Introducing weak affinity chromatography to drug discovery with focus on fragment screening

Linnaeus University Dissertations
No 124/2013

INTRODUCING WEAK AFFINITY CHROMATOGRAPHY TO DRUG DISCOVERY WITH FOCUS ON FRAGMENT SCREENING

MINH-DAO DUONG-THI

LINNAEUS UNIVERSITY PRESS

Introducing weak affinity chromatography to drug discovery with focus on fragment screening

Doctoral dissertation, Department of Chemistry and Biomedical Sciences, Linnaeus University, Kalmar, Sweden, 2013

Cover image: Structure of a human α-thrombin with a compound docked into the active site (upper). The image is generated by Schrödinger Suite 2010, where red and blue areas are negatively and positively charged, respectively. The X-ray structure of the thrombin is downloaded from Protein Data Bank (PDB ID: 3DA9). Bottom: Chromatograms of three fragments at different retentions on a thrombin affinity column. The compounds in the background are fragments from the TimTec library.

ISBN: 978-91-87427-14-5
Published by: Linnaeus University Press, S-351 95 Växjö
Printed by: Elanders Sverige AB, 2013

Abstract

Duong-Thi, Minh-Dao (2013). *Introducing Weak Affinity Chromatography to Drug Discovery with Focus on Fragment Screening.* Linnaeus University Dissertations No 124/2013. ISBN: 978-91-87427-14-5. Written in English.

Fragment-based drug discovery is an emerging process that has gained popularity in recent years. The process starts from small molecules called fragments. One major step in fragment-based drug discovery is fragment screening, which is a strategy to screen libraries of small molecules to find hits. The strategy in theory is more efficient than traditional high-throughput screening that works with larger molecules. As fragments intrinsically possess weak affinity to a target, detection techniques of high sensitivity to affinity are required for fragment screening. Furthermore, the use of different screening methods is necessary to improve the likelihood of success in finding suitable fragments. Since no single method can work for all types of screening, there is a demand for new techniques. The aim of this thesis is to introduce weak affinity chromatography (WAC) as a novel technique for fragment screening.

WAC is, as the name suggests, an affinity-based liquid chromatographic technique that separates compounds based on their different weak affinities to an immobilized target. The higher affinity a compound has towards the target, the longer it remains in the separation unit, and this will be expressed as a longer retention time. The affinity measure and ranking of affinity can be achieved by processing the obtained retention times of analyzed compounds.

In this thesis, WAC is studied for fragment screening on two platforms. The first system comprised a 24-channel affinity cartridge that works in cooperation with an eight-needle autosampler and 24 parallel UV detector units. The second system was a standard analytical LC-MS platform that is connected to an affinity column, generally called WAC-MS or affinity LC-MS. The evaluation criteria in studying WAC for fragment screening using these platforms were throughput, affinity determination and ranking, specificity, operational platform characteristics and consumption of target protein and sample. The model target proteins were bovine serum albumin for the first platform, thrombin and trypsin for the latter. Screened fragments were either small molecule drugs, a thrombin-directed collection of compounds, or a general-purpose fragment library. To evaluate WAC for early stages of fragment elaboration, diastereomeric mixtures from a thrombin-directed synthesis project were screened.

Although both analytical platforms can be used for fragment screening, WAC-MS shows more useful features due to easy access to the screening platform, higher throughput and ability to analyze mixtures. Affinity data from WAC are in good correlation with IC_{50} values from enzyme assay experiments. The possibility to distinguish specific from non-specific interactions plays an important role in the interpretation of WAC results. In this thesis, this was achieved by inhibiting the active site of the target protein to measure off-site interactions. WAC proves to be a sensitive, robust, moderate in cost and easy to access technique for fragment screening, and can also be useful in the early stages of fragment evolution.

In conclusion, this thesis has demonstrated the proof of principle of using WAC as a new tool to monitor affinity and to select hits in fragment-based drug discovery. This thesis has indicated the primary possibilities, advantages as well as the limitations of WAC in fragment screening procedures. In the future, WAC should be evaluated on other targets and fragment libraries in order to realize more fully the potential of the technology.

Keywords: *affinity LC-MS, fragment-based drug discovery, fragment screening, high throughput, mass spectrometry, stereoisomer, enantiomer, thrombin, weak affinity chromatography, WAC, WAC-MS.*

POPULÄRVETENSKAPLIG SAMMANFATTNING

Utvecklingen av ett läkemedel bygger i regel på att man har kännedom om vilka proteiner/enzymer som deltar i de sjukdomsframkallande processerna. Med denna kunskap försöker läkemedelskemisten att designa en molekyl som kan reglera aktiviteten på de sjukdomsrelaterade proteinerna (målproteiner), så att sjukdomens symptom försvinner eller minskar. Ett relativt nytt sätt att utveckla ett läkemedel är att använda sig av så kallad fragmentbaserad screening. Denna teknik bygger på att man har ett bibliotek med många små molekyler (fragment) som representerar ett brett urval av kemiska strukturer. Genom att söka igenom (screena) fragmentbiblioteket efter molekyler som binder till olika målproteiner så kan man hitta strukturer som sedan kombineras eller utvecklas till läkemedelskandidater.

På grund av sin ringa storlek så binder fragment alltid med låg affinitet. Det är därför nödvändigt att använda metoder som kan detektera dessa svaga bindningar mellan fragmentmolekylen och målproteinerna, när fragmentscreeningen utförs. Oftast används en kombination av olika tekniker för att hitta användbara fragment. Nya, kompletterande metoder som passar för fragmentscreening är av stort intresse. Målsättningen med denna avhandling var att introducera tekniken ”Weak Affinity Chromatography” (WAC) för fragmentscreening och utvärdering av vidareutvecklade fragment. Detta är en vätskekromatografisk metod som bygger på att målproteinet binds till bärarmaterialet i en kolonn. Därefter injiceras en liten volym med fragment eller andra små molekyler. De molekyler som inte binder till målproteinet i kolonnen passerar snabbt igenom kolonnen medan de molekyler som kan binda kommer att fördröjas.

Två olika analytiska plattformar för WAC har utvärderats. Det första systemet bestod av 24 parallella kolonner kombinerat med 24 parallella UV detektorer. Detta system kan analysera 8 prover samtidigt (som triplikat). Det andra systemet använde endast en kolonn tillsammans med ett vanligt analytiskt vätskekromatografiskt system med en masspektrometrisk detektor (WAC-MS). Modellproteiner i studierna var bovint serumalbumin för den första plattformen och trombin/trypsin för den andra. Både läkemedel, fragment och en utvald samling molekyler med känd affinitet användes för att utvärdera de två plattformarna i termer av reproducerbarhet, konsumtion av protein och fragment samt screeningshastighet.

Även om båda plattformarna uppfyllde de nödvändiga kriterierna har WAC-MS en rad fördelar. Den största fördelen är att den tekniska plattformen är allmänt tillgänglig och metoden har också en högre kapacitet och förmåga att identifiera molekyler i en blandning. Detta gör att orena prover, blandningar av stereoisomerer och blandningar av fragment kan analyseras. Möjligheten att skilja specifika och icke-specifika bindningar till målproteinet spelar en viktig roll i utvärderingen av resultatet. Detta utfördes genom att hämma det aktiva bindningsstället på målproteinet genom permanenta eller reversibla bindare. WAC-MS har visat sig ha god korrelation med andra metoder och den största begränsningen i metoden är molekyler med högre affinitet såsom vidareutvecklade fragment (läkemedelskandidater) inte kan analyseras eftersom de tar extremt lång tid att eluera.

Denna avhandling visar att WAC-MS är en robust teknik för fragmentscreening och utvärdering av vidareutvecklade fragment. Dess fulla potential kan dock utvärderas först när fler målproteiner och fragmentbibliotek har testats.

LIST OF PUBLICATIONS

This thesis is based on the following articles, which are referred to in the text by Roman numerals. All published papers are reproduced with permission from the respective publishers.

I. Ohlson S, **Duong-Thi MD**, Bergström M, Fex T, Hansson L, Pedersen L, Guazotti S, Isaksson R (2010) Toward high-throughput drug screening on a chip-based parallel affinity separation platform. J. Sep. Sci., 33, 2575-81.

II. Duong-Thi MD*, Meiby E*, Bergström M, Fex T, Isaksson R, Ohlson S (2011) Weak affinity chromatography as a new approach for fragment screening in drug discovery. Anal. Biochem., 414, 138-46.
**These authors contributed equally*

III. Duong-Thi MD, Bergström M, Fex T, Isaksson R, Ohlson S (2013) High-throughput fragment screening by affinity LC-MS. J. Biomol. Screen., 18(2), 160-171.

IV. **Duong-Thi MD**, Bergström M, Fex T, Svensson S, Ohlson S, Isaksson R (2013) Weak affinity chromatography for evaluation of stereoisomers in early drug discovery. Accepted for publication in J. Biomol. Screen.
DOI: 10.1177/1087057113480391.

ABBREVIATIONS

[L]	ligand (analyte) molar concentration
[R]	receptor (target) molar concentration
[RL]	receptor-ligand complex molar concentration
3-ABA	3-amino benzamidine
4-ABA	4-amino benzamidine
ac	alternating current
ADME(T)	absorption, distribution, metabolism, excretion (and toxicity)
amu	atom mass unit
APCI	atmospheric pressure chemical ionization
API	atmospheric pressure ionization
APPI	atmospheric pressure photo ionization
Asp	aspartate
B_{max}	total number of moles of target in a column (=B_{tot})
BSA	bovine serum albumin
B_{tot}	total number of moles of target in a column
BZA	benzamidine
CEfrag	fragment screening by capillary electrophoresis
CV	coefficient of variation
dc	direct current
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
ESI	electrospray ionization
F	mobile phase flow rate
f_b	fraction of bound receptor; receptor occupancy
FBDD	fragment-based drug discovery
FDA	Food and Drug Administration of the United States
Gly	glycine
H	plate height
His	histidine
HP(L)AC	high-performance (liquid) affinity chromatography
HPLC	high-performance liquid chromatography

HSA	human serum albumin
HTS	high-throughput screening
IC_{50}	half maximal inhibitory concentration
Id	internal diameter
ITC	isothermal titration calorimetry
K	Kelvin
K_A	association constant
K_D	dissociation constant
k_{off}	off-rate
k_{on}	on-rate
LC	liquid chromatography
LC-MS	liquid chromatography coupled with mass spectrometer
LE	ligand efficiency
logP	logarithm of water-n-octanol partition
m/z	mass to charge ratio
MS	mass spectrometry/mass spectrometer
MW	molecular weight
NHA	number of heavy atoms
NME	new molecular entities
NMR	nuclear magnetic resonance
PBS	phosphate buffer saline, pH=7.4
PMSF	phenylmethylsulfonyl fluoride
PPACK	D-phenylalanyl-L-prolyl-L-arginine chloromethyl ketone
Pro	proline
Q_{max}	total number of moles of target in a column (=B_{tot})
R	universal gas constant
RF	radio frequency
RNA	ribonucleic acid
Ro3	rule of three
Ro5	rule of five
R_{tot}	total number of moles of target in a column (=B_{tot})
RU	response/resonance unit
SAR	structure-activity relationship
SIM	selected ion monitoring

SPR	surface plasmon resonance
T	absolute temperature
t'_R	adjusted retention time
TINS	target immobilized NMR screening
t_M	void time
t_R	apparent retention time
Trp	tryptophan
t_{spec}	specific retention time
Tyr	tyrosine
u	linear flow velocity of mobile phase
US	United States
USD	US dollar
UV	ultraviolet
V_M	void volume
V_R	retention volume
WAC	weak affinity chromatography
WAC-MS	weak affinity chromatography coupled with mass spectrometry; affinity LC-MS
ΔG^o	change in standard Gibbs free energy
ΔH^o	change in standard enthalpy
ΔS^o	change in standard entropy
η	critical ratio
σ_M^2	variance of non-retained peak
σ_R^2	variance of retained peak

TABLE OF CONTENTS

1. INTRODUCTION

The pharmaceutical industry has been receiving criticism for low throughput of new drugs despite heavy investment (Paul *et al.*, 2010; Mullard, 2011). Most of the time, selected molecules do not exert expected properties when they enter the human body. Therefore, reducing the cost, shortening time and increasing output for drug discovery are challenging. Although most potential drugs have failed at the clinical stage, improvements of the hit and lead identification steps would greatly contribute to the success of future drugs.

The fragment-based drug discovery (FBDD) process is a hit identification strategy that screens small molecules (molecular weight (MW) ≤ 300) for affinity to target, such as an enzyme (Scott *et al.*, 2012; Baker, 2013). Hits obtained are subsequently elaborated towards drug candidates that have higher potency, selectivity and appropriate physico-chemical properties. The technique is believed to have more advantages than conventional screening of drug-size molecules. As fragments are combinations of fewer atoms compared to larger molecules, the possible number of fragments is exponentially fewer than the number of available larger substances. Consequently, fragment screening covers the chemical space more efficiently. In addition, hits from fragment screening are small in size, which provides medicinal chemists with more room to develop them into drugs. Hit rates from fragment screening are often high (typically 0.5–10 %, but the rate also depends on the target and screening library) (Chen & Shoichet, 2009; Winter *et al.*, 2012) because fragments can reach and bind to an active site of the target more easily due to reduced steric hindrance (Hann *et al.*, 2001). Fragment hits, however, are intrinsically weak in affinity (dissociation constant, K_D, is higher than μM) and therefore, it is a challenge for screening techniques to see them.

The weak affinity range of hits often gives false results in fragment screening. At low affinity, specific interactions are difficult to be distinguished from promiscuous binding which causes false positives. In other situations, false positives may be caused by aggregation, reactive chemicals or interference to

detection signals. On the other hand, fragments with weak affinity may be missed from detection and therefore create false negatives. In general, each screening method will have its own pattern of artifacts, and no single technique can be used for all kinds of target proteins and fragments. To handle these above mentioned problems, fragment screening operations often employ a variety of approaches which complement each other. Consequently, there is an interest for new technologies to be introduced for fragment screening.

Weak affinity chromatography (WAC) is a high-performance liquid chromatography (HPLC)-based technique, first developed in the 1980s, to separate carbohydrate antigens by monoclonal antibodies (Ohlson *et al.*, 1988; Bergström *et al.*, 1998). In an attempt to introduce WAC to fragment screening and possibly to the early hit development stage in FBDD, this thesis characterizes the technique as a sensitive, robust and easy-to-access affinity determination tool. The evaluation criteria focus on performance, material consumption and throughput of the technique.

2. THE DRUG DISCOVERY PROCESS

2.1. History

The drug discovery process has evolved through many steps marked by important events. In ancient times, people used herbs, minerals or animal parts as drugs to cure diseases. Then, the drug components were mixtures of undefined substances. With the isolation of morphine at the beginning of the 19th century by Sertürner (Sertürner, 1805, 1817), the science of pharmacology began to take shape with the focus on monitoring the effects of single substances on the body (Michne, 2010). In 1828, the success with urea synthesis by Wöhler established the science of synthetic organic chemistry (Wöhler, 1828). The drug discovery mainstream then turned to the synthesis of analogs of naturally originated actives. In the early 20th century, Paul Ehrlich proposed the concept of the "magic bullet" which hypothesizes that drug effects result from specific interactions with targets (Michne, 2010). Specificity enables a drug selectively kill parasites, microorganisms or cancer cells, but keeps the healthy host cells undamaged (Drews, 2000). Many years later in the 1980s, achievements in genomics have offered great opportunities for drug discovery and development (Rankovic *et al.*, 2010). From the work on structural molecular biology (X-ray crystallography) and genetic engineering, many therapeutic targets has been identified and produced at a fast pace. At the same time, parallel synthesis and combinatorial chemistry that can generate high-throughput syntheses flourished (Michne, 2010). Following the trend, high-throughput screening in drug discovery emerged and developed. However, the resulting outputs from pharmaceutical industries still fall short of the high expectations (Betz, 2005; Macarron, 2006; Paul *et al.*, 2010).

2.2. Current status

For decades, the output of new molecular entities (NME) approved by the United States Food and Drug Administration (FDA) has declined

considerably. The cost to develop a NME, on the contrary, has passed a billion USD in 2010 and continues to increase. As a rough estimate in the US alone, the annual cost for research and development in major pharmaceutical companies is more than 50 billion USD, but they can produce only about 20 new drugs per year (Paul *et al.*, 2010; Mullard, 2011). One of the explanations for this low output is probably the more stringent drug regulations from regulatory authorities and the increasing expectations on efficacy and safety from patients and the general public. However, the strategies to increase output also need improvement.

2.3. Stages and strategies in the drug discovery process

The modern drug discovery process includes four major steps with different characteristics and requirements: target identification and validation, hit and lead identification, lead optimization, and clinical trials (Michne, 2010; Hoelder *et al.*, 2012). The hit and lead identification can be divided into two steps, which are hit generation and hit optimization (figure 1). Any failures in the drug discovery become more costly as the drug candidates progress along the pipeline.

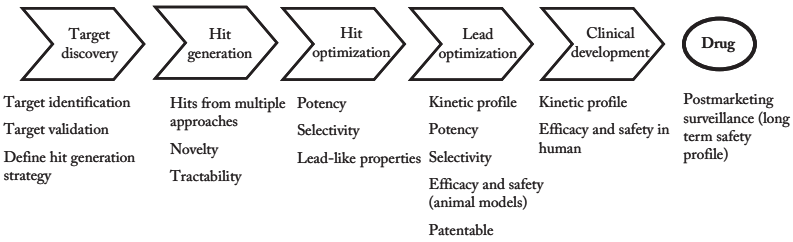

Figure 1. The target-based drug discovery pipeline. Adapted from Michne, 2010 with permission from John Wiley & Sons.

Drug discovery and development can either be based on effects of the drug candidate on the phenotypes, or focused on elucidation of the disease mechanism and drug target structure (genotypes). The mainstream of drug discovery today is the latter, also called target-based, and concentrates on finding the effect of chemicals on well-defined targets. To find a treatment for an illness, first the disease mechanisms have to be elucidated. Targets for the disease are identified and validated followed by hit identification. The next step is the hit-to-lead process and lead optimization. The last step includes clinical trials which are distributed into phases with different objectives. Among the different steps in the drug discovery process, target validation and clinical trials play especially important roles in the success of the whole process

(Michne, 2010; Hoelder *et al.*, 2012). Target validation is important because the quality of a target will affect all the steps in the drug discovery pipeline. Clinical trials are particularly important because they build up the safety profile and dosage of the drug for humans.

Target identification and validation

The identification and validation of the target greatly affect the efficiency of a drug discovery process. The quality of this very first step influences the possible attrition of a drug candidate which occurs later in the clinical stage (Hoelder *et al.*, 2012). In order to be useful, a target should be "druggable", meaning that it can bind to other molecules and that these interactions can modify its function in a therapeutically meaningful way (Imming *et al.*, 2006; Edfeldt *et al.*, 2011). For new targets, where no information about binding to other molecules is available, the target is often tested for "ligandability", meaning that it can bind to other molecules. A ligandable target may not be druggable, but a druggable target has to be ligandable first (Edfeldt *et al.*, 2011). Most of targets are proteins, but other structures such as nucleic acids (DNA, RNA, ribosomes) can also be drug targets (Ecker & Griffey, 1999; Imming *et al.*, 2006; Moumné *et al.*, 2012).

Each research unit has different strategies for carrying out drug discovery. Normally at the end of the target validation and identification process, a strategy for hit generation is defined. When the screening method is selected, suitable molecular libraries can be ready for use in the next step of the hit and lead generation.

Hit and lead identifications

The hit and lead identifications are to select and prioritize compounds to progress further. The process can be incorporated with a target validation procedure, where molecules are screened to detect binding to a target or causing an effect such as changing the functions of a cell. For a known druggable target, hits and leads can be generated either by the screening of many substances and/or by structure-based drug design. The latter approach needs more elaborated knowledge about the structure of the target and mode of interaction of the known leads.

There have been many approaches to screen for hits and leads (Rankovic *et al.*, 2010). When high-throughput screening (HTS) was introduced in the early 1990s, the number of tested substances and the speed of screening were prioritized (Macarron, 2006). However, it was early recognized that many of these resource-intensive screenings produced little output in terms of new drug candidates. In the late 1990s, screening was performed with more

consideration on the quality of the hits and leads (Macarron, 2006; Rankovic *et al.*, 2010). Although the efficiency of drug discovery has improved, the output from the whole process is still meager. An explanation for the low hit rate and hit quality in conventional HTS is that relatively large and complex molecules were selected to screen (Hann *et al.*, 2001). A possible solution to these problems is to reduce the size of the screened molecules to smaller fragments. Fragment screening since then has been introduced and is now a well-established screening strategy (Früh *et al.*, 2010; Hubbard & Murray, 2011; Sun *et al.*, 2011; Lee *et al.*, 2012; Murray *et al.*, 2012; Baker, 2013). This thesis focuses on the hit identification step by the use of fragment screening. A more detailed presentation of fragment screening will follow later in the text.

Selected hits should demonstrate reproducible binding activity in the assays, and they should be able to develop chemically (Michne, 2010). Preliminary ADME (absorption, distribution, metabolism, and excretion) profiling is also helpful when selecting which hits to develop. For patentable reasons, hits are only attractive for further development when they are novel. Hits that have been modified or elaborated towards more potency and selectivity will become leads (Michne, 2010). The process to identify suitable hits in a drug development project often needs one year and it may take another 1.5 years to develop the hit to a lead (Paul *et al.*, 2010).

Lead optimization

Once hits are identified, they are modified and developed towards more potency and selectivity with the goal to obtain leads. The resulting leads are further optimized towards increased potency, selectivity and favorable physico-chemical properties such as to facilitate absorption. The control of physico-chemical properties is important as higher affinity leads tend to be of larger size and higher lipophilicity (Teague *et al.*, 1999). The establishment of proper leads usually requires extensive chemical syntheses.

The optimized leads are then taken to pre-clinical studies to investigate details of their mode of action, including animal disease models, extensive ADME characterizations and preliminary safety profiles (Jorgensen, 2012). The results are drug candidates that proceed to clinical trials. A full lead optimization project including pre-clinical studies may take three years on average to complete (Paul *et al.*, 2010).

Clinical trials

Being at the last stage in the drug discovery and development pipeline, attrition in clinical trials is expensive not only because the trials involve costly tests on humans, but also because failure in this stage means that previous

efforts are fruitless. Unfortunately, most attrition of drug candidates occurs during clinical trials (Paul *et al.*, 2010).

Clinical trials are generally divided into successive phases with different study purposes. Phase I tests are conducted on a small number of healthy volunteers to probe primary safe dosages and to study ADMET (absorption, distribution, metabolism, excretion and toxicity) profiles. The purpose of Phase II is to evaluate more thoroughly drug safety issues and most importantly, the primary efficacy upon hundreds of patients that potentially need the tested drug. Phase III is a large-scale trial on thousands of subjects to establish efficacy and detect rare side-effects (DiMasi *et al.*, 2003). Drug candidates may obtain marketing approval after phase III if they show an adequate safety profile, efficacy and significant advantages compared to other available treatments. Post marketing surveillance, sometimes called the phase IV trial, is the subsequent step to monitoring long-term events. It is carried out by collecting voluntary reports on adverse reactions of the drug. Sometimes approved drugs have to be withdrawn from the market in phase IV due to major toxicity problems.

3. FRAGMENT-BASED DRUG DISCOVERY (FBDD)

FBDD originates from the concept that Jencks proposed in 1981, where he suggested that binding energy could be gained from connecting small weak binders into larger molecules (Jencks, 1981). These small molecules or fragments with simple functionalities do not possess drug-like characteristics, but can be sub-units of larger, drug-like molecules. The first attempt to support this theory was published in 1985 (Nakamura & Abeles, 1985), but the strategy only began to catch attention in 1996 when a successful "SAR (structure-activity relationship) by NMR" approach (Shuker *et al.*, 1996) was published. Since then, FBDD has developed into a well-established strategy in drug discovery (Erlanson *et al.*, 2004; Chessari & Woodhead, 2009; Baker, 2013). In 2011, the first drug obtained by FBDD was approved by the FDA for treatment of late-stage melanoma (vemurafenib from Plexxikon), which proved the rationality of the FBDD strategy (Tsai *et al.*, 2008; Bollag *et al.*, 2010).

Traditionally, the description of a good drug candidate has been based on the "rule of five" (Ro5), which was coined by Lipinski and colleagues (Lipinski *et al.*, 1997). It stated that a molecule with more than 5 hydrogen-bond donors, a MW higher than 500, a logP higher than 5, or more than 10 hydrogen bond acceptors, would have poor oral absorption and permeation in the body (i.e. low bioavailability) (Lipinski *et al.*, 1997). Later, Veber and colleagues added that a molecule with a number of rotatable bond not more than 10 and polar surface area not more than 140 $Å^2$ would have a higher probability of good absorption (Veber *et al.*, 2002). However, there are many exceptions, and the rules can only be considered as guidelines in the discovery and development of drugs (Lipinski *et al.*, 1997; Abad-Zapatero, 2007). Based on the Ro5, the criteria to select compounds to screen in conventional HTS and in FBDD were established. In traditional HTS, drug discovery starts with compounds having a MW of 450, logP from –3.5 to +4.5, no more than 4 rings, no more

than 5 hydrogen bond donors or 8 hydrogen bond acceptors (Oprea *et al.*, 2001). In FBDD, the "rule of three" coined by Congreve and colleagues from Astex (Congreve *et al.*, 2003) defines fragments as compounds with a MW less than 300, number of hydrogen bond donors ≤ 3, number of hydrogen bond acceptors ≤ 3 and logP ≤ 3. The difference between conventional HTS and fragment screening is illustrated in figure 2 (Hajduk *et al.*, 2011).

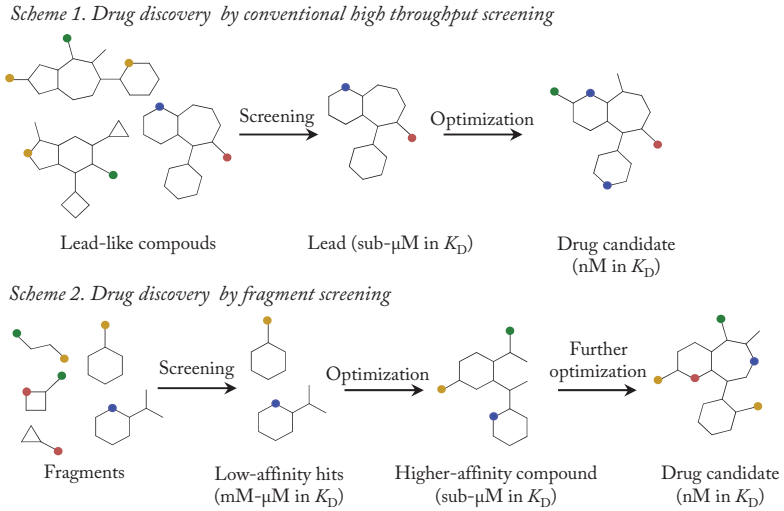

Figure 2. Conventional HTS and fragment screening in drug discovery. The figure is adapted from Hajduk et al., 2011 by permission from Macmillan Publishers.

One advantage with fragment libraries is that they cover the chemical space more efficiently compared to HTS libraries. A rough estimate of the number of combinations that is possible from 30 heavy atoms, which is the average for a MW = 400 compound, is about 10^{60} molecules (Bohacek *et al.*, 1996). The combination possibilities do, however, decrease exponentially with the number of heavy atoms in a molecule. Another calculation estimated the increase of chemical space to approximately eight times when adding one heavy atom (Fink & Reymond, 2007; Hubbard & Murray, 2011). Therefore, a fragment library with a few thousand compounds can cover the same chemical space as a billion-compound conventional HTS library (Hubbard & Murray, 2011). Another advantage with small size compounds is that they will suffer less from steric hindrance that, in theory, should make it easier to obtain a hit (Hann *et al.*, 2001).

Fragment screening will produce hits with dissociation constants (K_D-values) in the range of μM to mM (Boyd *et al.*, 2012), while the expected K_D-values of HTS hits are below 1 μM (Bohacek *et al.*, 1996). The fragment hits then

have to be elaborated by merging, linking or growing to achieve higher potency and selectivity. The two drug discovery approaches may, however, produce drugs with similar affinities and physico-chemical properties. The average affinity of marketed oral drugs is about 20 nM (Overington *et al.*, 2006), which can be a good reference for both approaches. On the other hand, with the speculation of transient affinity drugs (Ohlson, 2008), the drug affinity may not need to reach the nanomolar range but could be considerably weaker. Consequently, the optimization process in FBDD can then be shortened and the average size of the drug molecules can be smaller.

3.1. Ligand efficiency

As fragments are small in size, the affinities of fragment hits are generally weaker than those of large molecule hits. However, if the binding energy is normalized to the number of non-hydrogen atoms, it can be shown that fragments often bind more efficiently than larger molecules. Andrews *et al.* introduced the energy normalization idea by calculating binding energy for functional groups (Andrews *et al.*, 1984). Later, Kuntz and colleagues investigated the binding energy per atom of different compounds (Kuntz *et al.*, 1999). Some years later, Hopkins and co-workers introduced the term ligand efficiency (LE) as a metric for evaluating the quality of binding (Hopkins *et al.*, 2004).

$$LE = \frac{-\Delta G^{\circ}}{NHA} = \frac{-RT\ln K_D}{NHA} \qquad \text{(Equation 1)}$$

LE is calculated by equation 1 in which ΔG° (kcalmol^{-1}) (1 cal = 4.18 J) is the standard free energy of the binding, R is the universal gas constant (R = 1.99 × 10^{-3} kcalK^{-1}mol^{-1}), T is the absolute temperature in Kelvin (K), K_D (M) is the dissociation constant, and NHA represents the number of non-hydrogen atoms (heavy atoms). More explanation of the binding energy ΔG° is found in chapter 5. The common unit for LE is kcalmol^{-1} per atom. As the average MW of each heavy atom is 13.286 (Hopkins *et al.*, 2004), a drug with a MW of 500 may have 38 NHA and an average affinity of about 10-20 nM (Hopkins *et al.*, 2004; Overington *et al.*, 2006), which results in an LE of about 0.3 kcalmol^{-1} per atom by equation 1. The optimization process normally cannot increase LE, and as a result, LE = 0.3 kcalmol^{-1} per atom is often considered as the lower cut-off for hits (Congreve *et al.*, 2008).

3.2. Tools for fragment screening

The most obvious difficulty in hit identification using FBDD is the weak affinity of a hit. According to equation 1, fragment hits with LE = 0.3 at an

MW of 100–300 (NHA from 8–23) will generate K_D values ranging from 17.5 mM to 9 μM. Therefore, in order to detect these small but efficient fragments, high sensitivity to affinity is a prerequisite for any technique involved in fragment hit detection. Consequently, primary screening techniques in FBDD are essentially different from traditional screening tools. Additional requirements for any screening technique are, for economical reasons, a high throughput and a low consumption of the target protein and library samples. The throughput in fragment screening is, however, less demanding compared to conventional HTS. The reason, as mentioned above, is that fragment libraries include much fewer compounds compared to HTS libraries.

A number of affinity-determining techniques are available for fragment screening, and each tool possesses its own advantages and pitfalls. A fragment that shows up as a hit by one screening technique is not necessarily a hit with another screening method (Kobayashi *et al.*, 2010; Wielens *et al.*, 2013). Furthermore, some kinds of fragments or targets are not suitable for a certain screening technique. For example, fragments that are difficult to ionize may challenge detection by mass spectrometers (MS) that use electrospray ionization (Annis *et al.*, 2004), large proteins are difficult to use in protein-observed NMR (Hubbard & Murray, 2011), and membrane-associated proteins cannot participate in thermal shift assays that use lipophilic dyes as assay indicators (Kranz & Schalk-Hihi, 2011). It is therefore generally advisable to use several screening methods that complement each other to achieve more reliable results in FBDD, and new techniques are always appreciated.

Nuclear magnetic resonance (NMR)

The NMR spectroscopic methodologies for screening include several approaches. The basic classifications of the methods are due to the object of the signal monitoring: protein-observed and ligand-observed NMR experiments. The protein-observed method, although being more complex, can provide detailed information of the binding to the protein (Scott *et al.*, 2012), which can guide in the fragment structure development and distinguish specific from non-specific interactions. Ligand-observed methods, on the other hand, are more common in screening due to their capability of high throughput and that they are not dependent on protein size and labeling. Normally, mixtures of 8–30 compounds can be monitored to improve throughput to about 1000 compounds per day (Hajduk *et al.*, 1999; Klages *et al.*, 2006; Hubbard & Murray, 2011). Ligand affinity from 100 nM to 10 mM can be detected in ligand-observed NMR (Neumann *et al.*, 2007; Moumné *et al.*, 2012), although there has been a report on detected affinity down to 10 nM (Mayer & Meyer, 1999). Information about the binding site is, however,

not acquired with this approach. FBDD employs both methods, where in principle ligand-observed NMR is used for screening and protein-observed NMR is employed for characterization of the binding.

The protein consumption requirement is often the subject of optimization in NMR spectroscopy. In the commonly used ligand-observed NMR screening methods, the protein consumption is in the range of 20–40 mg for a few thousands of fragments (Hubbard & Murray, 2011). One effort to reduce protein consumption is to immobilize the target on a supporting resin, also called target immobilized NMR screening (TINS) (Vanwetswinkel *et al.*, 2005; Früh *et al.*, 2010). The immobilization of the target makes it possible to do flow injections of fragment mixtures with a single sample of protein, which reduces the protein consumption to 3–5 mg protein to screen a fragment library. In ^{19}F NMR, the protein consumption for screening a few thousands of fragments has been reported to be about 1 mg (Dalvit & Vulpetti, 2012).

X-ray crystallography

Another method that was applied early in fragment screening is X-ray crystallography (Nienaber *et al.*, 2000; Lesuisse *et al.*, 2002; Sharff & Jhoti, 2003; Hartshorn *et al.*, 2005). Like protein-observed NMR, this method provides information of the protein-ligand complex, which is helpful in guiding fragment development. In most FBDD programs, the ability to obtain a crystal structure with identified hits is considered crucial.

Protein crystals are normally grown separately before soaking into a cocktail of fragments. The cocktail may contain from 1–8 fragments, which should be diverse in shapes to facilitate an electron density reading. In rare cases, fragment cocktails containing up to 100 compounds have been reported (Nienaber *et al.*, 2000). The cocktail soaking method, although able to give high throughput, faces some technical difficulties. First, the protein crystal has to survive the soaking with the fragment cocktail, then it should have an orientation that exposes the active site for binding (Hubbard & Murray, 2011). If the binding of a fragment induces a large conformational change of the protein, the crystal may crack and no data can be obtained (Scott *et al.*, 2012). In the difficult cases, co-crystallization with fragment cocktails can be used as an alternative strategy (Jhoti *et al.*, 2007).

X-ray crystallography for fragment screening has advantages as well as shortcomings. Although difficulties in growing crystals for some proteins exist, there is no limit for the protein size, which can be a problem in protein-based NMR. The obtained structure information from crystallography always has some uncertainties (Davis *et al.*, 2003; DePristo *et al.*, 2004), but it is often more reliable than the information from NMR experiments. A major

drawback is that protein consumption is normally high and throughput is low (Jhoti *et al.*, 2007). To tackle these problems, efforts to use the protein more efficiently by micro-fluidic platforms have been reported (Lau *et al.*, 2007). One example of improving throughput is pixel array detectors that have been reported to examine 345 crystals in 24 h (Wasserman *et al.*, 2012). The time needed to obtain the crystals was, however, not reported. Another drawback is that crystallography only provides results as a yes or no answer; no affinity data can be extracted. False positive hits are rarely obtained with the technique, but false negative results are common due to that the solid state of the protein crystal can interfere with binding. Errors in reading the electron density are also common (Davis *et al.*, 2008). Today X-ray crystallography is still a resource-intensive technology even though new techniques have been developed.

Surface plasmon resonance (SPR)

SPR is another powerful label-free technology for detecting and characterizing biomolecular interactions. Together with NMR and X-ray crystallography, SPR is among the most used biophysical methods in fragment screening today. The prominent advantage of SPR is the very low consumption of protein, only about 25–50 μg for a screening campaign (Dalvit, 2009; Perspicace *et al.*, 2009; Navratilova & Hopkins, 2010). Screening speed of SPR can reach 1000 compounds per day (Boyd *et al.*, 2012).

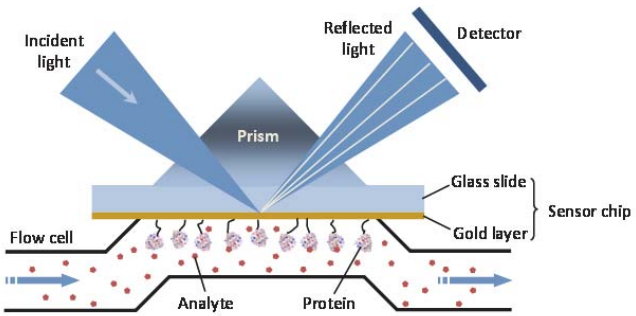

Figure 3. A scheme of the fluidic system in SPR instrumentation. The figure is adapted from Biacore Handbook (Biacore, 2008) with permission from GE Healthcare.

SPR instruments appear in many configurations, but the fluidic system is in principle the same as described in figure 3. It is comprised of a flow cell and a sensor chip with its surface covered by a thin layer of gold. The chip surface is divided into different channels that can have a parallel, serial, or array arrangement. One of the binding species is immobilized to the gold surface, using suitable chemistry. The active target is normally immobilized in one

channel and the other channels serve as references. The reference channels can contain active-site inhibited protein, or an irrelevant target protein, or just be blank. In another experimental design, different channels contained different targets to improve the screening throughput (Elinder *et al.*, 2011).

As SPR sensors rely on changes in reflective index upon binding, small-sized and weakly-binding molecules such as fragments will induce only small changes in the refractive index. The binding of fragments are therefore a challenge for detection by SPR. Furthermore, upon immobilization, the protein may lose activity due to unfavorable orientation, conformational changes or loss of structural integrity. The effect is most obvious with membrane proteins, which often have a complex structure and require the mimic of their natural environment to be fully functional. Despite these difficulties, successful SPR experiments with membrane proteins have been reported (Maynard *et al.*, 2009; Rich *et al.*, 2011; Seeger *et al.*, 2012).

SPR is one of very few techniques that can provide the kinetics (k_{on} and k_{off}) of an interaction as well as the affinity (K_D) (Myszka *et al.*, 1998; Papalia *et al.*, 2008). Kinetic data can be extracted from the shape of the sensorgram, and the dissociation constant can be deduced from the k_{on} and k_{off} values. However, due to the fast on-rate and off-rate of weak binders as fragments, it is often difficult to extract such data in fragment screening. The alternative is to measure steady state responses at many concentrations and fit the collected responses into a binding model to derive K_D. This approach is, however, time-consuming because many injections are needed. To deal with the problem, the SPR manufacturer SensiQ has recently developed an injection method in which binding data of many concentrations is acquired by a continuous dilution of a single injection (Rich *et al.*, 2010; Quinn, 2012). These advances in injection techniques promise a higher throughput in fragment screening by SPR.

Other technical challenges in fragment screening using the SPR technique include non-specific interactions that arise due to a combination of binding to the surface matrix and to the protein surface outside the active site. These problems might be controlled by using a reference surface with an inactive form of the protein or irrelevant protein, or adding a competitive agent into the mobile phase (Perspicace *et al.*, 2009; Navratilova & Hopkins, 2010). Another problem is compound aggregations that can cause complicated responses especially when high concentrations of compounds are used (Giannetti *et al.*, 2008). A major source of false positives in SPR is the presence of dimethyl sulfoxide (DMSO) in the fragment solutions. DMSO is the most popular storage solvent for compound libraries, so its co-existence with analytes is often unavoidable. A calibration curve for DMSO is often required to subtract the DMSO effect in SPR samples.

Mass spectrometry (MS)

Affinity-based mass spectrometry (MS) is another technology for screening small molecules in drug discovery. This methodology divides into branches with different operational modes. MS can detect binding by measuring the increased MW of a protein-ligand complex either directly (Maple *et al.*, 2012) or after size exclusion chromatography (AS-MS) (Annis *et al.*, 2004). In other configurations, the MS is hyphenated with other techniques, such as, affinity capillary electrophoresis (ACE-MS) (Mironov *et al.*, 2012) and frontal affinity chromatography (FAC-MS) (Chan *et al.*, 2003; Ng *et al.*, 2007) to identify the binding species. The WAC technique, which is the focus of this thesis (see section 6), can also be coupled with MS to form WAC-MS. The combination of WAC with MS allows analysis of mixtures containing many compounds, thereby increasing throughput with minimal protein and analyte consumption. WAC-MS was used in **papers II**, **III** and **IV** of this thesis.

MS is a technology that detects charged atoms or molecules by separating them in electrical/magnetic fields. It can be used to identify compounds in samples and to elucidate the structure of molecules. There are many set-ups of MS, but they all share a common instrumental scheme (figure 4).

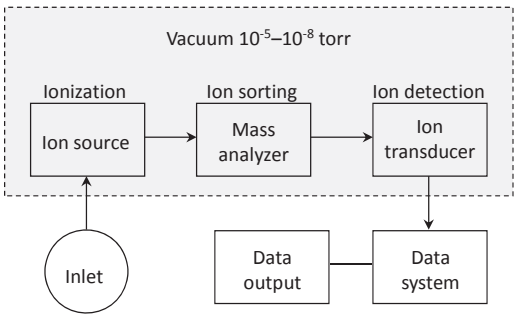

Figure 4. Organization of different units in a MS. The figure is adapted from Skoog et al., 2007 with modifications.

MS is commonly used in combination with techniques that can separate the components of a sample, such as, capillary electrophoresis or liquid chromatography. In these configurations, samples enter the MS system as liquid solutions. Since MS only recognizes charged species, the sample needs to be ionized and transferred into a gaseous phase before advancing into the mass analyzer where the ions are sorted according to their mass-to-charge ratio (*m/z*). Depending on the instrument configuration, there may be only one or a couple of mass analyzers successively connected. The sorted ions are finally recognized by the ion detector equipped with a signal amplifier. The

whole system, except the inlet and data processor, is placed under vacuum to avoid contamination from air molecules. The major parts that distinguish the different MS instruments are the ion source and the mass analyzer, which are discussed in the following paragraphs.

The common ion sources used in combination with a liquid sample inlet are atmospheric pressure electrospray ionization (API-ES, or ESI), atmospheric pressure chemical ionization (APCI) and atmospheric pressure photo-ionization (APPI) (Kostiainen & Kauppila, 2009). All three techniques are soft ionization methods that work at ambient pressure by creating a spray of droplets. The techniques differ in how the charge is transferred to the sample species, but the result is that gaseous ions are created from the components in the sample, while the solvent is evaporated from the droplets. The composition of the mobile phase is important for sample species to be able to receive a proton in positive ionization or to expel a proton in negative ionization mode. A common problem in sample ionization is that dominant ions in mobile phase may suppress the ionization, and therefore hampers the detection of species that are less abundant.

In ESI, the ionizing agent is a high electrical field which charges the substances in the sample spray (Dole *et al.*, 1968). In APCI, the sample is vaporized by high temperature and then ionized by a corona discharge needle reacting with other gas phase ion-molecules that assist in transferring protons or electrons from the reaction environment. The APPI ion source works by the same principle as APCI, but uses a vacuum ultraviolet lamp to produce photons that assist in ionization (Kostiainen & Kauppila, 2009). All mentioned techniques can ionize a wide range of molecules, but ESI is the most used method. Ion source selection and analytical conditions are of great importance in LC-MS because they directly affect the MS sensitivity.

The mass analyzer component in an MS also varies. The most common is an instrument that uses a quadrupole unit (figure 5). A quadrupole comprises four parallel rod-shaped electrodes. The rods are connected pairwise to a direct-current electric source (dc). One pair of rods has a negative charge, while the other pair has a positive. An alternating-current voltage (ac) at radio frequency (RF) is added to each pair of rods with an 180° phase difference. The combination of varying repulsive and attractive forces from the electromagnetic field caused by the quadrupole guides the trajectory of the sample ions when they enter the space between the electrodes. Only ions within a suitable range of mass-to-charge ratio can travel the correct way and reach the detector. Other ions will be deflected, neutralized by hitting the electrodes and/or pumped away. By varying the strength of the dc and ac but keeping their proportion constant, a range of ion masses can be detected (Miller & Denton, 1986; Leary & Schmidt, 1996; Skoog *et al.*, 2007).

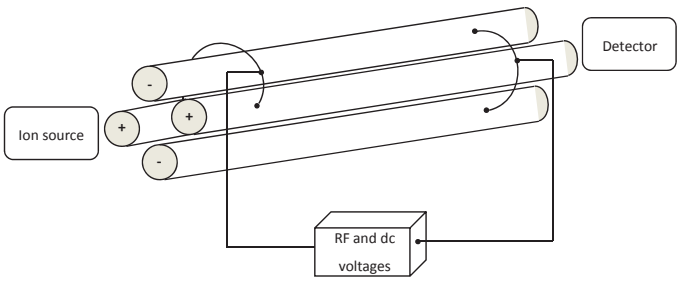

Figure 5. Configuration of a quadrupole mass filter. Adapted from Miller and Denton, 1986 with permission from American Chemical Society.

Other affinity detection techniques

There are many other techniques that are applied in fragment screening. The most used complementary method is probably the thermal shift assay. In this method, the denaturation or aggregation (unfolding) temperature of a protein is monitored (Senisterra *et al.*, 2006; Kranz & Schalk-Hihi, 2011). When a compound binds to a protein, it stabilizes the protein and therefore increases the unfolding temperature. Recording this change in unfolding temperatures can provide the affinity of an interaction. Thermal shift screening is a plate-based, high-throughput technique but results are not always reproducible and false negatives are frequent (Larsson *et al.*, 2011; Scott *et al.*, 2012).

Isothermal titration calorimetry (ITC) is another method for screening (Torres *et al.*, 2010). The technique measures the change in heat production or absorption upon binding. One can use ITC to determine the overall changes in free energy (and thus the affinity), enthalpy, and entropy of the binding. The shortcomings of the method with common ITC instruments are high consumption of protein, low throughput and low sensitivity (Ladbury, 2010; Torres *et al.*, 2010). Improvements in protein consumption and throughput have been achieved by using array calorimetry, but the low sensitivity is still a challenge (Torres *et al.*, 2010).

Plate-based functional screening, often used in conventional high throughput screening, can also be applied in fragment screening. The low affinity of fragments often requires the screening to be performed at high concentration, which may lead to more artifacts due to problems in fragment aggregation and signal interference (McGovern *et al.*, 2002; Albert, 2010; Hubbard & Murray, 2011). Functional screening, however, is more likely to provide hits that inhibit the target function rather than molecules that only bind to it. This

approach can perform screening without pre-knowledge of the target and is able to discover new phenotypes as well as new targets.

Affinity capillary electrophoresis also contributes to methods for fragment screening (CEfrag) (Austin *et al.*, 2012). Affinities are detected by monitoring changes in the electrophoretic mobility of the protein, ligand or protein – ligand complex due to binding. The interactions occur in solution, and the protein target does not need to be immobilized. The protein consumption in CEfrag is low, but the screening speed is rather slow.

Finally, computational methods can be involved in almost all phases of drug discovery and development (Ghose *et al.*, 1998; Ertl *et al.*, 2000; Kontoyianni & Rosnick, 2012). Although there are diverse opinions on its accuracy in screening (Klebe, 2006; Enyedy & Egan, 2008; Chen & Shoichet, 2009), it is the least costly tool and gives high throughput. When handled properly, *in silico* screening is helpful in scaling down the screening library. In library design, computational methods may be used to guide the synthesis of directed libraries and select fragments to put in a library. In addition, it is an integral part of the hit evaluation process.

3.3. Fragment library design

The design of a library for screening is of utmost importance for the outcome of a screening campaign. Many of the considerations discussed below can be applied for both conventional HTS and fragment screening libraries with some exceptions. A low MW is generally desirable in both HTS and FBDD libraries. Traditional HTS libraries previously contained large, lead-like molecules (MW less than 460), but suggestions of a lower level (MW ≤ 350) for HTS libraries have been proposed (Teague *et al.*, 1999; Nadin *et al.*, 2012). On the other hand, criteria for fragment libraries sometimes extend the upper limit of the MW to 350, which makes HTS and fragment libraries overlapping (Congreve *et al.*, 2008). However, fragment screening can now be considered for even smaller compounds with a MW not larger than 250. Compounds with MW from 250 to 350 are called scaffolds (Card *et al.*, 2005; Hubbard & Murray, 2011). As discussed in chapter 3.1, "the number of heavy atoms" (NHA) can be used to describe a molecule, which is often regarded as more accurate as it is not biased by heavy weight atoms, such as, iodine, bromine, etc.

The common requirements for compounds in all types of libraries are high aqueous solubility, high purity and stability, and low toxicity. A high aqueous solubility is a prerequisite for almost all screening methods, although it can restrict and narrow the diversity of the library (Stockman & Dalvit, 2002).

Purity and identity criteria of library compounds are important, especially in functional screening where impurities can have dramatic effects in activity assays. Other common physico-chemical considerations for library components are lipophilicity, numbers of hydrogen bonding donors and acceptors, polar surface area and number of rotatable bonds (Lipinski *et al.*, 1997; Congreve *et al.*, 2003). Compounds should also be suitable for chemical elaboration, be feasible to synthesize or should be commercially available.

The library design is also influenced by the applied screening method. For example, NMR approaches, X-ray crystallography, and WAC-MS can deal with mixtures, thereby creating less demand on purity. On the other hand, mixture screening requires that the included compounds do not react with each other. As a common criterion, fragment libraries used in NMR experiments possess higher aqueous solubility compared to general-purpose libraries. Furthermore, libraries with fluorine-labeled compounds are needed for ^{19}F-NMR experiments.

An important consideration in library design is the diversity of the included compounds. The level of diversity depends on the library purpose, style and experience of the designer. A focused, or directed library, is built based on fragment leads of a certain target family, and therefore possess common structures that are expected to fit into the target active site or exert functional activities (Stockman & Dalvit, 2002). Using this type of library narrows the focus only to known chemotypes. This strategy may be efficient in finding hits but may not be able to find novel binding entities. Another type of fragment library is designed for working with a wide range of targets (Chen & Hubbard, 2009). The construction of these general libraries concentrates more on chemical diversity and often needs the assistance of computational tools. This kind of library can cover a large chemical space with a few thousand members, and they can produce hits of many different chemical types; therefore novel hit structures can be achieved. However, the selection process is still rather subjective. Even more varieties of chemotypes in a fragment library may be achieved by diversity-oriented synthesis approaches that increase three-dimensional components including chirality (Hung *et al.*, 2011). This build-up of a library, however, does not strictly control physico-chemical properties of fragments and may raise concern about the unexpected behavior of the components.

Stereoisomerism in FBDD

It is well known that stereoisomers are distinct substances in terms of interacting with a stereoisomeric species such as proteins. Drugs that differ only in configurations can have very different therapeutic uses. One example is the pair quinine and quinidine (figure 6). These two stereoisomers differ in

geometries at two stereocenters, which results in different therapeutic indications: quinine is used for anti-malarial purposes while quinidine acts mainly as an anti-arrhythmic agent.

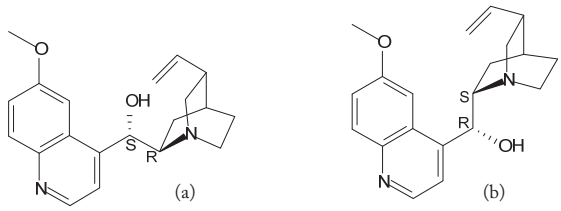

Figure 6. Structures of quinine (a) and quinidine (b) as an example of a stereoselective pair of drugs. The shown configurations indicate where differences take place. R and S in the structures indicate the different configurations.

In drug discovery and development, stereoisomerism is unavoidable but beneficial because it introduces more complexity and subsequently increases diversity of the library (Hung *et al.*, 2011). As chiral selective synthesis is costly and difficult, stereoisomeric mixtures are often present in fragment libraries. In most cases the determination of the affinity of individual stereoisomers is impossible, and therefore neglected. As a result, if a stereoisomeric mixture is identified as a hit, both stereoisomers may enter further fragment elaboration even though only one of them is useful. In addition, if the affinity differs greatly between the stereoisomers in a mixture, the presence of the weaker affinity may obscure the screening output, which might cause false negatives. Information of the affinity of individual stereoisomers is of the same reasons important in hit elaboration.

There are a few methods that can recognize affinity differences from the individual isomers in a stereoisomeric mixture. The ITC method can detect the phenomenon, in favorable situations, through the shape of the titration curve if the difference between the isomers is 50 to 200 fold (Fokkens & Klebe, 2006). Separation-based affinity determination methods, such as the WAC technique presented in this thesis, however, are advantageous for this application, which is discussed in greater detail in **paper IV** and section 9.2.

4. TARGETS IN DRUG DISCOVERY

Druggable targets for small organic molecules are typically proteins. The quality of a protein target is one of the important determinants of the overall drug discovery success, as discussed in 2.3. It is estimated that there are 20000–25000 protein-encoding genes in the human genome (Collins *et al.*, 2004). About 10–30 % of the proteins are druggable, and approximately 2–5 % is suitable protein targets for small organic molecules (Hopkins & Groom, 2002; Betz, 2005). Another source of a considerable number of therapeutic relevant proteins is found in microorganisms that constitute important targets for anti-infective drugs. However, only a fraction of all therapeutically relevant targets have been explored (Zheng *et al.*, 2006; Edwards *et al.*, 2011) and only about four new targets were introduced each year in the 1990s (Hopkins & Groom, 2002).

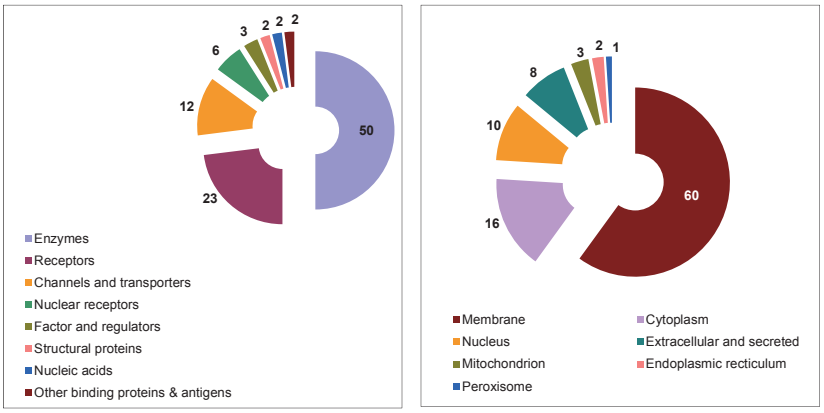

Figure 7. Percentages of target classes (left) and locations (right). Data were taken from Zheng et al., 2006 with permission from Elsevier.

The most commonly used targets for approved drugs are enzymes (50 % of all targets), receptors (23 %), channels and transporters (12 %), and nuclear receptors (6 %) (figure 7, left) (Zheng *et al.*, 2006). Enzymes contribute with a significant part of the target population, and they are mostly explored for anti-infective, cancer and cardiovascular treatments. More than half of enzyme targets for cardiovascular diseases are proteases (Betz, 2005).

The group of membrane-associated proteins occupies about 60 % of all the approved targets, and therefore these targets draw much attention from the drug discovery community (Overington *et al.*, 2006; Zheng *et al.*, 2006) (figure 7, right). From a FBDD perspective, membrane proteins are especially attractive, but they are difficult to deal with due to their lipophilicity and their demand of a suitable environment to function.

4.1. Studied protein targets

Thrombin

Thrombin (EC.3.4.21.5) is an important target in drug discovery for cardiovascular diseases, and it belongs to the enzyme family of serine proteases. It is further grouped into the trypsin-like/chymotrypsin-like superfamily due to its similarity in structure at the active site with trypsin/chymotrypsin (Tyndall *et al.*, 2005). As a protease, thrombin hydrolyzes peptide bonds of substrates. In this work, thrombin was used in **papers II**, **III** and **IV** as a model of a protein target.

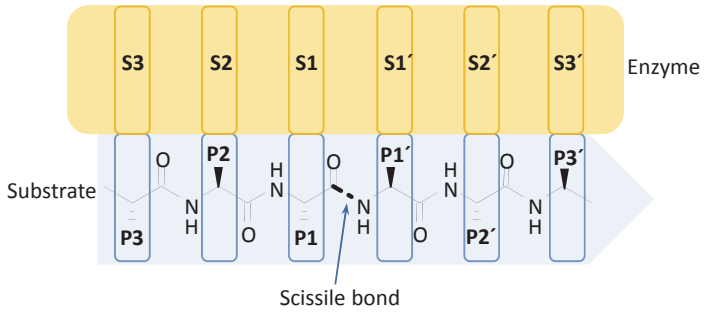

Figure 8. Binding scheme of a substrate ligand into the binding pocket of a protease. The figure is adapted from Schechter and Berger, 1967 with modifications.

The active site of a protease is often large and comprises of several sub-sites; each sub-site is a binding area that corresponds to an amino acid residue of the substrate (figure 8). Schechter and Berger have set up a naming system for the

sub-sites and binding residues, based on their relative position compared to the cleavable peptide bond (Schechter & Berger, 1967). Substrate residues located on the amino terminus side are denoted as P followed by non-primed numbers that start from the residue closest to the scissile bond. The other direction is the primed-numbering side. The binding sites on the enzyme are labeled correspondingly with the letter S.

Trypsin-like proteases share a common structure of two β-barrels that form two adjacent domains carrying the catalytic triad at the interface. The catalytic triad is highly preserved among the group members and includes Ser195, His57 and Asp102 (Perona & Craik, 1997). Together with the oxyanion hole, which is built up by the backbone of Ser195 and Gly193, the catalytic triad forms an active site cleft. In action, the hydroxyl group of Ser195 works in coordination with His57 and Asp102 to function as a nucleophile that attacks and breaks the peptide scissile bond of the substrate (figure 9) (Hedstrom, 2002). Among all the serine residues of the protease enzyme, only serine in the catalytic triad possesses a catalytic ability due to its advantageous location, where it receives not only full support from His57 and Asp102, but also the substrate at a right position for the reaction to occur. The specificity of an individual substrate of proteases is achieved by the topology of the binding site close to the catalytic triad (Hedstrom, 2002). Thrombin has insertion loops that define its narrow specificity (figure 10) (Bode *et al.*, 1989).

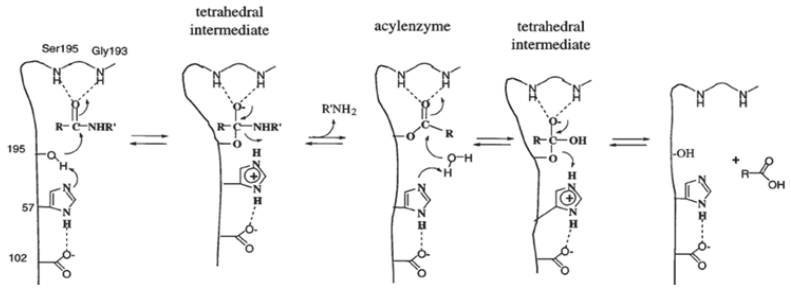

Figure 9. Suggested mechanism of action for serine proteases. The figure is from Hedstrom, 2002 with permission from American Chemical Society.

Human α-thrombin is a water soluble protein that consists of two disulfide-bridged polypeptide chains A and B of 36 and 259 amino acid residues, respectively (Butkowski *et al.*, 1977; Bode *et al.*, 1989). It is produced in the liver in the form of the zymogen, prothrombin (Hall, 2011). Upon autolysis or degradation of thrombin by other enzymes, β- and γ-thrombin are formed together with other degrading products (Fenton *et al.*, 1977; Boissel *et al.*, 1984). These products have lost almost all biological activity, but maintain the binding capacity to small ligands (Fenton *et al.*, 1977; Bode *et al.*, 1989).

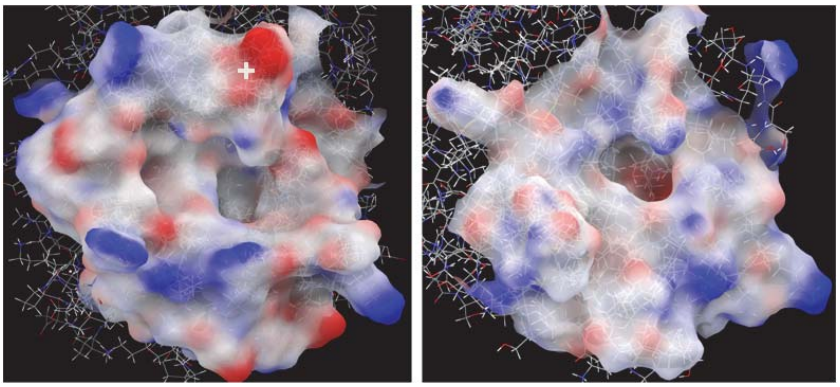

Figure 10. Active site surfaces of human thrombin (PDB ID 1K22) (left) and bovine trypsin (PDB ID 1K1P) (right) generated by Maestro (v.9.2, Schrödinger Suite 2012). The surface colors describe electrostatic potential with red for negative and blue for positive charges. The cross on the thrombin structure locates the selective insertion loop Tyr60A-Pro60B-Pro60C-Trp60D which is not present in the trypsin.

Thrombin plays an important regulatory function in the coagulation cascade. When blood vessel damage occurs, the body responds by a series of chemical reactions that will convert pro-thrombin into thrombin and start the coagulation process. In the coagulation cascade, thrombin also works as a fibrinolytic/anti-fibrinolytic agent according to the situation (Siller-Matula *et al.*, 2011). On the surface of thrombin, in addition to the active site, there are many binding sites for different kinds of molecules. The sodium binding site is located just about 15 Å from the catalytic triad where it changes the thrombin conformation upon binding with the sodium ion, thereby regulating thrombin functions (Di Cera *et al.*, 1995). Exosite I, or the hirudin binding site, mediates the substrate, co-factor and inhibitor recognition. Exosite II is a heparin binding site and regulates thrombin activity towards protease-activated receptors, which play a role in platelet aggregation (Bode, 2006). Although there are many binding sites that regulate thrombin function, which could be possible target sites for drug discovery, inhibition of thrombin at the active site to prevent its clotting activity is the most desired approach.

Despite considerable efforts to develop small molecule anti-coagulants that act on thrombin, there are only two drugs of this kind available (Sinauridze *et al.*, 2011). They are argatroban (Schwarz *et al.*, 1997) for intravenous administration and dabigatran etexilate (Stangier *et al.*, 2007) that can be used orally. Another drug is melagatran which was approved for clinical use as a thrombin inhibitor (Brighton, 2004) but was withdrawn from the market in 2006 due to severe liver toxicity (AstraZeneca, 2006). All three drugs have an

amidine moiety as the common structure (figure 11). This functional group is an established anchor that mimics the arginine or guanidine residue of a natural substrate. The moiety binds to Asp189 which is located deep at the bottom of the S1 pocket in the thrombin active site (Dullweber *et al.*, 2001).

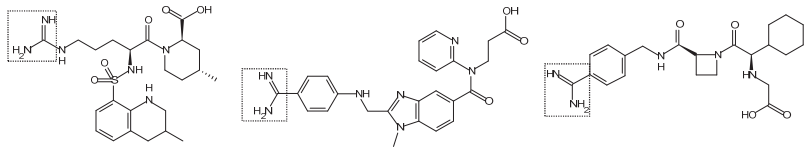

Figure 11. Anticoagulants that act on thrombin. Amidine structure is shown in boxes.

Trypsin

Trypsin (EC 3.4.21.4), as mentioned above, is another serine protease in the same super-family as thrombin. The two proteins share many similar features in structure, but their functions are completely different. Thrombin is active in the coagulation cascade, while trypsin works mainly as a digestive enzyme. However, except the insertion loops, which are present in thrombin and are responsible for its selectivity, the active sites of the two proteins are quite similar (Hilpert *et al.*, 1994). Trypsin is the least selective enzyme in the family and like thrombin; it cleaves peptide bonds at the carboxyl side of the amino acids lysine or arginine (Olsen *et al.*, 2004). The enzyme is produced in the pancreas in the form of the inactive pro-enzyme trypsinogen and becomes active when it is cleaved either by enterokinase secreted by the intestinal mucosa, or by itself when contact occurs with the active form (Hall, 2011). Trypsin was employed as a model protein in **paper II** together with thrombin.

Albumin

Serum albumin is a transport protein that plays a crucial role in drug ADME. Apart from ligand binding and transport functions, serum albumin also regulates the colloid osmotic pressure and capillary membrane permeability, scavenges free radicals and has anti-oxidant and circulatory protective properties (Fanali *et al.*, 2012).

Albumin has binding sites for fatty acids, thyroxin, bilirubin, drugs and other species (Bujacz, 2012). There are two major binding sites for drugs, called Sudlow sites I and II, which are located on sub-domains IIA and IIIA of the protein, respectively (Sudlow *et al.*, 1975; Kragh-Hansen *et al.*, 2002). Site I is more flexible and binds to dicarboxylic acids or bulky heterocyclic molecules with a negative charge in the middle, such as, warfarin, phenylbutazone and

salicylate. Site II, also called the indole-benzodiazepine site, binds to aromatic carboxylic acids with a negatively charged acidic group at one end of the molecule distant from a hydrophobic center. Examples of such molecules are diazepam, ibuprofen and L-tryptophan (Kragh-Hansen *et al.*, 2002).

Bovine serum albumin (BSA) was used as a model protein in **paper I**. Although BSA exhibits only about 76 % identity with human serum albumin (HSA) (Bujacz, 2012), it shares many features with HSA. The structure of serum albumin includes three helical domains, each comprising two sub-domains; all are arranged into a heart-shaped albumin molecule (Bujacz, 2012; Majorek *et al.*, 2012). The sub-domain IB contains the regions where significant differences between HSA and BSA are found (Majorek *et al.*, 2012).

5. INTERACTIONS BETWEEN DRUG AND TARGET

Interactions always exist between two molecules when they are in close proximity of each other. They can be strong or weak, attractive or repulsive and depend on various properties, such as, size, shape, hydrophobicity and charge patterns of the interacting pair. In drug discovery, the focus is to define the interactors that modify (activate or inhibit) the behavior of a target in a way that can be translated into a therapeutic use. Although there are many other features that constitute a drug, the characteristics of the interaction with the target and off-targets are the most obvious that researchers have to focus on in the early stages of drug discovery.

The common parameters to characterize an interaction between two species are affinity and kinetics. Affinity is often expressed as an association or a dissociation constant termed K_A (M^{-1}) or K_D (M), respectively. Kinetics is described by on-rate and off-rate of the interactions (k_{on}; $M^{-1}s^{-1}$ and k_{off}; s^{-1}, respectively). The full mechanism of the interaction may be complex to clarify, but the outcome can be easily measured by the concentrations of the interacting species at equilibrium. High concentration of the product complex compared to the reactant concentrations at equilibrium means the complex is stable, or of high affinity. This will subsequently translate into a high value of K_A or a low value of K_D. In affinity chromatography, compounds at a K_D of about 1 μM can be eluted by isocratic mobile phases and they are detectable (**paper IV**). Therefore, 1 μM of K_D may be a reasonable borderline for defining a weak interaction.

Considering the simplest interaction between a small molecule ligand (L) and a target receptor (R), the reaction can be described as in scheme 1 and equation 2, in which the brackets express molar concentrations at equilibrium:

$$R + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} RL \qquad \text{(Scheme 1)}$$

$$K_A = \frac{k_{on}}{k_{off}} = \frac{[RL]}{[R][L]} = \frac{1}{K_D} \qquad \text{(Equation 2)}$$

Other parameters that are important in the characterization of the drug-target interaction are the fraction of the bound receptor or the target occupancy and target residence time. Target occupancy relates to affinity and concentration of the drug at the site of action, and it correlates with the strength of the pharmacology effect. On the other hand, target residence time is the time that a drug binds to a target, and thereby can exert its action. Drug activity therefore directly depends on target residence time. The residence time depends on the off-rate of the interaction (Copeland *et al.*, 2006; Tummino & Copeland, 2008; Lu & Tonge, 2010), and it is calculated as the reciprocal of k_{off}. As residence time directly influences the effect of a drug, the characterization of the kinetic parameters, especially the off-rate of an interaction, is essential in drug discovery and development. In this thesis, a pre-study of kinetics by WAC is reported in section 9.2.

The target occupancy (f_b), or fraction bound of a target to the ligand in equilibrium studies can be determined by equation 3:

$$f_b = \frac{[RL]}{[R] + [RL]} \qquad \text{(Equation 3)}$$

From equation 2, the free concentration of the receptor [R] can be expressed as in equation 4:

$$[R] = \frac{K_D \times [RL]}{[L]} \qquad \text{(Equation 4)}$$

After combining equations 3 and 4, equation 5 is deduced:

$$f_b = \frac{1}{\frac{K_D}{[L]} + 1} = \frac{[L]}{K_D + [L]} \qquad \text{(Equation 5)}$$

It can be seen from equation 5 that the bound fraction of receptor is 50 % of its total concentration (f_b = 0.5) when the concentration of free ligand [L] is equal to the K_D. This characteristic is useful to determine the K_D from ligand

binding assays at equilibrium, such as, frontal chromatography (see more details about this technique in chapter 8). Equation 5 can also be used in inhibitory or competition studies, in which it indicates the concentration of the free ligand [L] needed in relation to the K_D to obtain a reasonable fraction of the occupied protein.

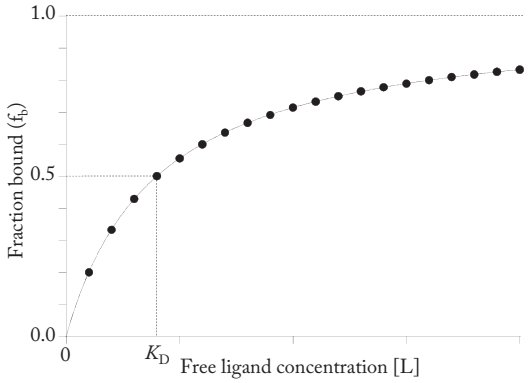

Figure 12. Relationship between fraction bound (f_b), free ligand concentration [L] and K_D.

By defining the total concentration of the receptor $[R]_o$ as the sum of receptor in complex with ligand [RL] and free receptor [R] concentration at equilibrium (equation 6) and combine this expression with equation 4, the so called Langmuir equation (Langmuir isotherm) is obtained (equation 7):

$$[R]_o = [R] + [RL] \qquad \text{(Equation 6)}$$

$$[RL] = \frac{[R]_o \times [L]}{K_D + [L]} \qquad \text{(Equation 7)}$$

The strength of an interaction between a ligand and a receptor can also be described by the standard Gibbs free energy (ΔG^o) ($kcalmol^{-1}$) (Atkins & de Paula, 2006); the more negative value of ΔG^o the stronger the binding. By definition, the ΔG^o is contributed by standard changes in the enthalpy (ΔH^o) ($kcalmol^{-1}$) and the entropy (ΔS^o) ($kcalK^{-1}mol^{-1}$) of the interaction (equation 8).

$$\Delta G^o = \Delta H^o - T\Delta S^o = -RT\ln K_A = RT\ln K_D \qquad \text{(Equation 8)}$$

By thermodynamic laws, in order for spontaneous binding to occur at a constant temperature and pressure, ΔG^o has to be negative so that the forming of the product (the complex) will drive the system to a lower energy point.

Typically hydrogen bonding, the hydrophobic effect, van der Waals forces and electrostatic interactions all contribute to the binding free energy (Kuntz *et al.*, 1999). An entropic advantage is often gained from increasing the molecular size and hydrophobicity, which is beneficial for binding affinity but disadvantageous regarding the physico-chemical properties of the molecule (ADME properties, see chapter 2.3). A favorable enthalpy interaction is usually connected to hydrogen bonding and electrostatic interactions with the active site of the target protein, which increases the affinity without sacrificing ADME properties of the molecule. The dominant enthalpic interacting compounds are for this reason more desirable in drug discovery (Ferenczy & Keserű, 2010). However, both these components of the free energy are intertwined, and optimization of a binding molecule in one component, such as enthalpy, may lead to a penalty of the other, such as entropy.

The selectivity or specificity of an interaction is an important parameter to estimate if a drug candidate interacts with a target at the active site (a specific interaction), or if it binds off-site (a non-specific interaction). The selectivity can be conveniently expressed as the ratio between $K_{\mathrm{Dspecific}}$ and $K_{\mathrm{Dnon\text{-}specific}}$. In other words, it describes the degree at which the interaction is specific towards the active site as compared to all other non-specific interactions to irrelevant sites. There is no clear relation between affinity and specificity. However, it is the general perception that higher affinities give better specificity, although this is not necessarily true. For example, a lower affinity interaction can show a higher specificity if non-specific interactions are virtually not present and/or escape detection (van Regenmortel, 1998).

6. WEAK AFFINITY CHROMATOGRAPHY (WAC)

Affinity chromatography was first introduced in the late 1960s for protein purification (Cuatrecasas *et al.*, 1968). It works by running a solution containing one type of interaction species (often a protein) through an affinity column that contains the corresponding binder (ligand or receptor), which is immobilized to a support (chromatography matrix). Typically strong interactions between the two species will trap the interacting molecules in the column. A following elution step by e.g. a drastic change in pH or ionic strength recovers the compound of interest in pure form. The supporting matrices are often made of soft carbohydrate-based materials. Over the years, the method has been refined and expanded towards different applications, many of which can be seen in modern biotechnology, such as, antibody purification (Ayyar *et al.*, 2012), isolation of plasma glycoproteins (Plavina *et al.*, 2007; Kullolli *et al.*, 2010), and characterization of protein-protein interactions (Dunham *et al.*, 2012). Due to its highly selective nature and efficiency, affinity chromatography today is the most popular technique for purification purposes. Modern formats and applications of affinity chromatography are summarized in a recent review by Hage and colleagues (Hage *et al.*, 2012).

A recent trend in affinity chromatography has been the introduction of high performance supports, such as, highly cross-linked carbohydrates, silica or synthetic polymer monoliths (Hage *et al.*, 2012). It all started with the first successful integration of affinity chromatography with HPLC to form the high-performance (liquid) affinity chromatography (HP(L)AC) platform (Ohlson *et al.*, 1978; Larsson *et al.*, 1983). The affinity column in HP(L)AC was packed with a high resolution supporting material, such as, silica of small particle size in the range of 5–10 µm giving high back pressures as in traditional HPLC. The use of efficient supports in combination with automation allows high-performance analyses, and HP(L)AC is primarily

used as an analytical method e.g. in the characterization of biomolecular interactions.

WAC is a high-performance affinity technique and, as the name suggests, it is based on weak interactions between the target and the ligand typically in the range of K_D from 1 μM to 10 mM (Ohlson *et al.*, 1988; Zopf & Ohlson, 1990). WAC is usually performed under mild and isocratic conditions. Useful applications of WAC and closely related techniques have been found in many fields, such as, in chiral separations (Hofstetter *et al.*, 2002; Ravelet *et al.*, 2004), and in the evaluation of protein-analyte interactions (Moaddel & Wainer, 2006; Bergström *et al.*, 2009; Yoo & Hage, 2009). WAC has also been characterized under overloaded conditions to predict its behaviors in non-ideal conditions (Hubble, 2001). This thesis introduces the application of WAC into FBDD with special emphasis on fragment screening.

In liquid chromatography, the retention factor (k´) is the ratio between the amount of substance in the stationary phase and the amount of substance in the mobile phase, respectively. It is related to affinity and elution volume as shown in the following equation 9 (Bergström, 2006):

$$k' = \frac{V_R - V_M}{V_M} = \frac{K_A R_{tot}}{V_M} = \frac{R_{tot}}{K_D V_M} \qquad \text{(Equation 9)}$$

where R_{tot} (also commonly known as B_{tot}, B_{max} or Q_{max}) is the total number of moles of the actively interacting receptor, e.g. the target protein in the stationary phase, V_R is the retention volume of the analyte and V_M is the mobile phase volume in the column, or void volume. It is important to state that equation 9 is only valid under a linear Langmuir isotherm, i.e. when the concentration of the ligand [L] is much less than K_D and can be ignored in the denominator in equation 7.

Equation 9 can be re-arranged into equations 10. The void volume and the retention volume (V_M and V_R, respectively) can be deduced directly from the void time (t_M), apparent retention time (t_R) and flow rate (F). The adjusted retention time (t'_R) can replace ($t_R - t_M$) to simplify the calculation:

$$K_D = \frac{R_{tot}}{V_R - V_M} = \frac{R_{tot}}{(t_R - t_M) \times F} = \frac{R_{tot}}{t'_R \times F} \qquad \text{(Equation 10)}$$

Equation 10 is frequently employed in WAC to calculate affinity. The peak apex is normally used as the point to measure the retention time of a compound. The various retention times in equation 10 are illustrated in figure 13.

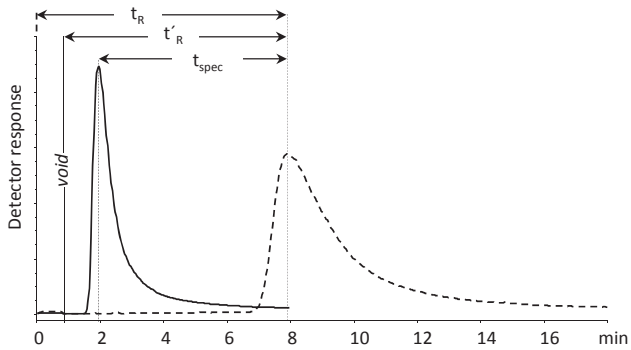

Figure 13. Measurements of t_R, t'_R and t_{spec} from chromatograms on an active (dashed-lined chromatogram) and an inhibited target (solid-lined chromatogram). The figure is from ***paper III****.*

Binding in a WAC column is composed of specific interactions with the active site of a target, as well as non-specific interactions. The non-specific interactions include with the supporting matrix, with off-sites of the target and linkages between the target and the matrix. Information on non-specific binding to the matrix can be obtained by performing WAC on reference columns, such as, a blank column with no immobilized target. A WAC column with a blocked-active site provides an estimation of the total amount of non-specific interactions (figure 13). Specific affinity is calculated based on the specific retention time (t_{spec}, see figure 13) as in equations 11:

$$K_{D\text{specific}} = \frac{R_{tot}}{F \times t_{spec}} \qquad \text{(Equation 11)}$$

A blocked-site protein can be obtained in several ways. One possibility is to block the site either by a permanent covalent binder (**papers II–IV**), or by adding a non-covalent binder with high affinity into the mobile phase (**paper II**). Another alternative is to use a recombinant protein with an inactivated active site. Although an inhibitor introduces a new structure in the active site that may cause irrelevant binding of the analytes, it is still probably the most relevant reference for measuring the total amount of non-specific interactions. The binding to irrelevant proteins can also be used to measure non-specific interactions, but this approach is less accurate since the interaction environment may not be related to the non-specific interactions of the target.

7. AIMS

The overall aim of this thesis is to characterize and develop WAC into a new approach for fragment screening and determination of affinity in early drug discovery. A further aim is to set up an efficient fragment screening procedure using WAC that can operate on standard instrumentation at high throughput and with minimum consumption of the target protein and samples.

To fulfill these aims, the applicability of WAC was evaluated by screening compound collections and a fragment library to select hits and estimate affinities using different HPLC configurations. Throughout the thesis, pros and cons as well as specifications of the WAC technology were defined.

8. METHODS

8.1. Preparation of the weak affinity stationary phase

Porous spherical silica micro-particles are used as the supporting matrix in this thesis. The quality criteria of the supporting matrix are satisfactory coupling yield and immobilization capacity combined with minimum non-specific interactions of the matrix. The choice of pore size involves consideration of the dimensions of the target to be immobilized, such as a protein, where smaller pores normally generate a larger surface area and immobilization capacity, but the pore size should also be large enough to accommodate the target. In **paper I**, silica particles with a diameter of 7 µm and a pore size of 300 Å were used; while the silica support in **papers II–IV** was 5 µm, 300 Å. The silica surface area in both cases was approximately 100 m^2/g.

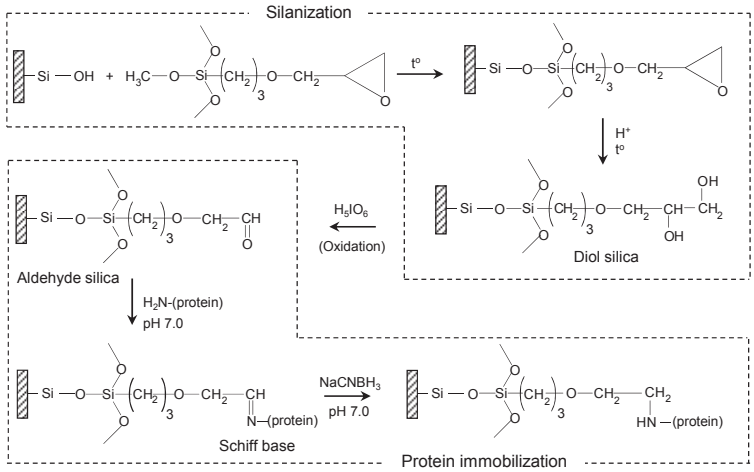

Figure 14. Protein immobilization chemistry. From the oxidation step, the procedure can be performed in situ. The figure is adapted from Bergström et al., 2009 with modifications.

The immobilization of water-soluble proteins proceeds in three steps (figure 14) (Hermanson *et al.*, 1992; Bergström *et al.*, 2009). First, the bare silica beads are silanized by (3-glycidyloxypropyl)-trimethoxysilane. The purpose of silanization is to cover the reactive surface of silanol functionalities, to introduce diol groups to the surface and to add a spacer arm to provide more flexibility for the coupling. The second step is a cleavage and oxidation of the adjacent diol groups by periodic acid, resulting in aldehyde groups on the silica particle. From the second step and onwards, reactions are carried out with the silica particles either in a batch (**paper I**) or pre-packed into HPLC columns (**papers II–IV**). The last step is coupling the protein on the aldehyde silica through protein residues containing primary amines, such as lysine residues. The immobilization reaction is reductive amination that forms a Schiff base as an intermediate, which is then reduced by sodium cyanoborohydride.

Determination of the total binding site in a column (R_{tot} or B_{tot})

The most reliable method to determine the total active site (R_{tot}) of an affinity column in WAC is using frontal chromatography (Kasai *et al.*, 1986). It is carried out by injecting large volumes of analyte solutions into the immobilized target column to achieve break-through curves. The break-through volume is considered as the volume necessary to saturate the binding sites in the column at a specific concentration, which then can be used to calculate the corresponding number of moles needed to saturate the column. By performing many injections of the analyte that cover a wide range of concentrations, the obtained number of moles can be plotted versus the analyte concentration and simulated into a saturation binding curve to extract R_{tot} and K_D values (figure 15).

Papers I and **III** estimated R_{tot} by frontal chromatography. In **paper IV**, R_{tot} of thrombin was estimated by comparing the retention time of known binders (3-aminobenzamidine, 3-ABA; 4-aminobenzamidine, 4-ABA; and benzamidine, BZA) on a thrombin column with the retention time of those substances on another column with known R_{tot}, which was previously determined by frontal chromatography. The inclusion of reference binders (internal standards) to calibrate for a loss in R_{tot} during a screening campaign is important to achieve more accurate affinity determinations. The R_{tot} values on the affinity columns in **paper II** were assumed to be 50 % of the immobilized protein content.

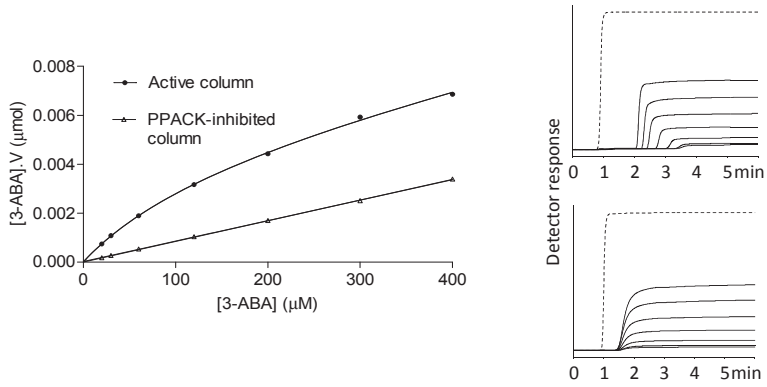

Figure 15. Frontal affinity chromatography data of 3-ABA from an active and an inhibited thrombin capillary column. Data were fitted into the Langmuir isotherm by GraphPad Prism (v.5.02, GraphPad Software, La Jolla, CA, USA). Corresponding frontal chromatograms on an active (upper) and a PPACK-inhibited (lower) column are depicted on the right. The broken lines in chromatograms are signals from the void marker DMSO. The figure is from ***paper III****.*

8.2. WAC platforms

The affinity separation platforms were either based on a 24-channel cartridge for parallel analysis (**paper I**) or single HPLAC columns (**papers I–IV**). The cartridge platform comprised 24 parallel capillary columns (dimension of each channel was 80 mm×0.5 mm id.; total volume 15.7 µL) that functioned together with an eight-needle autosampler and 24 separate individual UV detectors (figure 16) (Brios µHPLC cartridge, Nanostream, Pasadena, CA, USA). Injections were carried out as three sets of eight samples into the cartridge. The HPLAC column on the other platform was either a micro column (35 mm×2.1 mm id.) or a capillary column (35 mm×0.5 mm id.) that was connected to an MS detector to form a WAC-MS system that allowed mixture analyses. The configuration of the WAC-MS was identical to a standard analytical LC-MS system. In MS settings for screening, selected ion monitoring (SIM) mode in detection was used, assuming that the molecular ions $[M+H]^+$ and $[M-H]^-$ were the dominant products of positive and negative ionizations. To simplify the data analysis, the possible adducts as well as the fragmentations and the less abundant isotopes were ignored. An UV detector was also connected to serve as a system control. When screening singletons in the physiological buffer (10 mM sodium phosphate and 150 mM sodium chloride, pH 7.4; PBS), the MS was disconnected and only the UV detector was used.

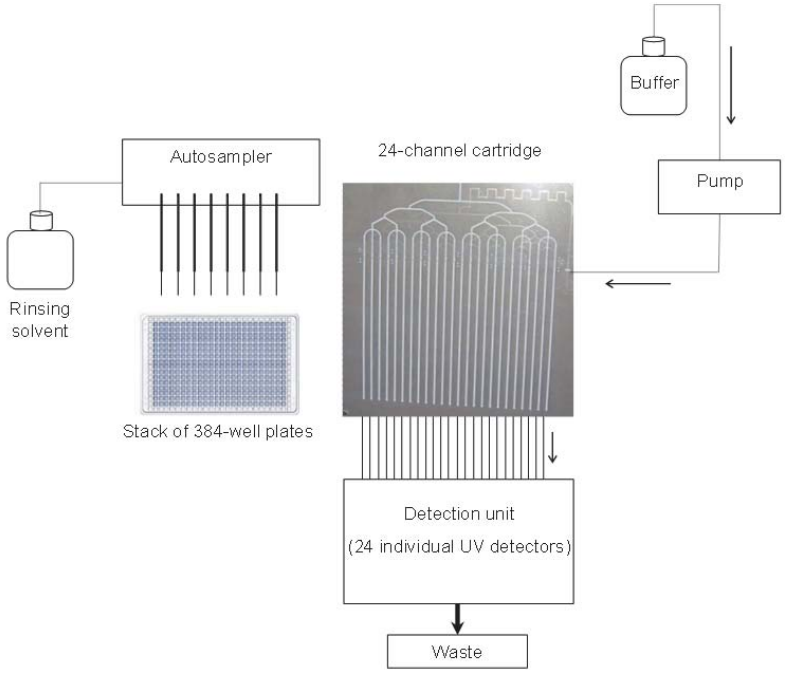

Figure 16. Schematic diagram of the Nanostream platform including a 24-channel cartridge that served as parallel affinity capillaries and a 24-detector unit. The figure is from ***paper I*** *with permission from John Wiley & Sons.*

In screening by WAC-MS, the mobile phase was ammonium acetate at a concentration of 10–20 mM, pH 6.8–7.0. Samples were often re-run in PBS to confirm affinities under physiologically mimicked conditions.

8.3. Computational methods

In **paper III**, the hit and reference substances were docked into a thrombin crystal to postulate binding poses. In **paper IV**, the fragment library was enriched by docking on two thrombin crystals. The docking scores on the two different thrombin structures were averaged and ranked; the highest ranking fragments were then selected. The computational tools were Glide for docking and other assistant modules from the Schrödinger Suite software.

9. RESULTS AND DISCUSSION

9.1. Overview of papers

Paper I describes a weak affinity platform that employs parallel capillaries for high-throughput analysis. The capillaries were combined with an eight-needle autosampler and parallel UV detector units. This paper demonstrates that WAC is robust and has potential to achieve high throughput by column parallelization. The immobilized protein model was BSA.

Paper II reports on a pilot screening study of a collection of amidines and similar structures on thrombin and trypsin columns. The protein targets were *in situ* immobilized on affinity columns. The column (35 mm×2.1 mm id.) was connected to a normal analytical HPLC and run at a flow rate of 0.2 mL/min. This paper is the first report on the use of WAC-MS in the analysis of sample mixtures. A good correlation of retention in a MS-compatible mobile phase, such as ammonium acetate, and the physiologically mimicked PBS was observed.

Paper III is the first report of how WAC-MS can be applied for high-throughput fragment screening. The capillary column (35 mm×0.5 mm id.) was used for the first time to reduce target consumption. The subjects for screening were 590 fragments from a commercial library (TimTec). The studied target protein was thrombin. The main findings are that WAC can perform screening at high throughput (3000–4000 compounds in 24 h) with low protein and sample consumptions. However, this paper raises questions about the effect of DMSO on detection and about the presence of false positives by off-active site interactions. Evaluations and solutions for these concerns are proposed in the paper. This paper plays an important role for this thesis, as it includes many relevant results and fulfills the main aim.

Paper IV demonstrates the capability of WAC to distinguish enantiomers and diastereomers in stereoisomeric mixtures by specific affinity differences from

the target. In addition, a procedure to elute tight binders without damaging the immobilized thrombin is developed. This paper confirms the major benefit of WAC to analyze mixtures of closely related compounds, such as stereoisomers.

9.2. Characteristics of WAC for fragment screening

Throughput

A feasible approach to increase the throughput of WAC is parallelization. One way to speed up analysis is by using parallel columns, for example, in the format of a 24-channel cartridge (**paper I**) (see figure 16). The Nanostream platform used in **paper I** is mainly applicable for UV detection, which can be manufactured as small detectors incorporated into a compact detection unit. This platform is able to inject a triplicate of eight samples at a time. Another way to speed up screening is parallelization by analyzing multi-compound mixtures and detecting them by MS (**papers II-IV**). This approach allows identification of the compounds in the mixture, which is not possible when using UV detection.

In the parallel capillary platform (**paper I**), BSA is used as the immobilized target protein. With a running cycle of 20 min, the system can reach a throughput of about 1700 samples in 24 h with satisfactory reproducibility. The coefficient of variation (CV) of 1800 runs of adjusted retention times and peak areas are less than 2 % and 6 %, respectively. **Papers II–IV** demonstrate another parallelization by analyzing mixtures with WAC-MS, as illustrated with immobilized thrombin, which reach higher throughput (figure 17). In **paper III**, mixtures containing up to 65 fragments are screened, which provide a throughput of about 3000–3500 compounds in 24 hours. If needed, the MS detector is capable of detecting many more compounds in a mixture. The weakness of WAC-MS, however, is the demand of using volatile mobile phases, which are typically non-physiological. If necessary, fragment hits can be confirmed in PBS by using e.g. UV detection, but that will slow down the screening. Improvements may be possible by post-column desalting of the mobile phase to allow running WAC-MS in PBS.

The bottle-neck of WAC-MS, however, is the data processing time. As there is no established software to handle automatically the chromatographic data in the manner required by WAC-MS for fragment screening, the manual data analysis consumes extensive amounts of time. This is obviously a point for improvement of WAC-MS in the future.

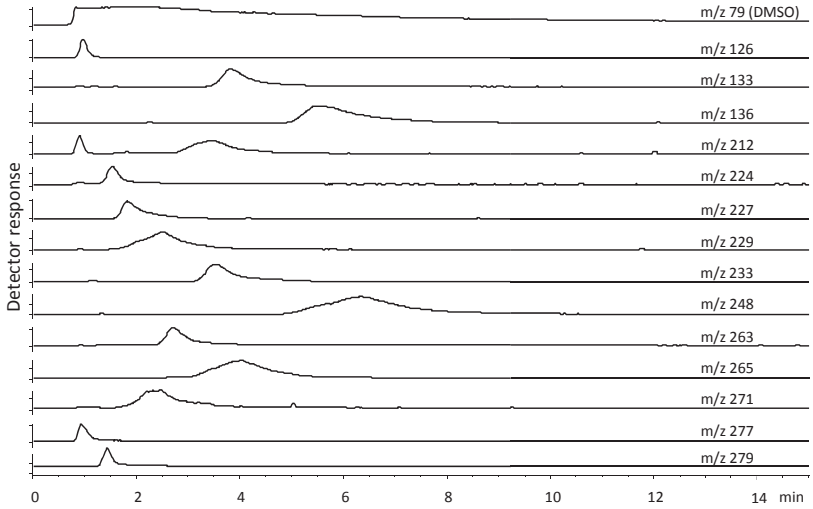

Figure 17. Extracted ion chromatograms of 15 compounds including DMSO in a mixture analyzed by an active thrombin column. The figure is from ***paper III****.*

Consumption of material

In the paralleled-capillary cartridge, the amount of protein needed for immobilization is less than 5 mg for 24 capillary columns. The injected amount of each analyte is in the nanogram range. In the WAC-MS system, when capillary columns (35 mm×0.5 mm id.) are used, consumption is less than 1 mg of the target protein and less than nanograms of the sample.

The affinity column has a high stability, which reduces the total target consumption in a screening campaign. The immobilized protein was for example still active after more than 60 days of storage; and lost only 17 % activity after more than 100 injections during 194 h of operation (**paper IV**). Part of the high stability is probably accounted for the immobilization process that stabilizes the protein and prevents autolysis degradation of the thrombin. The target consumption in WAC is relatively low as compared to that in NMR or X-ray crystallography, but a little higher than that in SPR. It is within reach in further developments to apply WAC on a miniaturized chip-based chromatography platform, where the amount of the target protein can be significantly reduced to possibly nanograms.

Operational platforms

The 24-capillary cartridge platform from Nanostream is no longer available on the market, and it remains to be seen if other parallel chromatography units will be commercialized and available for screening purposes. On the other hand, LC-MS can be considered as a commodity in analytical laboratories and is widely used in academic and industrial labs. As WAC-MS can easily be applied to standard LC-MS systems, it is expected that this technology will be widely accessible for use and evaluation. The operations of both systems are equally simple, and produce similar robustness and repeatability due to that both being based on HPLC.

Off-active site interactions

A key feature of screening with WAC-MS is to estimate any non-specific interactions, since it relates to off-active site binding to the target or to the matrix as discussed in chapter 6. It is interesting to note that a considerable number of compounds retained by WAC turned out as non-specific binders (**papers III** and **IV**). This is possibly explained by their small size facilitating their transient binding to other spots outside the active site of the protein target.

The off-site binding fragments can be of interest for identifying allosteric binding compounds, but might be a challenge for the screening when an active-site binder is the aim. As discussed in chapter 6, this problem can be solved by using an active-site modified protein in a complementary screening. One observation when using an inhibitor to block the active site is that the blocking efficiency varies depending on the type of inhibitor. In this thesis, the thrombin is blocked either covalently by phenylmethylsulfonyl fluoride (PMSF) or D-phenylalanyl-L-prolyl-L-arginine chloromethyl ketone (PPACK) (figure 18), or non-covalently by the high affinity inhibitor melagatran (see figure 11). Although PMSF is a covalent inhibitor, its blocking capacity is less efficient than the non-covalent inhibitor melagatran (**paper II**). This may be due to the small size of PMSF and therefore some small molecules may be able to squeeze into the S1 binding pocket, or interact with other pockets in the active site. On the contrary, melagatran and PPACK are substrate mimicking molecules equipped with amidine or arginine moieties, which can anchor them to the bottom of the S1 binding pocket of the thrombin. The remaining parts of the inhibiting molecule would spread along and occupy other pockets, eliminating the possibilities of binding to the active site for other molecules. As a result, melagatran and PPACK blocked the active site more efficiently than PMSF did. It is therefore important to select a well-functioning inhibitor to reduce the occurrence of false negatives in the screening.

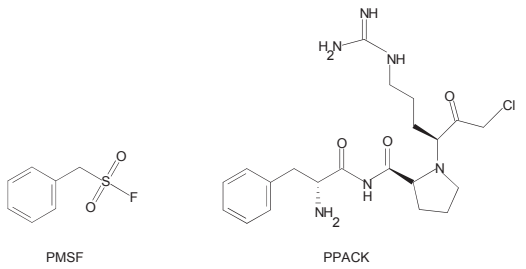

Figure 18. PMSF and PPACK: Two thrombin covalent inhibitors.

Affinity determination and ranking

The WAC-MS method is useful to determine the affinity of a hit. For example in **paper III**, screening identified one thrombin hit that has an amidine moiety in the structure, which is an established binding anchor for serine proteases. The calculated K_D value is 142 μM (LE = 0.53). As shown by the LE value, the binding of the hit, although rather weak, is efficient because of the small molecular size. The accuracy of finding a hit by WAC is also demonstrated in **papers II** and **IV**, which show that K_D-values by WAC agree well with IC_{50} from enzymatic assays. Although the comparisons are performed on only small sets of closely related compounds and therefore should be taken with caution, this information encourages the use of WAC for larger scale of screening.

Table 1. Affinity ranges measured by WAC in comparison to other fragment screening methods.

Screening method	Affinity range (K_D) (M)	References
Ligand-observed NMR	10^{-2}–10^{-8}	Mayer & Meyer, 1999; Neumann *et al.*, 2007; Moumné *et al.*, 2012
SPR	10^{-3}–10^{-11}	Myszka *et al.*, 1998; Hubbard & Murray, 2011
WAC	10^{-2}–10^{-6}	Bergström *et al.*, 2012 and **paper IV**

WAC can detect affinities in the range from 1 μM to 10 mM (Bergström *et al.*, 2012). This is comparable with other fragment screening methods, such as ligand-observed NMR and SPR (table 1).

In our experiments, the robustness of the system is estimated by triplicate injections of 14 compounds on both PPACK-inhibited thrombin and blank diol silica columns. The average coefficient of variation (CV) of the adjusted retention time (t'_R) is 14 %. This information may provide a rough estimation of possible random error in the affinity determination by WAC.

Screening of stereoisomers

Paper IV focuses on evaluating the ability of WAC to screen and separate stereoisomers in a fragment library and in synthesis mixtures of small molecules used in early SAR studies. WAC is able to identify individual stereoisomers that have affinities differing by a factor of three times or more in the sub-milimolar range. Figure 19 illustrates the separations of enantiomers in a racemic mixture by a thrombin affinity column. The affinities (K_D) of the enantiomers are 1.07 mM and 0.16 mM, respectively (**paper IV**).

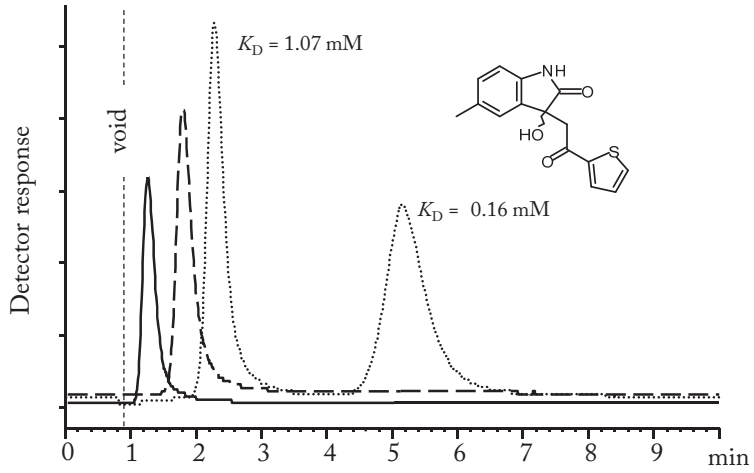

Figure 19. Overlaid chromatograms of the enantiomers ST058742 from the TimTec fragment library on a diol silica column (solid line), a PPACK-inhibited thrombin column (dashed line) and an active thrombin column (dotted line). The void is marked by DMSO (chromatogram is not shown). The figure is adapted from ***paper IV****.*

The advantage of WAC is that it is an affinity-based separation method; hence the readout is directly related to the affinity. Stereoisomeric molecules, such as enantiomers, are easily identified individually by WAC as long as they

are recognized differently by the target and thereby express differences in affinity. It is seen that a change in the configuration of one stereoisomeric center in a small molecule can improve the binding strength more than 50 times (**paper IV**). Stereoisomeric screening of fragments is a novel feature of WAC-MS, and it may encourage researchers to pay more attention to the role of stereoisomeric configurations of compounds early in drug development.

Kinetic determination

In drug discovery, the suitability of a compound to be developed into a drug candidate does not only rely on its affinity of binding to the target. Affinity merely reflects the equilibrium state, and there is increasing evidence that the kinetics of an interaction (on-rate (k_{on}) and off-rate (k_{off})) are equally important, as discussed in chapter 5 (Copeland *et al.*, 2006; Swinney, 2009). There are several methods to estimate kinetics of an interaction, including SPR, but also affinity chromatography offers possibilities (Schiel & Hage, 2009). Here the band broadening method is most suitable for compounds with weak to moderate affinities, and therefore it should be suitable for WAC. There are two sub-types of band broadening methods for kinetic determination: plate height (Loun & Hage, 1996) and peak profile (Talbert *et al.*, 2002) approaches. The peak profile method is more direct for estimating k_{on} and k_{off} values and it was evaluated in a preliminary study.

An important consideration when carrying out kinetic measurements by WAC is to operate at suitable flow rates. The van Deemter equation expresses the plate height as in equation 12 (van Deemter *et al.*, 1956):

$$H = A+(B/u)+(C\times u) \qquad \text{(Equation 12)}$$

where H is the plate height, A is the Eddy diffusion parameter, B is the longitudinal diffusion coefficient, C is related to the mass transfer of the analyte between the mobile and the stationary phase, and u is the linear mobile phase velocity. The "C term" is due to the mass transfer including the binding of analytes to the immobilized target. At high flow rates the "C term" will dominate the total H, and therefore the system is considered to be under kinetic control. For that reason, high flow rates are needed to evaluate properly the rate constants in WAC.

Another parameter that is used to evaluate if the chromatographic conditions can be used to calculate kinetics is the critical ratio $\eta = (\sigma_R/t_R)/(\sigma_M/t_M)$ (Talbert *et al.*, 2002), in which σ_R and σ_M are variances (peak width) of retained and non-retained peaks, respectively. The critical ratio reflects if chromatography is under equilibrium or kinetic control. When $\eta \leq 1$, chromatography is close to equilibrium conditions, and information on the

kinetics of binding is hard to obtain. When $\eta > 1$, kinetic effects dominate the band broadening, and therefore kinetic calculations should be possible.

In the preliminary attempts to determine if the kinetic rate constants can be deduced by WAC, capillary columns with thrombin and PPACK-inhibited thrombin were evaluated to determine kinetic constants for weakly-binding fragments (data not shown). However, as the critical ratio (η) was less than 1 even at the highest allowable flow rate, no meaningful data was obtained due to overshadowing of binding kinetics by other band-spreading effects. A solution for the problem may be to seek more efficient support matrices, for example, by using smaller silica particles with a solid core that can be run under high flow rates to reduce mass transfer effects.

9.3. Perspectives on WAC-MS for fragment screening

Advantages and challenges

The major advantageous feature of WAC-MS is its ability to screen mixtures. This characteristic allows WAC to screen fragment libraries with a high throughput. Applications may also be found in screening crude synthesis mixtures, in which WAC can save time and labor due to its lower demand of purification. WAC can also estimate the affinity of individual enantiomers in racemic mixtures, where individual purification of each enantiomer may be difficult and costly. ITC could possibly detect affinity differences among stereoisomers (Fokkens & Klebe, 2006), but WAC is more powerful because it can separate the substances and measure the affinity of each individual compound in one single experiment.

Another advantage of WAC, in contrast to other screening techniques such as SPR, NMR, ITC or functional assays, is that DMSO, which is included in almost all libraries, has a minor effect on the immobilized protein even at high concentrations. The reason is that DMSO is non-retained and quickly moves out of the column. The same reasoning applies for other organic solvents that may be used for sample storage, such as methanol.

The challenges in screening by WAC-MS are mainly connected to the detection of fragments by the MS. For example, there are always compounds that do not ionize well in the MS. Furthermore, the presence of DMSO and other compounds in a mixture can suppress the ionization of compounds with a weaker ionization ability to be protonated (positive mode) or de-protonated (negative mode), as discussed in section 3.2. Diluting the sample to decrease the DMSO concentration in combination with optimizing the MS

instrumentation can improve the detection to some extent, but some problems in the detection of some fragments are probably unavoidable. Therefore, it may be useful to test and exclude compounds that are not seen by the MS before building up a library for WAC-MS screening.

Another challenge is that tight binders can escape detection if they elute very late. The detection of tight binders will prolong the analysis time in WAC, and a tight binder peak can be difficult to distinguish from the baseline due to late eluting peaks being very wide. To deal with extreme tight binders which may be mistaken for non-detectable compounds, a simple elution procedure of running a small volume of acid through the column is reported in **paper IV**. The procedure has only been evaluated on thrombin columns, and it remains to be seen how the method works with other protein targets.

Design of fragment libraries

Most compound libraries are stored in DMSO. This solvent interferes with the MS detection as observed in **papers III** and **IV**. To reduce the DMSO influence, dilution of the final injection mixture with water to reach a DMSO concentration of less than 5 % is desired. This means that a low concentration of fragments will generate an even lower concentration in the injected solution, which challenges the MS detection. Consequently, the original fragment concentration in the library is a factor that limits the screening throughput. A high concentration of fragments in the original library is therefore favorable in WAC, as well as in other fragment screening methodologies.

The MW diversity of the library also influences the potential number of compounds that can be put into one mixture, and subsequently it affects the throughput. If the MWs of compounds in the library have a narrow range, many fragments will appear as identical masses, especially when using a quadrupole MS having a resolution of about 1 amu as the detection mean. Using a high resolution MS instrument could aid in the identification to some extent, but ideally one mass should encode for one compound in a mixture. If several fragments with identical MW values are included in a mixture, all fragments having the same MW as a potential hit have to be re-analyzed as singletons in a secondary screening, which may increase the workload.

When designing mixtures for screening, attention should also be paid to compounds containing abundant isotopes, such as chlorine, bromine and sulfur. The additional *m/z* values of these compounds have to be considered to avoid overlapping *m/z* values in the fragment mixtures. Chlorine has two stable isotopes ^{35}Cl and ^{37}Cl in a 3:1 ratio; bromine includes almost equal amounts of ^{79}Br and ^{81}Br, and sulfur comprises ^{32}S, ^{33}S and ^{34}S in 100:1:4

proportions. Besides these abundant isotopes, common atoms in organic compounds, such as carbon, nitrogen, oxygen and hydrogen, also have stable isotopes,^{13}C, ^{15}N, ^{18}O and ^{2}H, but at much lower abundance (about 1 % or less) (Skoog *et al.*, 2007). Problems from these latter isotopes are therefore only pronounced when dealing with large compounds that contain many atoms.

A proposed WAC-MS procedure for fragment screening

As in other screening methods, false positives may arise in WAC screening due to non-specific interactions. Consequently, complementary screening using a target with an inhibited active site is necessary, as discussed in chapter 6. There are two alternatives for performing the complementary (secondary) screening (figure 20). Option A is to pick out the most promising fragments in the initial (primary) screening and to run these potential hits as singletons. Promising fragments can be compounds with the longest retention. The secondary screening is performed by running singletons in PBS, a physiological buffer. MS detection is not compatible with PBS, and therefore other detection methods have to be implemented (e.g. UV detection). This strategy is in principle applied in **paper III**. Option B is to run the whole library twice, first on the active protein column and then on the inhibited target using mixtures and MS detection in both cases. Each option has its own strengths and weaknesses.

Option A, although is rather elaborate, should give more reliable results because of the physiological buffer (PBS) used in the secondary step. Furthermore, running singletons avoids the competitive effect that may arise when many compounds in a mixture compete for the same binding site. However, if Option A only selects the most retained fragments in the primary screening as in **paper III**, it may miss compounds that bind weakly but specifically and efficiently to the target. A proposed solution is to extend the secondary screening to include all fragments that have primary LE greater than 0.3 $kcalmol^{-1}$ per atom (figure 20). However, the strategy can be laborious as the new cut-off may add more fragments to the time-consuming singleton analyses.

In Option B, the benefit is that it, in theory, would provide results faster and with less work compared to Option A. However, there are concerns about the analysis of mixtures and using MS detection in both primary and secondary screenings. The presence of a high number of compounds in the mixtures may complicate the readout and subsequently prolong the analysis time of data. Competitive binding may occur, although this is not a great issue when the screening purpose is only to rank and select the best fragments. In addition, retention may be different in the MS compatible buffer compared to PBS.

However, the obtained hits, which in most cases are few and manageable, can be confirmed by analyzing them as singletons in PBS. The data in **paper II** indicate that binding in a MS-compatible buffer (ammonium acetate 10 mM, pH 7.0) correlates well with binding in PBS. However, as the number of tested compounds is small and lacks diversity, the result should be taken with caution.

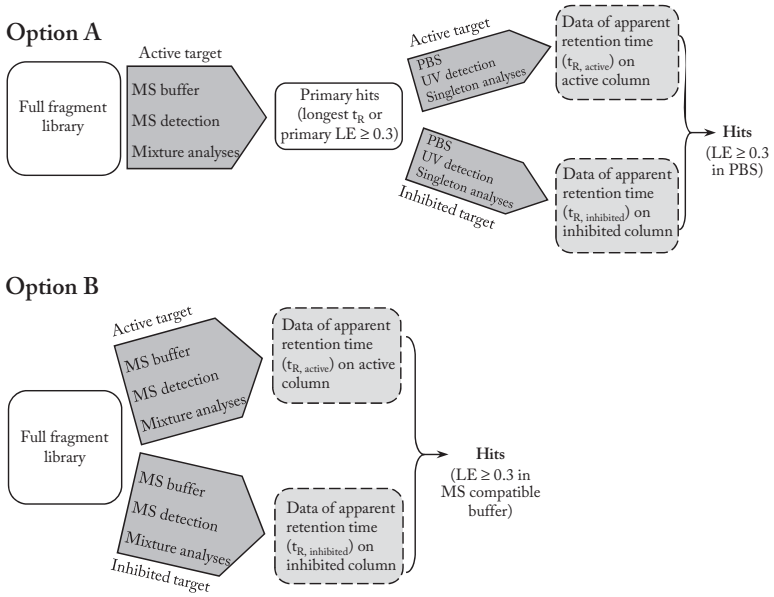

Figure 20. Proposed WAC-MS screening procedure. Arrows indicate screening steps. Option B may need another screening round to confirm the hits in PBS on both active and inhibited targets.

The prioritization of one option over the other depends on the screening library (number of compounds, concentrations, solvent, etc.), the screening platform, target availability and the preference of the operator.

Paper IV used *in silico* screening (Glide docking) to select enantiomeric fragments of the TimTec library that should have a high probability to be a hit. However, this procedure resulted in only one hit from the 30 highest ranked fragments, which shows that the hits from virtual screening are not always transferred into hits in an experiment.

10. CONCLUSIONS AND FUTURE PERSPECTIVES

In drug discovery, FBDD is considered to be a promising alternative to conventional HTS. As it is a relatively new strategy, approved drugs from the strategy are still limited in numbers. In the years to come, the power as well as the shortcomings of FBDD will be fully recognized. This thesis focuses on an introduction of WAC as a novel technology for fragment screening and for early stage hit optimization in FBDD.

WAC is realized for fragment screening by two chromatography platforms. One is a platform equipped with parallel capillaries and UV detectors (Nanostream), and the other is a standard LC-MS system employing capillary columns (WAC-MS or affinity LC-MS). As WAC is based on HPLC, it inherits the accuracy and robustness of an automated HPLC system. Both WAC chromatography systems in this study show good applicability for screening with characterization of analytes binding to the protein target at high throughput and low consumption of the target protein and the fragment sample. The K_D values obtained from WAC correlate well with the IC_{50} values of enzymatic assays.

The affinity LC-MS system possesses several distinct advantages. The most prominent aspect is that it comprises a standard LC-MS system, which can be utilized for other purposes when not used for screening. Furthermore, WAC-MS can analyze mixtures of analytes which lead to considerably higher throughput, as the MS can distinguish compounds according to their MW. Since mixture analysis is the key to high throughput in WAC-MS, it places some demands on the fragment library and the design of the mixture. Another key feature of screening mixtures is the ability to screen stereoisomers, such as racemates without further purification. However, physiological buffers like PBS are not compatible with MS detection, and this fact may introduce

deviations in the affinity determination with WAC-MS compared to other methods.

Off-site interactions, which can lead to false positives in the primary fragment screening, are an inherent feature of many fragment screening methods, including WAC. Fortunately, there are a few approaches available in WAC and other methods to quantify the binding to the target, for example, by doing active site modifications or competition experiments. Therefore orthogonal screening approaches should be frequently applied to assure that accurate hits are obtained. WAC-MS also proves to be useful in early drug development when fragment hits are identified and are subject to evolution. The affinities of evolved compounds in this stage can still be detected by WAC-MS (sub-micromolar in K_D). Again, the mixture resolving ability of WAC-MS may help to pick out interesting compounds from mixtures, including synthetic and natural mixtures, making purification efforts obsolete and saving considerable resources in man-power and equipment.

Within the scope of this thesis, WAC is only evaluated on soluble "easy" protein targets of serine proteases and serum albumin. It is indeed important to apply WAC on other targets to realize the potential of the technique more fully. Since membrane proteins represent a majority of drug targets, but are considered to be "difficult" to work with, it would be valuable to see if WAC can be applied to these "tricky" proteins for fragment screening. Knowledge of evaluating membrane proteins can be extracted from many attempts to use membrane proteins by NMR (Früh *et al.*, 2010), SPR (Maynard *et al.*, 2009; Rich *et al.*, 2010; Seeger *et al.*, 2012) as well as chromatography (Moaddel & Wainer, 2006). Another area for development of the WAC-technology is to optimize the data acquisition software to save time with the data processing and thereby increase the throughput. Finally, the immobilized target can be further scaled down to chip size (nanoliter volumes) to reduce the protein and the sample consumption to nanograms or less.

Although WAC shows promising results for screening on a few targets, it remains to be seen how it can be used for operations on various targets, with different molecular libraries and in different environments. Of special interest is to see its potential to achieve information on the kinetics of binding analytes to the target. Validation of WAC with other established screening methods is also necessary to see how well they correlate.

In conclusion, this thesis demonstrates that WAC is a potential novel tool in drug discovery for fragment screening and in early stages of hit development. Apart from being a high-throughput technology, it provides a simple way to achieve information on the affinity of analytes present in complex matrices.

11. ACKNOWLEDGEMENTS

My four years working on this thesis have given me the chance to meet and receive support from many people. Thank you all! I am able to name just a few here.

First of all, I want to thank Prof. Sten Ohlson for bringing me to Sweden and for his invaluable advice and mentorship during my research. Thank you and your wife, Karin, for receiving me as your guest in the summer and Christmas of my first year. Your non-academic advice was also very precious to me.

I would also like to show my greatest gratitude to Maria Bergström for her incredible patience and excellent guidance, and especially for always being a shoulder for me to lean on on many occasions. Thank you for helping me with my new surroundings, explaining Swedish culture to me and taking me to many interesting places.

I would like to send special thanks to Prof. Roland Isaksson, who has always been a good source of knowledge.

Thanks to Tomas Fex at Astra&Zeneca for very fruitful collaborations. You contributed to important parts of my thesis. Also thanks to Susanne S. at Chalmers University for the synthesis of the compounds in paper IV.

Thanks to Prof. Carl-Fredrik Mandenius and Gunnar Bergström at Linköping University for their collaboration with the SPR study. Gunnar, I am grateful that you and your family have been unbelievably open and friendly toward me.

Special thanks to Sten, Maria, Tomas, Anneli and Roland for critically reviewing my thesis. Thanks to Maria for helping with the summary in Swedish. Thanks to Susan Canali for the proofreading.

Thanks to Elinor, Ulf, Lennart H. and especially Anneli for being so friendly and supportive. Thank you, Elinor, for "rescuing" me at the airport.

Thanks to Professors Michael Lindberg and Sven Tågerud for organizing the BOB seminars, which gave me the opportunity to practice my presentation skills.

Thanks to all the administrators, especially Lena H., Ymke M., Leif E. and Therese G., for all their help with paperwork. I am deeply grateful to Ymke for her dedication in helping me even outside the office hours when I was stranded in San Francisco. Thanks to Prof. Ian Nicholls for your kind support in every aspect of administration.

Stefan and Bosse are thanked for always being available with a helping hand. I am very grateful for the "scenic view" writing room that you arranged for me! Thanks to Ann, Birgitta and Ing-Marie for keeping Norrgård clean and tidy.

To my friends and colleagues in Norrgård: Boel, BIM, Kjell, Kerstin S., Annette, Marlene, Håkan, Jesper, Lasse, Stina C., Malin, Dieter, Per, Maria A., Kerstin G., Camilla, Johanna, Nina, Mikael A., Anna G., and many more… thank you all for the friendly atmosphere that you contributed. Thanks for trying to speak English or translate the conversations so that I could understand the coffee break talks.

To the "foreigners" in Norrgård: Carlo, Ran, Hongzhen, Palani and Priya, Karthik and Sumathi, Selva, Saroj, Baskar, Subu, Ke, Shan, Junli, Pacho and Karolina, Vahid and Tahereh, Shirin, Lorena R., etc., I am grateful for your friendship. I have the feeling that I have friends everywhere in the world now!

红珍，我在卡尔马的第一个朋友。非常感谢你一如既往的帮助和支持

To my friends in the harbor: Lee, Nayani and Sarath, Devi, Asim, Norin and Muee, Anders, Fabio, Qiaoyu, Xiaofen, Zhen, Sarfraz, Frédéric and more… Thank you very much for your friendship!

หลี,ขอบคุณสำหรับมิตรภาพและกำลังใจที่มีให้กันเสมอฉันหวังว่าสิ่งนี้จะอยู่ในทุกๆที่ที่เราไป

නයනී, සරත් ඔබ හා එක්ව ගෙවූ කාලය සොඳුරුයි... ඒ ආදරයට හද පිරි තුති. හැමදාමත් ඔබ මා මතකයේ රැඳෙනු ඇත...

Thanks to Eva and Mats Å. for always being my friends. I will never forget the help that Eva offered me when I was not used to the lab yet, and the atmosphere at the Christmas market where you both took me to.

I appreciate very much the friendship outside the university from the Phan family (chi Hue, anh Hung, Hilda and Henning), Ketsuda and Anders, Taneda and our Thai friends. Thank you for opening your homes to me and providing me with the feeling of being home, which was very comforting. Thanks to Bodil, Johan and friends for great Midsummer celebrations and many other cultural activities that helped me understand more about Swedes.

Lastly, I want to thank my family for their unconditional love, support, and understanding throughout the years.

12. REFERENCES

Abad-Zapatero, C. (2007) A Sorcerer's apprentice and the rule of five: from rule-of-thumb to commandment and beyond. *Drug Discovery Today*, **12**, 995-997.

Albert, J.S. (2010) Fragment-based lead discovery. In: *Lead Generation Approaches in Drug Discovery* eds. Z. Rankovic and R. Morphy), pp. 106 - 134. John Wiley & Sons, NJ, USA.

Andrews, P.R., Craik, D.J. & Martin, J.L. (1984) Functional-group contributions to drug receptor interactions. *Journal of Medicinal Chemistry*, **27**, 1648-1657.

Annis, D.A., Nazef, N., Chuang, C.C., Scott, M.P. & Nash, H.M. (2004) A general technique to rank protein-ligand binding affinities and determine allosteric versus direct binding site competition in compound mixtures. *Journal of the American Chemical Society*, **126**, 15495-15503.

AstraZeneca (2006) *AstraZeneca decides to withdraw Exanta™*. Available at: http://www.astrazeneca.com/Media/Press-releases/Article/20060214--AstraZeneca-Decides-to-Withdraw-Exanta (accessed 01st March 2013).

Atkins, P. & de Paula, J. (2006) *Physical chemistry for the life sciences*. Oxford University Press, Oxford, UK.

Austin, C., Pettit, S.N., Magnolo, S.K., Sanvoisin, J., Chen, W.J., Wood, S.P., Freeman, L.D., Pengelly, R.J. & Hughes, D.E. (2012) Fragment screening using capillary electrophoresis (CEfrag) for hit identification of heat shock protein 90 atpase inhibitors. *Journal of Biomolecular Screening*, **17**, 868-876.

Ayyar, B.V., Arora, S., Murphy, C. & O'Kennedy, R. (2012) Affinity chromatography as a tool for antibody purification. *Methods*, **56**, 116-129.

Baker, M. (2013) Fragment-based lead discovery grows up. *Nature Reviews Drug Discovery*, **1**, 5-7.

Bergström, M. (2006) *Analytical affinity separations based on weak protein-carbonhydrate interactions*. PhD thesis, University of Kalmar, Sweden. ISBN: 91-89584-58-9.

Bergström, M., Lundblad, A., Påhlsson, P. & Ohlson, S. (1998) Use of weak monoclonal antibodies for affinity chromatography. *Journal of Molecular Recognition*, **11**, 110-113.

Bergström, M., Liu, S., Kiick Kristi, L. & Ohlson, S. (2009) Cholera toxin inhibitors studied with high-performance liquid affinity chromatography: A robust method to evaluate receptor-ligand interactions. *Chemical Biology & Drug Design*, **73**, 132-141.

Bergström, M., Åström, E., Påhlsson, P. & Ohlson, S. (2012) Elucidating the selectivity of recombinant forms of Aleuria aurantia lectin using weak affinity chromatography. *Journal of Chromatography B - Analytical Technologies in the Biomedical and Life Sciences*, **885**, 66-72.

Betz, U.A.K. (2005) How many genomics targets can a portfolio afford? *Drug Discovery Today*, **10**, 1057-1063.

Biacore (2008) Biacore sensor surface handbook BR-1005-71 Edition AB.

Bode, W. (2006) Structure and interaction modes of thrombin. *Blood Cells Molecules & Diseases*, **36**, 122-130.

Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S.R. & Hofsteenge, J. (1989) The refined 1.9 Å crystal structure of human alpha thrombin interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO (European Molecular Biology Organization) Journal*, **8**, 3467-3476.

Bohacek, R.S., McMartin, C. & Guida, W.C. (1996) The art and practice of structure-based drug design: A molecular modeling perspective. *Medicinal Research Reviews*, **16**, 3-50.

Boissel, J.P., Lebonniec, B., Rabiet, M.J., Labie, D. & Elion, J. (1984) Covalent structures of beta and gamma autolytic derivatives of human alpha-thrombin. *Journal of Biological Chemistry*, **259**, 5691-5697.

Bollag, G., Hirth, P., Tsai, J., Zhang, J.Z., Ibrahim, P.N., Cho, H.N., Spevak, W., Zhang, C., Zhang, Y., Habets, G., Burton, E., Wong, B., Tsang, G., West, B.L., Powell, B., Shellooe, R., Marimuthu, A., Nguyen, H., Zhang, K.Y.J., Artis, D.R., Schlessinger, J., Su, F., Higgins, B., Iyer, R., D'Andrea, K., Koehler, A., Stumm, M., Lin, P.S., Lee, R.J., Grippo, J., Puzanov, I., Kim, K.B., Ribas, A., McArthur, G.A., Sosman, J.A., Chapman, P.B., Flaherty, K.T., Xu, X.W., Nathanson, K.L. & Nolop, K. (2010) Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature*, **467**, 596-599.

Boyd, S.M., Turnbull, A.P. & Walse, B. (2012) Fragment library design considerations. *Wiley Interdisciplinary Reviews-Computational Molecular Science*, **2**, 868-885.

Brighton, T.A. (2004) The direct thrombin inhibitor melagatran/ximelagatran. *Medical Journal of Australia*, **181**, 432-437.

Bujacz, A. (2012) Structures of bovine, equine and leporine serum albumin. *Acta Crystallographica Section D-Biological Crystallography*, **68**, 1278-1289.

Butkowski, R.J., Elion, J., Downing, M.R. & Mann, K.G. (1977) Primary structure of human prethrombin-2 and alpha-thrombin. *Journal of Biological Chemistry*, **252**, 4942-4957.

Card, G.L., Blasdel, L., England, B.P., Zhang, C., Suzuki, Y., Gillette, S., Fong, D., Ibrahim, P.N., Artis, D.R., Bollag, G., Milburn, M.V., Kim, S.H., Schlessinger, J. & Zhang, K.Y.J. (2005) A family of phosphodiesterase inhibitors discovered by cocrystallography and scaffold-based drug design. *Nature Biotechnology*, **23**, 201-207.

Chan, N.W.C., Lewis, D.F., Rosner, P.J., Kelly, M.A. & Schriemer, D.C. (2003) Frontal affinity chromatography-mass spectrometry assay technology for multiple stages of drug. discovery: applications of a chromatographic biosensor. *Analytical Biochemistry*, **319**, 1-12.

Chen, I.J. & Hubbard, R.E. (2009) Lessons for fragment library design: analysis of output from multiple screening campaigns. *Journal of Computer-Aided Molecular Design*, **23**, 603-620.

Chen, Y. & Shoichet, B.K. (2009) Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nature Chemical Biology*, **5**, 358-364.

Chessari, G. & Woodhead, A.J. (2009) From fragment to clinical candidate-a historical perspective. *Drug Discovery Today*, **14**, 668-675.

Collins, F.S., Lander, E.S., Rogers, J., Waterston, R.H. & Int Human Genome Sequencing, C. (2004) Finishing the euchromatic sequence of the human genome. *Nature*, **431**, 931-945.

Congreve, M., Carr, R., Murray, C. & Jhoti, H. (2003) A rule of three for fragment-based lead discovery? *Drug Discovery Today*, **8**, 876-877.

Congreve, M., Chessari, G., Tisi, D. & Woodhead, A.J. (2008) Recent developments in fragment-based drug discovery. *Journal of Medicinal Chemistry*, **51**, 3661-3680.

Copeland, R.A., Pompliano, D.L. & Meek, T.D. (2006) Opinion - Drug-target residence time and its implications for lead optimization. *Nature Reviews Drug Discovery*, **5**, 730-739.

Cuatrecasas, P., Wilchek, M. & Anfinsen, C.B. (1968) Selective enzyme purification by affinity chromatography. *Proceedings of the National Academy of Sciences of the United States of America*, **61**, 636-643.

Dalvit, C. (2009) NMR methods in fragment screening: theory and a comparison with other biophysical techniques. *Drug Discovery Today*, **14**, 1051-1057.

Dalvit, C. & Vulpetti, A. (2012) Technical and practical aspects of ^{19}F NMR-based screening: toward sensitive high-throughput screening with rapid deconvolution. *Magnetic Resonance in Chemistry*, **50**, 592-597.

Davis, A.M., Teague, S.J. & Kleywegt, G.J. (2003) Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. *Angewandte Chemie International Edition*, **42**, 2718-2736.

Davis, A.M., St-Gallay, S.A. & Kleywegt, G.J. (2008) Limitations and lessons in the use of X-ray structural information in drug design. *Drug Discovery Today*, **13**, 831-841.

DePristo, M.A., de Bakker, P.I.W. & Blundell, T.L. (2004) Heterogeneity and inaccuracy in protein structures solved by X-ray crystallography. *Structure*, **12**, 831-838.

Di Cera, E., Guinto, E.R., Vindigni, A., Dang, Q.D., Ayala, Y.M., Wuyi, M. & Tulinsky, A. (1995) The Na^+ binding-site of thrombin. *Journal of Biological Chemistry*, **270**, 22089-22092.

DiMasi, J.A., Hansen, R.W. & Grabowski, H.G. (2003) The price of innovation: new estimates of drug development costs. *Journal of Health Economics*, **22**, 151-185.

Dole, M., Mack, L.L. & Hines, R.L. (1968) Molecular beams of macroions. *Journal of Chemical Physics*, **49**, 2240-2249.

Drews, J. (2000) Drug discovery: A historical perspective. *Science*, **287**, 1960-1964.

Dullweber, F., Stubbs, M.T., Musil, D., Stürzebecher, J. & Klebe, G. (2001) Factorising ligand affinity: A combined thermodynamic and crystallographic study of trypsin and thrombin inhibition. *Journal of Molecular Biology*, **313**, 593-614.

Dunham, W.H., Mullin, M. & Gingras, A.C. (2012) Affinity-purification coupled to mass spectrometry: Basic principles and strategies. *Proteomics*, **12**, 1576-1590.

Ecker, D.J. & Griffey, R.H. (1999) RNA as a small-molecule drug target: doubling the value of genomics. *Drug Discovery Today*, **4**, 420-429.

Edfeldt, F.N., Folmer, R.H. & Breeze, A.L. (2011) Fragment screening to predict druggability (ligandability) and lead discovery success. *Drug Discovery Today*, **16**, 284-7.

Edwards, A.M., Isserlin, R., Bader, G.D., Frye, S.V., Willson, T.M. & Yu, F.H. (2011) Too many roads not taken. *Nature*, **470**, 163-165.

Elinder, M., Geitmann, M., Gossas, T., Källblad, P., Winquist, J., Nordström, H., Hämäläinen, M. & Danielson, U.H. (2011) Experimental validation of a fragment library for lead discovery using spr biosensor technology. *Journal of Biomolecular Screening*, **16**, 15-25.

Enyedy, I.J. & Egan, W.J. (2008) Can we use docking and scoring for hit-to-lead optimization? *Journal of Computer-Aided Molecular Design*, **22**, 161-168.

Erlanson, D.A., McDowell, R.S. & O'Brien, T. (2004) Fragment-based drug discovery. *Journal of Medicinal Chemistry.*, **47**, 3463-3482.

Ertl, P., Rohde, B. & Selzer, P. (2000) Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *Journal of Medicinal Chemistry*, **43**, 3714-3717.

Fanali, G., di Masi, A., Trezza, V., Marino, M., Fasano, M. & Ascenzi, P. (2012) Human serum albumin: From bench to bedside. *Molecular Aspects of Medicine*, **33**, 209-290.

Fenton, J.W., Fasco, M.J., Stackrow, A.B., Aronson, D.L., Young, A.M. & Finlayson, J.S. (1977) Human thrombins - production, evaluation, and properties of alpha-thrombin. *Journal of Biological Chemistry*, **252**, 3587-3598.

Ferenczy, G.G. & Keserű, G.M. (2010) Thermodynamics guided lead discovery and optimization. *Drug Discovery Today*, **15**, 919-932.

Fink, T. & Reymond, J.-L. (2007) 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 discovery. *Journal of Chemical Information and Modeling*, **47**, 342-353.

Fokkens, J. & Klebe, G. (2006) A simple protocol to estimate differences in protein binding affinity for enantiomers without prior resolution of racemates. *Angewandte Chemie International Edition*, **45**, 985-989.

Früh, V., Zhou, Y.P., Chen, D., Loch, C., Eiso, A.B., Grinkova, Y.N., Verheij, H., Sligar, S.G., Bushweller, J.H. & Siegal, G. (2010) Application of fragment-based drug discovery to membrane proteins: Identification of ligands of the integral membrane enzyme DsbB. *Chemistry and Biology*, **17**, 881-891.

Ghose, A.K., Viswanadhan, V.N. & Wendoloski, J.J. (1998) Prediction of hydrophobic (lipophilic) properties of small organic molecules using fragmental methods: An analysis of ALOGP and CLOGP methods. *Journal of Physical Chemistry A*, **102**, 3762-3772.

Giannetti, A.M., Koch, B.D. & Browner, M.F. (2008) Surface plasmon resonance based assay for the detection and characterization of promiscuous inhibitors. *Journal of Medicinal Chemistry*, **51**, 574-580.

Hage, D.S., Anguizola, J.A., Bi, C., Li, R., Matsuda, R., Papastavros, E., Pfaunmiller, E., Vargas, J. & Zheng, X.W. (2012) Pharmaceutical and biomedical applications of affinity chromatography: Recent trends and developments. *Journal of Pharmaceutical and Biomedical Analysis*, **69**, 93-105.

Hajduk, P.J., Galloway, W.R. & Spring, D.R. (2011) A question of library design. *Nature*, **470**, 42-43.

Hajduk, P.J., Gerfin, T., Boehlen, J.M., Häberli, M., Marek, D. & Fesik, S.W. (1999) High-throughput nuclear magnetic resonance-based screening. *Journal of Medicinal Chemistry*, **42**, 2315-2317.

Hall, J.E. (2011) *Guyton and Hall textbook of medical physiology*, 12th edn. Saunders Elsevier, Philadelphia, PA, USA.

Hann, M.M., Leach, A.R. & Harper, G. (2001) Molecular complexity and its impact on the probability of finding leads for drug discovery. *Journal of Chemical Information and Computer Sciences*, **41**, 856-864.

Hartshorn, M.J., Murray, C.W., Cleasby, A., Frederickson, M., Tickle, I.J. & Jhoti, H. (2005) Fragment-based lead discovery using X-ray crystallography. *Journal of Medicinal Chemistry*, **48**, 403-413.

Hedstrom, L. (2002) Serine protease mechanism and specificity. *Chemical Reviews*, **102**, 4501-4523.

Hermanson, G.T., Mallia, A.K. & Smith, P.K. (1992) *Immobilized affinity ligand techniques*. Academic Press, San Diego, CA, USA.

Hilpert, K., Ackermann, J., Banner, D.W., Gast, A., Gubernator, K., Hadváry, P., Labler, L., Müller, K., Schmid, G., Tschopp, T.B. & Vandewaterbeemd, H. (1994) Design and synthesis of potent and highly selective thrombin inhibitors. *Journal of Medicinal Chemistry*, **37**, 3889-3901.

Hoelder, S., Clarke, P.A. & Workman, P. (2012) Discovery of small molecule cancer drugs: Successes, challenges and opportunities. *Molecular Oncology*, **6**, 155-176.

Hofstetter, O., Lindstrom, H. & Hofstetter, H. (2002) Direct resolution of enantiomers in high-performance immunoaffinity chromatography under isocratic conditions. *Analytical Chemistry*, **74**, 2119-2125.

Hopkins, A.L. & Groom, C.R. (2002) The druggable genome. *Nature Reviews Drug Discovery*, **1**, 727-730.

Hopkins, A.L., Groom, C.R. & Alex, A. (2004) Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today*, **9**, 430-431.

Hubbard, R.E. & Murray, J.B. (2011) Experiences in fragment-based lead discovery. *Fragment-Based Drug Design: Tools, Practical Approaches, and Examples* (ed. by L.C. Kuo), pp. 509-531. Elsevier Academic Press Inc, San Diego, CA, USA.

Hubble, J. (2001) Competition and intrinsic displacement effects in overloaded weak affinity chromatography. *Separation Science and Technology*, **36**, 541-554.

Hung, A.W., Ramek, A., Wang, Y.K., Kaya, T., Wilson, J.A., Clemons, P.A. & Young, D.W. (2011) Route to three-dimensional fragments using diversity-oriented synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 6799-6804.

Imming, P., Sinning, C. & Meyer, A. (2006) Drugs, their targets and the nature and number of drug targets. *Nature Reviews Drug Discovery*, **5**, 821-834.

Jencks, W.P. (1981) On the attribution and additivity of binding-energies. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences*, **78**, 4046-4050.

Jhoti, H., Cleasby, A., Verdonk, M. & Williams, G. (2007) Fragment-based screening using X-ray crystallography and NMR spectroscopy. *Current Opinion in Chemical Biology*, **11**, 485-493.

Jorgensen, W.L. (2012) Challenges for academic drug discovery. *Angewandte Chemie International Edition*, **51**, 11680-11684.

Kasai, K.I., Oda, Y., Nishikata, M. & Ishii, S.I. (1986) Frontal affinity chromatography: Theory for its application to studies on specific interactions of biomolecules. *Journal of Chromatography B: Biomedical Sciences and Applications*, **376**, 33-47.

Klages, J., Coles, M. & Kessler, H. (2006) NMR-based screening: a powerful tool in fragment-based drug discovery. *Molecular Biosystems*, **2**, 319-331.

Klebe, G. (2006) Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today*, **11**, 580-594.

Kobayashi, M., Retra, K., Figaroa, F., Hollander, J.G., Ab, E., Heetebrij, R.J., Irth, H. & Siegal, G. (2010) Target immobilization as a strategy for NMR-based fragment screening: Comparison of TINS, STD, and SPR for fragment hit identification. *Journal of Biomolecular Screening*, **15**, 978-989.

Kontoyianni, M. & Rosnick, C.B. (2012) Functional prediction of binding pockets. *Journal of Chemical Information and Modeling*, **52**, 824-833.

Kostiainen, R. & Kauppila, T.J. (2009) Effect of eluent on the ionization process in liquid chromatography-mass spectrometry. *Journal of Chromatography A*, **1216**, 685-699.

Kragh-Hansen, U., Chuang, V.T. & Otagiri, M. (2002) Practical aspects of the ligand-binding and enzymatic properties of human serum albumin. *Biol Pharm Bull*, **25**, 695-704.

Kranz, J.K. & Schalk-Hihi, C. (2011) Protein thermal shifts to identify low molecular weight fragments. *Fragment-Based Drug Design: Tools, Practical Approaches, and Examples* (ed. by L.C. Kuo), pp. 277-298. Elsevier Academic Press Inc, San Diego, CA, USA.

Kullolli, M., Hancock, W.S. & Hincapie, M. (2010) Automated platform for fractionation of human plasma glycoproteome in clinical proteomics. *Analytical Chemistry*, **82**, 115-120.

Kuntz, I.D., Chen, K., Sharp, K.A. & Kollman, P.A. (1999) The maximal affinity of ligands. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 9997-10002.

Ladbury, J.E. (2010) Calorimetry as a tool for understanding biomolecular interactions and an aid to drug design. *Biochemical Society Transactions*, **38**, 888-893.

Larsson, A., Jansson, A., Åberg, A. & Nordlund, P. (2011) Efficiency of hit generation and structural characterization in fragment-based ligand discovery. *Current Opinion in Chemical Biology*, **15**, 482-488.

Larsson, P.O., Glad, M., Hansson, L., Månsson, M.O., Ohlson, S. & Mosbach, K. (1983) High-performance liquid affinity-chromatography. *Advances in Chromatography*, **21**, 41-85.

Lau, B.T.C., Baitz, C.A., Dong, X.P. & Hansen, C.L. (2007) A complete microfluidic screening platform for rational protein crystallization. *Journal of the American Chemical Society*, **129**, 454-455.

Leary, J.J. & Schmidt, R.L. (1996) Quadrupole mass spectrometers: An intuitive leak at the math. *Journal of Chemical Education*, **73**, 1142-1145.

Lee, A.C.H., Ramanujulu, P.M., Poulsen, A., Williams, M., Blanchard, S., Ma, D.M., Bonday, Z., Goh, K.L., Goh, K.C., Goh, M.K., Wood, J. & Dymock, B.W. (2012) Thieno 3,2-d pyrimidin-4(3H)-one derivatives as PDK1 inhibitors discovered by fragment-based screening. *Bioorganic & Medicinal Chemistry Letters*, **22**, 4023-4027.

Lesuisse, D., Lange, G., Deprez, P., Bénard, D., Schoot, B., Delettre, G., Marquette, J.P., Broto, P., Jean-Baptiste, V., Bichet, P., Sarubbi, E. & Mandine, E. (2002) SAR and X-ray. A new approach combining fragment-based screening and rational drug design: Application to the discovery of nanomolar inhibitors of Src SH2. *Journal of Medicinal Chemistry*, **45**, 2379-2387.

Lipinski, C.A., Lombardo, F., Dominy, B.W. & Feeney, P.J. (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, **23**, 3-25.

Loun, B. & Hage, D.S. (1996) Chiral separation mechanisms in protein-based HPLC columns. 2. Kinetic studies of (R)- and (S)-warfarin binding to immobilized human serum albumin. *Analytical chemistry*, **68**, 1218-1225.

Lu, H. & Tonge, P.J. (2010) Drug-target residence time: critical information for lead optimization. *Current Opinion in Chemical Biology*, **14**, 467-474.

Macarron, R. (2006) Critical review of the role of HTS in drug discovery. *Drug Discovery Today*, **11**, 277-279.

Majorek, K.A., Porebski, P.J., Dayal, A., Zimmerman, M.D., Jablonska, K., Stewart, A.J., Chruszcz, M. & Minor, W. (2012) Structural and immunologic characterization of bovine, horse, and rabbit serum albumins. *Molecular Immunology*, **52**, 174-182.

Maple, H.J., Garlish, R.A., Rigau-Roca, L., Porter, J., Whitcombe, I., Prosser, C.E., Kennedy, J., Henry, A.J., Taylor, R.J., Crump, M.P. & Crosby, J. (2012) Automated protein-ligand interaction screening by mass spectrometry. *Journal of Medicinal Chemistry*, **55**, 837-851.

Mayer, M. & Meyer, B. (1999) Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angewandte Chemie International Edition*, **38**, 1784-1788.

Maynard, J.A., Lindquist, N.C., Sutherland, J.N., Lesuffleur, A., Warrington, A.E., Rodriguez, M. & Oh, S.H. (2009) Surface plasmon resonance for high-throughput ligand screening of membrane-bound proteins. *Biotechnology Journal*, **4**, 1542-58.

McGovern, S.L., Caselli, E., Grigorieff, N. & Shoichet, B.K. (2002) A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *Journal of Medicinal Chemistry*, **45**, 1712-1722.

Michne, W.F. (2010) Lead discovery: the process. In: *Lead Generation Approaches in Drug Discovery* eds. Z. Rankovic and R. Morphy), pp. 1-17. John Wiley & Sons, Inc, NJ, USA.

Miller, P.E. & Denton, M.B. (1986) The quadrupole mass filter - basic operating concepts. *Journal of Chemical Education*, **63**, 617-622.

Mironov, G.G., Logie, J., Okhonin, V., Renaud, J.B., Mayer, P.M. & Berezovski, M.V. (2012) Comparative study of three methods for affinity measurements: Capillary electrophoresis coupled with UV detection and mass spectrometry, and direct infusion mass spectrometry. *Journal of the American Society for Mass Spectrometry*, **23**, 1232-1240.

Moaddel, R. & Wainer, I.W. (2006) Development of immobilized membrane-based affinity columns for use in the online characterization of membrane bound proteins and for targeted affinity isolations. *Analytica Chimica Acta*, **564**, 97-105.

Moumné, R., Catala, M., Larue, V., Micouin, L. & Tisné, C. (2012) Fragment-based design of small RNA binders: Promising developments and contribution of NMR. *Biochimie*, **94**, 1607-1619.

Mullard, A. (2011) 2010 FDA drug approvals. *Nature Reviews Drug Discovery*, **10**, 82-85.

Murray, C.W., Verdonk, M.L. & Rees, D.C. (2012) Experiences in fragment-based drug discovery. *Trends in Pharmacological Sciences*, **33**, 224-32.

Myszka, D.G., Jonsen, M.D. & Graves, B.J. (1998) Equilibrium analysis of high affinity interactions using BIACORE. *Analytical Biochemistry*, **265**, 326-330.

Nadin, A., Hattotuwagama, C. & Churcher, I. (2012) Lead-oriented synthesis: A new opportunity for synthetic chemistry. *Angewandte Chemie International Edition*, **51**, 1114-1122.

Nakamura, C.E. & Abeles, R.H. (1985) Mode of interaction of beta-hydroxy-beta-methylglutaryl coenzyme-a reductase with strong binding inhibitors - compactin and related-compounds. *Biochemistry*, **24**, 1364-1376.

Navratilova, I. & Hopkins, A.L. (2010) Fragment screening by surface plasmon resonance. *ACS Medicinal Chemistry Letters*, **1**, 44-48.

Neumann, T., Junker, H.D., Schmidt, K. & Sekul, R. (2007) SPR-based fragment screening: Advantages and applications. *Current Topics in Medicinal Chemistry*, **7**, 1630-1642.

Ng, W., Dai, J.R., Slon-Usakiewicz, J.J., Redden, P.R., Pasternak, A. & Reid, N. (2007) Automated multiple ligand screening by frontal affinity chromatography-mass spectrometry (FAC-MS). *Journal of Biomolecular Screening*, **12**, 167-174.

Nienaber, V.L., Richardson, P.L., Klighofer, V., Bouska, J.J., Giranda, V.L. & Greer, J. (2000) Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nature Biotechnology*, **18**, 1105-1108.

Ohlson, S. (2008) Designing transient binding drugs: A new concept for drug discovery. *Drug Discovery Today*, **13**, 433-439.

Ohlson, S., Lundblad, A. & Zopf, D. (1988) Novel approach to affinity chromatography using weak monoclonal antibodies. *Analytical Biochemistry*, **169**, 204-208.

Ohlson, S., Hansson, L., Larsson, P.O. & Mosbach, K. (1978) High performance liquid affinity chromatography (HPLAC) and its application to the separation of enzymes and antigens. *FEBS Letters*, **93**, 5-9.

Olsen, J.V., Ong, S.E. & Mann, M. (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Molecular & Cellular Proteomics*, **3**, 608-614.

Oprea, T.I., Davis, A.M., Teague, S.J. & Leeson, P.D. (2001) Is there a difference between leads and drugs? A historical perspective. *Journal of Chemical Information and Computer Sciences*, **41**, 1308-1315.

Overington, J.P., Al-Lazikani, B. & Hopkins, A.L. (2006) How many drug targets are there? *Nature Reviews Drug Discovery*, **5**, 993-996.

Papalia, G.A., Giannetti, A.M., Arora, N. & Myszka, D.G. (2008) Thermodynamic characterization of pyrazole and azaindole derivatives binding to p38 mitogen-activated protein kinase using Biacore T100 technology and van't Hoff analysis. *Analytical Biochemistry*, **383**, 255-264.

Paul, S.M., Mytelka, D.S., Dunwiddie, C.T., Persinger, C.C., Munos, B.H., Lindborg, S.R. & Schacht, A.L. (2010) How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery*, **9**, 203-214.

Perona, J.J. & Craik, C.S. (1997) Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *Journal of Biological Chemistry*, **272**, 29987-29990.

Perspicace, S., Banner, D., Benz, J., Müller, F., Schlatter, D. & Huber, W. (2009) Fragment-based screening using surface plasmon resonance technology. *Journal of Biomolecular Screening*, **14**, 337-349.

Plavina, T., Wakshull, E., Hancock, W.S. & Hincapie, M. (2007) Combination of abundant protein depletion and multi-lectin affinity chromatography (M-LAC) for plasma protein biomarker discovery. *Journal of Proteome Research*, **6**, 662-671.

Quinn, J.G. (2012) Evaluation of Taylor dispersion injections: Determining kinetic/affinity interaction constants and diffusion coefficients in label-free biosensing. *Analytical Biochemistry*, **421**, 401-410.

Rankovic, Z., Jamieson, C. & Morphy, R. (2010) High throughput screening approach to lead discovery. In: *Lead Generation Approaches in Drug Discovery* eds. Z. Rankovic and R. Morphy), pp. 22-62. John Wiley & Sons, Inc, NJ, USA.

Ravelet, C., Michaud, M., Ravel, A., Grosset, C., Villet, A. & Peyrin, E. (2004) Streptavidin chiral stationary phase for the separation of adenosine enantiomers. *Journal of Chromatography A*, **1036**, 155-160.

Rich, R.L., Errey, J., Marshall, F. & Myszka, D.G. (2011) Biacore analysis with stabilized G-protein-coupled receptors. *Analytical Biochemistry*, **409**, 267-272.

Rich, R.L., Quinn, J.G., Morton, T., Stepp, J.D. & Myszka, D.G. (2010) Biosensor-based fragment screening using FastStep injections. *Analytical Biochemistry*, **407**, 270-277.

Schechter, I. & Berger, A. (1967) On size of active site in proteases .I. Papain. *Biochemical and Biophysical Research Communications*, **27**, 157-162.

Schiel, J.E. & Hage, D.S. (2009) Kinetic studies of biological interactions by affinity chromatography. *Journal of Separation Science*, **32**, 1507-1522.

Schwarz, R.P., Becker, J.C.P., Brooks, R.L., Hursting, M.J., Joffrion, J.L., Knappenberger, G.D., Kogan, T.P., Kogan, P.W. & McKinney, A.A. (1997) The preclinical and clinical pharmacology of Novastan (argatroban): A small-molecule, direct thrombin inhibitor. *Clinical and Applied Thrombosis-Hemostasis*, **3**, 1-15.

Scott, D.E., Coyne, A.G., Hudson, S.A. & Abell, C. (2012) Fragment-based approaches in drug discovery and chemical biology. *Biochemistry*, **51**, 4990-5003.

Seeger, C., Christopeit, T., Fuchs, K., Grote, K., Sieghart, W. & Danielson, U.H. (2012) Histaminergic pharmacology of homo-oligomeric beta 3 gamma-aminobutyric acid type A receptors characterized by surface plasmon resonance biosensor technology. *Biochemical Pharmacology*, **84**, 341-351.

Senisterra, G.A., Markin, E., Yamazaki, K., Hui, R., Vedadi, M. & Awrey, D.E. (2006) Screening for ligands using a generic and high-throughput light-scattering-based assay. *Journal of Biomolecular Screening*, **11**, 940-948.

Sertürner, F. (1805) No title. *Journal der Pharmacie fuer Aerzte und Apotheker*, **13**, 229-243.

Sertürner, F. (1817) Ueber das morphium, eine neue salzfähige grundlage, und diemekonsäure, als hauptbestandtheile des opiums. *Annalen der Physik*, **55**, 56-89.

Sharff, A. & Jhoti, H. (2003) High-throughput crystallography to enhance drug discovery. *Current Opinion in Chemical Biology*, **7**, 340-345.

Shuker, S.B., Hajduk, P.J., Meadows, R.P. & Fesik, S.W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR. *Science*, **274**, 1531-1534.

Siller-Matula, J.M., Schwameis, M., Blann, A., Mannhalter, C. & Jilma, B. (2011) Thrombin as a multi-functional enzyme Focus on in vitro and in vivo effects. *Thrombosis and Haemostasis*, **106**, 1020-1033.

Sinauridze, E.I., Romanov, A.N., Gribkova, I.V., Kondakova, O.A., Surov, S.S., Gorbatenko, A.S., Butylin, A.A., Monakov, M.Y., Bogolyubov, A.A., Kuznetsov, Y.V., Sulimov, V.B. & Ataullakhanov, F.I. (2011) New synthetic thrombin inhibitors: Molecular design and experimental verification. *PLoS ONE*, **6**, 12.

Skoog, D.A., Holler, J.F. & Crouch, S.R. (2007) *Principles of instrumental analysis*, 6th edn. Thomson Brooks/Cole, CA, USA.

Stangier, J., Rathgen, K., Stähle, H., Gansser, D. & Roth, W. (2007) The pharmacokinetics, pharmacodynamics and tolerability of dabigatran etexilate, a new oral direct thrombin inhibitor, in healthy male subjects. *British Journal of Clinical Pharmacology*, **64**, 292-303.

Stockman, B.J. & Dalvit, C. (2002) NMR screening techniques in drug discovery and drug design. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **41**, 187-231.

Sudlow, G., Birkett, D.J. & Wade, D.N. (1975) Characterization of 2 specific drug binding-sites on human-serum albumin. *Molecular Pharmacology*, **11**, 824-832.

Sun, C.H., Petros, A.M. & Hajduk, P.J. (2011) Fragment-based lead discovery: challenges and opportunities. *Journal of Computer-Aided Molecular Design*, **25**, 607-610.

Swinney, D.C. (2009) The role of binding kinetics in therapeutically useful drug action. *Current Opinion in Drug Discovery and Development*, **12**, 31-39.

Talbert, A.M., Tranter, G.E., Holmes, E. & Francis, P.L. (2002) Determination of drug-plasma protein binding kinetics and equilibria by chromatographic profiling: Exemplification of the method using L-tryptophan and albumin. *Analytical chemistry*, **74**, 446-452.

Teague, S.J., Davis, A.M., Leeson, P.D. & Oprea, T. (1999) The design of leadlike combinatorial libraries. *Angewandte Chemie International Edition*, **38**, 3743-3748.

Torres, F.E., Recht, M.I., Coyle, J.E., Bruce, R.H. & Williams, G. (2010) Higher throughput calorimetry: opportunities, approaches and challenges. *Current Opinion in Structural Biology*, **20**, 598-605.

Tsai, J., Lee, J.T., Wang, W., Zhang, J., Cho, H., Mamo, S., Bremer, R., Gillette, S., Kong, J., Haass, N.K., Sproesser, K., Li, L., Smalley, K.S.M., Fong, D., Zhu, Y.L., Marimuthu, A., Nguyen, H., Lam, B., Liu, J., Cheung, I., Rice, J., Suzuki, Y., Luu, C., Settachatgul, C., Shellooe, R., Cantwell, J., Kim, S.H., Schlessinger, J., Zhang, K.Y.J., West, B.L., Powell, B., Habets, G., Zhang, C., Ibrahim, P.N., Hirth, P., Artis, D.R., Herlyn, M. & Bollag, G. (2008) Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 3041-3046.

Tummino, P.J. & Copeland, R.A. (2008) Residence time of receptor-ligand complexes and its effect on biological function. *Biochemistry*, **47**, 5481-5492.

Tyndall, J.D.A., Nall, T. & Fairlie, D.P. (2005) Proteases universally recognize beta strands in their active sites. *Chemical Reviews*, **105**, 973-999.

van Deemter, J., Zuiderweg, F. & Klinkenberg, A. (1956) Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography. *Chemical Engineering Science*, **5**, 271-289.

van Regenmortel, M.H.V. (1998) From absolute to exquisite specificity. Reflections on the fuzzy nature of species, specificity and antigenic sites. *Journal of Immunological Methods*, **216**, 37-48.

Vanwetswinkel, S., Heetebrij, R.J., van Duynhoven, J., Hollander, J.G., Filippov, D.V., Hajduk, P.J. & Siegal, G. (2005) TINS, target immobilized NMR screening: An efficient and sensitive method for ligand discovery. *Chemistry and Biology*, **12**, 207-216.

Veber, D.F., Johnson, S.R., Cheng, H.Y., Smith, B.R., Ward, K.W. & Kopple, K.D. (2002) Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, **45**, 2615-2623.

Wasserman, S.R., Koss, J.W., Sojitra, S.T., Morisco, L.L. & Burley, S.K. (2012) Rapid-access, high-throughput synchrotron crystallography for drug discovery. *Trends in Pharmacological Sciences*, **33**, 261-267.

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. (2013) Parallel screening of low molecular weight fragment libraries: do differences in methodology affect hit identification? *Journal of Biomolecular Screening*, **18**, 147-59.

Winter, A., Higueruelo, A.P., Marsh, M., Sigurdardottir, A., Pitt, W.R. & Blundell, T.L. (2012) Biophysical and computational fragment-based approaches to targeting protein-protein interactions: applications in structure-guided drug discovery. *Quarterly Reviews of Biophysics*, **45**, 383-426.

Wöhler, F. (1828) Ueber künstliche bildung des harnstoffs. *Annalen der Physik und Chemie*, **88**, 253-256.

Yoo, M.J. & Hage, D.S. (2009) Evaluation of silica monoliths in affinity microcolumns for high-throughput analysis of drug-protein interactions. *Journal of Separation Science*, **32**, 2776-2785.

Zheng, C.J., Han, L.Y., Yap, C.W., Xie, B. & Chen, Y.Z. (2006) Progress and problems in the exploration of therapeutic targets. *Drug Discovery Today*, **11**, 412-420.

Zopf, D. & Ohlson, S. (1990) Weak-affinity chromatography. *Nature*, **346**, 87-88.