4 °C, more than 60% of the immobilized A_{2A}-StaR2 still remained competent for radioligand binding (Figure 1). In contrast, over the same time period, the A_{2A}R-WT lost 75% activity, despite the fact that it was in a native environment.

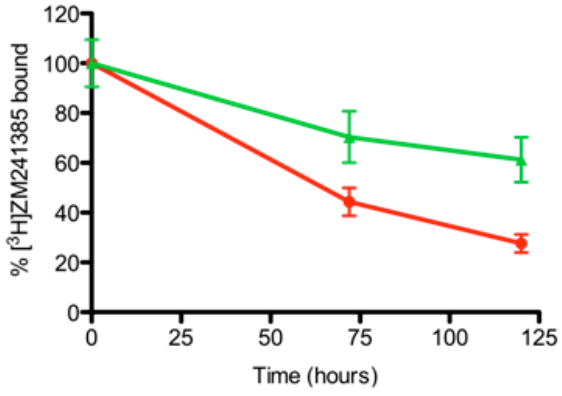


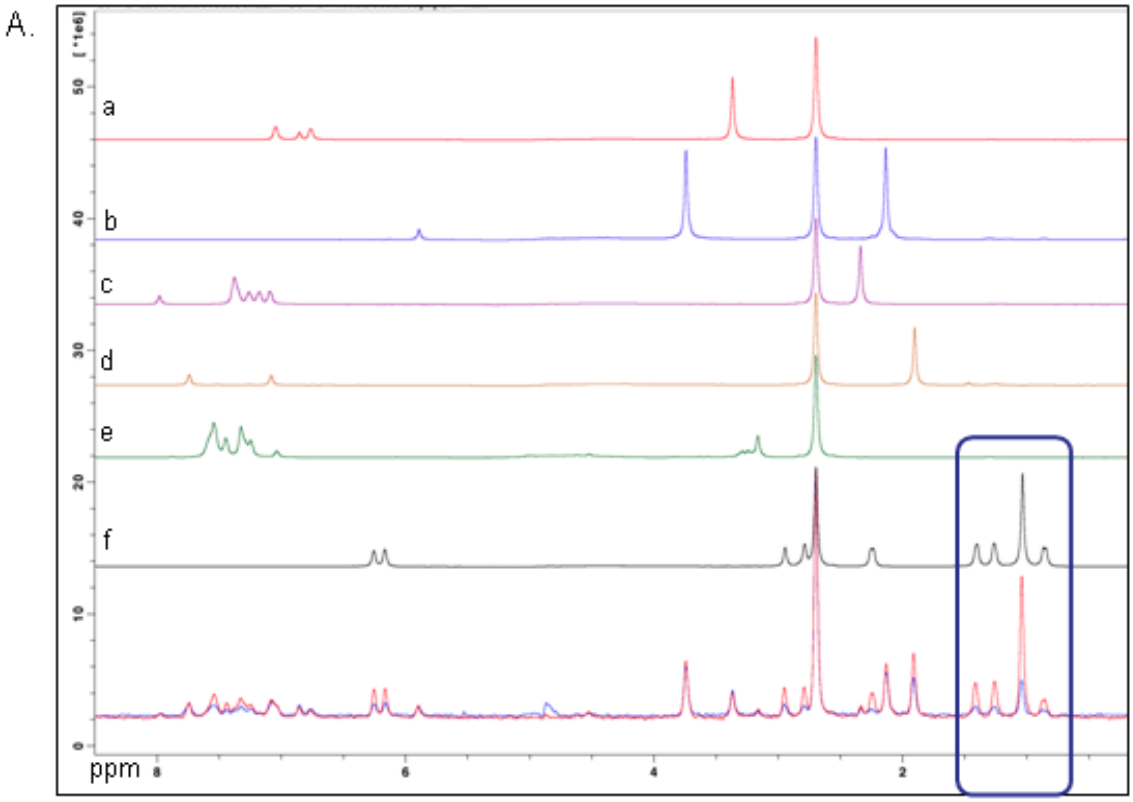
Figure 1. Functional immobilization and stability of micelle-solubilized A_{2A}-StaR2 (green) compared to A_{2A}R-WT in HEK cell membranes (red). A_{2A}-StaR2 was covalently bound to Sepharose resin (see text for details). Functionality was assessed by binding of [^{3}H]ZM241385 upon storage of the protein at 4 °C for the indicated time. Specific [^{3}H]ZM241385 binding at day 0 was set as 100% for both A_{2A}-StaR2 and A_{2A}R-WT.

Fragment Screening

Having determined the functionality and stability of immobilized A_{2A}-StaR2, we set out to screen a collection of fragments for binding to the receptor. A set of 531 fragments was selected for their chemical and shape diversity.[25] The TINS technique uses a reference to cancel out non-specific binding, thereby reducing the amount of false positives. Previously, we had determined that the *E. coli* outer membrane protein OmpA, a b-barrel structure that can be solubilized in various detergents, was an appropriate reference for micelle-solubilized membrane proteins[14,15] and therefore, we have used it for the current study as well. The fragments were injected into the dual cell sample holder[26] in mixtures of between 3 and 8 compounds. A spatially selective, one-dimensional ^{1}H Hadamard experiment[27] was used to acquire the NMR spectrum of the compounds in solution. Since the NMR relaxation of a spin is approximately 1,000 times more efficient in the solid state than in solution, binding of a ligand to the immobilized protein causes the resonances of that molecule to further broaden into the kHz range. As a result, ligand binding is detected as a simple reduction in the amplitude of the NMR signals of that molecule in the presence of the target in comparison with those in the presence of the reference (Figure 2A).

In order to differentiate hits from non-hits, we analyzed a profile of the entire screen (Figure 2B). In this plot, the T/R ratio (the weighted average of the amplitude of well resolved resonances in the presence of target divided by that in the presence of reference) for each fragment has been bucketed and the number of fragments in each bucket is presented as a histogram. Similarly to the β_1 adrenergic receptor,[15] the profile shows a peak at a T/R ratio of 1 indicating that the bulk of the fragments do not bind or bind non-specifically to both target and reference. Further, very few compounds have a T/R > 1 recapitulating that OmpA has minimal specific small molecule binding capacity and is therefore a good reference. In contrast, there are large numbers of fragments with a T/R ratio significantly less than 1, indicating that they preferentially bind to the A_{2A}-StaR2. In order to make a selection of positives, a discontinuity in the histogram is

chosen and that value is used as a cutoff. Such a discontinuity can be found at a T/R ratio of about 0.7. Since the goal was to characterize how the interaction of small molecules with the micelle solubilized GPCR is reflected in the TINS assay, we opted for a very conservative hit selection (i.e. a larger than normal number of compounds were designated as hits), which resulted in 94 fragments defined as $A_{2A}R$ hits. Application of the same criteria to OmpA binding identified six compounds as hits.



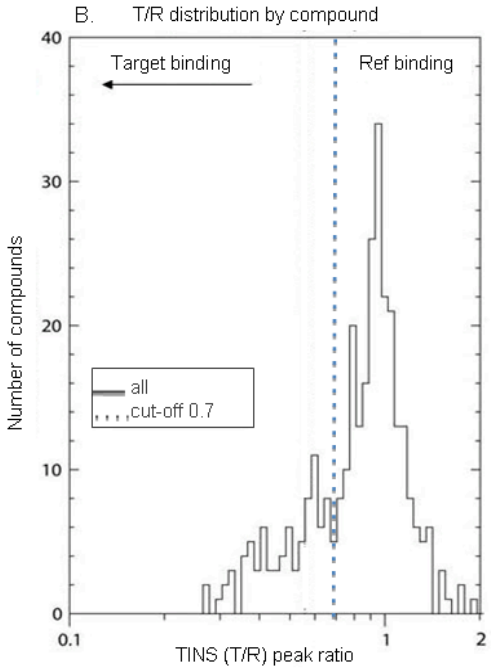


Figure 2. (A) An example of a TINS experiment applied to a mixture of 6 fragments. The ^{1}H NMR spectrum of each component in the mixture is presented, where extensive line broadening has been applied to simulate the linewidth of resonances in the heterogeneous TINS assay (a-f). The spectrum of the mixture in the presence of the reference protein (red) and A_{2A}-StaR2 (blue) are overlaid. The reduction of peaks in the presence of A_{2A}R indicates that primarily f binds. (B) Profile of the complete screen. The ratio of peaks in the presence of target & reference (T/R) for each compound has been binned and the number of compounds in each bin is shown. The cut-off for hit selection is shown.

Pharmacological Characterization of TINS Hits:

1). Probing for Orthosteric Site Ligands. A biophysical assay such as TINS simply discovers fragments that specifically bind to the target. We wished to validate this binding and determine whether or not the fragments exhibited a biochemical activity. For these studies we used HEK293 cell membranes transiently expressing the wild-type A_{2A} receptor. Initially we assessed the ability of each of the 94 fragment hits to displace the high affinity, orthosteric inverse agonist [^{3}H]ZM241385 in an equilibrium binding assay. This assay primarily detects ligands binding to the orthosteric site (i.e. same site as [^{3}H]ZM241385). Five of the fragment hits inhibited binding of ZM241385 by ≥ 30% at 500 μM, suggesting a biologically relevant interaction between the fragments and the wild-type A_{2A}R. Subsequent dose-response curves of these five hits were well

behaved, with IC_{50} values ranging from 70 μM to 1.9 mM (Table 1 and Figure 3). Analysis of the dose-response curves revealed that the Hill coefficient for each of the five compounds is not significantly different from unity which, in conjunction with the TINS data, indicates that these fragments reversibly bind $A_{2A}R$ with a 1:1 stoichiometry. As expected for compounds binding to the orthosteric site, the five hits also significantly inhibited the binding of [^{3}H]NECA, an adenosine receptor agonist, with IC_{50} values ranging from 80 μM to 7.7 mM (Table 1). Achieving receptor subtype selectivity is desirable but can be a challenge for fragments targeting the orthosteric site. Perhaps not unsurprisingly, the five compounds also displaced [^{3}H]DPCPX, a well-characterized orthosteric antagonist for the closely related adenosine A_1 receptor subtype. Moreover, the potency of each of the five compounds in the A_1R assay was similar to the $A_{2A}R$. It is not unexpected at this stage in hit finding for fragments to show binding to two closely related receptors. During a chemistry lead optimization program, such fragment hits could be optimized for selectivity using a variety of structure-based design approaches. Many known adenosine receptor ligands have been developed based on the adenosine[28] and xanthine[29] chemotypes. Newer adenosine receptor antagonists have modifications based on tricyclic derivatives of triazolopyrimidine,[30] including Preladenant[31] which has been evaluated in clinical trial (phase III) for the treatment of Parkinson's disease, and exhibit high selectivity for $A_{2A}R$. As shown in Table 1, the five fragment $A_{2A}R$ orthosteric ligands represent a diversity and novelty of chemical scaffolds. A hallmark of FBDD is that it generates ligands in which most of the atoms contribute to target binding, a characteristic termed ligand efficiency. Indeed, two of the fragments in Table 1 have high ligand efficiency (LEAN[32] (pIC_{50}/heavy atom count) ≥ 0.3), and are potentially suitable as starting points for lead generation, even with their low potency.

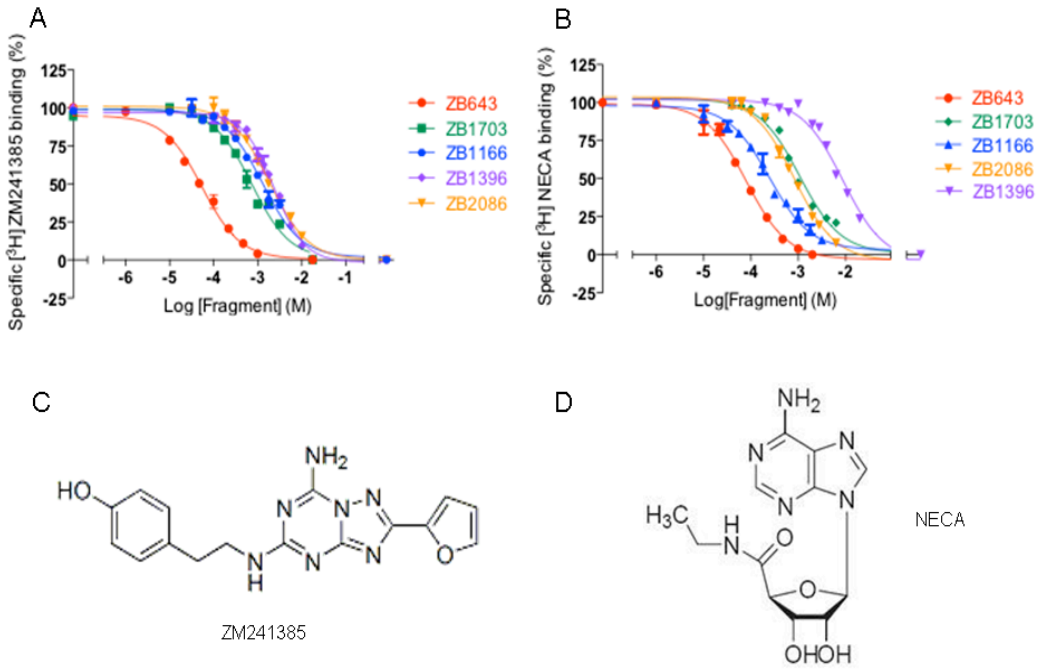
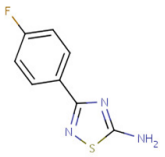
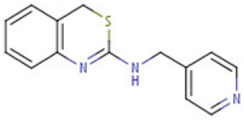
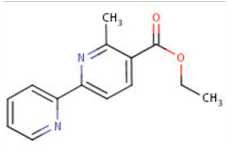
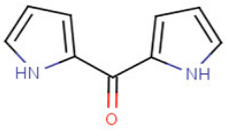
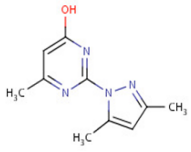


Figure 3. Representative dose-response curves for displacement of binding of the radiolabeled $A_{2A}R$ inverse agonist ZM241385 (A) and agonist NECA (B) by A_{2A} hits identified in the TINS screen. Each curve has been determined in triplicate in three independent experiments. The structures of the radiolabeled ligands are shown in C and D.

2). Probing for Allosteric Site Ligands. We next looked whether any of the 94 fragment hits bind to allosteric site(s) on the $A_{2A}R$. In order to do so, we employed an assay that measures the effect of a fragment on the dissociation rate (k_{off}) of an orthosteric ligand, but is not sensitive for competitive binding at the orthosteric site. Each of the 94 fragment hits were screened for their effect on the off rate of the inverse agonist [^{3}H]ZM241385 and the agonist [^{3}H]NECA dissociating from wild-type A_{2A} receptor using a single time point at which 50% of the radioligand has dissociated under control conditions. Three fragments significantly (at least 30%) increased k_{off} of the orthosteric ligand and hence are negative allosteric modulators (NAMs), while four significantly decreased k_{off} and hence are positive allosteric modulators (PAMs) of both ZM241385 and NECA (Table 2).

Table 1. Characterization of fragment hits by equilibrium radioligand displacement assays on wild-type A_{2A} and A_1 receptors.

ID	Fragment structure	Displacement of [^{3}H]ZM241385 on $A_{2A}R$[a]			Displacement of [^{3}H]NECA on $A_{2A}R$[a]			% displacement of [^{3}H]DPCPX on A_1R[b]
		IC_{50} (mM) (% displacement)	Hill Slope[c]	Ligand Efficiency[d] (LEAN)	IC_{50} (mM) (% displacement)	Hill Slope[c]	Ligand Efficiency[d] (LEAN)	
ZB643		0.07 ± 0.02 (91%)	-0.87 ± 0.12	0.43	0.08 ± 0.03 (90%)	-1.02 ± 0.12	0.42	92%
ZB1703		0.79 ± 0.32 (65%)	-1.23 ± 0.19	0.23	1.07 ± 0.04 (57%)	-1.11 ± 0.15	0.22	52%
ZB1166		1.32 ± 0.26 (40%)	-1.07 ± 0.24	0.22	0.25 ± 0.02 (68%)	-0.92 ± 0.09	0.27	48%
ZB2086		1.73 ± 0.47 (34%)	-1.07 ± 0.26	0.30	0.76 ± 0.05 (37%)	-1.04 ± 0.07	0.35	21%
ZB1396		1.88 ± 0.26 (30%)	-1.13 ± 0.21	0.23	7.71 ± 0.04 (26%)	-1.24 ± 0.24	0.19	22%

[a] Displacement of specific [^{3}H]ZM241385 or [^{3}H]NECA binding in HEK293 cell membranes transiently expressing human adenosine A_{2A}Rs expressed as IC_{50} ± SEM (n=3) or % displacement of specific binding at 0.5 mM fragment concentration (n=2). [b] % displacement of specific [^{3}H]DPCPX binding in CHO cell membranes stably expressing human adenosine A_1Rs at 0.5 mM fragment concentration. [c] Hill slope ± SEM (n=3). [d] LEAN ligand efficiency is -logIC_{50} divided by the heavy atom count.[32]

In addition, another four fragments allosterically enhanced either NECA binding or ZM241385 binding (Table 2). Subsequently, full dissociation curves of [^{3}H]ZM241385 and [^{3}H]NECA were measured in the absence (control) or presence of 2 mM of each of the active fragment hits. As shown in Table 2 and Figures 4A and 4B, the intrinsic dissociation rate of the radioligand was 0.011 ±

0.002 min^{-1} for [^{3}H]ZM241385 and 0.015 ± 0.002 min^{-1} for [^{3}H]NECA. Most of the fragment hits showed at least 2-fold modulation of radioligand dissociation. For example, in the presence of 2 mM ZB335, the k_{off} of [^{3}H]NECA was decreased 2.9-fold to 0.0051 ± 0.0009 min^{-1}. In combination with the lack of competitive binding, this data strongly suggests that ZB335 is acting at an allosteric site. Similarly, addition of 2 mM ZB418, resulted in an increase of the k_{off} of [^{3}H]ZM241385 by 3.5-fold to 0.039 ± 0.001 min^{-1}, suggesting this fragment is a NAM. Interestingly, ZB418 showed a hill slope about -2 in the equilibrium displacement assay for both [^{3}H]ZM241385 and [^{3}H]NECA which can also be indicative of an allosteric mechanism.[33] Importantly, the allosteric modulation by the fragment hits was concentration-dependent (Figure 4C and Table 3). In contrast to the fragments that were competitive with orthosteric ligands, the most potent allosteric modulators, such as ZB335 and ZB418, exhibited considerably reduced effects on the A_1R, suggesting that they have some A_{2A} subtype specificity.

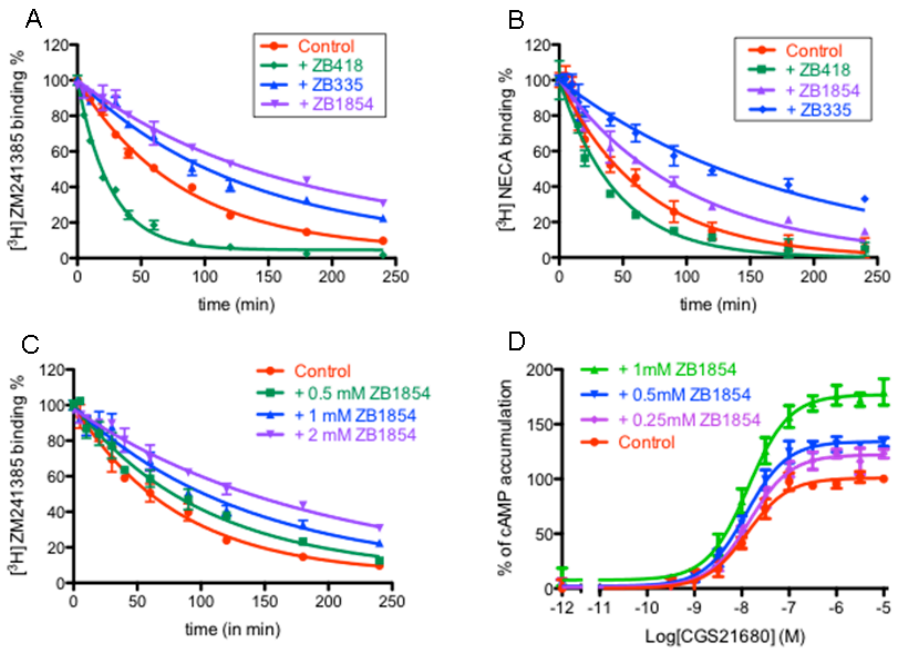
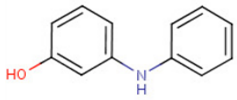
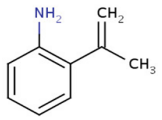
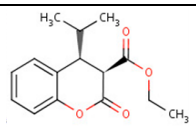
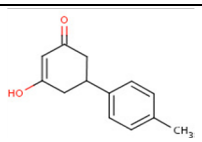
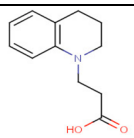
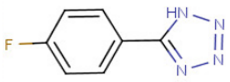
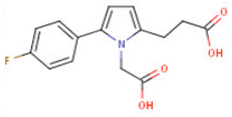
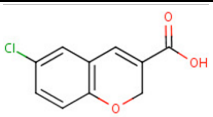
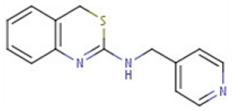
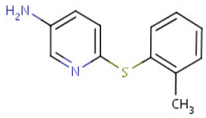
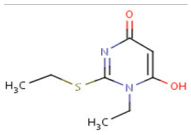


Figure 4. Allosteric effects of fragment hits on the $A_{2A}R$. A) [^{3}H]ZM241385 and B) [^{3}H]NECA kinetic dissociation assay. Fragments were assayed for their effect on the dissociation rate of two $A_{2A}R$ orthosteric ligands: [^{3}H]ZM241385 and [^{3}H]NECA. Dissociation was initiated by the addition of 10 μM ZM241385 or NECA in the absence (control) or presence of 2 mM of the indicated

fragment. C) Dissociation kinetics of [^{3}H]ZM241385 binding to A_{2A}R. Dissociation was initiated by the addition of 10 μM ZM241385 mixed with buffer (control) or the indicated concentration of the fragment. D) The effect of increasing concentration of the fragment ZB1854 on the cAMP production induced by the A_{2A}R specific agonist CGS21680 in HEK293 cells transiently expressing the human A_{2A} receptor. The % of cAMP accumulation is normalized to the effect of CGS21680 at 10 μM.

Table 2. Characterization of fragment hits by kinetic radioligand dissociation assays on the wild-type A_{2A} and A_1 receptors.

ID	Fragment structure	k_{off} (min^{-1}) or % allosteric modulation at 2 mM[a]			
		$A_{2A}R$[b]		A1R[c]	
		[^{3}H] ZM241385	[^{3}H] NECA	[^{3}H] DPCPX	[^{3}H] CCPA
Control		0.011 ± 0.002 (0%)	0.015 ± 0.002 (0%)	0%	0%
Negative Allosteric Modulators for both [^{3}H]ZM241385 and [^{3}H]NECA binding					
ZB418		0.039 ± 0.001 (-36%)	0.022 ± 0.003 (-27%)	-11%	-15%
ZB2044		0.019 ± 0.001 (-21%)	0.032 ± 0.003 (-32%)	-30%	-22%
ZB1153		0.017 ± 0.004 (-18%)	0.021 ± 0.002 (-26%)	-23%	-8%
Positive Allosteric Modulators for [^{3}H]ZM241385 and/or [^{3}H]NECA binding					
ZB335		0.0071 ± 0.0002 (20%)	0.0051 ± 0.0009 (41%)	16%	-6%
ZB1854		0.0063 ± 0.0005 (31%)	0.0095 ± 0.0005 (18%)	1%	-29%
ZB391		0.0066 ± 0.0006 (25%)	0.0085 ± 0.0005 (24%)	2%	-41%

ZB1250		0.0062 ± 0.0004 (30%)	0.0089± 0.0007 (22%)	3%	-14%
ZB268		0.0061 ± 0.0005 (31%)	(8%)	17%	-43%
ZB1703		14%	0.0079 ± 0.0006 (30%)	-8%	48%
ZB114		12%	0.0081 ± 0.0006 (28%)	11%	22%
ZB337		17%	0.0082 ± 0.0005 (26%)	3%	-4%

[a] k_{off} ± SEM (n=3), % allosteric modulation (n=2). [b] The values of the kinetic dissociation rate constants were obtained by analysis of the exponential dissociation curve of [^{3}H]ZM241385 or [^{3}H]NECA bound to human adenosine A_{2A}Rs. [c] % allosteric modulation of [^{3}H]DPCPX and [^{3}H]CCPA binding at human adenosine A_1Rs in the absence (control; 0 %) or presence of 2 mM of the fragment.

Table 3. Effect of ZB1854 on the dissociation of [^{3}H]ZM241385 from A_{2A}R.

Condition	k_{off} (min^{-1}) [a]	Shift [b]
Control ([^{3}H] ZM241385)	0.0121 ± 0.0010	-
+ 0.5 mM ZB1854	0.0102 ± 0.0009	1.2
+ 1 mM ZB1854	0.0072 ± 0.0011	1.7
+ 2 mM ZB1854	0.0058 ± 0.0015	2.1

[a] k_{off} ± SEM (n=3). [b] The shift is defined as the ratio of the k_{off} values in the absence (control) and presence of ZB1854, respectively.

Finally, we determined whether the fragments hits were active in a cell-based assay. We therefore investigated the effect of fragments on A_{2A}R activation by the agonist CGS21680 in a cyclic AMP accumulation assay (Table 4 and Figure 4D). Addition of ZB1854 at three concentrations did not cause a shift in potency of CGS21680 (EC_{50} = 14.1 ± 0.3 nM). However, increasing

concentrations of ZB1854 did result in a dose-dependent increase of the maximal effect (E_{max}) of CGS21680. The change in E_{max} is consistent with the previous data suggesting that ZB1854 acts *via* an allosteric mechanism.[33,34] For other compounds, such as ZB418 (NAM) and ZB335 (PAM) which were tested at these high concentrations, we observed some cell toxicity and therefore it was not possible to confirm their allosteric effects in the cellular assay.

Table 4. Receptor activation by CGS21680 in the presence of ZB1854.

ID	cAMP accumulation assay [a]	
	EC_{50} (nM)	E_{max} (%) [b]
CGS21680	14.1 ± 0.3	100 ± 2
CGS21680 + 0.25 mM ZB1854	14.1 ± 0.3	120 ± 3
CGS21680 + 0.5 mM ZB1854	13.8 ± 0.4	134 ± 9
CGS21680 + 1 mM ZB1854	14.0 ± 0.3	177 ± 13

[a] cAMP accumulation in HEK cells that transiently express the human adenosine $A_{2A}R$. [b] Maximal effect of CGS21680 in the absence or presence of modulator, where CGS21680 in the absence of modulator was set at 100 %. Values are means ± SEM (n=3).

3. Discussion

Traditionally, small molecule modulators of GPCRs have been discovered through pharmacological studies. More recently functional, cell-based assays have been coupled to high-throughput screening to discover novel molecules with selected activities such as positive allosteric modulation.[35,36] However, for a number of reasons such molecules have failed to gain marketing approval. The combination of structure-based drug discovery and fragment-based drug discovery has demonstrated the ability to generate approved drugs with excellent pharmacological properties[37] against soluble targets. It is becoming apparent that FBDD/SBDD can now be applied to GPCRs with great potential advantages.[5]

Use of a direct binding assay, such as TINS, selects for fragments that specifically bind to the target, regardless of the site. In principle it is possible to discover ligands with many different biological activities. Indeed here we have

found orthosteric ligands (although we have not investigated their pharmacology; Table 1), positive allosteric modulators and negative allosteric modulators (Table 2), thereby substantiating the claim. One of the orthosteric ligands, ZB643, constitutes a scaffold in a series of thiadiazole derivatives previously found to interact with other adenosine receptor subtypes.[38] Whether ZB643 can be a lead for further exploration needs careful consideration as other thiadiazoles have been reported to be protein modifiers due to their reactivity.[39] Moreover, this study is one of the first reports of allosteric modulators for the adenosine A_{2A} receptor, with indeed previously unknown chemistry.[40] Giorgi *et al* have provided evidence of allosteric enhancement in a class of substituted 8-azaadenines before.[41] As ZB1854 potentiates the action of an adenosine A_{2A} receptor agonist, its therapeutic potential may be in wound healing and in combating inflammation.[42] Given their apparent low potency, it is not clear whether the allosteric modulators would have been discovered by other methods. However, fragments typically bind in the μM to mM range from where they are developed to the potency ultimately required for *in vivo* efficacy.

The present study used a biophysical technique to discover novel fragment ligands. However, the various biochemical activities of these ligands were characterized by traditional radioligand binding methods and a functional assay. Likely these methods missed multiple, valid fragment ligands simply because their potency was too low to register in the selected assay. In order to take true advantage of the biophysical screening technique, it will be necessary to implement biophysical techniques to characterize the binding mode and discriminate orthosteric from allosteric ligands. This can be accomplished using a competition binding assay in TINS as we have demonstrated for the β_1 adrenergic receptor.[15] Furthermore, the Biophysical Mapping technique developed by Zhukov *et al*,[24] offers the possibility to refine the binding site to significantly better resolution than competition binding studies.

Evolving weakly binding hits to lead-like compounds remains a significant challenge for FBDD. It is well established that with the availability of high-resolution 3D structures of target-fragment complexes, the success and

efficiency of fragment elaboration is quite high.[9] One particular advantage of StaR GPCRs is that they are amenable to X-ray crystallographic analysis and indeed structure based drug design is beginning to be applied to this target class.[5,43] In addition to X-ray, StaRs may enable NMR analysis of protein-small molecule complexes. The NMR resonance assignment and structure elucidation of sensory rhodopsin has recently been accomplished.[44] Although the latter was a significant achievement, it is plausible that the enhanced thermostability and reduced conformational exchange of StaRs may enable resonance assignment, at least for putative allosteric sites at or near the extracellular loop region. If so, the ligand binding site can be mapped at low resolution using chemical shift perturbation techniques.[45] The tremendous increase in crystal structures of GPCRs in various states has had a positive impact on the ability to model small molecule-GPCR interactions.[46] The combination of biophysical mapping and/or NMR-based structural constraints with modeling should also provide considerable aid to the fragment elaboration process.

4. Methods

Materials

Tritiated 4-(2-[7-Amino-2-(2-furyl)-[1,2,4]-triazolo-[2,3-a]-[1,3,5]-triazin-5-yl-amino] ethyl)phenol ([^{3}H]ZM241385, specific activity 30 Ci mmol^{-1}) and tritiated 5'-N-ethylcarboxamido adenosine ([^{3}H]NECA, specific activity 33 Ci mmol^{-1}) were purchased from ARC Inc. (St. Louis, USA). Unlabelled ZM241385, NECA and 2-[p-(2-carboxyethyl)phenethylamino]-5'-N-ethyl-carboxamido adenosine (CGS21680) were purchased from Sigma-Aldrich (Steinheim, Germany). All the compounds in the fragment library are commercially available. A 1D ^{1}H spectrum of each fragment has been collected[25] and is consistent with the manufacturer's claim of > 95% purity. Bovine serum albumin (BSA, fraction V) was obtained from Sigma (St. Louis, MO, USA) and BCA protein assay reagent was from Pierce Chemical Company (Rockford, IL, USA). Human embryonic kidney 293 cells

transiently expressing the wild-type $A_{2A}R$ (HEK293 $A_{2A}R$) were generated as described below. All other chemicals were of analytical grade and obtained from standard commercial sources.

Protein Purification

For biophysical studies, the inverse agonist stabilized A_{2A}-$StaR_2$,[47] including a C-terminal His-10 tag, was expressed in Trichoplusia ni (Tni) cells and purified as previously described.[18] OmpA, including an N-terminal His-6 tag, was expressed in BL21(DE3) cells under the control of the T7 promoter and purified as previously described.[48] Successful refolding of OmpA from inclusion bodies was monitored by SDS-PAGE analysis.[48] Both $A_{2A}R$ and OmpA were buffer-exchanged to PBS (0.1 M sodium phosphate, 0.15 M sodium Chloride, pH 7.4) containing 0.47% n-dodecylphosphocholine (DPC) for OmpA or 0.1% n-decyl-β-D-maltoside (DM) for $A_{2A}R$.

Protein Immobilization

DM-solubilized $A_{2A}R$ and DPC-solubilized OmpA were immobilized on Actigel ALD resin (Sterogene, Carlsbad, CA, USA) via Schiff's base chemistry using the manufacturer's protocol. After overnight incubation at 4 °C, residual unreacted aldehydes on the resin were blocked by addition of 50 mM d_{11}-Tris buffer. Quantification of immobilized protein was monitored by absorption of the supernatant at 280 nm before and after immobilization, and by SDS-page gel with a known standard curve and band volume analysis. This data indicated that a final concentration of 50 μM of immobilized $A_{2A}R$ and 65 μM OmpA was achieved (nmol protein/ml settled bed volume), equating to an 80% and 75% yield respectively. Subsequently, the buffer of both protein samples was exchanged to PBS containing 0.01% n-dodecyl-β-D-maltopyranoside (DDM) for ligand screening experiments.

Target Immobilized NMR Screening

Immobilized, DDM solubilized $A_{2A}R$ (50 μM) and OmpA (65 μM) were each packed into a separate cell of a dual-cell sample holder.[26] Mixes of the 531 fragments were made by 200-fold dilution of a 100 mM stock of each compound in d_6-DMSO such that the final DMSO concentration was never greater than 5%. Upon injection of each mix into the dual-cell sample holder, flow was stopped and spatially selective Hadamard spectroscopy was used to acquire a 1D ^1H spectrum of each sample separately.[27] A CPMG T2 filter of 80 ms was used to remove residual broad resonances from the sepharose resin. To maintain the proper fold of each protein, the screen was performed at 15 ^oC, and 0.01% DDM was included in the buffer (PBS in D_2O) used to wash the fragment mixes from the sample holder.

Pharmacological Characterization of TINS Hits

Radioligand binding and kinetic dissociation studies on the adenosine A_{2A} receptor:

HEK293 cells were maintained in culture in DMEM supplemented with 10% newborn calf serum at 37 ºC in a moist, 7% CO2 atmosphere and passaged twice weekly. Cells were transfected with plasmids containing wild-type $A_{2A}R$ construct using the calcium phosphate precipitation method[49] and harvested after 48 hours. Cells were pelleted, re-suspended in 20 ml of ice-cold 50 mM Tris-HCl buffer, pH 7.4. An ULTRA-TURRAX® was used to homogenize the cell suspension. The cytosolic and membrane fractions were separated using a high-speed centrifugation step of 100,000 x g for 20 min at 4 ºC. This process was repeated twice and subsequently the pellet was re-suspended in 50 mM Tris-HCl buffer, pH 7.4, the protein concentration was determined by the BCA method[50] and samples were aliquoted and stored at - 80 ºC until further use.

Firstly, all fragment hits from the TINS screen were assayed in a [^3H]ZM241385 displacement assay at 500 μM. Fragments with more than 30%

displacement of [^{3}H]ZM241385 were titrated to determine an IC_{50} curve. For displacement binding studies, membranes from cells transiently expressing A_{2A}R-WT (20 µg proteins/well) were incubated in a total volume of 100 µl Tris-HCl buffer (50 mM, pH 7.4) and 5% DMSO with fragments at either a single final concentration of 500 µM or 10 × 0.25 log unit dilutions and a final [^{3}H]ZM241385 (an inverse agonist of the A_{2A}R) concentration of 3 nM or 20 nM of [^{3}H]NECA (an agonist of the A_{2A}R).[42] Non-specific binding was determined using a final concentration of 10 µM ZM241385 and represented less than 10% of the total binding. Total binding was determined in the presence of buffer and was set at 100% in all experiments, whereas non-specific binding was set at 0%. After incubation for 60 min (with [^{3}H]ZM241385) or 150 min (with [^{3}H]NECA) at 25 °C, assays were terminated by rapid filtration through 96-well GF/B UniFilter plates (PerkinElmer) followed by washing with 7 × 0.25 ml ice-cold Tris-HCl buffer (50 mM, pH 7.4). Plates were dried, 25 µl P-E Microscint 20 was added per well and bound radioactivity measured using a Packard Microbeta counter (PerkinElmer). Data were analyzed using GraphPad Prism v5, normalized as '% specific binding' from which the IC_{50} values were calculated. Experiments were performed three times in duplicate, unless stated otherwise. Experimental ligand efficiency (LEAN)[32] was defined as $LE = -RT\ln(IC_{50})/HAC$, where HAC is the number of heavy atoms in the compounds.

Secondly, all TINS fragment hits were assayed in a kinetic A_{2A}R-WT orthosteric radioligand dissociation assay using a single time point at which 50% of the radioligand has dissociated (under control conditions). Fragments that resulted in more than 30% increased or decreased radioligand dissociation (i.e. negative or positive allosteric modulation, respectively) were subsequently analyzed at multiple time points to determine the complete dissociation curve. For kinetic dissociation experiments, 3 nM [^{3}H]ZM241385 or 20 nM [^{3}H]NECA was pre-equilibrated with membranes (30 µg proteins/well) for 3 hours on ice; dissociation was initiated by addition of 10 µM ZM241385 in the absence (control) or presence of 2 mM test compound in a total volume of 5 µl. The

amount of radioligand (~50%) still bound to the receptor was measured after 50 min of dissociation. The amount of specific radioligand binding obtained under control conditions was set at 0% for both positive and negative allosteric modulators (PAMs & NAMs); the total binding (t = 0 min) was set at 100% for PAMs, whereas non-specific binding (in the presence of 10 μM ZM241385) was set at -100% for NAMs. The amount of remaining radioligand bound to the receptor was measured at various time intervals for a total of 4 hours. Incubations were terminated and samples were obtained as described above. The dissociation data were fit globally in GraphPad Prism v5 to determine k_{off}. Experiments were performed three times in duplicate, unless stated otherwise.

Selectivity assays using CHO cell membranes stably expressing human adenosine A_1 receptor:

The receptor subtype selectivity of compounds that were positive in the kinetic dissociation assay with A_{2A}R-WT using [^{3}H]NECA, was tested in [^{3}H]DPCPX[42] (an antagonist of the A_1R) and [^{3}H]CCPA[42] (an agonist of the A_1R) kinetic dissociation assays using wild type human adenosine A_1 receptors stably expressed in CHO cell membranes as described previously in the literature.[51]

Characterization of allosteric modulation of orthosteric ligand activity by intracellular cAMP assay:

HEK293 cells were grown and transfected as described above. Experiments were performed 48h after transfection. The cells were harvested and centrifuged two times at 275 x g for 5 min. For cyclic AMP production and determination, 3000 cells/well were used in 384-well Optiplates (PerkinElmer). The cells were incubated for 45 min at 37 °C with either the A_{2A}R agonist CGS21680 alone or CGS21680 together with test compounds in different concentrations. The assay buffer used was PBS with the addition of 5 mM HEPES, 0.1% BSA, 50 μM rolipram, 50 μM cilostamide and 0.8 IU/ml adenosine deaminase. Basal activity was determined in the presence of assay buffer and was set at 0% in all experiments. Maximal receptor activity was determined in the

presence of 10 μM CGS21680 and was set at 100% in all experiments. Cells were then lysed and the amount of cAMP produced was quantified using a LANCE® ultra cAMP kit (PerkinElmer) according to the instructions of the manufacturer. Following addition of the detection mixture, plates were left for 1 hour at room temperature prior to reading using an EnVision® plate reader (PerkinElmer Life Sciences).

Fragment Collection

The fragment collection consists of about 1,500 compounds derived from a variety of vendors. The philosophy used to design the collection as well as the physicochemical properties have been described.[25] For quality control purposes as well as for formatting the library for TINS, a 1D ^{1}H NMR spectrum of each compound was acquired and compared to the expected spectrum based on the structure provided by the vendor. Compounds whose purity was less than the 95% indicated by the vendor were rejected. A randomly selected sub-set of the complete collection was chosen to provide proof of principle while reducing the amount of subsequent characterization work to a practical level.

References

1. Filmore, D. (2004) It's a GPCR world. Modern Drug Discov. 7, 24–27.
2. Jacoby, E., Bouhelal, R., Gerspacher, M., and Seuwen, K. (2006) The 7 TM G-protein-coupled receptor target family. ChemMedChem 1, 761–782.
3. Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006) How many drug targets are there? Nat. Rev. Drug Discov. 5, 993–996.
4. Houston, J. G., Banks, M. N., Binnie, A., Brenner, S., O'Connell, J., and Petrillo, E. W. (2008) Case study: impact of technology investment on lead discovery at Bristol-Myers Squibb. 1998-2006. Drug Discov. Today 13, 44–51.
5. Congreve, M., Langmead, C. J., Mason, J. S., and Marshall, F. H. (2011) Progress in structure based drug design for G protein-coupled receptors. J. Med. Chem. 54, 4283–4311.
6. Gribbon, P., and Sewing, A. (2005) High-throughput drug discovery: what can we expect from HTS? Drug Discov. Today 10, 17–22.
7. Congreve, M., Carr, R., Murray, C., and Jhoti, H. (2003) A "Rule of Three" for fragment-based lead discovery? Drug Discov. Today 8, 876–877.
8. Rees, D. C., Congreve, M., Murray, C. W., and Carr, R. (2004) Fragment-based lead discovery. Nat. Rev. Drug Discov. 3, 660–672.
9. Hajduk, P. J., and Greer, J. (2007) A decade of fragment-based drug design: strategic advances and lessons learned. Nat. Rev. Drug Discov. 6, 211–219.
10. Wang, Y.-S., Strickland, C., Voigt, J. H., Kennedy, M. E., Beyer, B. M., Senior, M. M., Smith, E. M., Nechuta, T. L., Madison, V. S., Czarniecki, M., McKittrick, B. a, Stamford, A. W., Parker, E. M., Hunter, J. C., Greenlee, W. J., and Wyss, D. F. (2010) Application of fragment-based NMR screening, X-ray crystallography, structure-based design, and focused chemical library design to identify novel mM leads for the development of nM BACE-1 (b-site APP cleaving enzyme 1) inhibitors. J. Med. Chem. 53, 942–950.
11. Wyss, D. F., Wang, Y., Eaton, H. L., Strickland, C., Voigt, J. H., Zhu, Z., and Stamford, A. W. (2012) Combining NMR and X-ray Crystallography in Fragment-Based Drug Discovery: Discovery of Highly Potent and Selective BACE-1 Inhibitors. Top. Curr. Chem. 317, 83–114.
12. Wells, J. a, and McClendon, C. L. (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature 450, 1001–1009.
13. Congreve, M., Chessari, G., Tisi, D., and Woodhead, A. J. (2008) Recent developments in fragment-based drug discovery. J. Med. Chem. 51, 3661–3680.
14. Früh, V., Zhou, Y., Chen, D., Loch, C., Ab, E., Grinkova, Y. N., Verheij, H., Sligar, S. G., Bushweller, J. H., and Siegal, G. (2010) Application of fragment-based drug discovery to membrane proteins: identification of ligands of the integral membrane enzyme DsbB. Chem. Biol. 17, 881–891.
15. Congreve, M., Rich, R. L., Myszka, D. G., Figaroa, F., Siegal, G., and Marshall, F. H. (2011) Fragment screening of stabilized G-protein-coupled receptors using biophysical methods. In Method. Enzymol. 493, 115–136.
16. Vanwetswinkel, S., Heetebrij, R. J., van Duynhoven, J., Hollander, J. G., Filippov, D. V., Hajduk, P. J., and Siegal, G. (2005) TINS, target immobilized NMR screening: an efficient and sensitive method for ligand discovery. Chem. Biol. 12, 207–216.
17. Magnani, F., Shibata, Y., Serrano-Vega, M. J., and Tate, C. G. (2008) Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor. Proc. Natl. Acad. Sci. U.S.A. 105, 10744–10749.
18. Robertson, N., Jazayeri, A., Errey, J., Baig, A., Hurrell, E., Zhukov, A., Langmead, C. J., Weir, M., and Marshall, F. H. (2011) The properties of thermostabilised G protein-coupled receptors (StaRs) and their use in drug discovery. Neuropharmacology 60, 36–44.
19. Serrano-vega, M. J., Magnani, F., Shibata, Y., and Tate, C. G. (2008) Conformational thermostabilization of the b1-adrenergic receptor in a detergent-resistant form, Proc. Natl. Acad. Sci. U.S.A. 105, 877–882.
20. Shibata, Y., White, J. F., Serrano-vega, M. J., Magnani, F., Amanda, L., Grisshammer, R., and Tate, C. G. (2009) Thermostabilisation of the neurotensin receptor NTS1. J. Med. Chem. 390, 262–277.
21. Lebon, G., Bennett, K., Jazayeri, A., and Tate, C. G. (2011) Thermostabilisation of an agonist-bound conformation of the human adenosine A(2A) receptor. J. Mol. Biol. 409, 298–310.
22. Rich, R. L., Errey, J., Marshall, F., and Myszka, D. G. (2011) Biacore analysis with stabilized G-protein-coupled receptors. Anal. Biochem. 409, 267–272.
23. Warne, T., Serrano-Vega, M. J., Baker, J. G., Moukhametzianov, R., Edwards, P. C., Henderson, R., Leslie, A. G. W., Tate, C. G., and Schertler, G. F. X. (2008) Structure of a b1-adrenergic G-protein-coupled receptor. Nature 454, 486–491.
24. Zhukov, A., Andrews, S. P., Errey, J. C., Robertson, N., Tehan, B., Mason, J. S., Marshall, F. H., Weir, M., and Congreve, M. (2011) Biophysical mapping of the adenosine A2A receptor. J. Med. Chem. 54, 4312–4323.
25. Siegal, G., Ab, E., and Schultz, J. (2007) Integration of fragment screening and library design. Drug Discov. Today 12, 1032–1039.
26. Marquardsen, T., Hofmann, M., Hollander, J. G., Loch, C. M. P., Kiihne, S. R., Engelke, F., and Siegal, G. (2006) Development of a dual cell, flow-injection sample holder, and NMR probe for comparative ligand-binding studies. J. Magn. Reson. 182, 55–65.
27. Murali, N., Miller, W. M., John, B. K., Avizonis, D. a, and Smallcombe, S. H. (2006) Spectral unraveling by space-selective Hadamard spectroscopy. J. Magn. Reson. 179, 182–189.

28. Clementina, M., and Giuseppe, S. (2010) A2A receptor ligands: past, present and future trends. Curr. Top. Med. Chem. 10, 902–922.
29. Müller, C. E., and Jacobson, K. A. (2011) Xanthines as adenosine receptor antagonists. In Methylxanthines. In Handbook of Experimental Pharmacology, B.B. Fredholm, ed. 200, 151–199.
30. Müller, C. E., and Jacobson, K. A. (2011) Recent developments in adenosine receptor ligands and their potential as novel drugs. BBA-Biomembranes 1808, 1290–1308.
31. Müller, C. E., and Ferré, S. (2007) Blocking striatal adenosine A2A receptors: a new strategy for basal ganglia disorders. Recent Pat. CNS Drug Discov. 2, 1–21.
32. May, P. C., Dean, R. a, Lowe, S. L., Martenyi, F., Sheehan, S. M., Boggs, L. N., Monk, S. a, Mathes, B. M., Mergott, D. J., Watson, B. M., Stout, S. L., Timm, D. E., Smith Labell, E., Gonzales, C. R., Nakano, M., Jhee, S. S., Yen, M., Ereshefsky, L., Lindstrom, T. D., Calligaro, D. O., Cocke, P. J., Greg Hall, D., Friedrich, S., Citron, M., and Audia, J. E. (2011) Robust central reduction of amyloid-β in humans with an orally available, non-peptidic β-secretase inhibitor. J. Neurosci. 31, 16507–16516.
33. Kenakin, T., and Christopoulos, A. (2011) Analytical pharmacology: the impact of numbers on pharmacology. Trends Pharmacol. Sci. 32, 189–196.
34. Keov, P., Sexton, P. M., and Christopoulos, A. (2011) Allosteric modulation of G protein-coupled receptors: a pharmacological perspective. Neuropharmacology 60, 24–35.
35. Urwyler, S. (2011) Allosteric Modulation of Family C G-Protein-Coupled Receptors: from Molecular Insights to Therapeutic Perspectives. Pharmacol. Rev. 63, 59–126.
36. Conn, P. J., Christopoulos, A., and Lindsley, C. W. (2009) Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. Nat. Rev. Drug Discov. 8, 41–54.
37. Yang, H., Higgins, B., Kolinsky, K., Packman, K., Go, Z., Iyer, R., Kolis, S., Zhao, S., Lee, R., Grippo, J. F., Schostack, K., Simcox, M. E., Heimbrook, D., Bollag, G., and Su, F. (2010) RG7204 (PLX4032), a selective BRAFV600E inhibitor, displays potent antitumor activity in preclinical melanoma models. Cancer Res. 70, 5518–5527.
38. van Muijlwijk-Koezen, J. E., Timmerman, H., Vollinga, R. C., Frijtag von Drabbe Künzel, J., de Groote, M., Visser, S., and IJzerman, A. P. (2001) Thiazole and thiadiazole analogues as a novel class of adenosine receptor antagonists. J. Med. Chem. 44, 749–762.
39. Göblyös A, de Vries H, Brussee J, IJzerman AP. (2005) Synthesis and biological evaluation of a new series of 2,3,5-substituted [1,2,4]-thiadiazoles as modulators of adenosine A1 receptors and their molecular mechanism of action. J. Med. Chem. 48, 1145-1151.
40. Göblyös, A., and Ijzerman, A. P. (2011) Allosteric modulation of adenosine receptors. Biochimica et biophysica acta, 1808, 1309–1318.
41. Giorgi, I., Biagi, G., Bianucci, A. M., Borghini, A., Livi, O., Leonardi, M., Pietra, D., Calderone, V., and Martelli, A. (2008) N6-1,3-diphenylurea derivatives of 2-phenyl-9-benzyladenines and 8-azaadenines: synthesis and biological evaluation as allosteric modulators of A2A adenosine receptors. Eur. J. Med. Chem. 43, 1639–1647.
42. Fredholm, B. B., Ijzerman, A. P., Jacobson, K. A., Linden, J., and Müller, C. E. (2011) International Union of Basic and Clinical Pharmacology. LXXXI. Nomenclature and Classification of Adenosine Receptors — An Update. Pharmacol. Rev. 63, 1–34.
43. Congreve, M., Andrews, S. P., Dore, A. S., Hollenstein, K., Hurrell, E., Langmead, C. J., Mason, J. S., Ng, I. W., Tehan, B., Zhukov, A., Weir, M., and Marshall, F. H. (2012) Discovery of 1,2,4-Triazine Derivatives as Adenosine A2A Antagonists using Structure Based Drug Design. J. Med. Chem. 55, 1898–1903.
44. Gautier, A., Mott, H. R., Bostock, M. J., Kirkpatrick, J. P., and Nietlispach, D. (2010) Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. Nat. Struct. Mol. Biol. 17, 768–774.
45. Schieborr, U., Vogtherr, M., Elshorst, B., Betz, M., Grimme, S., Pescatore, B., Langer, T., Saxena, K., and Schwalbe, H. (2005) How much NMR data is required to determine a protein-ligand complex structure? ChemBioChem 6, 1891–1898.
46. Istyastono, E. P., Nijmeijer, S., Lim, H. D., Stolpe, A. V. D., Roumen, L., Kooistra, A. J., Vischer, H. F., Esch, I. J. P. D., Leurs, R., and Graaf, C. D. (2011) Molecular Determinants of Ligand Binding Modes in the Histamine H4 receptor: Linking ligand-based Three-Dimensional quantitative Structure−Activity Relationship (3D-QSAR) Models to in Silico Guided Receptor Mutagenesis Studies. J. Med. Chem. 54, 8136–8147.
47. Doré, A. S., Robertson, N., Errey, J. C., Ng, I., Hollenstein, K., Tehan, B., Hurrell, E., Bennett, K., Congreve, M., Magnani, F., Tate, C. G., Weir, M., and Marshall, F. H. (2011) Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthines XAC and caffeine. Structure 19, 1283–1293.
48. Arora, a, Rinehart, D., Szabo, G., and Tamm, L. K. (2000) Refolded outer membrane protein A of Escherichia coli forms ion channels with two conductance states in planar lipid bilayers. J. Biol. Chem. 275, 1594–1600.
49. Kingston, R. E., Chen, C. A., Okayama, H., and Rose, J. K. (2003) Calcium Phosphate Transfection. Curr. Protocols Mol. Biol. 63, 9.1.1–9.1.11.
50. Mallia, A. K., Frovenzano, M. D., Fujimoto, E. K., Olson, B. J., Klenk, D. C., and Company, P. C. (1985) Measurement of Protein Using Bicinchoninic Acid. Anal. Biochem. 150, 76–85.
51. Nieuwendijk, A. M. C. H. V. D., Pietra, D., and Heitman, L. (2004) Synthesis and Biological Evaluation of 2,3,5-Substituted [1,2,4]Thiadiazoles as Allosteric Modulators of Adenosine Receptors. J. Med. Chem. 47, 663–672.

Chapter 4

Complementarity Between *in silico* and Biophysical Screening Approaches in Fragment-Based Lead Discovery Against the A_{2A} Adenosine Receptor

D. Chen, A. Ranganathan, A. P. IJzerman, G. Siegal, J. Carlsson. Complementarity between *in silico* and biophysical screening approaches in fragment-based lead discovery against the A_{2A} adenosine receptor. *J. Chem. Inf. Model.* 2013, 53, 2701-2714.

ABSTRACT

Fragment-based lead discovery (FBLD) is becoming an increasingly important method in drug development. We have explored the potential to complement NMR-based biophysical screening of chemical libraries with molecular docking in FBLD against the A_{2A} adenosine receptor ($A_{2A}R$), a drug target for inflammation and Parkinson's disease. Prior to an NMR-based screen of a fragment library against the $A_{2A}R$, molecular docking against a crystal structure was used to rank the same set of molecules by their predicted affinities. Molecular docking was able to predict four out of the five orthosteric ligands discovered by NMR among the top 5% of the ranked library, suggesting that structure-based methods could be used to prioritize among primary hits from biophysical screens. In addition, three fragments that were top-ranked by molecular docking, but had not been picked up by the NMR-based method, were demonstrated to be $A_{2A}R$ ligands. While biophysical approaches for fragment screening are typically limited to a few thousand compounds, the docking screen was extended to include 328,000 commercially available fragments. Twenty-two top-ranked compounds were tested in radioligand binding assays, and 14 of these were $A_{2A}R$ ligands with K_i values ranging from 2 to 240 μM. Optimization of fragments was guided by molecular dynamics simulations. The results illuminate strengths and weaknesses of molecular docking and demonstrate that this method can serve as a valuable complementary tool to biophysical screening in FBLD.

1. Introduction

The adenosine receptors belong to the family of G protein-coupled receptors (GPCRs) and are expressed both in the central nervous system (CNS) and the periphery. The four adenosine receptor subtypes (A_1, A_{2A}, A_{2B}, A_3) are responsible for a wide range of physiological processes by acting on different signaling pathways. The A_{2A} and A_{2B} adenosine receptors increase cAMP levels by coupling to G_s whereas the A_1 and A_3 subtypes signal via G_i and decrease cAMP levels.[1] A_{2A} adenosine receptor ($A_{2A}R$) activation occurs in response to stress or cell damage and protects tissues by controlling inflammation. For this reason, $A_{2A}R$ agonists are candidates for development of anti-inflammatory drugs.[2] In the CNS, the $A_{2A}R$ is expressed in the basal ganglia and counter-balances the action of the D_2 dopamine receptor. The two receptors form heterodimers in the cell membrane and $A_{2A}R$ activation leads to a decrease in D_2 dopamine receptor signaling.[3] Since Parkinson's disease is characterized by a loss of dopamine receptor activity, $A_{2A}R$ antagonists have been studied intensively as drug candidates for Parkinson's disease and several compounds are currently in clinical trials.[1]

The rapidly increasing number of crystal structures for G protein-coupled receptors (GPCRs) has made structure-based approaches an attractive strategy for drug design against these pharmaceutically important targets.[4,5] Twelve high-resolution structures, with bound antagonists or agonists, have been determined for the $A_{2A}R$ since the first one was solved in 2008.[6-9] Structure-based computational ligand discovery based on these structures has been surprisingly successful. Two recent molecular docking screens against these structures have resulted in the discovery of several novel $A_{2A}R$ antagonists.[10,11] The "hit rates" in these screens, *i.e.* the percentages of active compounds among those selected for experimental evaluation, were 35 and 41%, which is remarkably good for virtual screening. Similar results were achieved in docking screens against the β_2 adrenergic[12,13], D_3 dopamine[14], CXCR4[15], and H_1 histamine receptors[16], which

further demonstrated the potential for structure-based drug design against these targets.

During the last decade, fragment-based lead discovery (FBLD) has steadily increased in popularity as an alternative to high-throughput screening (HTS).[17,18] The idea behind FBLD is to identify ligands that are about half the size of a drug. The increased coverage of chemical space in fragment libraries makes it possible to identify ligands from screens of only a few thousand compounds. Ligands from fragment screens are often of low affinity and lack selectivity, but can serve as a starting point for the development of new lead compounds. There are now a large number of successful examples of where fragment ligands identified by crystallography, surface plasmon resonance (SPR), or nuclear magnetic resonance (NMR) techniques have been used to generate high-affinity leads.[17] In fact, several compounds that originate from fragment screens are now in clinical trials and the first drug developed from FBLD was approved in 2011.[19] Fragment screens by NMR and SPR have also been carried out against GPCRs[20,21], but this is still challenging because of difficulties in expressing these proteins in large quantities and their inherent instability outside the cell membrane. The use of *in silico* fragment-based screening has also been explored successfully, both for GPCRs[16,22] and soluble proteins[23,24], but structure-based approaches have been questioned in this context[25]. Most molecular docking programs were developed for and benchmarked on drug-like compounds and thus may have difficulties to rank fragment ligands, which typically are of low affinity and only occupy subpockets of a binding site.

In this work, we explored fragment-based ligand discovery by carrying out two prospective docking screens against a crystal structure of the $A_{2A}R$. The questions that we wanted to address were if molecular docking could be used to discover fragment ligands and if *in silico* methods in combination with biophysical screening methods could improve the efficiency of FBLD. In a first step, we computationally docked a set of 500 fragments to an $A_{2A}R$ crystal structure and ranked these compounds by affinity prior to an NMR-based screen of the same library. As the entire library was screened both *in silico* and experimentally, our

blind prediction of the results made it possible to evaluate the utility of molecular docking in FBLD. The comparison of the computational and experimental screens for the small fragment library showed that a majority of the orthosteric ligands were among the top-ranked compounds in the docking screen. In a second step, we extended the docking screen to include several hundred thousand commercially available molecules and tested top-ranked compounds experimentally. The results from this second screen allowed us to directly compare the fragment ligands and hit-rates from biophysical and structure-based computational screening approaches. Fourteen fragment ligands were discovered from the docking screen of this larger set of commercially available molecules. The structure-activity relationships (SAR) were explored for three novel $A_{2A}R$ ligands, guided by molecular docking and molecular dynamics (MD) simulations.

2. Results

Ranking of a Fragment Library by Molecular Docking

To test the performance of molecular docking in FBLD against the $A_{2A}R$, parallel experimental and computational screens of an in-house library of 500 chemically diverse fragments were carried out. Prior to screening the fragment library using an NMR-based approach, the same set of molecules was computationally docked to the orthosteric site of an antagonist-bound crystal structure of the $A_{2A}R$[7] using the program DOCK3.6[26-28]. Each molecule was sampled in, on average, 18700 orientations and 16 conformations. For each conformation that did not clash with the receptor, a physics-based scoring function was used to evaluate the complementarity of the fragment to the orthosteric site. The docking score was calculated from the sum of the receptor-ligand electrostatic and van der Waals interaction energy, corrected for desolvation of the ligand. The best scoring conformation for each fragment was used to rank these molecules by their predicted affinity to the $A_{2A}R$. In total, 487 of the 500 fragments were successfully scored in the orthosteric binding site. The

average computation time per docked molecule was 10 seconds and, in total, only 80 minutes were required to screen the library.

Comparison of NMR-Based and Molecular Docking Screens of a Fragment Library

Given the low affinity and specificity of typical fragment ligands, the molecular docking technique can be expected to have difficulties to rank such compounds by affinity. As every library molecule was tested for binding to the $A_{2A}R$ in the NMR-based screen, this gave a unique opportunity to assess the ability of molecular docking to discriminate between fragment ligands and experimentally verified non-binders.

The details of the methodology and experiments used to screen the fragment library against a thermostabilized mutant[29] of the $A_{2A}R$ by NMR have been presented in separate publications.[21,30] The results of the NMR-based screen of the fragment library are presented briefly in this section to enable a comparison to the docking predictions. Five hundred fragments, which were selected to represent a diverse set of molecules in terms of shape and chemistry[31], were screened using the target immobilized NMR screening (TINS) technique[30]. The TINS screen resulted in 94 primary hits, which were further evaluated in radioligand displacement assays. Five molecules showed more than 30% displacement of the radioligand at 500 μM and had well-behaved dose response curves, which corresponds to a hit-rate of 5% for the secondary screen. The five hits are shown in Table 1 and the K_i values for these compounds ranged from 14 to 600 μM. Another key metric for judging the potential of a fragment ligand is its ligand efficiency (LE). The LE value was calculated as $RTlnK_i/N_i$, where N_l is the number of heavy atoms.[32,33] The best LE value for the fragments discovered in the NMR-based screen was 0.53 kcal mol^{-1} $atom^{-1}$, which is in the range considered to be promising for further optimization (typically >0.35 kcal mol^{-1} $atom^{-1}$). Additional experiments revealed that 11 of the initial hits were either positive or negative allosteric modulators of the $A_{2A}R$.[21] The remaining 79 hits were not found to displace radioligands from the orthosteric site or

allosterically modulate agonist or antagonist binding. These compounds were considered to be less potent fragment ligands, bind to regions that do not influence ligand binding to the $A_{2A}R$ orthosteric site, or be false positives.

To assess how well the docking was able to predict the results of the fragment screen, we generated receiver operating characteristic (ROC) curves for the identified ligands based on the predicted ranking for the compounds in the library. As the docking calculations only sampled fragment orientations in the orthosteric site, we did not expect to identify any allosteric ligands. For this reason, the ligands were divided into an orthosteric and allosteric group, which made it possible to directly compare the computational and NMR-based screening results. The ROC curves for the ortho- and allosteric ligands are shown in Figure 1. To quantify the ability of DOCK3.6 to identify ligands, we evaluated enrichment factors at different false positive rates ($ROC\_EF_{x\%}$, where x denotes the false positive rate).[34] For example, the $ROC\_EF_{5\%}$ value was calculated as the fraction of true positives (ligands) identified when 5% of the false positives (non-binders) had been found in the ranked list divided by the fraction of false positives at that point, i.e. 0.05 in this case.[35] No enrichment of orthosteric ligands was found at 0.5%-1% false positive rate. One ligand was identified at a 2% false-positive rate, corresponding to an enrichment factor of 10. Four out of five orthosteric ligands were identified in the top 5% of the ranked database, which corresponded to an $ROC\_EF_{5\%}$ equal to 16, i.e. a 16-fold improvement over random selection (Figure 1, red curve). The fifth orthosteric ligand, compound **4**, was not scored at all, which was due to the sampling scheme used in DOCK3.6. If the sampling parameters were (retrospectively) adjusted to determine a score for compound **4**, it was ranked as ~160 and would hence not have affected early enrichment. The enrichment of orthosteric ligands from the subset of 94 primary hits in the NMR-based screen was also calculated. The enrichment in this case was remarkably good, ranking four of the ligands in the very top of the database (Figure 1, gray curve, $ROC\_EF_{0.5\%}$ = 120). Finally, as may have been expected, the enrichment of allosteric modulators was poor,

as reflected by that the ROC curve was closer to random enrichment (Figure 1, blue curve). Predicted binding poses for two of the orthosteric ligands from the NMR-based screen, compounds **1** and **5**, are shown in Figure 2. Both compounds interact with Asn253$^{6.55}$ of the $A_{2A}R$, a residue that been shown to be critical for recognition of both agonists and antagonists.[36]

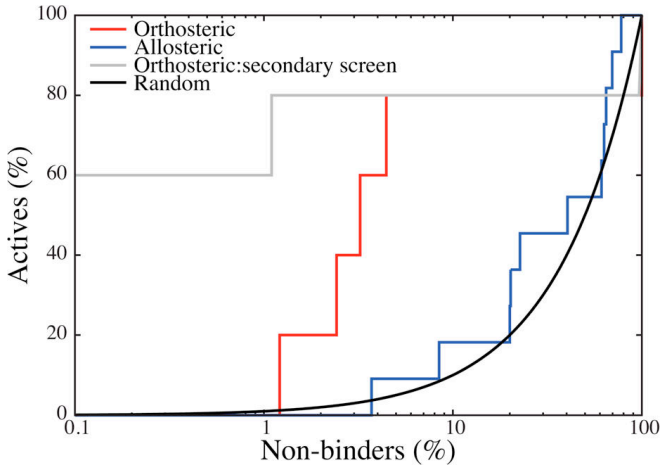


Figure 1. ROC curves for docking enrichment of orthosteric (red line) and allosteric (blue line) ligands discovered in the NMR-based screen. The enrichment of orthosteric ligands among the 94 molecules from the NMR-based screen that were selected for follow-up radioligand assays is also shown (grey line). The black line represents the curve expected from random enrichment.

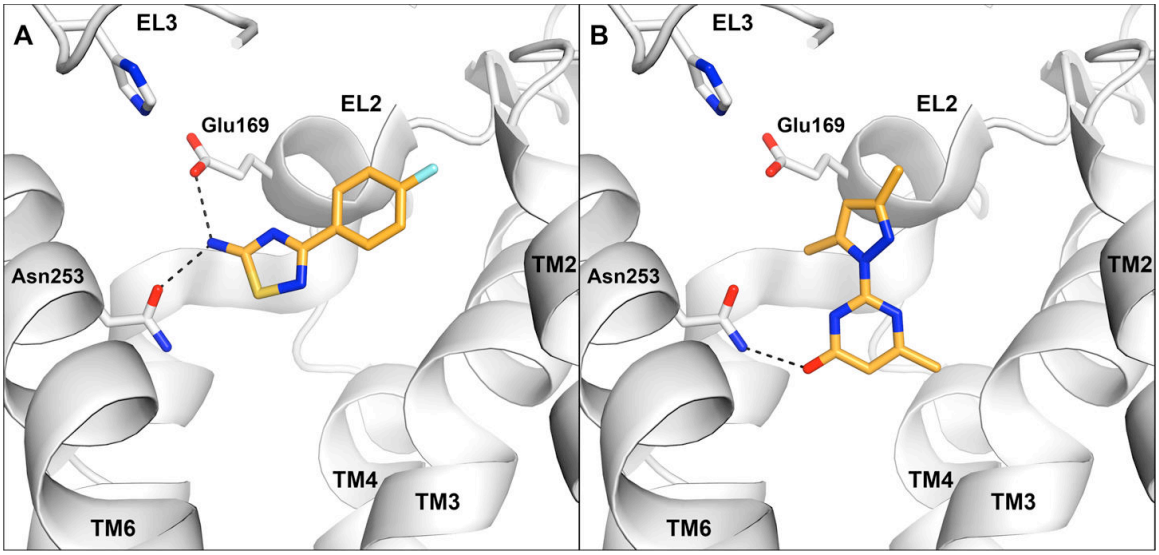


Figure 2. Predicted binding modes for two ligands (**1** and **5**) that were discovered in the NMR-based screen. The binding site is shown in white ribbons with selected side chains shown in sticks. The ligand is depicted with orange carbon atoms. Black dotted lines indicate hydrogen bonds. This figure was generated with PyMOL (version 1.4.1).

After assessing the ability of molecular docking to predict the results of the NMR-based screen, the top-ranked fragments from the docking calculations were analyzed in more detail. It was encouraging that 80% of the orthosteric ligands were among the top 50 compounds of the ranked library. However, this also suggested that there were 46 top-ranked inactive compounds for which the docking algorithm had predicted similar or better binding energies. To try to understand the origin of these false positives from docking, the pose for each of the top-ranked 50 compounds was inspected visually. If general problems with the scoring function could be identified, this information could be valuable for further development of the molecular docking algorithm. Two observations were made from this analysis. First of all, there were several cases where polar groups of the fragments or the receptor did not form hydrogen bonds, but were still desolvated to a large extent, which should be energetically unfavorable. Whereas the ligand desolvation is taken into account by the scoring function in DOCK3.6, the receptor desolvation energy contribution is ignored.[37] This may be a reasonable approximation for lead- or drug-like compounds that fill the entire binding site, resulting in a constant desolvation energy contribution. Fragments, on the other hand, often occupy subpockets of the binding site and in these cases the desolvation term could vary significantly and thus be crucial for accurate ranking.[38] The second observation was that there were several top-ranked fragments for which no obvious problems could be identified. Instead, we would likely have predicted several of them to be ligands in a prospective screen against the $A_{2A}R$. To test if any of these were false negatives from the NMR-based screen, we selected five compounds from the top 10% of the ranked library for experimental re-screening in radioligand displacement assays. These five molecules had all been screened by the NMR-based method, but were not among the 94 compounds that were tested in follow-up assays. Three out of five compounds showed >50% displacement of radioligand binding at 500 μM. The dose-response curves demonstrated that they were orthosteric ligands with K_i values equal to 17.6, 20.6 and 128 μM (Compounds **6-8**, Table 1). Compound **6** was ranked as number 34 in the docking screen and had the highest LE of all

$A_{2A}R$ fragment ligands discovered by docking or NMR (LE = 0.56 kcal mol^{-1} $atom^{-1}$) and the predicted binding mode for this compound is shown in Figure 3A. Given the potency of these three fragments, it is possible that they have relatively slow off-rates rendering them less sensitive to ligand observed NMR techniques such as TINS.

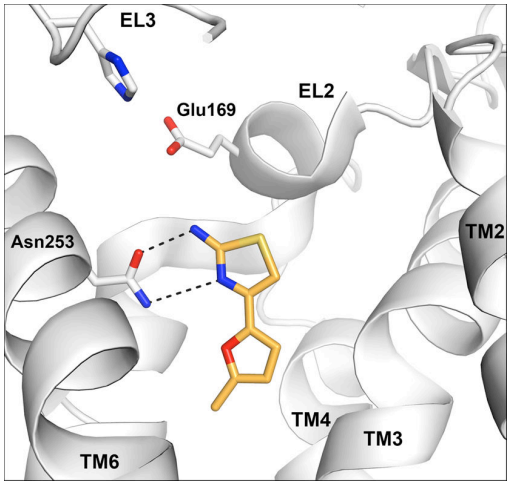


Figure 3. Predicted binding mode for a fragment ligand (compound **6**) from the small fragment library that was discovered by docking. The binding site is displayed as a white cartoon with selected side chains shown in sticks. The ligand is depicted with orange carbon atoms. Black dotted lines indicate hydrogen bonds. This figure was generated with PyMOL (version 1.4.1).

Molecular Docking Screen of Commercially Available Fragments Against an $A_{2A}R$ Crystal Structure

Although the positive enrichment of $A_{2A}R$ ligands in the fragment library was encouraging, a more interesting test of docking performance would be to extend the number of molecules screened to all commercially available or even all synthetically feasible fragments, *e.g.* the 166 billion compounds in the GDB-17.[39] One major strength of the *in silico* approach comes from the ability to screen large libraries and reduce the fraction that is tested experimentally to a small number of chemicals. This would enable a more realistic comparison of a typical experimental fragment-based screen to a computational structure-based approach. Screening of a larger chemical library could be also advantageous because of better coverage of chemical space, which gives the opportunity to discover ligand chemotypes that were not available to the NMR-based screen. On the other hand, ranking several hundred thousand disparate fragments by affinity is a tremendous challenge and, as only a few hundred top-ranked

compounds are considered for experimental testing, one may actually expect that molecular docking would perform worse for screening of large chemical libraries.

To test the utility of large-scale docking screens in FBLD, we docked 328,000 commercially available fragments from the ZINC library[40] to an $A_{2A}R$ crystal structure using DOCK3.6. The same receptor structure, sampling scheme, and physics-based scoring function as in the first screen were used. The docking of the library was completed in seven hours on 222 cores at a local computer cluster and each molecule was sampled in, on average, 12800 orientations and 40 conformations. The top-ranked 500 compounds, corresponding to 0.15% of the screened library, were considered for experimental testing. There was essentially no overlap of compounds between this set of molecules and the library screened by the NMR-based method. Only one molecule from the small library screened by NMR would have ranked among the top 500 in the docking screen. Twenty-two top ranked molecules from the docking screen of large fragment library (0.007%) were selected for experimental testing based on their complementarity to the orthosteric site and availability from commercial vendors.

The twenty-two predicted fragment ligands were first tested in radioligand assays at 500 μM and 14 molecules showed significant radioligand displacement. Subsequent dose-response curves were well behaved (Figure 4), and the K_i values ranged from 2 to 240 μM (Table 2, compounds **9-22**). With the same affinity cut-off as in the NMR-based screen (K_i < 600 μM), the hit-rate from the docking screen was 64%. Although it is difficult to compare screens against different GPCRs, our results were similar to the fragment-based docking campaign against the histamine H_1 receptor, which achieved an unprecedented hit-rate of 73% and identified several potent ligands.[16] Compared to the two previous docking screens of lead-like libraries against the $A_{2A}R$,[10,11] we achieved higher hit-rates, but the affinities of the discovered ligands were typically lower, as expected for fragment screening. To enable a comparison between the hits from our fragment screen and those discovered from docking of lead-like libraries against the same crystal structure, LE values (at 310 K) were calculated for each study. The LE values for the fragments from our *in silico* screen ranged from 0.29

to 0.56 kcal mol^{-1} $atom^{-1}$ with an average equal to 0.40, which was similar to the results of *Carlsson et al.* (0.29−0.44 kcal mol^{-1} $atom^{-1}$, average = 0.37)[10] and *Katrich et al.* (0.32−0.52 kcal mol^{-1} $atom^{-1}$, average = 0.41).[11]

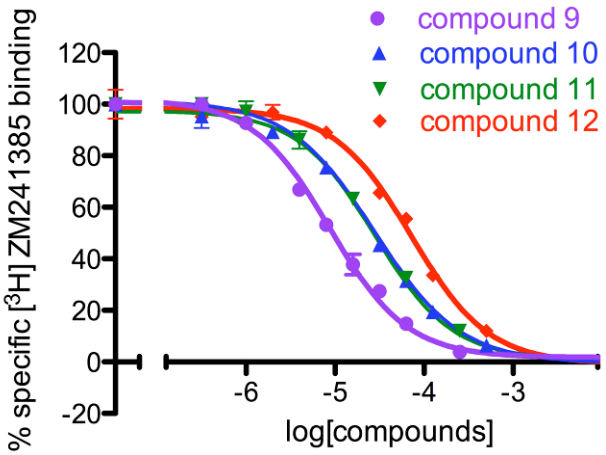


Figure 4. Representative concentration-effect curves for displacement of binding of the radiolabeled $A_{2A}R$ inverse agonist [^{3}H]ZM241385 by compounds **9**, **10**, **11**, and **12**.

The NMR-based and docking screens are summarized in Figure 5. Five orthosteric ligands were identified in the NMR-based screen, corresponding to a hit-rate of 5% from the secondary screen, and 14 fragment ligands or 64% of those tested experimentally were discovered by docking. Compared to the NMR-based screen, the hits from the $A_{2A}R$ docking screen had lower K_i values. Four of the hits from the docking screen were more potent than the compound with the highest affinity from the NMR-based screen. A similar trend was observed for the LE values. Ten compounds from the docking screen had LE values >0.35 kcal mol^{-1} $atom^{-1}$. In comparison, two compounds from the NMR-based screen reached this level. To investigate the novelty of the discovered fragment, we calculated the pairwise Tanimoto similarity with extended chemical fingerprints for four atoms (T_c, ECFP4) of each hit to the thousands of known $A_{2A}R$ ligands in the ChEMBL14[41] database. The Tanimoto coefficient (T_c) quantifies the two-dimensional chemical similarity between two molecules by a value between 0 and 1. A T_c value close to zero suggests no chemical similarity between a pair of molecules, whereas a value equal to one represents two identical molecules. For

each discovered ligand, the highest T_c value among the previously characterized $A_{2A}R$ ligands is presented in Tables 1 and 2. Some of the hits from the NMR-based and docking screens were relatively similar to known adenosine receptor ligands, *e.g.* compounds **4** and **9** ($T_c > 0.4$), but in both cases relatively novel chemotypes for the $A_{2A}R$ were also discovered (*e.g.* compounds **2, 5, 11** and **14**), as reflected by their lower Tanimoto coefficients (T_c = 0.28-0.33). Predicted binding modes for compounds **10, 11, 12,** and **13** are shown in Figure 6 and 7. As shown in Figure 6, these compounds were predicted to bind deep in the orthosteric binding site. The key interactions for the fragments are hydrogen bonds to Asn253$^{6.55}$ and Glu169$^{5.30}$ together with π-stacking against the side chain of Phe168$^{5.29}$ (superscripts denote Ballesteros-Weinstein numbers[42]), which is consistent with the ligand interactions observed in several of the twelve available $A_{2A}R$ crystal structures (Figure 7).[6-9]

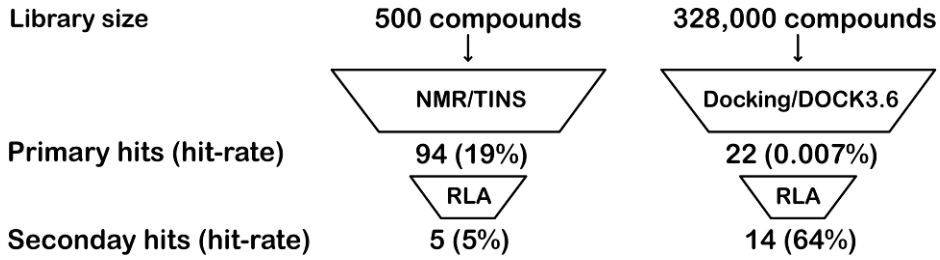


Figure 5. Summary of the discovery of orthosteric ligands from the NMR-based and docking screens. Primary screens were carried out using NMR-based (NMR/TINS) and docking (Docking/DOCK3.6) screens. Secondary screens were performed using radioligand binding assays (RLA).

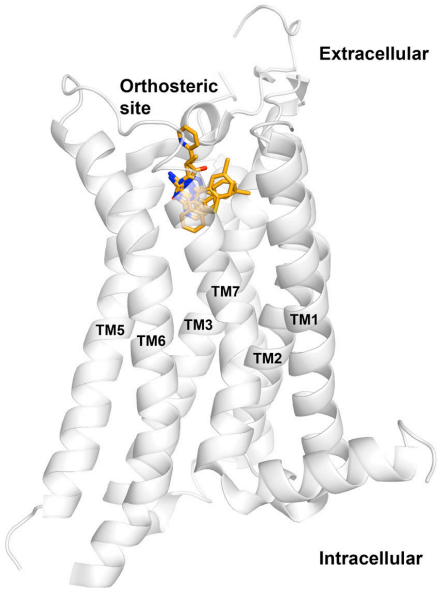


Figure 6. Predicted binding modes for compounds **10**, **11**, **12**, and **13** in the orthosteric site of the $A_{2A}R$ crystal structure. The $A_{2A}R$ is displayed as a white cartoon, and the four ligands are depicted with orange carbon atoms. This figure was generated with PyMOL (version 1.4.1).

Structure-Activity Relationships and Molecular Dynamics Simulations for Analogs of Compounds 10, 11, and 12

To explore the SAR for the ligands and investigate how the fragments could be further optimized, we docked commercially available analogs of the three potent compounds **10, 11,** and **12** to the $A_{2A}R$ orthosteric site. A series of 23 compounds were selected based on their predicted interactions in the orthosteric site and were tested in radioligand binding assays. Molecular docking in combination with MD simulations was used to guide analog selection and to investigate the predicted binding modes of the screening hits.

Compound **10** was one of the most potent fragment ligands from the screen with a K_i value of 6.4 μM. The compound also had a promising LE value equal to 0.46 kcal mol^{-1} $atom^{-1}$. The binding mode of compound **10** remained stable over one ns of MD simulation with no significant conformational changes in the orthosteric site. The exocyclic amine and triazole group formed strong interactions with $Asn253^{6.55}$ in the simulation, whereas the phenyl group was deeply buried in a hydrophobic pocket created by $Val84^{3.32}$, $Leu85^{3.33}$, and $Trp246^{6.48}$. MD simulations of compound **10b** suggested that it substituents that

extended further towards the extracellular loops could be accommodated by the binding pocket. Six analogs of compound **10** were tested to further investigate the binding mode of the ligand (Table 3). Possibilities for optimization of this fragment were first explored by replacing the ethyl with a methyl and propyl substituent (**10a** and **10b**). Reducing the alkyl chain to a methyl led to a 2-fold reduction of affinity, whereas the propyl substituent did not affect the potency. The retention of the potency by the compound with the longer substituent agreed with the predicted interactions with hydrophobic side chains in the opening of the orthosteric site, *e.g.* Ile274$^{7.44}$ and Met270$^{7.35}$. This suggests that further optimization could be achieved by combining **10b** with fragments that extend further toward the extracellular loops. Substitutions on the phenyl ring in the para-position with a methyl-, chlorine-, or methoxy-group led to reductions of the binding affinity in all cases (**10c-e**). The decrease in affinity for the largest substituent was consistent with the prediction that this group was located in an enclosed pocket, in agreement with the predicted binding mode (Figure 7A). Replacing the phenyl group of compound **10b** with a pyridine ring led to a ten-fold decrease in binding affinity to 50 μM (**10f**). To investigate the reasons for this large reduction of affinity, we calculated the difference in free energy of binding for these compounds using MD free energy calculations. MD simulations were used to alchemically transform compound **10b** to **10f** in the $A_{2A}R$ orthosteric site and in aqueous solution by employing the free energy perturbation (FEP) technique. From these calculations, the difference in binding free energy between the two ligands can be calculated from a thermodynamic cycle.[41] The calculated free energy difference between compounds **10b** and **10f** was +1.2 ± 0.1 kcal/mol, in close agreement with the experimental 10-fold loss of affinity caused by the pyridine group. Inspection of the MD trajectories suggested that the reason for the calculated reduction of binding affinity was a loss of favorable solvent interactions for the pyridine group in the bound state (Figure 8).

Compound **11** was one of the more novel fragment ligands identified in the docking screen (T_c = 0.28) and relatively potent with a K_i value of 6.3 μM. Based on molecular docking and MD simulations of selected fragments, eleven

compounds were selected for experimental testing in radioligand binding assays, which are summarized in Table 4. This congeneric series of fragments explored variations on the pyridine group of compound **11**, which was predicted to interact with residues in the opening of the orthosteric site. Compound **11a**, for which the nitrogen of the pyridine ring was moved to the ortho position, was slightly more potent than the initial hit. On the contrary, replacing the pyridine with a phenyl led to a 3-fold reduction of the K_i value to 18 μM (**11b**). For compounds **11c-i,** the effects of polar and non-polar substituents on the phenyl ring of compound **11b** were explored. In the ortho-position, polar (**11c**), non-polar (**11d**), and halogen substituents (**11e-f**) improved binding affinity with K_i values ranging from 5.5 to 8.8 μM. In meta-position, a methyl substituent (**11g**) did not affect affinity relative to compound **11b**, whereas an increase in polarity at this position was found to improve the affinity (**11a**). In para-position, a polar substituent (**11h**) led to a 2-fold improvement of affinity whereas a methyl group was not tolerated and reduced the K_i value to 60 μM (**11i**). The observed effects on affinity for ortho- and meta- substitutions on the phenyl ring could be explained by the interactions with residues in the opening of the orthosteric site, as this position is close to both polar and non-polar groups in the opening of the orthosteric site. However, the changes in affinity for polar and non-polar substitutents in the para-position could not be explained by our model because this position was typically solvent exposed in the predicted binding mode and in the MD simulations (Figure 7B). This may be related to interactions with flexible parts of extracellular loop two, e.g. two rotamers of Glu169$^{5.30}$ that have been observed in crystal structures.[6-9] Two additional compounds were designed by combining hydroxyl substituents in ortho- and para-postions with a methoxy group in meta-postion (compounds **11j** and **11k**, respectively). The most potent compound, **11k**, had an affinity of 2.4 μM, which is eight-fold better than compound **11b** and a 2-fold improvement compared to the initial hit from the docking screen. The MD simulation of compound **11k** suggested that the methoxy substituent complements the shape of the orthosteric site very well, which is shown in Figure 9.

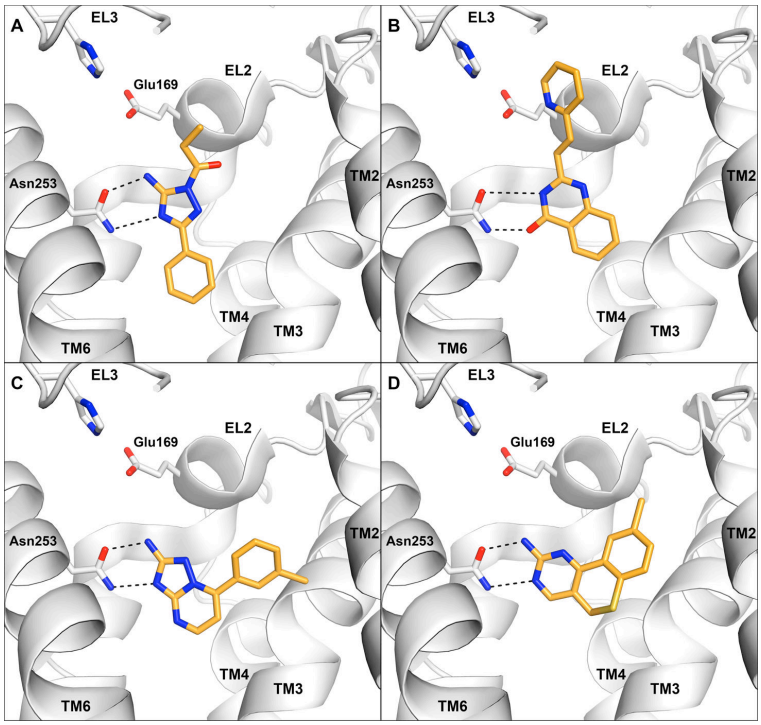


Figure 7. Predicted binding modes for four ligands discovered in the molecular docking screen: (A) **10** (B) **11**, (C) **12**, (D) **13**. The binding site is shown in white ribbons with selected side chains shown in sticks. The ligand is depicted with orange carbon atoms. Black dotted lines indicate hydrogen bonds. This figure was generated with PyMOL (version 1.4.1).

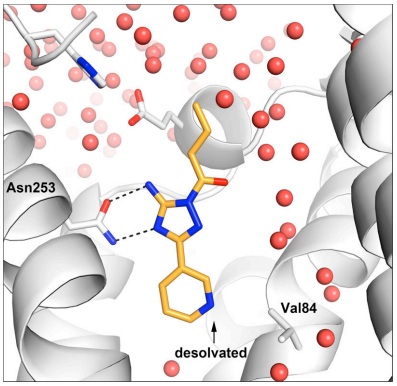


Figure 8. Representative molecular dynamics snapshot of the $A_{2A}R$ in complex with compound **10f**. The binding site is shown in white ribbons and selected side chains shown in sticks. The ligand is depicted with orange carbon atoms. Water molecules are shown as red spheres. Black dotted lines indicate hydrogen bonds. This figure was generated with PyMOL (version 1.4.1).

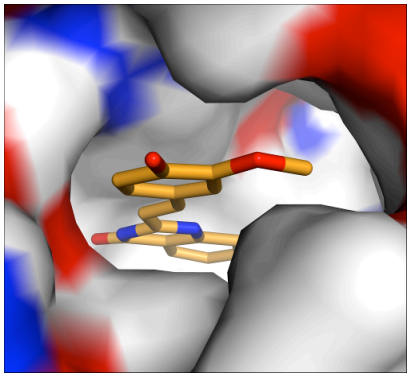


Figure 9. Representative snapshot from molecular dynamics simulation of the $A_{2A}R$ in complex with compound **11k**. The extracellular part of the $A_{2A}R$ is shown in surface representation and the ligand is depicted with orange carbon atoms. This figure was generated with PyMOL (version 1.4.1)

The affinity of compound **12** was 6.8 μM and this fragment had a LE value of 0.43 kcal mol^{-1} $atom^{-1}$. In the predicted binding mode of this ligand, it is almost completely buried in the orthosteric site with the 1,2,4-triazolo[1,5-a]pyrimidine-2-amine group hydrogen bonding with Asn253 and forming stacking interactions with the side chain of Phe168 (Figure 7C). The phenyl group of compound **12** occupied part of the sub-pocket where the ribose ring of $A_{2A}R$ agonists is located in crystal structures. MD simulations of this fragment suggested that the compound likely explores multiple conformations in the site. In these simulations hydrogen bonds to either Asn253 or Glu169 were conserved, but the phenyl ring explored several subpockets. Six additional compounds in this series were tested experimentally in radioligand binding assays and the results are summarized in Table 5. The compound without any substituents on the phenyl ring was 15-fold less potent than the initial hit, K_i = 101 μM (**12a**), demonstrating that the Cl in meta-position of compound **12** was essential for binding. In the next step, a Cl in para-position was tested, but this compound was also inactive. Subsequently, a methyl group was added to the pyrimidine group, which resulted in a 50-fold improvement of affinity to 2 μM and an increase of the LE to 0.47 kcal mol^{-1} $atom^{-1}$ (**12c**). Substituents in the para-position on the phenyl group on this compound (**12d-f**) did not further improve the affinity of the series. The predicted binding mode for compound **12c** was in good agreement with the observed changes in affinity for the series. The methyl substituent that improved the affinity of compound **12c** to 2.1 μM was positioned in a hydrophobic pocket formed by Val84$^{3.32}$, Leu85$^{3.33}$, and Trp246$^{6.48}$. Almost complete burial of the compound in the orthosteric site was likely to be responsible for its high efficiency. (LE=0.47 kcal mol^{-1} $atom^{-1}$, Figure 10)

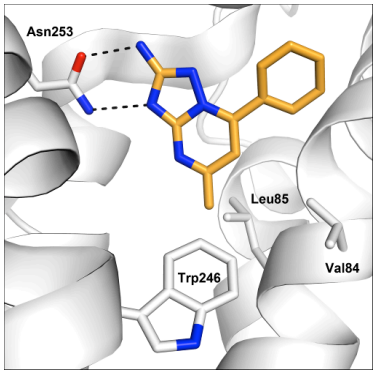


Figure 10. Predicted binding modes for compound **12c**. The binding site is shown in white ribbons with selected side chains shown in sticks. The ligand is depicted with orange carbon atoms. Black dotted lines indicate hydrogen bonds. This figure was generated with PyMOL (version 1.4.1).

Selectivity and Efficacy for the Discovered Fragment Ligands

One advantage of fragment-based screening is that high hit-rates are often achieved, but the ligands that are discovered are typically not selective for their target. To test the selectivity profile for the fragment ligands that emerged from the *in silico* screen, all compounds were also tested against the A_1R subtype in radioligand binding assays. Two of the more potent fragments ligands were also tested in cyclic-AMP (cAMP) assays. For the fourteen fragments discovered in the docking screen (Table 2), no significant selectivity for the $A_{2A}R$ was found. On the contrary, several ligands actually had better affinities for the A_1R, with an average 2-fold selectivity for this subtype. The maximal A_1 selectivity among the more potent compounds was a 4-fold difference in affinity for compound **15**. A similar trend was observed for the analogs of compounds **10, 11,** and **12** (Tables 3, 4, and 5). For the two smaller fragments, **10** and **12**, no significant selectivity for A_{2A} over the A_1 subtype was found for any of the analogs. Several of the analogs of compound **11** approached the size of lead-like compounds, and it was also among these compounds that some selectivity for the $A_{2A}R$ was observed. The most potent compound in the series (**11k**), had an affinity of 2.4 μM at the $A_{2A}R$, but it was a 4-fold weaker ligand for the A_1 subtype.

Two of the more potent fragments from the docking screen, compounds **10** and **11**, were also tested for their behavior in cAMP assays. Addition of 10 μM of each compound resulted in a rightward shift of the agonist (CGS21680) concentration-effect curve as expected for antagonists (Figure 11). This result is in agreement with previous docking screens of lead-like libraries against $A_{2A}R$ crystal structures in the inactive state that, to our knowledge, only have resulted in the discovery of antagonists.[10,11,44]

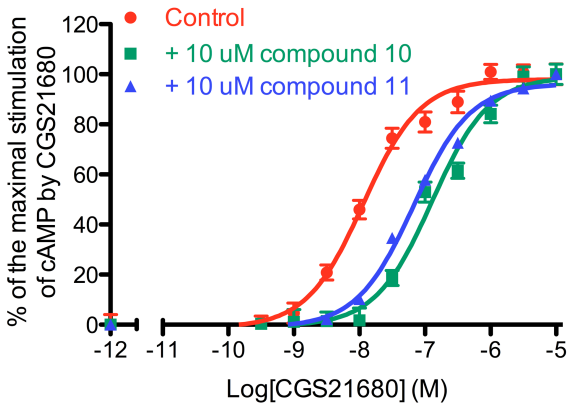


Figure 11. Functional assay based on measuring the concentration-dependent production of cAMP by $A_{2A}R$ agonist CGS21680 for compounds **10** and **11**.

3. Discussion

The revolution in GPCR structural biology during the last years has made it possible to take advantage of structure-based methods in ligand discovery against these pharmaceutically important targets. During the same period, FBLD has become widely used in drug discovery, but there are only a few examples of the use of experimental[20,21,45,46] and *in silico*[16,47] fragment screens against GPCRs. To assess the potential benefits of complementing biophysical screens against GPCRs with structure-based discovery of fragment ligands, we carried out two prospective molecular docking screens against the $A_{2A}R$, a drug target for development of drugs against inflammation and Parkinson's disease.[2,3] The parallel screen of a small chemical library against the $A_{2A}R$ allowed us to evaluate the ability of molecular docking to distinguish fragment ligands of the receptor. A second comparison was made possible by the large-scale screen of commercially available fragments and subsequent experimental testing of top-ranked compounds.

The first encouraging result was the ability of molecular docking to identify all but one of the hits from the NMR-based screen among the 30 top-ranked compounds. The results are similar to the parallel docking and HTS screen of 200,000 drug- like compounds against Cruzain by *Ferreira et al.*, in which several inhibitors were among the top-ranked compounds.[48] It should be noted that significantly higher enrichment of ligands has been reported by *Katritch et al.* in a benchmark of molecular docking against the same $A_{2A}R$ crystal structure.[11] As only lead- to drug-like compounds were used in these retrospective screens, this result could suggest that it is more challenging for docking to identify fragment ligands. However, it could also reflect differences in the scoring functions used and how the non-binders (decoys) were selected. The main finding from the two *in silico* screens that were undertaken in this work was the ability of the molecular docking to complement the biophysical screen for fragment ligands of the $A_{2A}R$. One of the main advantages of fragment-based screening compared to

HTS is that the high hit-rates provide numerous starting points for development of lead compounds.[17,18] However, prioritizing between a large number of primary hits from biophysical screens, a total of 94 from the NMR-based method analyzed here, is often difficult and time-consuming.[21] The successful prediction of the NMR-based screening results strongly suggests that it is advantageous to carry out docking screens in parallel and use the results to select hits for follow-up testing to reduce the number of false positives. This approach would also provide a starting point for structure-based optimization of fragments. Interestingly, a recent comparison between fragment-based screens with different biophysical approaches showed that different sets of ligands are identified using the two methods.[49] To some extent, this likely reflects challenges in detecting low-affinity fragments and suggests that a combination of several orthogonal methods could reduce the number of false negatives and positives from biophysical screens. Our results demonstrate that molecular docking also could play such a role. Several fragments that were top-ranked by docking, but had not been identified by the NMR-based screen, were micromolar ligands with high ligand efficiencies.

The parallel screen also demonstrated major weaknesses of molecular docking screens. As in a majority of all structure-based virtual screens, the receptor volume searched by the docking algorithm was reduced to one pocket on the protein surface. The consequence of this approximation was that allosteric modulators could not be identified. On the contrary, the NMR-based screen led to one of the first discoveries of compounds with positive and negative allosteric effects on ligand binding to the $A_{2A}R$.[2,3] The obvious solution would be to extend the searched volume to include the entire receptor surface, but this would be too computationally expensive for screening of large chemical libraries. Multiple pockets could be considered, but it is not yet clear where the allosteric ligands bind to the $A_{2A}R$, and it would also be a major challenge for docking scoring functions to rank molecules in sites of different sizes and chemical properties. Modulation of receptors via allosteric interactions may provide novel opportunities for development of completely new classes of drugs against the $A_{2A}R$.[50]

The second docking screen unleashed the full potential of the structure-based approach. Several hundred thousand fragments were docked to the $A_{2A}R$ crystal structure to identify compounds that complement this receptor conformation. As only a few dozen compounds are evaluated experimentally in this case, more extensive testing can be carried out for each compound, which reduces the risk for false negatives. As in earlier *in silico* screens against $A_{2A}R$ and other GPCR crystal structures, the hit-rates were unusually high.[10−16] It may be the combination of a highly ligandable orthosteric site[48] and bias in chemical libraries toward GPCR chemotypes that make structure-based screens against these receptors so successful.[10,12] If this is true, the successes for deep and enclosed binding sites, which are characteristic for the adenosine and aminergic receptors, may not be transferrable to GPCRs with more shallow and solvent exposed pockets. The observation that biophysical screens of fragments generally yield higher hit-rates than those that rely on lead- or drug-like libraries was also demonstrated to be true for *in silico* screens. Several ligands from the docking screen were also more potent than expected from the typical fragment screens, both in comparison to screens against soluble proteins and the NMR-based screen against the same target.[20,21,45]

Successful FBLD relies on the possibility to optimize the initial hits, and, in this step, access to atomic resolution structures provides valuable information. However, in the case of predicting affinities for a series of analogs, subtle structural changes of the receptor may be responsible for differences in potency, which may require a more accurate representation of protein−ligand interactions than molecular docking can provide.[51] Initial attempts to use more rigorous methods in lead optimization of GPCR ligands have been encouraging, and such approaches will likely also play an important role in FBLD against these receptors.[52,53] An increasing amount of computational resources has made it possible to use atomistic simulations to compute the relative affinities for large series of protein−ligand complexes.[54,55] The small size and relatively good availability of force field parameters for fragment-like molecules make them ideal cases for MD simulations and free energy calculations. Here, three new scaffolds

from the *in silico* screen were further explored via analogs and MD simulations of several complexes. The MD simulations provided a deeper understanding of the SAR for analogs of the docking hits, and free energy calculations were used to explain changes in binding affinity due to changes in ligand structure. Two of the scaffolds were optimized to an affinity of 2 μM. In particular these compounds, as well as the 15 other discovered fragments that were left unexplored, provide new starting points for development of new lead compounds against this medically important target.

Drug design against GPCRs also involves identifying compounds with a specific efficacy and selectivity profile. The fragments that were tested for efficacy were all antagonists, and this was consistent with that the screen was carried out against a crystal structure in the inactive state. Fragment ligands are typically not selective for their targets, and this is also a general challenge for the $A_{2A}R$ because of the strong sequence conservation in the orthosteric site among the four subtypes.[1] None of the ligands discovered in the NMR-based or docking screens showed significant selectivity for $A_{2A}R$ over the closely related A_1 subtype, which likely reflected that the fragment hits explored the conserved core of the binding site. If anything, the styryl moiety in compound series **11** had been found as a chemical substructure contributing to A_{2A} receptor selectivity before, and indeed in this series there was some selectivity for this receptor subtype too.[56] The refined binding modes for the most potent compounds now provide starting points for optimization of selectivity. In this case, future crystal structures and models of other adenosine receptor subtypes will be crucial for optimizing the selectivity of fragment ligands discovered in biophysical and docking screens.[22,57,58]

4. Conclusions

The picture that emerges from this study is that fragment-based ligand discovery via biophysical and computational structure-based screens are highly

complementary approaches. Both the NMR- and molecular docking based screens identified novel $A_{2A}R$ fragment ligands with no overlap between the two sets of hits. Molecular docking was shown to be a highly efficient method for the identification of fragments ligands, as reflected by the 64% hit-rate achieved in this screen. On the other hand, the unbiased experimental screen of a relatively small fragment library unexpectedly discovered allosteric modulators of the $A_{2A}R$, which will provide new opportunities for drug development against this receptor. The opportunities to combine biophysical and *in silico* screening methods with structure determination for GPCRs will likely position FBLD as one of the principal approaches for drug development against this pharmaceutically important class of targets.

5. Methods

Molecular Docking Calculations

All docking calculations were carried out with the program DOCK3.6[26-28] using a 2.6 Å resolution crystal structure of the human $A_{2A}R$ in complex with an antagonist (PDB accession: 3EML[7]). The receptor was prepared by removing all non-protein atoms and the intracellular T4-lysozyme insertion. The protonation states of ionizable residues of all Asp, Glu, Arg and Lys residues were set to the most probable in aqueous solution at pH 7. The protonation states of the histidines were set by manual inspection (His278$^{7.43}$: protonated in delta position, His250$^{6.52}$: protonated in epsilon position, His264$^{6.66}$: double protonated).

The flexible ligand sampling algorithm in DOCK3.6 superimposes atoms of the docked molecule onto binding site matching spheres, which indicate putative ligand atom positions. Sixty matching spheres were used, and these were either based on the atoms of the crystallographic ligand or positioned manually.[27,28] The spheres were also labeled for chemical matching based on the local receptor environment.[59] The degree of ligand sampling was determined by the bin size, bin size overlap, and distance tolerance. These three parameters were set to 0.4 Å, 0.2 Å, and 1.5 Å, respectively, for both the binding site matching spheres and

the docked molecules. An initial filter discarded conformations that had more than one overlapping heavy atom with the receptor. An overlap between two atoms was defined as a distance shorter than 2.3 and 2.6 Å for polar and non-polar atoms, respectively, and was calculated from a precalculated grid. For ligand conformations passing this steric filter, a physics-based scoring function was used to evaluate the fit to the binding site. For the best scoring conformation of each docked molecule, 100 steps of rigid-body minimization were carried out.[27,28] The score for each conformation was calculated as the sum of the receptor−ligand electrostatic and van der Waals interaction energy, corrected for ligand desolvation. These three terms were evaluated from precalculated grids. The three-dimensional map of the electrostatic potential in the binding site was prepared using the program Delphi.[60] In this calculation, partial charges from the united atom AMBER force field[61] were used for all receptor atoms except the side chain amide of Asn253$^{6.55}$, for which the dipole moment was increased to favor hydrogen bonding to this residue.[10] The program CHEMGRID was used to generate a van der Waals grid, which is based on a united atom version of the AMBER force field.[62] The desolvation penalty for a ligand conformation was estimated from a precalculated transfer free energy of the molecule between solvents of dielectrics 78 and 2. The desolvation energy was obtained by weighting the transfer free energy with a scaling factor that reflects the degree of burial of the ligand in the receptor binding site.[37,63]

Prior to the DOCK3.6 calculation, each docked library was prepared for docking by pre-generating up to 600 conformations using the program OMEGA.[64] Partial atomic charges and transfer free energies were calculated using AMSOL[63,66] and van der Waals parameters were derived from an all-atom AMBER potential.[67] Two libraries were screened. The first was an in-house fragment library of 500 compounds.[31] The second was the ZINC fragment-like library[40] of 328,000 commercially available molecules (molecular weight < 250, LogP < 3.5, and rotatable bonds < 5).

Similarity Calculations

Similarity calculations for the fragments were carried out using Screenmd program from Chemaxon.[68] We calculated the maximum Tanimoto coefficient with ECFP4 fingerprints between each discovered fragment ligand to the 6509 compounds with recorded activity against the human $A_{2A}R$ in the ChEMBL14[41] database.

Molecular Dynamics Simulations and Free Energy Calculations

The molecular dynamics simulations (MD) and binding free energy calculations for the $A_{2A}R$ were performed using a recently published high-resolution crystal structure (PDB accession code: 4EIY[9], 1.9 Å). In the first step, a hydrated POPC membrane was first equilibrated around the $A_{2A}R$ structure without any ligand using periodic boundary conditions in GROMACS[69] using the OPLSAA force field,[70] TIP3P waters,[71] and Berger lipids.[72] In this simulation, all protein atoms were tightly restrained to their initial coordinates, and the hydrated membrane was equilibrated for 40 ns at 300 K. All other MD simulations and free energy calculations of the protein-ligand complexes were carried out starting from the membrane equilibrated $A_{2A}R$ system using spherical boundary conditions in the program Q.[73] All ligand force field parameters were obtained using Schrodinger's program hetgrp_ffgen.[74] The simulations were carried out at a constant temperature of 310 K in a sphere of 18 Å radius centered on the ligand. All protein, water, and ligand atoms within 18 Å of the center of the sphere were explicitly included in the simulations. All atoms outside the sphere were tightly restrained to their initial coordinates and excluded from nonbonded interactions. Asp, Glu, His, Lys, and Arg residues within the spherical system were protonated as in the docking calculation. All other ionizable residues close to the sphere edge or further away than 18 Å were set to their neutral state. The SHAKE[75] algorithm was applied to all solvent bonds and angles and the water molecules at the sphere surface were subjected to radial and polarization restraints.[73,76] A nonbonded cutoff of 10 Å was used for all atoms except the ligand, for which no cutoff was applied. Long-range electrostatic interactions were treated with the

local reaction field (LRF) multipole expansion method.[77] The time step was set to 1 fs, and nonbonded pair lists were updated every 25 steps.

For the docking hits and analogs that were studied with MD simulations, each protein-ligand complex was equilibrated for 810 ps. During the equilibration, harmonic restraints on the protein and ligand atoms were gradually released, which was followed by one ns of unrestrained simulation. To compute the relative free energy of binding from a thermodynamic cycle, the ligands have to be transformed to each other in water and the $A_{2A}R$ binding site. The free energy of transforming ligand **10b** to **10f** was computed using the free energy perturbation (FEP) technique, and the calculation was carried out via several intermediates.[78] The potentials governing the intermediate states are defined by $U_m = \lambda_m U_A + (1 - \lambda_m) U_B$ where *A* and *B* represent two different ligands, respectively, and λ_m is a mapping parameter which varies from $\lambda_1 = 0$ to $\lambda_n = 1$. The free energy difference between states *A* and *B* can be calculated by summing up the free energy differences of the *n* intermediate states using

$$\Delta G_{A\to B}^{FEP} = -kT\sum_{m=1}^{n-1}\ln\left\langle e^{-(U_{m+1}-U_m)/kT}\right\rangle_m$$

where $\langle ... \rangle_m$ represents an ensemble average on the potential U_m, which is calculated from MD simulations.[43] The free energy was computed in two steps. First the van der Waals parameters of the phenyl ring of compounds **10b** were transformed to the pyridine of compound **10f**. In a second step, the charges for the same group were changed to those representing pyridine. These calculations were carried out both for the ligand in aqueous solution and in the $A_{2A}R$ orthosteric site. The calculations in aqueous solution were carried out by solvating the ligand in a water droplet of radius 18 Å. In these simulations a weak harmonic restraint was applied to keep the ligand close to the center of the sphere. Each FEP calculation comprised a 700 ps heating scheme through eight steps of equilibration. The transformation of partial charges and LJ parameters were carried out with 11 λ states (from λ=0 to λ=1) each, with a production run time of 500ps for each lambda point and energies were extracted every 50 steps.

Cell Culture and Membrane Preparation

HEK293 cells were maintained in culture in DMEM supplemented with 10% newborn calf serum at 37 ºC in a moist, 7% CO_2 atmosphere and passaged twice weekly. Cells were transfected with plasmids containing wild-type A_{2A}AR construct using the calcium phosphate precipitation method and harvested after 48 hours. Cells were pelleted, re-suspended in 20 ml of ice-cold 50 mM Tris-HCl buffer, pH 7.4. An ULTRA-TURRAX® was used to homogenize the cell suspension. The cytosolic and membrane fractions were separated using a high-speed centrifugation step of 100,000 x g for 20 min at 4 ºC. This process was repeated twice and subsequently the pellet was re-suspended in 50 mM Tris-HCl buffer, pH 7.4, the protein concentration was determined by the BCA method[79] and samples were aliquoted and stored at - 80 ºC until further use.

Radioligand Binding Studies

Tritiated 4-(2-[7-Amino-2-(2-furyl)-[1,2,4]-triazolo-[2,3-a]-[1,3,5]-triazin-5-yl-amino] ethyl)phenol ([^{3}H]ZM241385, specific activity 30 Ci/mmol) and tritiated 1,3-dipropyl-8-cyclopentyl-xanthine ([^{3}H]DPCPX, specific activity 120 Ci/mmol) were purchased from ARC Inc. (St. Louis, USA). Unlabelled ZM241385 and -cyclopentyladenosine (CPA) were purchased from Sigma-Aldrich (Steinheim, Germany). The tested compounds were purchased from five different vendors (Asinex, ChemBridge, Vitas-M, Pharmeks and KeyOrganics). The identity of all tested compounds was verified by ^{1}H-NMR and HPLC/MS was used to assess the purity for all of the molecular docking screening hits.

For the A_{2A}Rs, membranes containing 20 µg of protein were incubated in a total volume of 100 µl Tris-HCl buffer (50 mM, pH 7.4), 10 mM $MgCl_2$, increasing concentrations of the test compounds (5% DMSO) and [^{3}H]ZM241385 at final concentration of 3 nM. Non-specific binding was determined using a final concentration of 10 µM ZM241385 and represented less than 10% of the total binding. For the A_1Rs, membranes containing 5 µg of protein were incubated in a total volume of 100 µl Tris-HCl buffer (50 mM, pH 7.4), [^{3}H]DPCPX (final concentration 1.6 nM) and increasing concentrations of the test compounds (5%

DMSO). Non specific binding was determined using a final concentration of 100 μM CPA. Total binding was determined in the presence of buffer and was set at 100% in all experiments, whereas non-specific binding was set at 0%. After incubation for 1 hour at 25 °C in a shaking water bath, assays were terminated by rapid filtration through 96-well GF/B UniFilter plates (PerkinElmer) followed by washing with 7 × 0.25 ml ice-cold Tris-HCl buffer (50 mM, pH 7.4). Plates were dried, 25 μl P-E Microscint 20 was added per well, and after 3 hours, bound radioactivity measured using a Packard Microbeta counter (PerkinElmer). Data were analyzed using GraphPad Prism v5, normalized as '% specific binding' from which the K_i values were calculated. Experiments were performed three times in duplicate, unless stated otherwise.

Cyclic AMP Accumulation Assay

HEK293 cells were grown and transfected as described above. Experiments were performed 48 h after transfection. The cells were harvested and centrifuged two times at 1.000 rpm for 5 min. For cyclic AMP production and determination, 1500 cells/well were used in 384-well Optiplates (PerkinElmer). The cells were incubated for 30 min at 37 °C with either the $A_{2A}R$ agonist CGS21680 alone or CGS21680 together with test compounds in different concentrations. The assay buffer used was PBS with the addition of 5 mM HEPES, 0.1% BSA, 50 μM rolipram, 50 μM cilostamide, and 0.8 IU/mL adenosine deaminase. Basal activity was determined in the presence of assay buffer and was set at 0% in all experiments. Maximal receptor activity was determined in the presence of 10 μM CGS21680 and was set at 100% in all experiments. Cells were then lysed, and the amount of cAMP produced was quantified using a LANCE ultra cAMP kit (PerkinElmer) according to the instructions of the manufacturer. Following addition of the detection mixture, plates were left for 1 h at RT prior to reading using an EnVision plate reader (PerkinElmer Life Sciences).

HPLC/MS Protocol

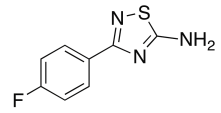
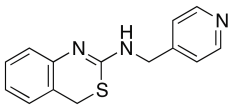
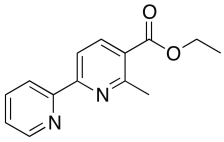
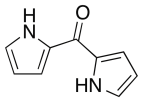
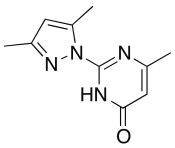
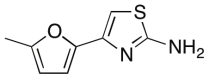
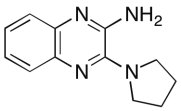
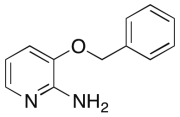
Analytical purity of the compounds **9-15**,**16-17**,**19**, and **22** was determined by high-pressure liquid chromatography (HPLC) with a Phenominex Gemini 3u C18 110A column (50 x 4.6 mm, 3 µm), measuring UV absorbance at 254 nm. Sample preparation and HPLC method is - unless stated otherwise - as follows: 0.3-0.8 mg of compound was dissolved in 1 mL of a 1:1:1 mixture of $CH_3CN/H_2O/tBuOH$ and eluted from the column within 15 minutes, with a three component system of $H_2O/CH_3CN/1\%$ TFA in H_2O, decreasing polarity of the solvent mixture in time from 80/10/10 to 0/90/10.

LC-MS analysis was performed on a Finnigan Surveyor HPLC system with a Gemini C18 50 × 4.60 mm column (detection at 200-600 nm), coupled to a Finnigan LCQ Advantage Max mass spectrometer with ESI using the same procedure as mentioned above.

Contributions

Anirudh Ranganathan and Jen Carlsson (Stockholm University) performed molecule docking screening and molecular dynamics simulations. Jacobus van Veldhoven (LACDR, Leiden) helped for HPLC/MS measurement.

Table 1. Fragment ligands and discovered in the NMR-based screen (**1-5**)[21] and false negatives from NMR-screen that were identified by the docking screen (**6-8**).

ID (Rank)[a]	Ligand structure	K_i (LE[c]) μM (kcal mol^{-1} atom^{-1})	T_c[d]
1 (26)		14 ± 2.3 (0.53)	0.33
2 (19)		183 ± 20 (0.29)	0.33
3 (7)		324 ± 29 (0.27)	0.39
4 (NS[b])		472 ± 34 (0.39)	0.42
5 (14)		586 ± 45 (0.31)	0.31
6 (34)		17.6 ± 1.5 (0.56)	0.38
7 (45)		20.6 ± 2.4 (0.42)	0.30
8 (13)		128 ± 11.3 (0.37)	0.44

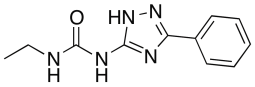
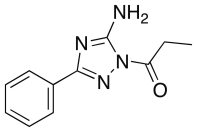
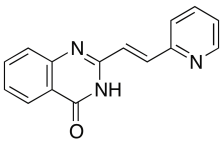
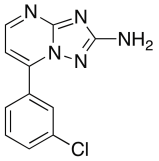
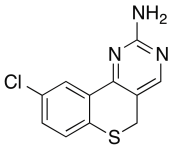
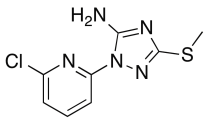
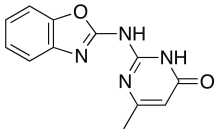
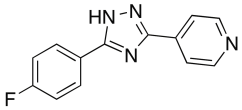
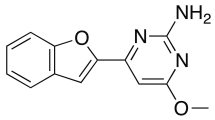
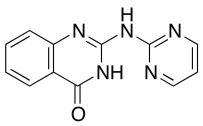
[a] Rank by docking screen of fragment library (500 compounds in total).

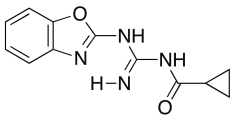
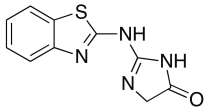
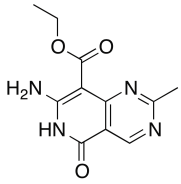
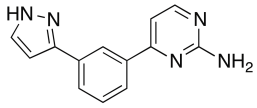
[b] Not scored in molecular docking screen.

[c] Ligand efficiency

[d] Tanimoto similarity (ECFP4 fingerprints) to A_{2A} adenosine receptor ligands in the ChEMBL14 database.

Table 2. Structures and experimental data for the fragment ligands discovered in the docking screen against the A_{2A} adenosine receptor.

ID (Rank)[a]	Ligand structure	K_i (LE[b]) μM or % displacement at 10μM (kcal mol^{-1} atom^{-1})		T_C[c]
		A_{2A}	A_1	
9 (27)		2.2 ± 0.2 (0.47)	0.9 ± 0.1	0.41
10 (348)		6.4 ± 0.6 (0.46)	4.0 ± 0.5	0.40
11 (224)		6.3 ± 0.5 (0.39)	11.1 ± 1.0	0.28
12 (288)		6.8 ± 0.4 (0.43)	1.2 ± 0.3	0.37
13 (468)		16 ± 3.0 (0.43)	8.0 ± 0.6	0.31
14 (499)		19.3 ± 4.2 (0.45)	13 ± 2	0.28
15 (233)		22.1 ± 3.1 (0.37)	5.5 ± 0.5	0.33
16 (386)		29.7 ± 6.1 (0.36)	22 ± 2	0.41
17 (235)		33.5 ± 6.0 (0.35)	10 ± 2	0.36
18 (10)		39.3± 5.3 (0.35)	7.2 ± 0.5	0.32

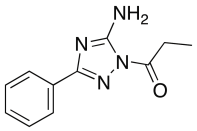
19 (492)		46.4 ± 4.7 (0.34)	40 ± 5	0.37
20 (82)		95 ± 5.4 (0.36)	80 ± 4	0.34
21 (203)		100 ± 7 (0.31)	104 ± 8	0.36
22 (206)		240 ± 8 (0.29)	173 ± 6	0.38

[a] Rank by docking screen of fragment library with 328,000 commercially available compounds.

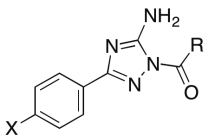
[b] Ligand efficiency

[c] Tanimoto coefficient (ECFP4 fingerprints) to A_{2A} adenosine receptor ligands in the ChEMBL14 database.

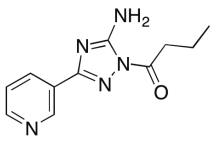
Table 3. Structures and experimental data for analogs of compounds **10**.



10, K_i=6.4 μM



10a-e

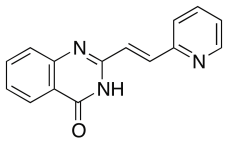
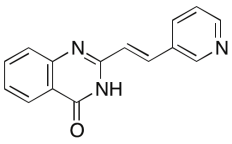
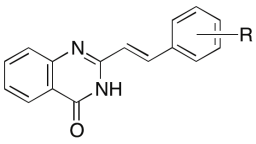


10f

ID	Ligand structure		K_i (LE[a]) displacement at 10μM (kcal mol^{-1} atom^{-1})	μM or %
	X	R	A_{2A}	A_1
10a	H	CH_3	11.0 ± 0.5 (0.47)	11.4 ± 1.0
10b	H	$(CH_2)_2CH_3$	5.1 ± 0.2 (0.44)	1.0 ± 0.2
10c	Cl	$(CH_2)CH_3$	17.0 ± 3.2 (0.40)	44 %
10d	CH_3	$(CH_2)CH_3$	10.6 ± 0.6 (0.41)	6.4 ± 0.3
10e	OCH_3	$(CH_2)_2CH_3$	42 ± 4 (0.33)	8.1 ± 0.3
10f	-	-	50 ± 5 (0.36)	47 %

[a] Ligand efficiency.

Table 4. Structures and experimental data for analogs of compound **11**.

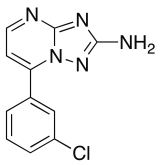




11, K_i=6.3 μM **11a** **11b-k**

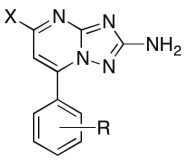
ID	Ligand structure	K_i (LE^a) μM or % displacement at 10 μM (kcal mol^{-1} $atom^{-1}$)	
	R	A_{2A}	A_1
11a	-	4.1 ± 0.3 (0.40)	20 ± 2
11b	H	18.0 ± 1.3 (0.35)	29 %
11c	2-OH	8.8 ± 0.6 (0.36)	26 %
11d	2-CH_3	7.8 ± 0.5 (0.36)	19 %
11e	2-Cl	7.0 ± 0.4 (0.37)	20 ± 3
11f	2-F	5.5 ± 0.3 (0.37)	3.5 ± 0.4
11g	3-CH_3	22 ± 1 (0.33)	24 %
11h	4-OH	7.4 ± 0.8 (0.36)	34 %
11i	4-CH_3	60 ± 5 (0.30)	18 %
11j	2-OH, 3-OCH_3	9.6 ± 0.5 (0.32)	25 %
11k	3-OCH_3, 4-OH	2.4 ± 0.2 (0.36)	9.6± 0.4

[a] Ligand efficiency

Table 5. Structures and experimental data for analogs of compound **12**.



12 K_i = 6.8 μM



12a-f

ID	Ligand structure		K_i (LE[a]) μM or % displacement at 10μM (kcal mol^{-1} $atom^{-1}$)	
	X	R	A_{2A}	A_1
12a	H	H	101 ± 6 (0.35)	2 %
12b	H	4-Cl	99 ± 5 (0.33)	17 %
12c	CH_3	H	2.1 ± 0.2 (0.47)	3.5 ± 0.3
12d	CH_3	4-OCH_3	8.7 ± 0.1 (0.40)	2.8 ± 0.2
12e	CH_3	4-CH_3	3.4 ± 0.4 (0.41)	2.4 ± 0.3
12f	CH_3	4-Cl	9.1 ± 0.2 (0.38)	2.5 ± 0.2

[a] Ligand efficiency

Reference

(1) Fredholm, B. B.; IJzerman, A. P.; Jacobson, K. A.; Linden, J.; Muller, C. E. International union of basic and clinical pharmacology. LXXXI. Nomenclature and classification of adenosine receptors-an update. Pharmacol. Rev. 2011, 63,1−34.
(2) Blackburn, M. R.; Vance, C. O.; Morschl, E.; Wilson, C. N. Adenosine receptors and inflammation. Handb. Exp. Pharmacol. 2009, 215−269.
(3) Sebastiao, A. M.; Ribeiro, J. A. Adenosine receptors and the central nervous system. Handb. Exp. Pharmacol. 2009, 471−534.
(4) Katritch, V.; Cherezov, V.; Stevens, R. C. Diversity and modularity of G protein-coupled receptor structures. Trends Pharmacol. Sci. 2012, 33,17−27.
(5) Congreve, M.; Langmead, C. J.; Mason, J. S.; Marshall, F. H. Progress in structure based drug design for G protein-coupled receptors. J. Med. Chem. 2011, 54, 4283−4311.
(6) Dore, A. S.; Robertson, N.; Errey, J. C.; Ng, I.; Hollenstein, K.; Tehan, B.; Hurrell, E.; Bennett, K.; Congreve, M.; Magnani, F.; Tate, C. G.; Weir, M.; Marshall, F. H. Structure of the adenosine A2A receptor in complex with ZM241385 and the xanthines XAC and caffeine. Structure 2011, 19, 1283−1293.
(7) Jaakola, V. P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y. T.; Lane, J. R.; IJzerman, A. P.; Stevens, R. C. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. Science 2008, 322, 1211−1217.
(8) Lebon, G.; Warne, T.; Edwards, P. C.; Bennett, K.; Langmead, C. J.; Leslie, A. G. W.; Tate, C. G. Agonist-bound adenosine A(2A) receptor structures reveal common features of GPCR activation. Nature 2011, 474, 521−525.
(9) Liu, W.; Chun, E.; Thompson, A. A.; Chubukov, P.; Xu, F.; Katritch, V.; Han, G. W.; Roth, C. B.; Heitman, L. H.; IJzerman, A. P.; Cherezov, V.; Stevens, R. C. Structural basis for allosteric regulation of GPCRs by sodium ions. Science 2012, 337, 232−236.
(10) Carlsson, J.; Yoo, L.; Gao, Z. G.; Irwin, J. J.; Shoichet, B. K.; Jacobson, K. A. Structure-based discovery of A2A adenosine receptor ligands. J. Med. Chem. 2010, 53, 3748−3755.
(11) Katritch, V.; Jaakola, V. P.; Lane, J. R.; Lin, J.; IJzerman, A. P.; Yeager, M.; Kufareva, I.; Stevens, R. C.; Abagyan, R. Structure-based discovery of novel chemotypes for adenosine A2A receptor antagonists. J. Med. Chem. 2010, 53, 1799−1809.
(12) Kolb, P.; Rosenbaum, D. M.; Irwin, J. J.; Fung, J. J.; Kobilka, B. K.; Shoichet, B. K. Structure-based discovery of beta2-adrenergic receptor ligands. Proc. Natl. Acad. Sci. U. S. A. 2009, 106, 6843−6848.
(13) Sabio, M.; Jones, K.; Topiol, S. Use of the X-ray structure of the beta2-adrenergic receptor for drug discovery. Part 2: Identification of active compounds. Bioorg. Med. Chem. Lett. 2008, 18, 5391−5395.
(14) Carlsson, J.; Coleman, R. G.; Setola, V.; Irwin, J. J.; Fan, H.; Schlessinger, A.; Sali, A.; Roth, B. L.; Shoichet, B. K. Ligand discovery from a dopamine D3 receptor homology model and crystal structure. Nat. Chem. Biol. 2011, 7, 769−778.
(15) Mysinger, M. M.; Weiss, D. R.; Ziarek, J. J.; Gravel, S.; Doak, A. K.; Karpiak, J.; Heveker, N.; Shoichet, B. K.; Volkman, B. F. Structure- based ligand discovery for the protein-protein interface of chemokine receptor CXCR4. Proc. Natl. Acad. Sci. U. S. A. 2012, 109, 5517−5522.
(16) de Graaf, C.; Kooistra, A. J.; Vischer, H. F.; Katritch, V.; Kuijer, M.; Shiroishi, M.; Iwata, S.; Shimamura, T.; Stevens, R. C.; de Esch, I. J.; Leurs, R. Crystal structure-based virtual screening for fragment-like ligands of the human histamine H(1) receptor. J. Med. Chem. 2011, 54, 8195−8206.
(17) Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J. Recent developments in fragment-based drug discovery. J. Med. Chem. 2008, 51, 3661−3680.
(18) Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. Nat. Rev. Drug Discovery 2007, 6, 211−219.
(19) Baker, M. Fragment-based lead discovery grows up. Nat. Rev. Drug Discovery 2012, 12,5−7.
(20) Congreve, M.; Rich, R. L.; Myszka, D. G.; Figaroa, F.; Siegal, G.; Marshall, F. H. Fragment screening of stabilized G-protein-coupled receptors using biophysical methods. Methods Enzymol. 2011, 493, 115−136.
(21) Chen, D.; Errey, J. C.; Heitman, L. H.; Marshall, F. H.; IJzerman, A. P.; Siegal, G. Fragment screening of GPCRs using biophysical methods: identification of ligands of the adenosine A(2A) receptor with novel biological activity. ACS Chem. Biol. 2012, 7, 2064− 2073.
(22) Tosh, D. K.; Phan, K.; Gao, Z. G.; Gakh, A. A.; Xu, F.; Deflorian, F.; Abagyan, R.; Stevens, R. C.; Jacobson, K. A.; Katritch, V. Optimization of adenosine 5′-carboxamide derivatives as adenosine receptor agonists using structure-based ligand design and fragment screening. J. Med. Chem. 2012, 55, 4297−4308.
(23) Chen, Y.; Shoichet, B. K. Molecular docking and ligand specificity in fragment-based inhibitor discovery. Nat. Chem. Biol. 2009, 5, 358−364.
(24) Huang, D. Z.; Caflisch, A. Library screening by fragment-based docking. J. Mol. Recognit. 2010, 23, 183−193.
(25) Hubbard, R. E.; Chen, I.; Davis, B. Informatics and modeling challenges in fragment-based drug discovery. Curr. Opin. Drug Discovery Dev. 2007, 10, 289−297.
(26) Irwin, J. J.; Shoichet, B. K.; Mysinger, M. M.; Huang, N.; Colizzi, F.; Wassam, P.; Cao, Y. Q. Automated docking screens: a feasibility study. J. Med. Chem. 2009, 52, 5712−5720.
(27) Lorber, D. M.; Shoichet, B. K. Flexible ligand docking using conformational ensembles. Protein Sci. 1998, 7, 938−950.
(28) Lorber, D. M.; Shoichet, B. K. Hierarchical docking of databases of multiple ligand conformations. Curr. Top. Med. Chem. 2005, 5, 739−749.
(29) Magnani, F.; Shibata, Y.; Serrano-Vega, M. J.; Tate, C. G. Co- evolving stability and conformational homogeneity of the human adenosine A2a receptor. Proc. Natl. Acad. Sci. U. S. A. 2008, 105, 10744−10749.

(30) Vanwetswinkel, S.; Heetebrij, R. J.; van Duynhoven, J.; Hollander, J. G.; Filippov, D. V.; Hajduk, P. J.; Siegal, G. TINS, target immobilized NMR screening: an efficient and sensitive method for ligand discovery. Chem. Biol. 2005, 12, 207−216.
(31) Siegal, G.; Ab, E.; Schultz, J. Integration of fragment screening and library design. Drug Discovery Today 2007, 12, 1032−1039.
(32) Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. The maximal affinity of ligands. Proc. Natl. Acad. Sci. U. S. A. 1999, 96, 9997−10002.
(33) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. Drug Discovery Today 2004, 9, 430− 431.
(34) Jain, A. N.; Nicholls, A. Recommendations for evaluation of computational methods. J. Comput.-Aided Mol. Des. 2008, 22, 133− 139.
(35) Nicholls, A. What do we know and when do we know it? J. Comput.-Aided Mol. Des. 2008, 22, 239−255.
(36) Kim, J.; Wess, J.; van Rhee, A. M.; Schoneberg, T.; Jacobson, K. A. Site-directed mutagenesis identifies residues involved in ligand recognition in the human A2a adenosine receptor. J. Biol. Chem. 1995, 270, 13987−13997.
(37) Mysinger, M. M.; Shoichet, B. K. Rapid context-dependent ligand desolvation in molecular docking. J. Chem. Inf. Model. 2010, 50, 1561−1573.
(38) Majeux, N.; Scarsi, M.; Caflisch, A. Efficient electrostatic solvation model for protein-fragment docking. Proteins: Struct., Funct., Genet. 2001, 42, 256−268.
(39) Ruddigkeit, L.; van Deursen, R.; Blum, L. C.; Reymond, J. L. Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17. J. Chem. Inf. Model. 2012, 52, 2864−2875.
(40) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: a free tool to discover chemistry for biology. J. Chem. Inf. Model. 2012, 52, 1757−1768.
(41) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res. 2012, 40, D1100−D1107.
(42) Ballesteros, J. A.; Weinstein, H. Integrated methods for the construction of three-dimensional models of structure−function relations in G protein-coupled receptors. Methods Neurosci. 1995, 25, 366−428.
(43) Kollman, P. Free-energy calculations - applications to chemical and biochemical phenomena. Chem. Rev. 1993, 93, 2395−2417.
(44) Langmead, C. J.; Andrews, S. P.; Congreve, M.; Errey, J. C.; Hurrell, E.; Marshall, F. H.; Mason, J. S.; Richardson, C. M.; Robertson, N.; Zhukov, A.; Weir, M. Identification of novel adenosine A(2A) receptor antagonists by virtual screening. J. Med. Chem. 2012, 55, 1904−1909.
(45) Navratilova, I.; Besnard, J.; Hopkins, A. L. Screening for GPCR ligands using surface plasmon resonance. ACS Med. Chem. Lett. 2011, 2, 549−554.
(46) Stoddart, L. A.; Vernall, A. J.; Denman, J. L.; Briddon, S. J.; Kellam, B.; Hill, S. J. Fragment screening at adenosine-A(3) receptors in living cells using a fluorescence-based binding assay. Chem. Biol. 2012, 19, 1105−1115.
(47) Sirci, F.; Istyastono, E. P.; Vischer, H. F.; Kooistra, A. J.; Nijmeijer, S.; Kuijer, M.; Wijtmans, M.; Mannhold, R.; Leurs, R.; de Esch, I. J.; de Graaf, C. Virtual fragment screening: discovery of histamine h(3) receptor ligands using ligand-based and protein-based molecular fingerprints. J. Chem. Inf. Model. 2012, 52, 3308−3324.
(48) Ferreira, R. S.; Simeonov, A.; Jadhav, A.; Eidam, O.; Mott, B. T.; Keiser, M. J.; McKerrow, J. H.; Maloney, D. J.; Irwin, J. J.; Shoichet, B. K. Complementarity between a docking and a high-throughput screen in discovering new cruzain inhibitors. J. Med. Chem. 2010, 53, 4891− 4905.
(49) Wielens, J.; Headey, S. J.; Rhodes, D. I.; Mulder, R. J.; Dolezal, O.; Deadman, J. J.; Newman, J.; Chalmers, D. K.; Parker, M. W.; Peat, T. S.; Scanlon, M. J. Parallel screening of low molecular weight fragment libraries: do differences in methodology affect hit identification? J. Biomol. Screening 2013, 18, 147−159.
(50) Goblyos, A.; IJzerman, A. P. Allosteric modulation of adenosine receptors. BBA, Biochim. Biophys. Acta, Biomembr. 2011, 1808, 1309− 1318.
(51) Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A critical assessment of docking programs and scoring functions. J. Med. Chem. 2006, 49, 5912−5931.
(52) Higgs, C.;Beuming, T.;Sherman,W.Hydration site thermodynamics explain SARs for triazolylpurines analogues binding to the A2A receptor. ACS Med. Chem. Lett. 2010, 1, 160−164.
(53) Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y. B.; Xu, H. F.; Shaw, D. E. Pathway and mechanism of drug binding to G-protein-coupled receptors. Proc. Natl. Acad. Sci. U. S. A. 2011, 108, 13118−13123.
(54) Carlsson, J.; Boukharta, L.; Aqvist, J. Combining docking, molecular dynamics and the linear interaction energy method to predict binding modes and affinities for non-nucleoside inhibitors to HIV-1 reverse transcriptase. J. Med. Chem. 2008, 51, 2648−2656.
(55) Jorgensen, W. L. Efficient drug lead discovery and optimization. Acc. Chem. Res. 2009, 42, 724−733.
(56) Jacobson, K. A.; Gallo-Rodriguez, C.; Melman, N.; Fischer, B.; Maillard, M.; van Bergen, A.; van Galen, P. J.; Karton, Y. Structure- activity relationships of 8-styrylxanthines as A2-selective adenosine antagonists. J. Med. Chem. 1993, 36, 1333−1342.
(57) Katritch, V.; Kufareva, I.; Abagyan, R. Structure based prediction of subtype-selectivity for adenosine receptor antagonists. Neuro- pharmacology 2011, 60, 108−115.
(58) Kolb, P.; Phan, K.; Gao, Z. G.; Marko, A. C.; Sali, A.; Jacobson, K. A. Limits of ligand selectivity from docking to models: in silico screening for A(1) adenosine receptor antagonists. PloS One 2012, 7, e49910.
(59) Shoichet, B. K.; Kuntz, I. D. Matching chemistry and shape in molecular docking. Protein Eng. 1993, 6, 723−732.
(60) Nicholls, A.; Honig, B. A rapid finite-difference algorithm, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. J. Comput. Chem. 1991, 12, 435−445.

(61) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. A new force-field for molecular mechanical simulation of nucleic-acids and proteins. J. Am. Chem. Soc. 1984, 106, 765−784.
(62) Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. Automated docking with grid-based energy evaluation. J. Comput. Chem. 1992, 13, 505− 524.
(63) Shoichet, B. K.; Leach, A. R.; Kuntz, I. D. Ligand solvation in molecular docking. Proteins: Struct., Funct., Genet. 1999, 34,4−16.
(64) Omega, OpenEye Scientific Software v 2.3.2; 2008. www. eyesopen.com.
(65) Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Model for aqueous solvation based on class IV atomic charges and first solvation shell effects. J. Phys. Chem. 1996, 100, 16385−16398.
(66) Li, J. B.; Zhu, T. H.; Cramer, C. J.; Truhlar, D. G. New class IV charge model for extracting accurate partial charges from wave functions. J. Phys. Chem. A 1998, 102, 1820−1831.
(67) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An all atom force-field for simulations of proteins and nucleic-acids. J. Comput. Chem. 1986, 7, 230−252.
(68) JChem, Chemaxon, v 5.11.4; 2012. http://www.chemaxon.com.
(69) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J. Chem. Theory Comput. 2008, 4, 435−447.
(70) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J. Am. Chem. Soc. 1996, 118, 11225−11236.
(71) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 1983, 79, 926−935.
(72) Berger, O.; Edholm, O.; Jahnig, F. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. Biophys. J. 1997, 72, 2002−2013.
(73) Marelius, J.; Kolmodin, K.; Feierberg, I.; Aqvist, J. Q: A molecular dynamics program for free energy calculations and empirical valence bond simulations in biomolecular systems. J. Mol. Graphics Modell. 1998, 16, 213−225.
(74) Hetgrp_ffgen, Schrodinger, 2011. http://www.schrodinger.com.
(75) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical- integration of Cartesian equations of motion of a system with constraints - molecular-dynamics of N-alkanes. J. Comput. Phys. 1977, 23, 327−341.
(76) King, G.; Warshel, A. A surface constrained all-atom solvent model for effective simulations of polar solutions. J. Chem. Phys. 1989, 91, 3647−3661.
(77) Lee, F. S.; Warshel, A. A local reaction field method for fast evaluation of long-range electrostatic interactions in molecular simulations. J. Chem. Phys. 1992, 97, 3100−3107.
(78) Brandsdal, B. O.; Osterberg, F.; Almlof, M.; Feierberg, I.; Luzhkov, V. B.; Aqvist, J. Free energy calculations and ligand binding. Adv. Protein Chem. 2003, 66, 123−158.
(79) Smith, P. K.; Krohn, R. I.; Hermanson, G. T.; Mallia, A. K.; Gartner, F. H.; Provenzano, M. D.; Fujimoto, E. K.; Goeke, N. M.; Olson, B. J.; Klenk, D. C. Measurement of protein using bicinchoninic acid. Anal. Biochem. 1985, 150,76−85.

Chapter 5

A New Strategy for Over-expression and Functional Reconstitution of G Protein-Coupled Receptors

— Human Adenosine A_{2A} Receptor

ABSTRACT

Structural information plays a key role in modern drug discovery, in combination with biophysical screening, enables hit identification and generates compounds with drug-like properties. An increasing number of GPCR structures enables virtual screening and structure-based drug discovery for these important pharmaceutical targets. Here, we aim to investigate solution NMR as a potential means to elucidate the structure of $A_{2A}R$ in complex with potential allosteric modulators discovered by previous fragment screening. For this, we overexpressed the recombinant receptor as inclusion bodies in the *E. coli* prokaryote system, using a human α_5 integrin as fusion partner. With this strategy, the recombinant adenosine A_{2A} receptor was produced and purified in large amounts. However, optimal refolding conditions to recover the isolated receptor in its native state have not yet been found. If successful, this novel approach could be powerful aid to investigate orthosteric ligand and allosteric modulators binding to GPCRs in general.

1. Introduction

G-protein coupled receptors (GPCRs) are integral membrane proteins that respond to a variety of ligands such as biogenic amines, peptides and hormones, and create intracellular responses via G proteins, β-arrestins, and other downstream effectors. GPCRs mediate the majority of transmembrane signals in living cells, and represent 30–40% of the current drug targets which have been developed for cardiovascular, metabolic, neurodegenerative, psychiatric, and oncologic diseases.[1] The adenosine A_{2A} receptor is one of four adenosine receptor subtypes (A_1, A_{2A}, A_{2B}, and A_3), belongs to the family A of GPCRs, and is activated by endogenous adenosine.[2] The A_{2A} receptor plays a role in regulating the inflammatory response and ischemic brain damage. Several A_{2A} receptor agonists and antagonists/inverse agonists have passed successfully into clinical trials: agonist developed by Biovitrum (BTV.115959, structure not disclosed) [3] may be used for treatment of inflammatory disease, whereas antagonists/inverse agonists such as preladenant and istradefylline, [4,5] may be used for treatment of Parkinson's disease as a non-dopaminergic approach.

To date, more than 70 high-resolution structures of GPCRs have been solved, several structures have been solved in their active-state conformations and as complexes with an agonist, an inverse agonist, an antibody and even a G-protein.[6] Such structures can provide detailed information on protein form and function at atomic level and are also useful for characterization of protein-ligand interactions. However, obtaining novel high-resolution GPCR structures in complex with multiple ligands is still very challenging, largely due to intrinsic protein flexibility, instability in detergent, and the requirement of specialized crystallization techniques like lipidic cubic phase crystallization.[7]

Most of the efforts in the structural biology and medicinal chemistry of adenosine receptors have concentrated on orthosteric ligands. More recently, allosteric modulators appear to be an emerging class of therapeutically interesting agents that bind to receptors at a topographically distinct site of orthosteric ligands.[8] These molecules have no intrinsic activity on their receptors,

they modulate the activity of the receptors when the orthosteric ligand is bound concurrently. Allosteric ligands that enhance the agonist activities are referred to as positive allosteric modulators (PAMs) while those that inhibit or decrease the activities are called negative allosteric modulators (NAMs). There are a number of therapeutic advantages of using allosteric modulators over orthosteric ligands, such as higher selectivity for the target, decreased toxicity or side effects, enhanced physiological specificity of action.[9] However, allosteric modulators have only been described for a few GPCRs. Designing specific allosteric ligands for GPCRs has been hampered by a lack of structural information on their binding sites. However, biophysical methods like NMR and SPR can directly measure the GPCR-ligand/modulator interaction. Use of NMR techniques could be valuable for future allosteric modulator design with resonance assignment for the putative allosteric sites, and an improved molecular dynamic understanding of functional activities of GPCRs.

Protein-observed NMR studies usually require isotope labelled protein at high concentration. *Escherichia coli* is the most widely used protein expression system for structure biology studies using NMR techniques. However, bacterial expression of GPCRs has been hampered by low levels of expression and cell toxicity caused by targeting the recombinant receptors to the bacterial inner membrane. Recent new techniques contribute to a breakthrough on the difficulties encountered in GPCR expression, purification and refolding, allowing preparation of active samples for further biophysical and structure studies. In this chapter, we combined several protein production strategies to prepare sufficient amount and functional samples for receptor–small molecule interaction determination using solution-state NMR. We took advantage of the stabilized receptor (StaR) strategy, which has been developed to overcome the instability and purification problems of GPCRs.[10] Several point mutations are made outside of known orthosteric sites, and increase thermostability of the receptor by locking it in one conformation (i.e., agonist or antagonist conformation).[11] The recombinant A_{2A} StaR was produced in inclusion bodies (IBs) in *E. coli*. IBs are mechanically stable and not toxic to the cell, so GPCR production level can be

exceptionally high. Then the receptor was purified in a mixture of sodium dodecyl sulfate (SDS) and urea solution in an inactive form. The receptor was subsequently folded to its native state by transferring it from SDS to a polymeric surfactant called amphipols (APols). APols are amphipathic polymers that can keep individual integrated membrane proteins soluble in aqueous solutions, and preserve their functionality.[12] It has been shown that several MPs are more stable in APols than in detergent micelles, and Apols are a highly efficient agent to fold MPs into their native state.[13] Solution NMR techniques have been successfully applied to GPCRs in complex with APols for protein-ligand interaction and conformational transition characterization.[14]

In chapter 3, we have shown that TINS is suitable to screen fragment libraries for target specific binding. We have been able to discover structurally-independent new chemical series with functional activity on the human $A_{2A}R$. In addition, one compound has shown to be active in an intracellular cAMP assay. It did not compete with radioactive orthosteric ligand binding, suggesting an allosteric mechanism of action.[15] Due to the low potency of this compound, it is not clear whether the high-resolution crystal structure of $A_{2A}R$ in complex with this allosteric modulator could be solved. However, APol-trapped A_{2A}-StaR may enable NMR analysis of protein-small molecule complexes. A StaR with high production yield, enhanced stability and reduced conformational exchange may enable NMR resonance assignment for the putative allosteric site near the extracellular loop domains.

As a first step, in this study, we have shown that the recombinant A_{2A} StaR can be produced in large quantities in inclusion bodies in *E.coli*, yielding 20-30 mg recombinant protein per liter of cell culture. The first bottleneck in the determination of the GPCR structure pathway has thus been solved. However, the protein fusion partner needs to be efficiently removed, then the purified isolated target protein has to be refolded into APols. Fuctional APol-trapped A_{2A}-StaR in complex with its ligand or allosteric modulator may enable solution NMR analysis.

2. Results and Discussion

Using a New Fusion Partner for Efficient GPCR Over-Expression in *E. coli* IBs

Expression of GPCRs in *E. coli* is generally targeted to insoluble IBs in the cytoplasm to avoid toxicity associated with expression in membrane.[16] However, when using a protein partner like maltose binding protein (MBP), some GPCRs can be produced in the *E. coli* inner membrane in the folded active state, albeit at very low yield.[17,18] For example, the production of MBP fused A_{2A} StaR in the *E. coli* inner membrane is very low, yielding only 0.4 mg purified protein per liter of cell culture (data not published). Here, we used an approach that had been successfully applied to expression of the leukotriene BLT1 and BLT2 receptors.[16,19] This approach is based on accumulation of the receptor protein in *E. coli* IBs as a fusion protein followed by *in vitro* refolding, which may eventually lead to their native state. A fragment of the extracellular β-propeller domain of the human α_5 integrin (α_5I) was selected as a fusion helper partner. The α_5I contains a high proportion of charged residues (-19) and higher fraction of β-turn-forming residues (31.6%) than the usual fusion partner like glutathione *S*-transferase (GST) or phospholipase C.[19] These two characteristics are important for efficient inclusion body formation.[20]

In this study, pET21b was used as expression vector and several constructs were designed. A His_{10} tag was present at either terminus (Figure 1A), and a thrombin cleavage site was inserted between the α_5I and A_{2A}R gene. Furthermore, to get better separation of cleaved A_{2A}R and non-cleaved fusion protein, the construct was optimized with His_{10} tag located in between the fusion partner and the receptor protein (Figure 1A), followed by two enzymatic cleavage sites: thrombin and Tobacco Etch Virus (TEV). In addition, a poly-glycine linker was introduced between the TEV cleavage site and the N-terminus of A_{2A}R, in order to gain more accessibility for the enzyme.

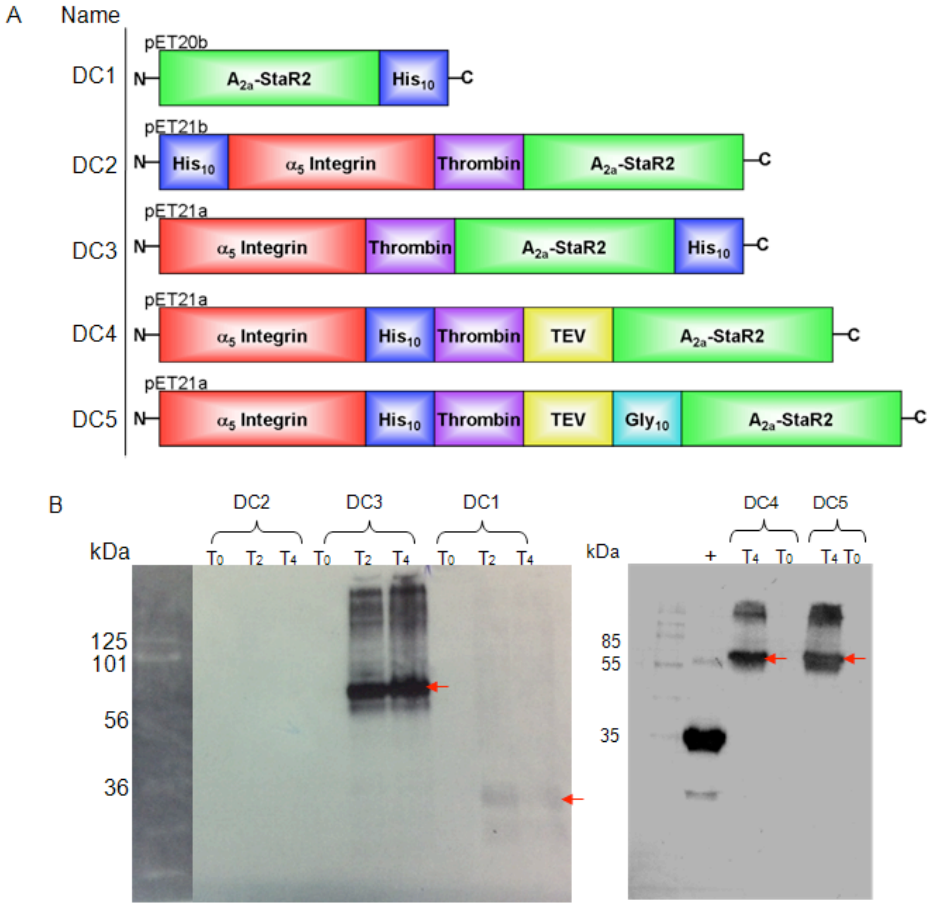


Figure 1. A) Schematic representation of designed constructs in pET20 expression vector in this study. And B) comparison of total expression levels of five different $A_{2A}R$ constructs expressed in Rosetta 2 (DE3) strain at 37 °C. Different samples were taken before IPTG induction (T_0), and after 2 hours (T_2) and/or 4 hours (T_4) IPTG induction. Anti $A_{2A}R$ Western blots were performed with total lysates.

GPCR Production and Purification

The total expression level and integrity of the different expression constructs was monitored by Western blot (WB) using anti-$A_{2A}R$ antibodies (Figure 1B). According to the intensity of the protein band, α_5I proved to be a very efficient fusion partner for GPCR over-expression in IBs. Without α_5I, minimal protein expression was observed. When the His_{10} tag was present at the N-terminus of the protein, expression was completely blocked. The apparent molecular weight on WB of each construct was compatible with the

corresponding calculated masses, 67 kDa for the construct **DC3** and 35 kDa for the $A_{2A}R$-His_{10}.

Expression of **DC3** in Rosetta 2 (DE3) strain was scaled up at 37 °C, cells were harvested after four hours IPTG induction. In cell lysates, IBs can be easily isolated from other cell constituents by centrifugation, then solubilized in denaturing buffer, a mixture of urea and SDS and subsequently purified by immobilized metal affinity chromatography (IMAC) on a Ni^{2+} NTA column. Typically, 30 mg of purified recombinant protein was recovered from one liter of cell culture (Figure 2B). Before refolding, the integrin fusion partner must be removed. The protein was dialyzed in an aqueous buffer to eliminate most SDS and urea, which could inhibit the thrombin cleavage. In addition to driving protein into IBs, the α_5I fusion can keep the receptor soluble after dialysis to remove detergents.[21] Unfortunately, in this case, the enzymatic cleavage was not efficient, only 30-40% of the α_5I partner was removed by thrombin cleavage from the recombinant protein. As illustrated in Figure 2B, inefficiently cleaved α_5I-$A_{2A}R$-His_{10} appeared as three major bands, corresponding to uncleaved recombinant protein at 67 kDa, the cleaved α_5I partner around 31 kDa and cleaved $A_{2A}R$ around 30 kDa. The cleaved $A_{2A}R$ and uncleaved protein cannot be further separated by a second IMAC purification, due to the His_{10} tag present at C-terminal for both cases. An additional gel filtration chromatography cannot separate them either (data not shown).

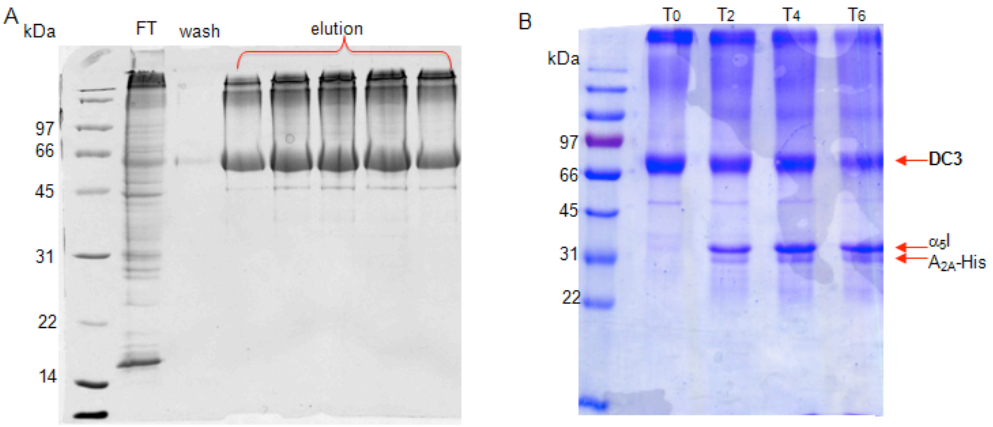


Figure 2. Coomassie-stained 12.5 % SDS-PAGE of construct **DC3** (A) after Ni^{2+} NTA column purification: FT flow through, wash, and elution peaks from Ni^{2+} NTA column and (B) thrombin

cleavage efficiency monitored after 0, 2, 4, and 6 hours incubation. Bands corresponding to the uncleaved α_5I-A_{2A}R (67 kDa), α_5I (31 kDa), and isolated A_{2A}R (35 kDa) are marked with red arrows.

To solve the inefficient cleavage of the fusion partner, a new construct was designed with a His_{10} tag present at the C-terminus of integrin, before the enzymatic cleavage sites. A TEV cleavage site was also added (Figure 1A). These modifications did not improve the enzymatic cleavage efficiency either by thrombin or TEV protease (Figure 3A). According to the 3D crystal structure, A_{2A}R has a very short N-terminal extension (6 amino acids) that projects out of its first transmembrane α-helix, suggesting that the enzymatic cleavage site might be shielded by the α-helix, thus not accessible for enzyme. Therefore, a 10 glycine linker was introduced between the TEV cleavage site and the N-terminus of A_{2A}R, which didn't affect the expression level of fusion protein (Figure 1B). After eliminating most SDS and urea by dialysis, the fusion protein was incubated with TEV protease at a 1:1 molar ratio for 2, 4 or 6 hours at room temperature. As shown on the SDS PAGE gel (Figure 3B), approximately 80% of fusion protein was cleaved. Immediately, the α_5I fusion was eliminated through a second IMAC step under denaturing condition, 0.8% SDS (Figure 3C). The A_{2A}R was kept soluble in high concentration of SDS solution, in which GPCR displays a significant α-helix content.[22] In the flow through of Ni NTA column, as shown in Figure 3C, a number of bands were observed on SDS PAGE, which might correspond to aggregation or precipitation of cleaved A_{2A}R. Proteins tend to precipitate during enzymatic cleavage, because of the low level of detergent present in the reaction. Although a high concentration of SDS was added during the second IMAC purification, it is difficult to re-dissolve the precipitates. Many of the membrane bound proteins show a smeared band or a dark band at the very top of the gel, when they tend to precipitate.

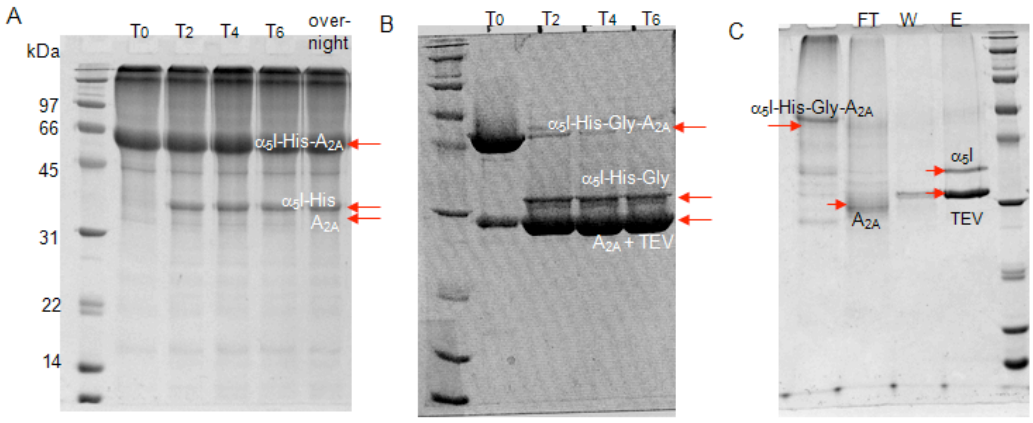


Figure 3. Coomassie-stained 12.5 % SDS-PAGE analysis of enzymatic cleavage efficiency of **DC4** (A) and **DC5** (B). Figure 3A, thrombin cleavage product after 2, 4, or 6 hours incubation, overnight incubation of fusion protein with enzyme did not improve the cleavage efficiency. Figure 3B, TEV cleavage product of **DC5** after 2, 4, or 6 hours incubation. Figure 3C, FT, flow through after Ni NTA column; W, wash of Ni NTA column; E, elution of Ni-NTA column.

In Vitro Refolding of $A_{2A}R$

SDS was used as an artificial chaperone to keep the GPCR soluble and to ensure that the GPCR does not aggregate before it reaches its native form.[23,24] *In vitro* refolding of GPCR can be achieved by exchange of SDS solubilized receptor into amphipols (A8-35) complex, with lipid supplementation. A8-35 was added at 1:5:1 protein/A8-35/asolectin mass ratio and SDS was removed by potassium chloride precipitation followed by dialysis. There was no significant protein loss through precipitation during amphipol-mediated folding. To compare Apol-mediated folding efficiency with that of detergent folding, 2.5%/0.5% DDM/CHS was added to the SDS solubilized receptor to induce folding. The functional fraction of $A_{2A}R$ was then assessed by [^{3}H] ZM241385 radioligand competition assay. Unfortunately, there is no [^{3}H] ZM241385 binding shown on the refolded receptor preparations, so both APol-trapped $A_{2A}R$ and the DDM solubilized $A_{2A}R$ were not folded in their native state.

3. Conclusions - What to Do Next?

Overexpression of GPCRs has been carried out in several different expression systems such as yeast, insect cells, and mammalian cells. Among all the expression systems developed so far, *E. coli* is considered as an ideal host for producing isotope labeled proteins for NMR structure studies. However, GPCR expression in *E. coli* usually leads to cell toxicity and low levels of expression. For this reason, a limited number of receptors have been produced as functional proteins in *E. coli* inner membrane. In this chapter, an alternative approach has been investigated to overexpress recombinant GPCR targeting IBs.

In order to increase the yield of expressed protein, an intergrin α_5I fragment was introduced as fusion partner in the expression vector, which helps addressing GPCRs to IBs. α_5I fused GPCR can accumulate in milligram amounts in IBs. Among the different constructs, N-terminal His_{10} tag presents a noticeable negative effect on protein expression. For purification purpose, a 10 glycine linker was introduced between the enzymatic cleavage site and the N-terminus of $A_{2A}R$, which made TEV recognition site more accessible for enzyme. However, pure and well isolated protein is critical for further successful refolding. 1) More efforts need to be put into optimization of enzymatic cleavage conditions. Efficient removal of the fusion partner could provide purer isolated GPCR for further folding step. 2) Re-solubilization of GPCR after TEV cleavage could be tested in high concentration of a harsh detergent such as 0.6-0.8% SDS and 6-8 M urea; and 3) The second IMAC purification conditions need to be further optimized. Moreover, the GPCR refolding preparation is not yet fully successfully set up. Refolding of the StaR GPCR in APols might not be feasible, because StaR was designed for making protein more stable in short chain detergent, like *n*-decyl-β-D-maltopyranoside (DM).[7] New constructs based on wild-type receptor can be followed.

An alternative method, directed evolution for high expression and high stability, has the potential to produce high levels of a thermostable neurotensin receptor 1 from *E. coli* for structural studies.[25] This *E. coli* expression strategy can be suitable for producing isotope-labeled proteins, and therefore might be a viable alternative to the present approach.

4. Materials and Methods

Construction of $A_{2A}R$ Fusions

The expression vector pET21a-α_5I, containing coding sequence for the α_5 integrin (α_5I) fragment (amino acid residues 231-514), was kindly provided by Jean-Louis Banère (Université de Montpellier). The coding sequence of A_{2A}-$StaR_2$ was commercially synthesized, and optimized for bacterial expression. For GPCR expressed alone, the A_{2A}-$StaR_2$ coding sequence was inserted into the BamHI/XhoI sites of the pET20b vector. For GPCR expressed as a fusion, the A_{2A}-$StaR_2$ coding sequence was inserted into the BamHI/XhoI sites of the pET21a-α_5I vector. A thrombin cleavage site or a Tobacco Etch Virus (TEV) cleavage site between the two regions coding for α_5I and A_{2A}-$StaR_2$ was introduced to enable removal of the fusion partner α_5I. A His_{10} tag was appended at the N-terminus or C-terminus to allow easy purification by immobilized metal affinity chromatography (IMAC).

Expression of $A_{2A}R$

For expression, *E.coli* Rosetta™ 2 (DE3) cells were transformed freshly each time with different expression constructs and grown overnight at 37 ºC on Luria-bertani (LB)-agar plates containing 100 μg/ml ampicillin. For each construct, an isolated colony was picked up from the plate and the bacteria were grown at 37 ºC in LB medium supplemented with 100 μg/ml ampicillin, 34 μg/ml chloramphenicol and 0.2 % glucose until an OD_{600} at 0.6. Protein expression was induced by adding 1 mM isopropylthio-β-D-galactoside (IPTG), and growth

continued at 37 ºC for 4 hours. Cells were harvested at 4 ºC by centrifugation at 4,000 x g for 20 min, and the cell pellet was resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 1M urea and one tablet of complete protease inhibitor cocktail) then stored at −80 ºC. Total expression was visualized by Western blot using anti-$A_{2A}R$ antibody according to standard protocols.

Purification of Recombinant $A_{2A}R$

The cell pellet was thawed on ice before adding 10 mg/ml lysozyme and 20 μg/ml DNase, then lysed by French Press. IBs were recovered at 4 ºC by centrifugation at 19,000 rpm for 45 min, washed with 40 ml of 20 mM Tris-HCl pH 8.0 buffer containing 1 M urea and 0.5% Triton X-100, respectively. Solubilization of IBs was carried out in 40 ml of solubilization buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 6 M urea, 0.4% SDS, 10% glycerol, 4 mM β-mercaptoethanol and protease inhibitors) overnight at 4 ºC on rotation. Solubilized proteins were clarified from insoluble material at 4 ºC by centrifugation at 19,000 rpm for 45 min and passed through a 0.22 μm filter.

Solubilized IBs were applied to an Ni^{2+} NTA column equilibrated with solubilization buffer. The column was then washed with 10 column volumes of solubilization buffer supplemented with 10 mM imidazole. The recombinant GPCR was eluted with 20 column volumes of elution buffer (solubilization buffer + 100 mM imidazole). Fractions of elution peak were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), the protein concentration was determined by absorption at 280 nm.

Thrombin Cleavage and TEV Cleavage of Recombinant $A_{2A}R$

The purified α_5I fused $A_{2A}R$ was extensively dialyzed in 3 X 1L of 20 mM Tris-HCl pH 8.0 and 150 mM NaCl using dialysis membrane with a 12-14 KDa molecular weight cut-off. The first two dialyses were carried out at room temperature to prevent SDS crystallization, and the last dialysis was transferred to 4 ºC to prevent protein degradation. The thrombin (Millipore) cleavage was done at room temperature using 10 cleavage units per mg of fusion protein,

supplementing the solution with 2.5 mM $CaCl_2$. The TEV cleavage was performed at room temperature using a TEV protease and cleavage buffer (25 mM Tris-HCl pH 8.0, 150 mM NaCl and 2 mM DTT). The optimized enzyme:fusion protein ratio and incubation time were determined by monitoring the samples on SDS-PAGE. The reaction was stopped by adding 6 M urea and SDS up to 0.4 %.

The cleaved sample was incubated with 6 ml 50% Ni^{2+} NTA superflow slurry for 3 hours at room temperature. The sample-resin suspension was transferred into a chromatography column. In the case of α_5I-His-TEV-A_{2A}R, the cleaved A_{2A}R was present in the flow through. For other constructs, the cleaved A2aR was bound to the Ni^{2+} NTA agarose, and eluted with elution buffer (25 mM Tris-HCl pH 8.0, 150 mM NaCl, 6 M urea, 0.4 % SDS and 100 mM imidazole). The purity of the protein was examined by SDS-PAGE. The purified and isolated A_{2A}R was pooled and dialyzed for 24 hours at room temperature in 3 X 1L of 50 mM Tris-HCl pH 8.0 and 0.8% SDS to eliminate urea and imidazole.

Refolding of A_{2A}R

The quantity of A_{2A}R was determined by absorption at 280 nm. For optimal refolding, the concentration of A_{2A}R should be between 0.3-0.5 mg/ml. The GPCR renaturation was carried out by precipitating dodecyl sulfate as a potassium salt (KDS). Asolectine was added, at a mass ratio of 1 g lipid per g GPCR, to SDS-solubilized A_{2A}R supplemented or not with 10 μM ZM241385. After 1 hour of incubation at room temperature, amphipol A8-35 was added to the sample, at the mass ratios of 5 or 10 g A8-35 per g GPCR. After 30 min of incubation, SDS was precipitated by addition of KCl. KCl was added to a final concentration equal to concentration of SDS + 150 mM. After 30 min incubation under vigorous stirring, the KDS precipitate was removed by centrifugation and the supernatant was dialyzed three times against 1 L of potassium buffer (30 mM potassium phosphate pH8.0 and 150 mM KCl). To determine whether amphipol could refold SDS denatured sample to an active receptor, a high affinity radioligand ([^{3}H]ZM241385) binding competency was performed. For the

detergent folding preparation, the SDS solublized receptor was diluted in the final buffer, with 25 mM Tris-HCl pH 8.0 and 0.1% SDS, before the addition with stirring of 2.5 % *n*-dodecyl-β-D-maltopyranoside (DDM), 0.5 % cholesteryl hemisuccinate (CHS), and 10 μM ZM241385. After 48 hours stirring at 4 °C, the unsolubilized and aggregated receptors were removed through Superdex 200 10/300 GL gel filtration column.

Western Blot

Proteins were transferred from a 12 % SDS-PAGE gel to Hybond PVDF membrane, which was blocked with 5 % non-fat milk powder in PBS buffer pH7.4 at room temperature. Then, the membrane was treated with anti-$A_{2A}R$ primary antibodies, epitope mapping to the third intracellular loop. Protein specific bands were visualized by ECL chemiluminescence.

Radioligand Binding Assays

For the equilibrium binding assay, 25 μl of protein sample was incubated with [^{3}H]ZM241385 at final concentrations of 1, 5 and 16 nM, in a total volume of 100 μl Tris-HCl buffer (50 mM, pH 7.4), 10 mM $MgCl_2$. Non-specific binding was determined using a final concentration of 10 μM ZM241385. After incubation for 1 hour at 25 °C in a shaking water bath, assays were terminated by rapid filtration through Whatman GF/B filters pre-socked in 0.3 % polyethyleneimine under reduced pressure with a Brandell harvester. Filters were washed by 3 × 2 ml ice-cold Tris-HCl buffer (50 mM, pH 7.4) and placed in scitillation vials. 3.5 ml of Emulsifer Safe was added, after 3 hours, radioactivity was counted in an LBK rack β scintillation counter.

Reference

1. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? Nat. Rev. Drug Discov. 5, 993–996 (2006).
2. Foord, S. M. et al. International Union of Pharmacology . XLVI . G Protein-Coupled Receptor List. Pharmacol. Rev. 57, 279–288 (2005).
3. Müller, C. E. & Jacobson, K. a. Recent developments in adenosine receptor ligands and their potential as novel drugs. BBA-Biomembranes 1808, 1290–1308 (2011).
4. Müller, C. E. & Ferré, S. Blocking striatal adenosine A2A receptors: a new strategy for basal ganglia disorders. Recent Pat. CNS Drug Discov. 2, 1–21 (2007).
5. Andrews, S. P. & Tehan, B. Stabilised G protein-coupled receptors in structure-based drug design: a case study with adenosine A2A receptor. Medchemcomm 4, 52 (2013).
6. Venkatakrishnan, a J. et al. Molecular signatures of G-protein-coupled receptors. Nature 494, 185–94 (2013).
7. Bennett, K. a et al. Pharmacology and Structure of Isolated Conformations of the Adenosine A2A Receptor Define Ligand Efficacy. Mol. Pharmacol. (2013). doi:10.1124/mol.112.084509
8. Keov, P., Sexton, P. M. & Christopoulos, A. Allosteric modulation of G protein-coupled receptors: a pharmacological perspective. Neuropharmacology 60, 24–35 (2011).
9. Conn, P. J., Christopoulos, A. & Lindsley, C. W. Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. Nat. Rev. Drug Discov. 8, 41–54 (2009).
10. Magnani, F., Shibata, Y., Serrano-Vega, M. J. & Tate, C. G. Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor. P. Natl. Acad. Sci. USA 105, 10744–10749 (2008).
11. Doré, A. S. et al. Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthines XAC and caffeine. Structure 19, 1283–1293 (2011).
12. Popot, J.-L. et al. Amphipols from A to Z. Annu. Rev. Biophys. 40, 379–408 (2011).
13. Dahmane, T., Damian, M., Mary, S., Popot, J.-L. & Banères, J.-L. Amphipol-assisted in vitro folding of G protein-coupled receptors. Biochemistry 48, 6516–6521 (2009).
14. Catoire, L. J. et al. Structure of a GPCR ligand in its receptor-bound state: leukotriene B4 adopts a highly constrained conformation when associated to human BLT2. J. Am. Chem. Soc. Am. Chem. Soc. 132, 9049–9057 (2010).
15. Kenakin, T. & Christopoulos, A. Analytical pharmacology: the impact of numbers on pharmacology. Trends Pharmacol. Sci. 32, 189–196 (2011).
16. Baneres, J.-L. et al. Structure-based Analysis of GPCR Function: Conformational Adaptation of both Agonist and Receptor upon Leukotriene B4 Binding to Recombinant BLT1. J. Mol. Biol. 329, 801–814 (2003).
17. Weiss, H. M. & Grisshammer, R. Purification and characterization of the human adenosine A2a receptor functionally expressed in Escherichia coli. Eur. J. Biochem. 269, 82–92 (2002).
18. White, J. F., Trinh, L. B., Shiloach, J. & Grisshammer, R. Automated large-scale purification of a G protein-coupled receptor for neurotensin. FEBS lett. 564, 289–293 (2004).
19. Arcemisbéhère, L. et al. Leukotriene BLT2 receptor monomers activate the G(i2) GTP-binding protein more efficiently than dimers. J. Biol. Chem. 285, 6337–6347 (2010).
20. Wilkinson, D. L., Harrison, R. G. Predicting the solubility of recombinant proteins in Escherichia coli. Biotechnology 9, 443–448 (1991).
21. Michalke, K. et al. Mammalian G protein-coupled receptor expression in Escherichia coli: II. Refolding and biophysical characterization of mouse cannabinoid receptor 1 and human parathyroid hormone receptor 1. Anal. Biochem. 401, 74–80 (2010).
22. Muller, I., Sarramégna, V., Renault, M., Lafaquière, V., Sebai, S., Milon, A., Talmont, F. The full-length mu-opioid receptor: a conformational study by circular dichroism in trifluoroethanol and membrane-mimetic environments. J. Membr. Biol. 223, 49–57 (2008).
23. Rozema, D. & Gellman, S. H. Artificial Chaperones: Protein Refolding via Sequential Use of Detergent and Cyclodextrin. J. Am. Chem. Soc. 117, 2373–2374 (1995).
24. Daugherty, D. L. Artificial Chaperone-assisted Refolding of Citrate Synthase. J. Biol. Chem. 273, 33961–33971 (1998).
25. Egloff, P. et al. Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. 111, E655–62 (2014).

Chapter 6

General Discussion

Integral membrane proteins including GPCRs and ion channels are key regulators of cellular function and account for up to two thirds of known druggable targets,[1] highlighting their critical pharmaceutical importance. Traditionally, new drugs against this class of proteins have been discovered through high-throughput screening. However, not all GPCRs are amenable to traditional screening and pharmacological evaluation methods. Recently, fragment-based drug discovery (FBDD) in combination with structure-based drug discovery (SBDD) has emerged as a powerful strategy to generate approved drugs against soluble targets such as kinases and proteases.[2–4] Now, FBDD/SBDD can be applied to GPCRs with great potential advantages.[5] In recent years, a number of FBDD techniques have been validated for use with GPCRs. Biophysical methods (Chapter 2 and 3) are usually employed for hit identification. The impressive growth in GPCR structure information leads to broad use of structure-based methods for hit discovery and optimization (Chapter 4). However, the dynamic nature of GPCRs, and standard issues associated with low level expression and instability during purification, made biophysical and structural characterization of GPCRs particularly difficult. New advances in protein stabilization by using different protein engineering methods and alternative solubilization strategy (Chapter 2, 3, and 5) have shown the potential to facilitate GPCR structural and biophysical studies. The goal of the work described in this thesis was to develop and implement efficient fragment screening methods to discover ligands of GPCRs with novel biological activities, and new advances in receptor production and stabilization to facilitate structural biology of GPCRs in the early stages of drug discovery. Below some commentary on the current state and perspectives of these approaches is provided.

Screening Techniques for Fragment-Based Drug Discovery

The first step in FBDD is the identification of fragment hits that have sufficiently high ligand efficiency and can be used as a starting point for hit-to-

lead optimization. Fragments are low-molecular weight compounds, making only few interactions with the target protein and displaying low binding affinity. This makes the interaction particularly difficult to detect by conventional biochemical assays. However, biophysical methods such as nuclear magnetic resonance (NMR), surface plasmon resonance (SPR) and X-ray crystallography are commonly used to identify fragment hits. Especially, NMR based methods and SPR provide sensitivity that is required for detection of low affinity ligands. As described in Chapters 2 and 3, we have established the efficacy of TINS for fragment screening against membrane proteins, including GPCRs. TINS can detect small molecules that bind to different binding sites on the target protein. It has been shown that it is possible to discover ligands with different biological activities in a single screen. In Chapter 2, the fragments identified by TINS were validated as specific binders of DsbB and they turned out to be either inhibitors of native ligand binding or inhibitors of protein-protein interactions. In Chapter 3, among the TINS hits, it is possible to find at once orthosteric ligands and multiple allosteric modulators of the $A_{2A}R$ with both positive and negative effects. Several of these ligands can be a good chemical start for further lead exploration. Use of a biophysical fragment screening method allows identification of very weak binders, such as the allosteric modulators, which are unlikely to be discovered by a traditional cell-based assay. Ligands with a wide range of affinities can be detected because of the sensitivity of TINS, thereby minimizing false negatives. TINS, as a direct binding assay, also allows discovery of ligands that bind to new binding sites that would be silent in radioligand binding assays. A reference control was included in the screen, eliminating false positives and non-specific protein binders, thus insuring the specificity of hits. Furthermore, a strong correlation has been noted between the hit rate in NMR fragment screening and the ligandability of a target. Nevertheless, an issue for TINS is that the parameters of the NMR experiment are optimized to detect weak binders. Tight binding ligands, which have slow off rates, will not generate as much NMR signal amplification as weak binders. Furthermore, a target protein which has a deeply

buried binding site, might not be suitable for detection of fragment ligands using TINS.

Contrary to TINS, SPR is capable of detecting binding of fragments with wide range affinity to protein targets. It should be applicable when TINS is less effective *i.e*. for tightly bound ligands. A number of SPR fragment screens against GPCRs, either thermostabilised[6,7] or solubilised wild type,[8,9] have been reported. In addition, SPR requires much less protein (25-50 μg) than NMR or X-ray crystallography. It is possible to immobilize receptors directly from crude solubilised extracts, avoiding loss of functionality during purification process. In order to run a high-throughput biophysical screening on large numbers of fragment compounds, it is important to separate real binding from non-specific binding. Successful SPR fragment screening requires parallel immobilization of the target protein, reference protein(s) and leaving one channel blank as a reference surface.[10] A known ligand must be available to assess the validity of the assay. The detectable minimal response level of fragment hits in the screen must be adjusted based on that of control compounds; this could be a limitation when there are no tool compounds available. Thereby, for many clinically validated targets, such as peptide and protein hormone binding receptors, hit identification using SPR methods is still challenging. An idea to conquer this limitation of SPR is to combine it with TINS. TINS can generate required tool compounds for SPR studies.

TINS and SPR have shown more advantages compared to traditional biochemical assay. Traditional biochemical assays are not generally applicable in FBDD due to their insufficient sensitivity to detect fragments with weak binding affinity to protein targets. Moreover, for peptide activated GPCRs, their native ligands, which bind with low nM affinities, are often too tight to be displaced by fragments that bind in the mM range. The high concentration of fragments required for screening may lead to a high level of false positives and negatives due to aggregation, compound reactivity, and artificial biological inhibition. Despite the disadvantages of biochemical assays, recent studies showed that good fragment hits could be identified by a screening with high

concentration of fragments.[2,11] The high level of false positives and negatives can be minimized by removing fragments with poor solubility, reactive properties, and aggregators from the fragment library.

Recent insight into GPCR structure has enabled the use of SBDD techniques to be applied to this target class. In Chapter 4, we have investigated the ability to identify hits in new areas of chemical space through *in silico* screening against GPCRs with solved crystal structures. One major strength of the *in silico* approach is the ability to screen large libraries (328,000 fragments are commercially available) and have better coverage of chemical space, which gives the opportunity to discover ligands with different chemotypes. We have shown in Chapter 4 that the docking-based screen was able to identify novel hits that did not overlap with those detected by NMR fragment screening. The present study showed that these two approaches are highly complementary and the combination of biophysical and *in silico* screening methods will be a powerful strategy for future fragment-based hit discovery against GPCRs. However, majority of GPCR crystal structures are in the inactive state, and the identification of agonists remains challenging.

Micelles vs. Nanodisc in Biophysical Fragment Screening

In order to detect binding of fragments to proteins, it is necessary to use highly sensitive biophysical methods such as NMR, SPR or X-ray crystallography, but the surfactants required to solubilize membrane proteins interfere with these methods. As membrane protein activity is enhanced in more bilayer-like environments, nanodiscs (NDs) offer the potential to enable a more generic approach to handling membrane proteins since they can be used to functionally solubilize a variety of membrane proteins.[12] The results described in Chapter 2 show that the use of NDs as a solubilization/immobilization system for ligand screening is effective and empty NDs represent a high-quality reference system to remove false positives, thereby ameliorating issues with nonspecific

binding of fragments to detergent micelles. The NMR spectra of the fragments during the TINS screen were of better quality in the presence of NDs as compared to detergent micelles. The signal-to-noise ratio of fragments using NDs was double to that in micelles, which suggested 30-40% compounds were non-specifically absorbed into the micelles. In addition, ligand screening in the presence of micelle-solubilized membrane proteins may bias the chemical nature of the fragment library toward more hydrophilic. This alternative solubilization method could enable more than just fragment screening for GPCRs that typically do not behave well in detergents. Several reports demonstrated a successful application of NDs for cell-free expression of membrane proteins.[13] NMR structural studies of less stable membrane proteins such as GPCRs may strongly benefit from the use of this non-conventional solubilization approach.[14] However, ND-embedded membrane proteins have relatively large size. The effective molecular weight of a protein-ND complex has important implications for the applicability of solution NMR. An increase in molecular size leads to wide ^{1}H resonance line widths, associated with low sensitivity and resonance overlap.[15]

NDs represent a relatively new membrane mimic, whereas liposomes (phospholipid bilayers) and bicelles represent a well-established, native-like environment for solution and solid state NMR studies. In addition, liposomes are likely to keep GPCRs in their biologically active conformation, thus increase receptor stability. Recently, high-resolution structural information of membrane proteins in liposomes under physiological conditions (temperature, pH, hydration), has become accessible, by using solid-state NMR techniques.[16,17] However, the limited solubility and the inaccessibility of the lipid vesicle interior may interfere with functional studies. Bicelles have shown their ability to preserve protein function and stability. Unlike liposomes, bicelles provide full access to both sides of membrane protein, which is important for the activity assay. Successful use of bicelles to reconstitute GPCRs has been reported.[18] Bicelles have been shown to be a suitable solubilization medium for both solution NMR and solid-state NMR studies,[15] because their spontaneous orientation in NMR magnetic fields. A limitation of using bicelles in solution NMR is their large

molecular weight, which usually leads to broad resonance lines. However, recent research has reported that bicelles solubilized sensory rhodopsin is accessible to NMR structure determination.[19] The observed sensitivity and resolution of NMR spectroscopy in non-detergent systems enables a detailed structural characterization. The use of non-detergent solubilization media should be strongly considered if protein stability needs to be increased and/or native structure can not be retained in a micelle environment.

Fragment Hit Validation

Validation of fragment hits from a biophysical screening assay is a critical step in the drug discovery process. A number of different approaches were used to quantitatively confirm hit binding and provide information on the binding mechanisms, such as competition binding TINS, biochemical assays, and SPR. The studies presented in Chapters 2 and 3 used a biophysical technique, TINS, to discover novel fragment ligands for membrane proteins. The biological activities of these ligands were characterized by enzymatic inhibition studies for DsbB (Chapter 2), and by radiochemical binding methods and cell-based functional assay for $A_{2A}R$ (Chapter 3). Biochemical assays, when possible, present an ideal hit validation and prioritization tool after a biophysical assay. However, many active fragment ligands can be missed by biochemical validation, simply due to their low potency. In order to minimize the false negatives, it is necessary to implement biophysical techniques to characterize the binding mode and discriminate orthosteric from allosteric ligands. A competition binding assay in TINS has been reported to validate and characterize the fragment hits for the β_1 adrenergic receptor.[6] The target protein is immobilized in both cells of the sample holder and the same mix is applied to both cells while a competitor (a known ligand for the target) is present only in one of the cells. Competition binding studies can be carried out using small molecules, protein, or DNA as competitors. This method allows rapid characterization of the ligand binding site

and binding mode. SPR assays are often used in FBDD as a highly quantitative, complementary technique to confirm and validate hits from the primary screen. SPR experiments are able to provide information of both binding affinity and kinetics to aid the prioritisation of fragment hits. In particular, the combination of TINS for primary fragment screening, and SPR for validation/ prioritization of hits has been shown to a powerful combination to rapidly find and evaluate weak binding fragment hits for further structural studies and elaboration projects.

Structure-Based Approaches for GPCR Drug discovery

Fragment based drug discovery (FBDD) involves the identification of low affinity fragments. Elaboration of lead-like compounds from weak binding fragment hits remains a significant challenge for FBDD.[20] The availability of high-resolution 3D structural information of target-fragment complexes makes structure based approaches a highly powerful strategy for fragment hit to lead elaboration. Both X-ray crystallography and NMR based techniques can provide atomic resolution structure of protein-ligand complexes, thereby allowing the optimization of molecules with increased affinity and selectivity for the receptor of interest. The field of GPCR structural biology has grown rapidly since the first high-resolution crystal structure of rhodopsin was obtained in 2000. At the time of writing, 70 high-resolution GPCR structures with different bound ligands and in different conformations have been reported. The great increase in crystal structures of GPCRs in different states has provided detailed insights into small molecule-GPCR interactions and highlighted key conformational changes in different receptors. New technical advances have been developed to overcome the obstacles associated with structure determination of GPCRs, such as i) overexpression of recombinant protein in different expression hosts; ii) increase of protein stability by introducing mutations, fusion partners and antibodies; iii) development of more native like environments for efficient solubilization. StaR® GPCRs developed by Heptares, enable biophysical (Chapter 3) and

crystallographic analysis, thus fragment and structure based drug design have been applied to this important pharmaceutical target.[7,21,22] However, obtaining small molecule-protein co-crystal structures is still challenging. As half of the GPCR crystal structures were solved with a resolution of approximately 3 Å, the interpretation of electron density maps can be ambiguous. Some amino acid residues like asparagine, glutamine and threonine might be uncertainly placed, due to undistinguishable nitrogen from oxygen in side-chain.[23] Interpretation of ligand binding site in crystal structures also becomes more difficult at low resolution. Poor electron density of non-covalently bound ligand can lead to incorrect orientation of ligand in the binding pocket. Moreover, obtaining novel GPCR structures in complex with different ligands is very challenging, primarily because of the requirement of highly potent ligands and new crystallization methods using a lipid environment for the protein. Complementary methods are required to enable the application of SBDD to GPCRs.

The biophysical mapping (BPM) method developed by Zhukov *et al*.[24] can generate information on ligand receptor interactions within the known binding pocket of GPCRs. In BPM, mutations are designed in the ligand binding site (side chain replacement of each residue) of multiple StaR® receptors, and the effect of mutations on binding affinity are compared across a matrix of ligands and mutants. This method can be used to refine the binding site and provide the key structural information on ligand-GPCR interactions in the absence of X-ray crystallography data. BPM of known ligands can predict the orientation of the key residues in the binding site. It can also be used to predict novel binding modes with different chemotypes as well as analogues within a particular chemical class.[25] This methodology facilitated the lead optimization of the hits identified by other screening methods like virtual screening.

Other structural biology techniques are starting to show promise, such as NMR based methods, which are sensitive to analyze weak protein-ligand interactions. They can be applied when X-ray crystallography is much less successful, such as weakly bound ligands, poorly diffracting crystals and low-resolution structures. For example, the structure of CXCR1 has been solved

using rotationally aligned solid state NMR in combination with molecular fragment replacement.[18] Interestingly, the uniformly $^{13}C/^{15}N$ labeled CXCR1 was expressed in *E. coli* in inclusion bodies and refolded in phospholipid bilayers.[26] While this structure provides information only about the protein backbone, solid state NMR has major potential to understand the ligand-receptor interaction. Recently technology of solution NMR has also made advances,[27,28] the structure elucidation of sensory rhodopsin II in phospholipid bicelles[19] and proteorhodopsin in detergent micelles[29] has recently been accomplished. Furthermore, there are a variety of NMR based methods that can be used to obtain structural information on ligand-protein interactions, such as chemical shift perturbation (CSP),[30] sparse nuclear Overhauser effect (NOE) based methods,[31] and paramagnetic NMR.[32] As described in Chapter 2, NMR based CSP analysis was applied to map the ligand binding sites at low resolution and predict the binding mode of two classes of DsbB inhibitors. A simple structural model proposed by CSP analysis is consistent with the results of the biochemical mode of action study. The stabilized GPCRs, with enhanced thermostability and reduced conformational exchange might enable resonance assignment for the putative allosteric sites, to further map the ligand binding site using CSP techniques.

Expression of GPCRs in *E. coli*

GPCRs generally express in very little amounts (<1 mg/L). A large number of expression systems such as bacteria, yeast, insect cells, and mammalian cells, have been evaluated in order to obtain large quantities of receptors necessary for structural studies. *E. coli* expression system presents several advantages compared with eukaryotes, including easy handling, inexpensive media, fast growth, high cell densities, and quick genetic modification. *E. coli* is suitable to produce fully isotope labeled protein, including deuteration, which is crucial for many NMR structural studies. A number of techniques have been developed, which enable efficient expression and correct folding of receptors: i) Expression of GPCRs as fusion proteins allows production of functional proteins

targeted to the inner membrane, such as the neurotensin receptor[33] and the adenosine A_{2A} receptor[34]. ii) An alternative approach is to overexpress the GPCR in inclusion bodies and subsequently refold it *in vitro*. Amphipols have been used to refold receptors like leukotriene receptors BLT1 and BLT2.[35] This method has been investigated to overexpress $A_{2A}R$ as described in Chapter 5. If the $A_{2A}R$ can be successfully refolded in amphipols, this could enable production of significant amounts of isotope labeled receptor for NMR structural biology studies. If so, the ortho- and allo- steric binding sites can be mapped at low resolution using CSP techniques.[36]

In most cases, an efficient refolding of a GPCR from IBs is still challenging. An alternative technique, CHESS (cellular high-throughput encapsulation, solubilization and screening), developed by Plückthun and colleagues, enables production of a reasonable quantity of functional protein in *E. coli*.[37] As *E. coli* is a suitable expression system for producing isotope labeled protein, CHESS provides the potential to apply NMR based structural biology to GPCRs for drug discovery. In addition, the improved stability of GPCR using directed evolution would facilitate biophysical fragment screening. However, in order to perform CHESS, many instruments are required and experiment set up is rather complex. Although it's a very powerful technique for GPCR production and stabilization, the CHESS intellectual property is legally protected and cannot be used for commercial purposes without authorization. To investigate the potential of this method, a collaborative drug discovery project between ZoBio and the lab of Prof. Andreas Pluckthün at the University of Zürich, will use CHESS technology to produce large quantities of thermostabilized receptor in *E. coli*. The GPCR will be engineered to attach paramagnetic tags at different sites on the protein. The pseudo-contact shift (PCS)[32] generated by the paramagnetic metal will be used to determine the binding site of small molecule ligands in rapid exchange with the GPCR.

To summarize, traditional GPCR drug discovery has focused on high-throughput screening of large compounds libraries using cell-based assays to discover novel molecules with selective activities such as agonist and positive allosteric modulators. Since the end of the 20th century, the impressive progress with structural biology of GPCR has been coming to fruition. The combination of fragment-based drug discovery and structure-based drug design applied to GPCRs has shown the ability to establish FDA approved drugs with improved physiochemical properties and selectivity profiles. The power of FBDD and SBDD may deliver important breakthroughs that likely contribute to discover more efficacious and safer medicines over the coming years.

Reference

1. Zheng, C., Han, L., Yap, C. W., Xie, B. & Chen, Y. Progress and problems in the exploration of therapeutic targets. Drug Discov. Today 11, 412–420 (2006).
2. Yang, H. et al. RG7204 (PLX4032), a selective BRAFV600E inhibitor, displays potent antitumor activity in preclinical melanoma models. Cancer Res. 70, 5518–5527 (2010).
3. Wada, C. K. The evolution of the matrix metalloproteinase inhibitor drug discovery program at abbott laboratories. Curr. Top. Med. Chem. 4, 1255–1267 (2004).
4. Park, C.-M. et al. Discovery of an orally bioavailable small molecule inhibitor of prosurvival B-cell lymphoma 2 proteins. J. Med. Chem. 51, 6902–6915 (2008).
5. Congreve, M., Dias, J. M. & Marshall, F. H. Structure-based drug design for G protein-coupled receptors. Prog. Med. Chem. 53, 1–63 (2014).
6. Congreve, M. et al. Fragment Screening of Stabilized G-Protein-Coupled Receptors Using Biophysical Methods. Method. Enzym. 493, 115–136 (2011).
7. Christopher, J. A et al. Biophysical Fragment Screening of the β1-Adrenergic Receptor: Identification of High Affinity Arylpiperazine Leads Using Structure-Based Drug Design. J. Med. Chem. 56, 3446–3455 (2013).
8. Iva Navratilova, Jeremy Besnard, A. L. H. Screening for GPCR Ligands Using Surface Plasmon Resonance. ACS Med. Chem. Lett. 2, 549–554 (2011).
9. Aristotelous, T. et al. Discovery of β 2 Adrenergic Receptor Ligands Using Biosensor Fragment Screening of Tagged Wild-Type Receptor. ACS Med. Chem. Lett. 4, 1005–1010 (2013).
10. Navratilova, I. & Hopkins, A. L. Fragment Screening by Surface Plasmon Resonance. ACS Med. Chem. Lett. 1, 44–48 (2010).
11. May, P. C., Dean, R. a, Lowe, S. L., Martenyi, F., Sheehan, S. M., Boggs, L. N., Monk, S. a, Mathes, B. M., Mergott, D. J., Watson, B. M., Stout, S. L., Timm, D. E., Smith Labell, E., Gonzales, C. R., Nakano, M., Jhee, S. S., Yen, M., Ereshefsky, L., Lindstrom, T. D., Calligaro, D. O., Cocke, P. J., Greg Hall, D., Friedrich, S., Citron, M., and Audia, J. E. (2011) Robust central reduction of amyloid-β in humans with an orally available, non-peptidic β-secretase inhibitor. J. Neurosci. 31, 16507–16516.
12. Bayburt, T. H. & Sligar, S. G. Membrane protein assembly into Nanodiscs. FEBS lett. 584, 1721–1727 (2010).
13. Lyukmanova, E. N. et al. Lipid-protein nanodiscs for cell-free production of integral membrane proteins in a soluble and folded state: comparison with detergent micelles, bicelles and liposomes. Biochim. Biophys. Acta 1818, 349–358 (2012).
14. Etzkorn, M. et al. Cell-free expressed bacteriorhodopsin in different soluble membrane mimetics: biophysical properties and NMR accessibility. Structure 21, 394–401 (2013).
15. Raschle, T., Hiller, S., Etzkorn, M. & Wagner, G. Nonmicellar systems for solution NMR spectroscopy of membrane proteins. Curr. Opin. Struct. Biol. 20, 471–479 (2010).
16. Serebryany, E., Zhu, G. A. & Yan, E. C. Y. Artificial membrane-like environments for in vitro studies of purified G-protein coupled receptors. Biochim. Biophys. Acta 1818, 225–233 (2012).
17. Park, S. H. et al. Local and Global Dynamics of the G Protein-Coupled Receptor CXCR1. Biochemistry 50, 2371–2380 (2011).
18. Park, S. H. et al. Structure of the chemokine receptor CXCR1 in phospholipid bilayers. Nature 491, 779–783 (2012).
19. Gautier, A., Mott, H. R., Bostock, M. J., Kirkpatrick, J. P. & Nietlispach, D. Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. Nat. Struct. Mol. Biol. 17, 768–774 (2010).
20. Hajduk, P. J. & Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. Nat. Rev. Drug Discov. 6, 211–219 (2007).
21. Congreve, M., Langmead, C. J., Mason, J. S. & Marshall, F. H. Progress in structure based drug design for G protein-coupled receptors. J. Med. Chem. 54, 4283–4311 (2011).
22. Congreve, M. et al. Discovery of 1,2,4-Triazine Derivatives as Adenosine A2A Antagonists using Structure Based Drug Design. J. Med. Chem. 55, 1898–1903 (2012).
23. Chilingaryan, Z., Yin, Z. & Oakley, A. J. Fragment-based screening by protein crystallography: successes and pitfalls. Int. J. Mol. Sci. 13, 12857–12879 (2012).
24. Zhukov, A. et al. Biophysical mapping of the adenosine A2A receptor. J. Med. Chem. 54, 4312–4323 (2011).
25. Andrews, S. P. & Tehan, B. Stabilised G protein-coupled receptors in structure-based drug design: a case study with adenosine A2A receptor. Med. Chem. Comm 4, 52–67 (2013).
26. Park, S. H. et al. Optimization of purification and refolding of the human chemokine receptor CXCR1 improves the stability of proteoliposomes for structure determination. Biochim. Biophys. Acta 1818, 584–591 (2012).
27. Catoire, L. J. et al. Solution NMR mapping of water-accessible residues in the transmembrane beta-barrel of OmpX. Eur. Biophys. J. 39, 623–630 (2010).
28. Langelaan, D. N. et al. Structural features of the apelin receptor N-terminal tail and first transmembrane segment implicated in ligand binding and receptor trafficking. Biochim. Biophys. Acta 1828, 1471–1483 (2013).
29. Reckel, S. et al. Solution NMR structure of proteorhodopsin. Angew. Chemie 50, 11942–11946 (2011).
30. Wang, B., Westerhoff, L. M. & Merz Jr, K. M. A Critical Assessment of the Performance of Protein-ligand Scoring Functions Based on NMR Chemical Shift Perturbations. J. Med. Chem. 50, 5128–5134 (2008).
31. Pellecchia, M. et al. Perspectives on NMR in drug discovery: a technique comes of age. Nat. Rev. Drug Discov. 7, 738–745 (2008).
32. Guan, J. et al. Small-Molecule Binding Sites on Proteins Established by Paramagnetic NMR Spectroscopy. J. Am. Chem. Soc. 135, 5859–5868 (2013).
33. Grisshammer, R., Duckworth, R. & Hendersont, R. Expression of a rat neurotensin receptor in Escherichia coli. Biochem. J. 576, 571–576 (1993).

34. Weiss, H. M. & Grisshammer, R. Purification and characterization of the human adenosine A2a receptor functionally expressed in Escherichia coli. Eur. J. Biochem. 269, 82–92 (2002).
35. Dahmane, T., Damian, M., Mary, S., Popot, J.-L. & Banères, J.-L. Amphipol-assisted in vitro folding of G protein-coupled receptors. Biochemistry 48, 6516–6521 (2009).
36. Schieborr, U. et al. How much NMR data is required to determine a protein-ligand complex structure? Chem. Bio. Chem 6, 1891–1898 (2005).
37. Egloff, P. et al. Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *Escherichia coli*. PNAS 111, E655-E662 (2014).

Summary

The present thesis starts with an overview of G protein-coupled receptors (GPCRs) (Chapter 1), which represent the largest protein family in the human proteome. Currently, more than 30% of all drugs on the market target this class of proteins, highlighting their pharmaceutical importance. Traditional GPCR drug discovery efforts rely on cell-based assays, combined with high-throughput screening. However, not all GPCRs are amenable to traditional screening and pharmacological evaluation methods. The chapter introduces the notion of fragment-based drug discovery (FBDD) and biophysical methods that are widely used for fragment discovery, such as nuclear magnetic resonance (NMR), surface plasmon resonance (SPR) and X-ray crystallography. Until recently, research in GPCR drug discovery was very challenging, mainly because of the instability of these proteins when extracted from their native membrane environment. The chapter also describes recent new advances in receptor stabilization and solubilization to overcome these limitations for biophysical studies.

The work described in Chapter 2 proves that target immobilized NMR screening (TINS) can be applied to identify small molecule hits on membrane proteins. The bacterial membrane enzyme disulphide bond forming protein B (DsbB) was chosen as a target and was immobilized on a sepharose resin and screened against the bacterial reference protein outer membrane protein A (OmpA), both solubilized in dodecylphosphocholine micelles. The reference protein is expected to account for non-specific binding of fragments to the solubilization media. Chapter 2 also describes an alternative solubilization method for TINS in aqueous buffer. The proteins were encapsulated into a lipid bilayer surrounded by two amphipathic scaffold proteins in a complex named the nanodisc (ND). Compared to detergent solubilized DsbB, the ND embedded DsbB showed enhanced activity in enzymatic assays and higher immobilization efficiency. NDs offer the potential to enable a more generic method to handle

membrane proteins. Furthermore, the use of NDs alleviated issues of non-specific partitioning of fragments into detergent micelles and improved the quality of NMR spectra of the fragments.

The work presented in Chapter 3 demonstrates the feasibility of combining TINS and stabilized GPCRs (StaRs®) to screen fragment libraries for target specific binding. StaRs® contain a small number of point mutations for both enhanced thermostability and conformational homogeneity. A moderate sized fragment library was screened for binding to an inverse agonist stabilized adenosine A_{2A} receptor (A_{2A}R), with OmpA as reference protein. Multiple TINS hits were validated by pharmacological assays with wild-type receptor. Among these hits, ligands with many different biological activities were discovered such as orthosteric ligands and positive and negative allosteric modulators. This study is one of the first reports of allosteric modulators for the adenosine A_{2A}R. Several of these ligands could be a good start for further lead exploration. It is likely that these allosteric modulators would be missed by other screening methods, due to their apparent low potency. TINS, as a direct binding assay, shows its great potential to detect multiple fragments that bind to different sites on the receptor, thus to discover ligands with novel biological activities, despite of their low potency.

In recent years, the impressive growth in GPCR structure information leads to broad use of structure-based methods for hit discovery and optimization. The work reported in Chapter 4 shows the potential to complement NMR-based fragment screening with molecular docking screens in FBDD against the A_{2A}R. The *in silico* screening approach can be extended to include 328,000 commercially available fragments and has better coverage of chemical space, which provides the opportunity to discover ligands with different chemotypes. This study showed that the molecular docking based screen was able to identify fragment hits that do not overlap with those discovered by NMR-based screening. The combination of biophysical and *in silico* screening methods with

structure determination for GPCRs will provide considerable aid to the drug elaboration process.

The dynamic nature of GPCRs and issues associated with low level expression and instability during purification made biophysical and structural characterization of GPCRs particularly difficult. In Chapter 5, a new methodology to overexpress GPCR genes in *Escherichia coli* yielding the protein in inclusion bodies is investigated. If the receptor can be successfully refolded in its native state using amphipols, it could enable the production of a sufficient amount of isotope labeled GPCR for NMR structural studies. In this case, the ligand binding sites can be mapped at low resolution using chemical shift perturbation techniques.

Chapter 6 presents the general discussion of the research described in this thesis and further perspectives in field of GPCR drug discovery. In recent years, a number of FBDD techniques have been validated for application with GPCRs, including some of the most challenging targets. The impressive progress with structural biology of GPCRs has been coming to fruition. This new era of FBDD and SBDD for GPCRs has the great potential to yield safe and effective drugs in the coming decade.

Samenvatting

In Hoofdstuk 1 van dit proefschrift wordt de grootste familie van eiwitten in het menselijk genoom, de G-eiwit gekoppelde receptoren (GPCRs), beschreven. GPCRs zijn belangrijke targets voor medicijnen en meer dan 30% van alle momenteel verkrijgbare geneesmiddelen grijpt dan ook aan op deze klasse van eiwitten. In het onderzoek naar nieuwe geneesmiddelen voor GPCRs worden doorgaans testsystemen met levende cellen gebruikt. Deze systemen zijn echter niet toepasbaar op alle GPCRs. In dit hoofdstuk introduceer ik daarom het gebruik van biofysische methodes, zoals *nuclear magnetic resonance* (NMR), *surface plasmon resonance* (SPR) en Röntgen-kristallografie, voor fragment-gebaseerde medicijn ontwikkeling (FBDD). Omdat GPCRs moeilijk te isoleren zijn uit het celmembraan is geneesmiddelenontwikkeling met gebruik van deze biofysische methoden voor deze klasse van eiwitten zeer uitdagend. In dit hoofdstuk worden de laatste ontwikkelingen op het gebied van receptorstabilisatie beschreven.

In Hoofdstuk 2 toon ik aan dat *target immobilized NMR screening* (TINS) kan worden gebruikt om kleine moleculen te vinden die affiniteit voor membraaneiwitten hebben. Het enzym *disulphide bond forming protein B* (DsbB) werd hiervoor geïmmobiliseerd op een dragermateriaal en stabiel in oplossing gehouden door middel van micellen. Een TINS screen met dit geïmmobiliseerde eiwit werd vervolgens uitgevoerd, waarbij *outer membrane protein A* (OmpA) als referentie-eiwit diende om te corrigeren voor eventuele aspecifieke binding van de kleine moleculen aan de matrix. Een alternatieve methode om eiwitten stabiel in oplossing te houden voor TINS, wordt tevens beschreven in Hoofdstuk 2. Hierbij werd gebruik gemaakt van synthetische membraanmodellen, zogenaamde *nanodiscs* (NDs). In het geval van DsbB resulteerde het gebruik van NDs in een hogere enzymatische activiteit dan van het door middel van micellen gestabiliseerde DsbB. Tevens kon er een hogere immobilisatiegraad bereikt worden door gebruik te maken van NDs. Door NDs te gebruiken treedt er

bovendien minder aspecifieke binding op van de kleine moleculen en wordt de kwaliteit van de NMR-spectra van de moleculen sterk verbeterd.

In Hoofdstuk 3 wordt beschreven hoe gestabiliseerde GPCRs (StaRs®) gebruikt kunnen worden in TINS. StaRs® zijn GPCRs waarin een klein aantal punt mutaties zijn aangebracht om de thermostabiliteit van deze eiwitten te verbeteren en het aantal conformaties van de GPCRs te beperken. Een TINS-screen tegen een inverse agonist gestabiliseerde STaR® van de adenosine A_{2A} receptor ($A_{2A}R$) werd uitgevoerd met een kleine hoeveelheid fragmenten, waarbij OmpA als referentie-eiwit diende. Meerdere hits uit deze TINS screen werden gevalideerd in farmacologische testsystemen, waarbij gebruik werd gemaakt van de niet-gemuteerde vorm van de A_{2A}-receptor. Dit resulteerde in de identificatie van liganden met uiteenlopende biologische activiteiten, waaronder allostere modulatoren van de A_{2A}-receptor. Verschillende van deze liganden zijn goede startpunten voor verdere optimalisatie naar grotere en meer potente verbindingen. Deze studie beschrijft als een van de eerste de identificatie van allostere modulatoren van de A_{2A}-receptor. Met behulp van TINS kunnen moleculen worden geïdentificeerd die met lage affiniteit, op iedere willekeurige locatie met een eiwit binden. TINS is dan ook zeer geschikt om binding van meerdere fragmenten tegelijkertijd met een receptor te identificeren, ondanks de lage bindingsaffiniteit van deze liganden. Waarschijnlijk zouden eerder genoemde allostere modulatoren van de A_{2A}-receptor niet gevonden zijn wanneer een andere, minder gevoelige, detectiemethode was gebruikt.

Hit identificatie- en validatie technieken, waarbij structurele informatie van het target nodig is (SBDD), worden steeds vaker toegepast op GPCRs. In Hoofdstuk 4 wordt beschreven hoe fragment-screening door middel van NMR, in combinatie met computer screening kan worden toegepast ter identificatie van liganden voor de A_{2A}-receptor. Door gebruik te maken van *in silico* technieken kunnen liganden worden gevonden in een collectie van 328.000 commercieel verkrijgbare fragmenten. Deze vertegenwoordigen een grote chemische diversiteit, wat het mogelijk maakt om liganden met nieuwe chemotypen te

identificeren. Deze studie toont aan dat met behulp van de *in silico* screen liganden konden worden geïdentificeerd die geen structurele gelijkenis hadden met fragmenten die door middel van NMR screening waren gevonden. Het combineren van biofysische- en *in silico* screening technieken met structuuropheldering zal in de toekomst aanzienlijk bijdragen aan de ontwikkeling van nieuwe GPCR geneesmiddelen.

Naast het flexibele karakter van GPCRs zijn lage cellulaire expressie en instabiliteit tijdens opzuiveren de grootste obstakels voor biofysische en structurele karakterisering van GPCRs. In Hoofdstuk 5 wordt een nieuwe methode beschreven waarin GPCRs tot over-expressie worden gebracht in *Escherichia coli* en waarbij de receptoren in onoplosbare vorm geproduceerd worden. Wanneer de receptoren uit deze zogenaamde *inclusion bodies* opnieuw gevouwen zouden kunnen worden tot een oplosbare vorm, biedt dit een mogelijkheid om voldoende isotoopgelabeld eiwit voor NMR-studies te produceren. De bindingsplaats van liganden aan de receptor kan vervolgens redelijk accuraat bepaald worden, door middel van het in kaart brengen van de “chemische verschuiving” van de pieken in het NMR spectrum.

In Hoofdstuk 6 wordt het onderzoek dat is beschreven in dit proefschrift bediscussieerd en wordt vooruitgekeken naar de geneesmiddelenontwikkeling van de toekomst in het veld van GPCRs. In de recente jaren zijn meerdere FBDD technieken ontwikkeld en succesvol toegepast op GPCRs, waaronder een aantal zeer uitdagende targets. De indrukwekkende vooruitgang in het veld van structurele biologie met GPCRs wordt dan ook duidelijk zichtbaar. In de toekomst ligt een grote rol weggelegd voor FBDD en SBDD technieken bij de ontwikkeling van veilige en effectieve geneesmiddelen die op GPCRs aangrijpen.

List of Publications

D. Chen, J. C. Errey, L. H. Heitman, F. H. Marshall, A. P. IJzerman and G. Siegal. Fragment screening of GPCRs using biophysical methods: identification of ligands of the adenosine A_{2A} receptor with novel biological activity. *ACS Chem. Biol.*, 2012, 7, 2064-2073.

D. Chen, A. Ranganathan, A. P. IJzerman, G. Siegal and J. Carlsson. Complementarity between *in silico* and biophysical screening approaches in fragment-based lead discovery against the A_{2A} adenosine receptor. *J. Chem. Inf. Model.* 2013, 53, 2701-2714.

V. Früh, Y. Zhou, **D. Chen**, C. Loch, E. Ab, Y. N. Grinkova, H. Verheij, S. G. Sligar, J. H. Bushweller and G. Siegal. Application of fragment-based drug discovery to membrane proteins: identification of ligands of the integral membrane enzyme DsbB. *Chem. Biol.* 2010, 17, 881−891.

Curriculum Vitae

Dan Chen was born on the 22nd April, 1980 in Dalian, China. She obtained her early education in the beautiful harbor city of Dalian.

After a few years of secondary study in Dalian University, she decided to take an adventure in the other side of the world, to embrace the romance of France and employ the language of Molière. First, she graduated in International Economics, the study of which started in China, before devoting herself to the world of biochemistry. In 2008, she completed her master degree in Structural Biochemistry from Paul Sabatier University in Toulouse. Her graduation internship was performed at Sanofi R&D Center in Paris, focusing on crystallization of medically relevant human kinases alone and in complex with their specific inhibitors.

In 2009, she started her PhD research in the Protein Chemistry group in Leiden University under the supervision of Dr. Gregg Siegal. The research focused on the biophysical characterization of membrane protein-small molecule interactions. Since January 2013, she is employed as Scientist at ZoBio B.V., a company specialized in fragment drug discovery.