Review

A new strategy for improved secondary screening and lead optimization using high-resolution SPR characterization of compound–target interactions

Walter Huber*
F. Hoffmann-La Roche Ltd, Basel, Switzerland

Biophysical label-free assays such as those based on SPR are essential tools in generating high-quality data on affinity, kinetic, mechanistic and thermodynamic aspects of interactions between target proteins and potential drug candidates. Here we show examples of the integration of SPR with bioinformatic approaches and mutation studies in the early drug discovery process. We call this combination ‘structure-based biophysical analysis’. Binding sites are identified on target proteins using information that is either extracted from three-dimensional structural analysis (X-ray crystallography or NMR), or derived from a pharmacore model based on known binders. The binding site information is used for in silico screening of a large substance library (e.g. available chemical directory), providing virtual hits. The three-dimensional structure is also used for the design of mutants where the binding site has been impaired. The wild-type target and the impaired mutant are then immobilized on different spots of the sensor chip and the interactions of compounds with the wild-type and mutant are compared in order to identify selective binders for the binding site of the target protein. This method can be used as a cost-effective alternative to high-throughput screening methods in cases when detailed binding site information is available. Here, we present three examples of how this technique can be applied to provide invaluable data during different phases of the drug discovery process. Copyright © 2005 John Wiley & Sons, Ltd.

Keywords: surface plasmon resonance; biosensor; protein interaction analysis; drug discovery; direct binding assay; Biacore

Received 15 March 2005; revised 4 April 2005; accepted 2 May 2005

INTRODUCTION

The identification of high-quality hits and lead compounds is crucial in the drug discovery process. An understanding of structure–activity relationships and mechanistic aspects of action are key in selecting the best chemical class for further optimization. It is therefore essential to obtain sufficient high-quality data on affinity, kinetic, mechanistic and thermodynamic aspects of an interaction between potential drug candidates and their targets.

High-throughput screening assays are routinely used for the identification of compounds which bind to the target, although a very large proportion of the hits identified are false positives that do not bind to the target binding site (Xiang et al., 2003; Gribbon and Sewing, 2003). Compounds produce false positive results for a number of reasons, including interference with the labelling method (fluorescence quenching), binding to the substrate or general (promiscuous) protein binding. There are also situations when high-throughput screening (HTS) is not applicable, such as when new target proteins have been identified from their similarity with established target proteins, but no natural binders (e.g. substrate) have been identified, thus making an inhibition-based labelling approach impossible. Similarly, HTS may be hindered if the substrate is known, but is not stable. This is often the case for isomerases, and was also true for the hydroxymethyl-pterin pyrophosphokinase (HPPK) study we report here.

Biophysical binding assays are essential tools in generating label-free, high-quality data on the interaction between target proteins and potential drug candidates. Owing to their moderate throughput and/or high target protein and compound consumption, these assays are generally only suited to characterization of a restricted number of compounds and are therefore not suited to systematic secondary screening of hits from HTS. Label-free biophysical screening methods include isothermal titration calorimetry (ITC), analytical ultrafiltration, nuclear magnetic resonance, mass spectrometry, size exclusion chromatography, electrophoresis and biosensors.

Direct binding assays based on surface plasmon resonance (SPR) protein interaction analysis technology (Fig. 1) have several advantages in comparison with other biophysical screening methods. The major advantages are low target consumption (a few µg protein/microplate compounds), the possibility to simultaneously use reference...
proteins (to discriminate between selective and promiscuous binders) and that it resolves affinity into on/off-rate constants, improving structure–activity relationship studies. The study of direct binding between low molecular weight compounds and target molecules was developed by Helena Danielson at Uppsala University, Sweden in collaboration with scientists at Biacore, AB (Markgren et al., 1998, 2000, 2001, 2002; Alterman et al., 2001; Hämäläinen et al., 2000; Karlsson et al., 2000). The use of optical biosensors in drug discovery has recently been reviewed (Cooper, 2002, 2003a,b; Myszka and Rich, 2000; Rich and Myszka, 2004; Karlsson, 2004; Lofás, 2004). Direct binding SPR data on affinity, kinetics and thermodynamics have been published for a number of different drug target systems, including HIV-1 protease (Shuman et al., 2003, 2004a,b), CMV protease (Geitmann and Danielson, 2004), HIV TAR-RNA (Davis et al., 2004), p38 MAP kinase (Davidson et al., 2004; Bukhtiyarova et al., 2004; Casper et al., 2004), oestrogen receptor alpha and beta (Rich et al., 2002), carbonic anhydrase (Myszka, 2004), PPAR (Yu et al., 2004), DNA/telomerase (Tediou et al., 2004; Carrasco et al., 2002, 2003; Harrison et al., 2003; Bailly et al., 2003), MDM2-p53 (Vassiljev et al., 2004), (pp60)Src (Lange et al., 2003) and VEGF (Ueda et al., 2003).

Despite these reports, the true value of resolving affinity into kinetic constants has not been fully recognized. This is mainly because kinetic data have not previously been available for larger series of lead compounds and for different scaffold classes. A study of the literature reveals, however, that concrete examples of a critical role for the kinetic resolution of drug–target binding properties have been available for more than a decade. In the study of saxitoxins (Hall et al., 1990), the authors found that it was not possible to understand structure–activity relationship before they had resolved the affinity into binding kinetics. They showed that an increased charge state of the toxins correlated with increases in on-rate. Elg et al. (1997) found that, for thrombin inhibitors, the on-rate correlated much better than affinity with the in vivo pharmacological effect in a rat blood coagulation model. This means that, for thrombin inhibitors, using affinity as a measure of binding strength is misleading from the critical perspective of pharmacological effect. Markgren et al. (2002) showed in their structure–activity relationships (SAR) study for HIV-1-protease that scaffolds clustered in the on/off-rate/KD-map and that the affinity of some of the more structurally rigid cyclic scaffold were limited by the instability of the complex formed (rapid off-rate). They could also show that compounds with virtually identical affinities could have off/on-rate ratios that differed by three orders of magnitude. Kroogsgard et al. (2004) showed that off-rate and heat capacity are responsible for the T cell activation, not the affinity.

In this paper, we show three examples in which binding site-modified proteins were used as reference targets in direct binding SPR assays aimed at identifying compounds with true target protein binding site specificity. This modification is achieved either by point mutations (hydroxymethyl-pterin pyrophosphokinase, HPPK; dihydronicotinamide-aldolase, DHNA) or by chemically blocking the active site with covalent inhibitors (dipeptidylpeptidase IV, DPP-IV). The examples presented represent different steps of the drug discovery process. The HPPK example shows the value of the technique for hit finding, since no HTS could be performed for this target due to the lack of a stable substrate. The process applied in this case is called ‘structure-based biophysical analysis’ (SBBA), an approach which was first applied to the study of DNA gyrase (Boehm et al., 2000). SBBA is based on the use of structural information for setting up a focused library and for directing mutations of the binding site of the wild-type (wt) target protein, enabling the simultaneous use of the wt and the mutant for parallel analysis in SPR studies. Two additional examples (DPP-IV and DHNA) are also presented, in which SBBA is used for lead selection and optimization during later stages of the drug discovery process. For DPP-IV, kinetic data for 160 compounds from three structural classes are used for the selection/prioritization of scaffolds during lead optimization. For
DHNA, SPR-data are used for analysing mechanistic aspects of inhibition and biological effects.

EXPERIMENTAL

The SPR studies were performed with Biacore 2000 (DHNA), Biacore 3000 (HPPK) and Biacore S51 (DPP-IV) instruments. The target proteins were immobilized using standard amine coupling using EDC/NHS coupling chemistry and CM5 sensor chips, or using biotinylated target proteins and streptavidin-coated SA sensor chips. HBS-EP or phosphate buffer saline pH 7.4 were used as running and sample buffers. The DMSO content in sample and running buffers varied between 2 and 5%, as appropriate.

Structure-based biophysical analysis

In bioinformatic approaches, three-dimensional structural information from X-ray crystallography or NMR and mutation studies are used to identify a binding site on the target protein. Binding site analysis leads to the generation of a pharmacophore hypothesis, which in principle outlines potential interactions between the binding site and a small ligand molecule. This pharmacophore hypothesis is used for an in silico pre-screening of large libraries to establish a focused library of a few hundred compounds. This step circumvents the need for an HTS assay.

The pharmacophore hypothesis can also be used to design mutants with altered binding sites. In general, these involve point mutations producing a change in one or two of the amino acids located in the binding site. These amino acids are assumed to be important in any proposed binding mechanism, or pharmacophore hypothesis, and different binding patterns are expected for wild-type and mutant proteins. Once the focused library has been established, it can be tested with a biophysical screen using the wild type and mutant proteins.

SBBA of HPPK—virtual screening and SPR

As no biological assay was available for HPPK, a focused library was generated, using in silico screening, and a direct binding assay had to be developed to screen these compounds for HPPK binding. The X-ray structure of a protein substrate analogue complex was determined in-house (Hennig et al., 1999) and the identification of the binding site and generation of a pharmacophore hypothesis was straightforward. The most prominent features of the binding site included polar interactions between the pterin ring of the substrate and asparagine 56, serine 43, lysine 44 and leucine 46, and non-polar interactions with phenylalanine 123 and tyrosine 54 (Fig. 2).

Two-hundred and fifty compounds were selected from an in-house library. These fully or partly matched the binding pattern identified through the pharmacophore hypothesis that was developed from structural information.

A mutant form of the protein designed to have impaired binding capability was generated by exchanging asparagine 56 with alanine. While the binding site had been successfully altered, and binding affinity for the substrate analogue was fully destroyed by this single point mutation, protein conformation remained virtually unchanged, as confirmed by circular dichroism spectroscopy.

The binding activity profiles of the wt and mutant proteins were elucidated using SPR. Proteins were immobilized on a Biacore sensor chip using standard amine coupling chemistry. The saturation signal for the binding of small molecules (molecular weight, MW ≈400 Da) was approximately 25 resonance units (RUs). Comparing this saturation response with the amount of protein immobilized indicated that 30% of the immobilized wt protein remained active. No binding was observed for the substrate analogue to the immobilized mutant protein.

Figure 3 shows binding curves of three representatives from the 250 compounds. This analysis enables binding and non-binding compounds to be distinguished. Compounds that bind show a time-dependent increase in response during the association phase when the surface is in contact with the target and shows therefore signals for report points 2 and 3. Compound A shows a rapid association and dissociation from the target and shows therefore signals for report points 1, 2, 3, and 4. Compound B does not bind to the target. Compound C depicts a strong binder with significant signals for report points 2–4.
theoretical saturated 1:1 binding level of Compound 5 indicates promiscuous binding as it binds twice the compounds (2, 3, 6, 14) show selective binding to the wt HPPK. wt HPPK (left bar) and to the impaired mutant (right bar). Report affinities, determined by SPR, were between 400 nM and hits belong to four different classes of compounds and their showed selective binding to the HPPK binding site. These direct binding assay identified a total of 15 compounds that cophore/scaffold classes were screened in this project. The ing affinity that takes no account of the degree of specificity. selectivity is generally a more valuable parameter for com- pound selection and development than absolute target bind- ing activity, bioavailability and metabolic profiles and decreased toxicity.

Validation of DPP-IV lead series using high-resolution kinetic analysis and modified binding site controls

SPR was also used to validate leads for DPP-IV (Thoma et al., 2003). In this assay, DPP-IV was immobilized on the sensor surface and the binding and dissociation of inhibitor compounds was monitored in real time by SPR.

Active site point mutants were prepared and used to probe for site specificity. For this project, another strategy was applied. The active site serine was used for the binding of a phosphinate ester (Augustyns et al., 1999). The covalent binding of this ester to the serine is irreversible and blocks the active site. The flow system design of the Biacore instrument allows this blocking reaction to be carried out on proteins already immobilized to a specific region of the sensor surface, enabling parallel analysis of blocked and unmodified targets in the same assay.

The outcome of such an experiment is shown in Fig. 5. The unmodified protein yields information on reversibility of binding along with kinetic data, with all interactions monitored in real time. From the time course of these reactions it is possible to extract the relevant kinetic data (association rate constant, $k_{on}$, and dissociation rate constant, $k_{off}$) and calculate the equilibrium binding constant, $K_D$ ($K_D = k_{off}/k_{on}$). The relevance of this high-resolution binding characterization was clearly validated by the lack of response seen with the modified protein, showing that the kinetic data obtained for wt DPP-IV represents compound binding to the active site (Fig. 5).

We routinely use this setup for further characterization of active inhibitors synthesized during the hit expansion step. The compounds could be assigned to three different structure classes, represented in the scheme below. In general, $K_D$ values correlate well with IC$_{50}$ values and confirm the biological results. The significant additional value of characterizing lead compounds with an alternative technique is readily demonstrated by looking at a so-called on/off/$K_D$ plot (Markgren et al., 2002), in which the two rate constants are plotted against each other and the diagonal iso-affinity lines correspond to the equilibrium dissociation constant value, $K_D$ (Fig. 6). Important information can be obtained from such plots. The plot in Fig. 6 clearly indicates clustering of the compounds in three major areas (A, B, C). These clusters contain mainly compounds from three different structural classes, cyanopyrrolidines (A), bazoquinolinizin (B) and pyrrolidinones (C) (Scheme 1). The cyanopyrrolidi- dines that cluster in the upper left corner exhibit fast association and slow dissociation rate constants. The compounds with highest affinities ($K_D$=1 nM) have on-rates around $10^6$ M$^{-1}$s$^{-1}$ and off-rates around $10^{-2}$ s$^{-1}$. This is probably the class from which a potential drug might
emerge if all other prerequisites for developing a successful compound are fulfilled. The benzoquinoline class in the upper right corner comprises compounds with fast on-rates but, in comparison to the first class, also fast off-rates. These compounds can be optimized mainly through structural modifications that slow down the dissociation process because it appears from these two compound classes that an increase in $K_D$ is mainly obtained by a decrease in $k_{off}$.

The third cluster (C) is somewhat diffuse and is mainly located in the lower left corner, indicating very slow on- and off-rates. Since such extremely slow on-rates have not previously been observed for low molecular weight ligands, additional analytical ultracentrifugation (AUC) experiments were performed to analyse protein, inhibitor and protein/inhibitor preparations. Solutions of these ligands showed very high sedimentation coefficients, indicating that these compounds exist mainly as very large aggregates and not as monomers. These aggregates have molecular weights larger than $10^6$ Da, which explains why slow association is seen in the SPR experiments. The slow association is due to the low amount of monomer in solution and to the fact that there is no real driving force for the molecules to leave these aggregates and to bind to the immobilized protein as a monomer. The kinetic of this process is very slow and is associated with a high activation barrier.

The ligands from cluster C behave totally differently when the protein is present in solution. When a DMSO solution of these ligands is dispensed into an aqueous solution containing the protein, no sedimentation of large compound aggregates or loss of material was seen in the AUC experiment. However, independently from the

---

**Figure 6.** On/off/$K_D$-map for 160 compounds binding to DPP-IV. (a) Association rate constant ($k_{on}$) plotted vs. dissociation rate constant ($k_{off}$) and the affinity, equilibrium dissociation constant ($K_D$) added as diagonal iso-affinity lines. The three structural classes (A = cyanopyrrolidines, B = benzoquinolizines, C = pyrrolidinones) cluster in different positions. Compounds 1 and 2 are highlighted to illustrate that compounds with almost identical affinities can exhibit very different kinetic properties. (b) Kinetic binding profiles of compounds 1 and 2. Although their affinities differ by only around 2-fold, compound 2 exhibits a much more rapid kinetic profile, with on- and off-rates around 100-fold faster.

---

**Scheme 1.** DPP-IV binders.

**Scheme 2.** DHNA binders.
concentration ratio of protein to ligand, all of the ligand was found to be bound to the protein. Obviously the presence of protein in solution reduces the concentration of free ligand such that no aggregates can be formed. Stoichiometric factors of up to 13 are determined. It is not surprising that enzymatic activity is suppressed for a protein that is fully covered with an unspecifically binding compound. The compounds appear as inhibitors in the enzymatic assay, but the inhibitory effect may not be due to the binding of the molecules to the active site. Consequently such compounds are not drug-like and the substance class C was rejected as a potential lead compound. Compounds exhibiting such behaviour have been classified recently in the literature as ‘promiscuous inhibitors’ or ‘frequent hitters’ (Ryan et al., 2003). These types of compounds are one of the main sources of false positives found in high-throughput screens. Since the behaviour of this compound class is independent of the protein (it is, for instance, similarly solubilized by human serum albumin), it will probably be active in every protein assay with a similar format.

Mechanistic aspects of inhibition of DHNA

The SBBA strategy was also used to design suitable mutant proteins for testing binding site activity in the case of dihydronopterin-aldolase (DHNA). DHNA is another protein of the bacterial folate pathway. The three-dimensional structure of DHNA, which was obtained in-house, shows that this aldolase exists as an octamer (Hennig et al., 1998). This quaternary structure was confirmed in solution by AUC. The protein is only active in this octameric form. A crystal structure of a substrate–protein complex enabled the design of mutants with impaired binding sites. Two mutants were constructed, with the glutamic acid 74 replaced by arginine or alanine. These mutations eliminated binding activity but did not change the folding or physico-chemical properties of the protein.

In the course of lead selection, structure–activity analysis showed very different activities for a series of compounds with similar structures (especially those with similar pharmacophores). L-Biopterin, for example, is a highly active compound, while D-neopterin is not (Fig. 7). The parts of the molecules responsible for the polar interactions with the protein are virtually identical in both compounds. In the process of optimizing such compounds for high activity, it is essential to know why such minor structural changes can influence the activity of a lead compound to such an extent. This problem was tackled using SPR assays, with the aldolase biotinylated and immobilized on a streptavidin-coated sensor chip surface. This enabled full regeneration of the surface and immobilization of a new batch of aldolase before each binding experiment. The experimental protocol included capture of the biotinylated aldolase, injection of the analyte solution, and regeneration of the surface with ethanolamine solution and short pulses of 50 mM sodium hydroxide and 3 M sodium chloride. This regeneration step removed all the aldolase from the surface without destroying the binding capacity of the streptavidin. After regeneration, fresh aldolase could be captured on the sensor chip.

The octameric form is only stable at salt concentration above 1 M and at basic pH (>8.0), conditions that are often found in bacteria. This experimental setup using captured target allows the evaluation of varying salt or pH on the cleavage of the octamer. The amount of protein loss due to the cleavage of the octamer could be measured directly.

The reason why one compound is active and the other inactive becomes obvious when comparing sensorgrams. For the active compound, a significant amount of aldolase is lost from the surface, a phenomenon shown by AUC experiments to be due to cleavage of octamers into tetramers. Such protein loss is not observed with the inactive compound. In independent experiments using tetrameric aldolase, it was shown that both compounds bind to the aldolase. The degree of octamer cleavage is concentration-dependent. Using the mutant protein with an impaired binding site as a reference, it was demonstrated that the compound must bind to this site to cleave octameric aldolase. Figure 7 shows the concentration-dependent cleavage of the wt protein by L-biopterin, the active compound. No cleavage is observed when the binding site is impaired, as in the mutant protein E74R-aldolase, or when wt aldolase is contacted with D-neopterin, the inactive compound in inhibition experiments.

In the case of DHNA, structure–activity interpretation has to consider areas not directly involved in forming polar interactions to binding site amino acids. It has to consider inhibitor moieties that protrude from the binding site to influence intramolecular interactions responsible for the formation of the octamer. This conclusion is confirmed by another series of compounds, again with identical moieties protruding into the binding site but with different fragments at the side. All bind to aldolase, but only those that cleave the octamer in a concentration-dependent manner are active in the biological assay.

DISCUSSION

In our laboratory the SPR- direct binding assays have been routinely used for 10 years and assays have been developed
for about 20 different low molecular weight compound protein target systems. The main use has been as a secondary screen for validation of hits from HTS. We have found that a high percentage of hits are false positives, e.g. they do not bind to the target or they are promiscuous binders. This does not mean that HTS is unreliable, but rather that it should be used in combination with a biophysical secondary screening method that identifies the true binders. The use of structural information about the binding site and pharmacophore hypothesis, as in SBBA, gives a significant improvement because an impaired mutant can be used for the identification of compounds that bind selectively to the targeted binding site. The use of impaired mutants as references significantly reduces the number of false positives. The selection process is further improved because promiscuous binders can be identified by the fact that they give signals far higher than expected from 1:1 binding.

We found that the extra work needed for preparing mutant target proteins is well worth the trouble and we now routinely use them when screening focused libraries or fragment libraries as well as in hit validation using direct SPR binding. The design of such mutants is based on the three-dimensional structure information and the respective pharmacophore hypothesis or binding model derived therefrom. Generally amino acids that are involved in polar interactions are replaced by amino acids that have altered donor/acceptor properties, have the opposite charge or have higher spatial demand and block the site by steric hindrance. It is important to check that the mutation has little or no influence on the protein conformation and the physicochemical properties. When using an irrelevant protein such as for instance HSA that is not closely related to the target protein such as a mutant, these indispensable properties of an ideal control are lost.

The examples described here show that increased information achieved by resolving the affinity into binding rate constants significantly improves the understanding of structure–activity relationships. In the DPP-IV example, we found that only one of the three scaffolds provided adequate binding properties. In addition a second scaffold possessed properties that justified further attempts at improving the kinetic behaviour of the compounds. The third structural class was not considered for further optimization attempts because of its promiscuous binding properties.

In the case of DHNA structural information on ligand binding site and binding model was not sufficient to understand inhibitory activity. The Biacore experiments clearly showed that efficient inhibitors must be able to bind to the active site and to cleave the octamer.

The major advantages with SPR direct binding assays in comparison with other biophysical screening methods are binding kinetic information and the very low consumption of the target molecule. The throughput fits well in a secondary screening situation when hits from HTS or virtual screens are going to be validated. The target consumption is much lower than with the other biophysical screening methods. For example with the same amount used to determine the affinity using ITC, NMR or AUC, an SPR assay can determine the affinities for 100–1000 substances. If the affinities of a few compounds are going to be measured then AUC and ITC are very useful methods with very short assay development times. However, these methods do not give the kinetic information useful for lead optimization and information on promiscuous binding, which is readily obtained by using reference proteins in SPR assays.

The approach described here is a powerful method for the identification of inhibitors aimed at a well-defined active binding site on the target protein. It may not, however, be equally applicable to all drug discovery scenarios. Allosteric inhibitors, for example, bind to regions that are spatially distinct from the active site and may be missed altogether using impaired binding site mutant as a reference protein.

There are also some complications with SPR direct binding assay-based secondary screening methods, which are related to the reuse of the target protein. The SPR assays need reasonably pure and active proteins as the detection principle is related to detection of mass measured as change in refractive index. Therefore care is needed in finding proper immobilization conditions and chemistries. In most cases simple amine coupling using EDC/NHS coupling chemistry works. However, there are proteins which are unstable in acidic conditions which are used in the preconcentration step. This problem can be minimized by mixing the target with the immobilization buffer immediately before the injection on the sensor chip. The use of an inhibitor or substrate to block the binding site has also been shown to increase the activity of the immobilized protein (Casper et al., 2004). Additionally, a number of other chemical and capturing-based immobilization options are also available.

In secondary screening there is little need for regeneration experiments since low affinity compounds dissociate very rapidly. During lead optimization, the need for proper regeneration methods is greater as the compounds have slower dissociation rates. When the protein is unstable and strong regeneration conditions cannot be used, increasing the assay temperature to increase the dissociation rate and using longer dissociation times are often effective. Another method is to use capture-based immobilization, where regeneration breaks the binding between the capturing molecule and the target (although this does increase target consumption).

**CONCLUSION**

Structure-based biophysical analysis is a powerful approach to the selection of lead compounds with potential for drug development. The use of mutated target proteins as reference in the SPR studies further improves the quality of the information obtained. The structure activity data, strengthened by the resolution of affinities into kinetic parameters, can provide researchers with the knowledge they need to reliably select the most promising compounds for further development. In our experience, SBBA is a powerful tool for improving the drug discovery process.

**Acknowledgements**

The author thanks Dr G Franklin and Dr M. Hämäläinen from Biacore AB for support and valuable discussions.
REFERENCES


Shuman CF, Markgren P-O, Hämäläinen MD, Danielsson UH. 2003. Elucidation of HIV-1 protease resistance by character-


