ARTICLE

OBC
www.rsc.org/obc

Exploring the active site of human factor Xa protein by NMR screening of small molecule probes

Lee Fielding,*[a] Dan Fletcher,[a] Samantha Rutherford,[a] Jasmit Kaur †[b] and Jordi Mestres ‡[b]

[a] *Department of Analytical Chemistry, Organon Laboratories Ltd, Newhouse, Lanarkshire, UK ML1 5SH. E-mail: l.fielding@organon.co.uk; Fax: +44 1698 736187; Tel: +44 1698 736182*

[b] *Department of Medicinal Chemistry, Organon Laboratories Ltd, Newhouse, Lanarkshire, UK ML1 5SH*

Received 26th August 2003, Accepted 13th October 2003
First published as an Advance Article on the web 31st October 2003

A collection of small molecules (MW < 350 Da) was screened for binding to human factor Xa using saturation transfer difference NMR spectroscopy to detect binding. The NMR screening experiments identified four hits. Binding isotherms constructed from NMR linewidth data showed that the binding affinities of the hits were all in the 30–210 μM range. Competition binding experiments showed that three of the ligands were displaced by a known μM inhibitor of factor Xa. The success of the method for identifying new ligands and the relevance of this information to the design of new factor Xa inhibitors are discussed.

Introduction

Some recent advances in the field of NMR spectroscopy that are of interest to the drug discovery enterprise are techniques that use NMR as a direct tool to detect small molecule–receptor interactions.[1] Technologies for using NMR as a biophysical screening tool are particularly relevant to lead discovery because they have the potential to identify interesting low affinity, low molecular weight compounds that may otherwise be overlooked by conventional bioassay based high throughput screening techniques.[2,3] NMR has the additional advantage of providing structural information that can then be exploited to guide the progression of those low affinity compounds to novel potent leads.

A methodology that looks to be particularly promising is that of screening modest collections of small molecules for low affinity compounds.[4,5] In this strategy a small but diverse library of low molecular weight soluble compounds is assembled. This library is generally composed of compounds representing fragments commonly found in drug molecules,[6] often combined with structural motifs present in known active molecules for the target of interest or a closely related target of the same family. Then, screening of the library against a protein target by NMR allows for detecting compounds with weak binding (K_D ~ μM–mM). This makes the technology very attractive at the early stages of lead discovery to identify, on the one hand, hit precursors that can be further optimised using structure-guided chemistry and, on the other hand, privileged fragments that can then be incorporated or enriched in the larger compound collections undergoing high-throughput screening.

Cardiovascular disease (myocardial infarction, stroke, deep vein thrombosis and pulmonary embolism *etc.*) is a major cause of mortality in the western world. The immediate cause of these disease conditions is an occlusive blood clot and hence antithrombotic therapy is a crucial component for the treatment and management of these diseases. The blood clotting process is a natural response to stop the loss of blood in higher organisms following vascular injury. Initially, platelets are activated and aggregate at the site of injury, then in a sequence of events termed the coagulation cascade, the blood thickens and fibrin formation leads to an insoluble clot. Therapy usually consists of a combination of anticoagulant and antiplatelet agents, and hence drug discovery activities have focused on finding orally active inhibitors of proteins in the coagulation cascade.[7]

Of all the proteins involved in the coagulation cascade, factor Xa is a prime target in cardiovascular drug discovery. Factor Xa is a trypsin like serine protease that converts prothrombin to catalytically active thrombin – the protein at the centre of events in the coagulation process, and is therefore key to several of the processes necessary to form stable clots. As such, factor Xa is an attractive target for new anticoagulant agents.[8,9]

Factor Xa (51 kDa) consists of a heavy chain (251 amino acids) and a light chain (139 amino acids), held together by a single disulfide bond. The heavy chain incorporates the catalytic triad composed of His-57, Asp-102 and Ser-195. The light chain contains two epidermal growth factor like domains and the chymotrypsin cleavage site (Tyr-44, Lys-45). The crystal structure of factor Xa is known,[10] and several crystal structures of factor Xa in complex with a variety of ligands have been deposited in the Protein Data Bank.[11]

Four binding pockets have been identified within the active site, labelled S1 to S4, and the S1 and S4 binding pockets are the important ones that are exploited by factor Xa inhibitors. The S1 pocket is a narrow cleft defined by Asp-189, Ala-190 and Gln-192 and favours positively charged moieties such as amine, guanidine and benzamidine. A hydrogen bond formed between the aspartate residue at the bottom of the cleft (Asp-189) and benzamidine based inhibitors is a very important piece of the SAR, but is not an absolute prerequisite for high affinity.[12] The S4 pocket is a shallow groove formed by aromatic residues Tyr-99, Phe-174 and Trp-215 and is sufficiently rich in π electrons that it is not only a hydrophobic pocket, but also forms a cation recognition site, thus favouring both hydrophobic and basic moieties in ligands.[13]

The present study reports the results obtained from a structure-based NMR screening approach to factor Xa lead generation. The small molecule screening collection was specifically designed to include both diverse fragments from known drugs and targeted fragments from known factor Xa inhibitors. The study was also conceived to test entirely NMR based methodologies from the initial screen to follow-up experiments designed to measure K_D and to validate the specificity of the hits. An overview of the strategy is shown in Scheme 1.

† *Present address*: OSI Pharmaceuticals, Watlington Road, Oxford, UK OX4 6LT.

‡ *Present address*: Universitat Pompeu Fabra, Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Spain.

DOI: 10.1039/b310265c

identification of target protein
↓
focussed library design
↓
assembly of compound collection
↓
screening (STD NMR)
↓
validation of hits
(competition binding, and K_D)
↓
small molecule leads

Scheme 1 Flow chart showing the sequence of the targeted library lead discovery process.

Results and discussion

Library design

Many factor Xa inhibitors have been reported in the last decade. In general terms factor Xa inhibitors can be classified into the following four classes, the order reflecting strategy progression with time – highly basic bisamidino factor Xa inhibitors, *e.g.* DX 9065a (**1**),[14] monoamidino factor Xa inhibitors, *e.g.* SN 429 (**2**),[15] weakly basic factor Xa inhibitors, *e.g.* RP 208815 which contains an aminoisoquinoline moiety (**3**),[16] and neutral factor Xa inhibitors, *e.g.* Eli Lilly compound **4**[17] (Scheme 2).

The goal of library design was to identify at least 40 low molecular weight soluble compounds from in-house chemical stores and the Available Chemicals Directory (ACD).[18] The selection strategy focussed on including compounds representing fragments from known drugs and compounds containing moieties that are likely to exhibit affinity based on known structures of factor Xa inhibitors. The primary selection procedure had a target of 50–70 compounds to allow for potential attrition due to lack of availability or poor water solubility.

An initial set of 6609 compounds readily available from stores was electronically filtered to retain compounds with molecular weights in the range 80–350 Da, *c*logP values between −2.0 and 3.0, less than 10 rotatable bonds and at least one heteroatom (N or O). The selection procedure then focussed on compounds containing frameworks commonly found in drugs,[6] leaving just 120 compounds from the original set. Finally, visual inspection removed compounds that might be expected to cause chemistry problems (*e.g.* solubility, stability) leaving 29 miscellaneous compounds to draw from stores.

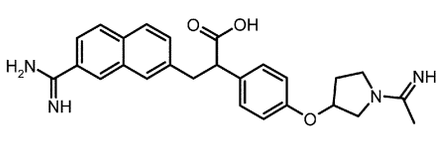
The remaining part of the library was conceived to represent fragments present in known factor Xa inhibitors. A total of 42 factor Xa ligands were retrieved from our in-house drugs database[19] and the structures were translated back to fragments by considering key reagents in the original synthesis of the drugs. This was not straightforward as many factor Xa inhibitors involve several synthetic steps and tracking back to reagents was not always possible. Manual multiple substructure searches were then performed to identify compounds most likely to exhibit weak affinity for factor Xa. For instance, the S1 binding moieties (benzamidines, guanidines and aminoisoquinolines), and the S4 binding moieties (fragments containing aromatic and basic functionalities) were preferred. This produced 281 possible candidates which were subjected to cluster analysis. Final visual inspection of the different clusters allowed for a sensible reduction of the number of compounds to 23, which were then selected for purchase from ACD suppliers.

Hence, at this stage the library was composed of 52 compounds: 29 compounds based on preferred drug frameworks (diverse set) and 23 compounds based on fragments present in known factor Xa inhibitors (focused set). Issues of material availability and solubility reduced the final library to 34 compounds. The composition of the library can be observed in Scheme 3.

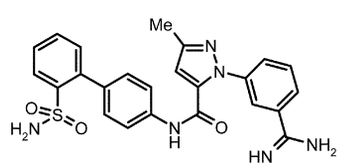
NMR screening

Many NMR methods will detect small molecule binding to target proteins. The methods are broadly classified according to whether the observed species is the ligand or the protein. Advantages to observing a change in the ligand signals rather than the protein signals include a straightforward detection of which components bind to the target, and no limit to the size of the protein target.

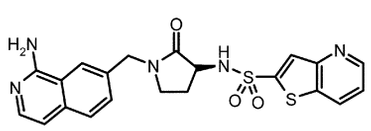
For this study it was decided to use the saturation transfer difference experiment for the screen,[20] Fig. 1. This technique can be used to provide binding information for almost any target, soluble or membrane bound, and it has the benefits that it is a one-dimensional technique, which keeps experimental time down. It has good sensitivity, reducing the amount of required protein to the 100 nM–100 μM range. The experiment takes advantage of the equilibrium condition where free (non bound) and bound forms of the ligand are in fast exchange on the NMR time scale. Under these conditions information about the bound ligand state is retained and is detectable by NMR even in the presence of a large excess of the non bound ligand. The magnetisation of protein protons is manipulated first by selective excitation of an upfield region of the ^{1}H spectrum. When this irradiation is maintained a process known as spin diffusion rapidly (< 200 ms) and effectively saturates all other spins of the macromolecule including any small molecule ligands that are associated with it. A spectral editing technique compares the difference in intensities between ligand signals and saturation



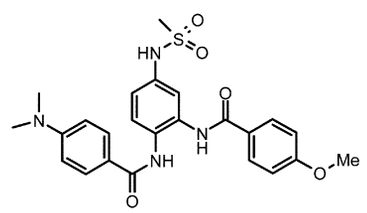
1 (K_i = 41 nM)



2 (K_i = 0.013 nM)

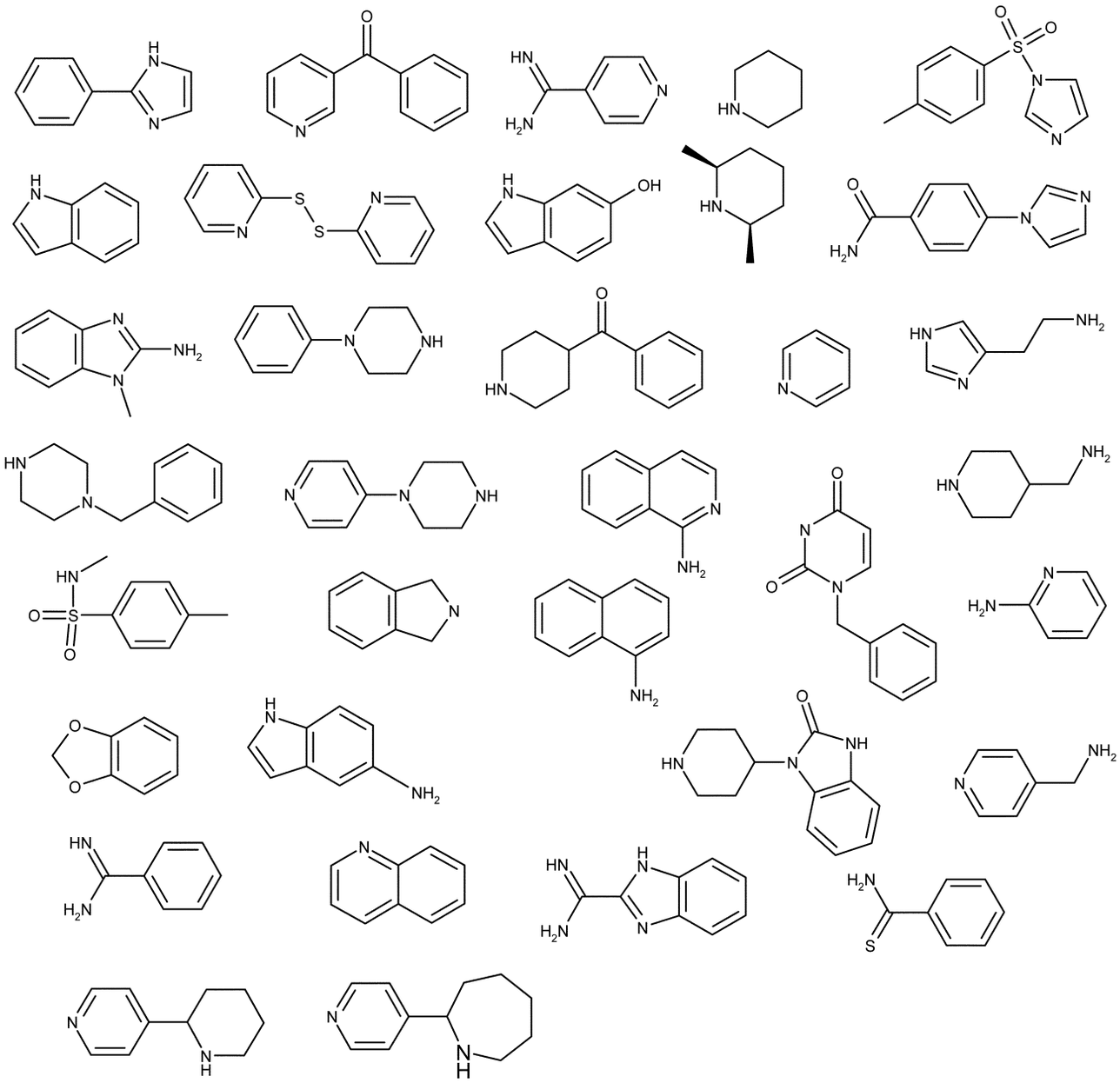


3 (K_i = 22 nM)



4 (K_i = 11 nM)

Scheme 2 Examples of known factor Xa inhibitors.



Scheme 3 The screening library.

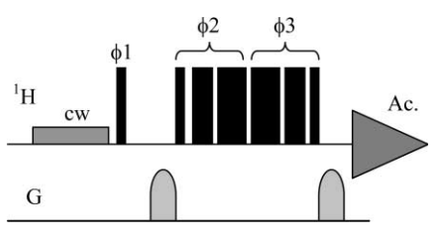


Fig. 1 The pulse program for the saturation transfer difference experiment. The reference spectrum for subtraction was achieved *via* phase cycling and altering the irradiation frequency of the continuous wave pulse between 0.8 ppm (on protein resonance) and 30 ppm (off resonance). The length of the continuous wave pulse was 1.6 s. The phase cycling is $\phi1 = (x, -x)$; $\phi2 = (x, x, y, y, -x, -x, -y, -y)$; $\phi3 = (-x, -x, -y, -y, x, x, y, y)$; ϕrec = $(x, x, -x, -x)$. The two gradient pulses have equal intensity and sign. Water suppression is achieved by a binomial 3–9–19 sequence. The delay between the 3–9–19 pulses was 0.15 ms.

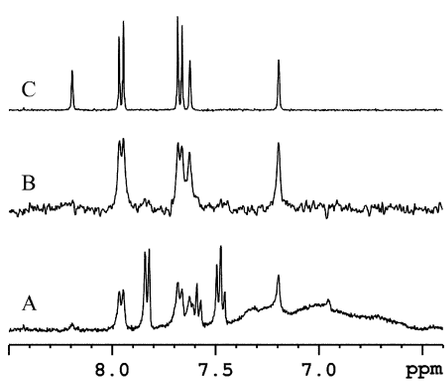
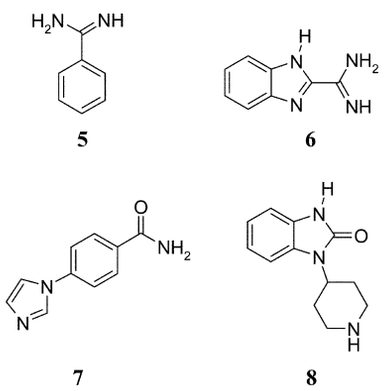


Fig. 2 A typical result from the STD-NMR screen. The proton spectrum of a mixture of protein and potential ligands (A), the saturation transfer difference spectrum minus the protein signals (B), and the reference spectrum of **7** (C). These spectra indicate that of the three small molecules in this sample only **7** binds to factor Xa.

transfer attenuated ligand signals, thus revealing the ligands that have been associated with protein. When a spin lock is applied to filter out the broad protein signal, the experiment cleanly detects only the small molecules that were bound, Fig. 2. The small molecules are usually screened as mixtures to reduce experimental time and to reduce the amount of protein required. Thus a straightforward visual inspection of the saturation transfer difference spectra differentiates between binding and non-binding compounds.

The saturation transfer difference screen indicated that 9 molecules out of the total of 34 were binding to factor Xa. The strength of the interaction of the hits was roughly classed as weak, medium or strong by inspection of the size of their saturation transfer peaks (assuming tighter binding produces more intense peaks), and the 4 ligands that produced the strongest responses were selected for further study, Scheme 4.



Scheme 4 Hits from the STD NMR screen.

Quantitation of binding affinity

The primary saturation transfer difference NMR screening data identified ligands with sub mM affinity for factor Xa and contain some qualitative information on binding affinity (strength of response). It is desirable to put further discussion of the interactions onto a quantitative basis by measuring the stability of the protein–ligand complexes. Many techniques are available for this determination,[21] but in the present case it was most suitable to use NMR methods.

NMR is applicable because spectra of ligand–protein mixtures carry quantitative information about the species distribution as well as the usual structural information. In order for this approach to work, it is necessary to observe an effect on the ligand spectrum caused by binding to the protein and to be able to measure the effect at several different ratios of ligand or protein.[22] The parameter chosen for this study was the linewidth of the ligand ^{1}H signals. The measurements at different ligand/protein ratios are necessary in order to deconvolute some missing data caused by fast dynamic averaging of the free ligand and weakly bound ligand signals.

For a molecule in fast exchange between two sites, any measurable NMR parameter is the population weighted average of that pertaining to the two different environments. So, for the spin lattice relaxation rate $1/T_2$ of a small molecule ligand in solution with a receptor

$$1/T_{2(\mathrm{obs})} = X_{(\mathrm{free})}1/T_{2(\mathrm{free})} + X_{(\mathrm{bound})}1/T_{2(\mathrm{bound})} \qquad (1)$$

where $X_{(\mathrm{free})}$ and $X_{(\mathrm{bound})}$ are the mole fraction (of ligand) in free solution and bound at the receptor, and $1/T_{2(\mathrm{free})}$ and $1/T_{2(\mathrm{bound})}$ are the respective spin lattice relaxation rates of the protons in the free and bound forms. To a good approximation the linewidth of an NMR line is directly related to the relaxation rate $1/T_2$, so the linewidth (ν – measured in Hz) can be substituted into equation (1) and we have a relation between some NMR observables and species distribution, and therefore K_D.

$$\nu_{(\mathrm{obs})} = X_{(\mathrm{free})}\nu_{(\mathrm{free})} + X_{(\mathrm{bound})}\nu_{(\mathrm{bound})} \qquad (2)$$

For formation of a 1 : 1 complex

$$K_D = [P][L]/[PL] \qquad (3)$$

where K_D is the stability constant and [P], [L] and [PL] are the equilibrium concentrations of protein, ligand and complex. After defining some further relationships between the equilibrium concentrations and the known total concentrations of protein $[P]_0$ and ligand $[L]_0$ it is possible to write quadratic equations that define the relationship between NMR observables, solution compositions and K_D. Because these relationships are non linear it is not possible to measure K_D from a single NMR data point. Instead, a series of experiments are performed at different concentrations of $[P]_0$ and $[L]_0$ and data analysis is accomplished by fitting a calculated curve to the data. Essentially the experiment is simple – the linewidth of an appropriate ligand signal is measured over a range of protein/ligand ratios and the observed data are fitted to a calculated curve that describes the system analytically. In practice, some care is required with issues such as manipulation of protein solutions, correct measurement of linewidths and validation of data treatments before the methodology becomes routine.

Results from a typical titration are shown in Fig. 3. This data is for 4-(1*H*-imidazol-1-yl)benzenecarboxamide **7**. The protein produces a broad envelope that can be seen to increase in intensity as the protein concentration increases and these signals are ignored. At low ligand ratios (top) the small molecule signals are broad and they are seen to sharpen as the ligand ratio is increased. This is because the spectrum is increasingly dominated by the sharp signals of the non-bound component as the ratio of unbound ligand to bound ligand increases. Any convenient signal can be chosen for the linewidth measurement. In this example the downfield doublet was chosen because it was well resolved and was readily simulated. The relationship between linewidth and $[L]_0$ is illustrated in Fig. 4, which also includes the calculated binding curve that defines K_D.

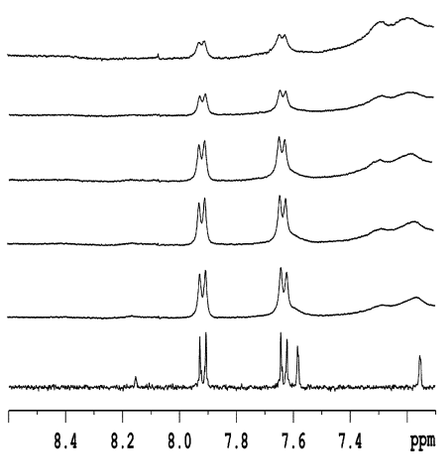


Fig. 3 Changes in the downfield region of the ^{1}H spectrum of **7** at different ratios of factor Xa. The lower trace is pure ligand ($[P]_0 = 0$), and the other traces, reading from top to bottom are at $[L]_0/[P]_0 = 2 : 1$, 4 : 1, 6 : 1, 8 : 1, and 10 : 1. The double doublet ($J = 8.6$ Hz) is the AA′BB′ pattern from the *para* substituted phenyl ring. The imidazole protons appear as broadened singlets because all of the expected *ortho* and *meta* couplings are less than 1.5 Hz and are unresolved. The H-2 signal at 8.15 ppm is attenuated by deuterium exchange with the solvent.[23]

There are several possible reasons for the poor fit between the calculated and experimental data in Fig. 4. The assumed 1 : 1 binding model may not be correct. Data points at the high $[L]_0/[P]_0$ ratio region of the titration may be skewed if the protein has non-specific lower affinity binding sites. The assumption that the observed linewidth is determined only by the weighted average of the bound and non-bound relaxation rates may be too simplistic. The observed NMR linewidth is more properly described as the sum of the weighted $1/T_2$, plus a contribution from magnet inhomogeneity, plus a possible exchange contribution.[24,25] The magnet inhomogeneity is negligible in this case. A contribution from exchange broadening seems unlikely given that there is no detectable dependency of chemical shift on $[P]_0$, but cannot be ruled out. We decided that the small number of data points in the titrations did not justify the use of more complex models and their additional degrees of freedom.

Table 1 Result of quantitative K_D work and competition binding experiments with the high affinity reference ligand TPAM (**9**)

Ligand	K_D/μM	$\nu_{(bound)}$	Displaced[a]
6	30	53	No
5	80	49	Yes
8	110	51	Yes
7	210	48	Yes

[a] See Fig. 5.

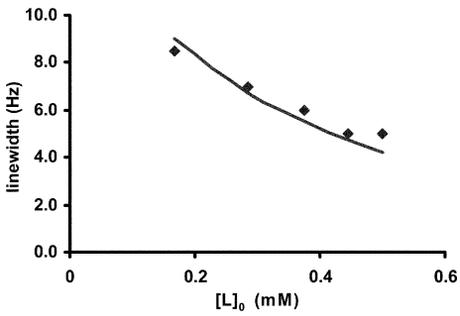
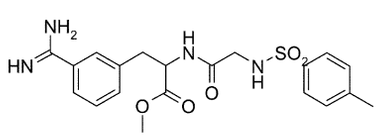


Fig. 4 Effect of factor Xa on the linewidth of aromatic protons of **7**. This presentation shows the linewidth as a function of ligand concentration. The protein concentration varies from *ca.* 50 to 83 μM. The solid line is the predicted curve for a 1 : 1 complex with K_D = 210 μM, $\nu_{(free)}$ = 1.7 Hz and $\nu_{(bound)}$ = 47.6 Hz.

The results of the quantitative binding experiments are summarised in Table 1. Note that as well as giving K_D, the analysis also gives $\nu_{(bound)}$, the linewidth of the ligand NMR signal when constrained in the protein complex. It is notable that the linewidths of all the ligands in complexed form are approximately the same at 50 ± 3 Hz. This is an encouraging self consistency in the data because $1/T_{2(bound)}$ is related to the size of the protein complex. This value is also similar to those quoted for other small molecules bound to proteins, *e.g.* serum albumins.[26]

Competition binding

The residual solutions from the K_D experiments contained approximately a 12 : 1 ratio of ligand to protein with ligand in fast exchange between bound and non-bound sites. We have already discussed how dynamic averaging of the sharp (*c.* 1–2 Hz linewidth) non-bound spectrum with the broad (*c.* 50 Hz) bound form broadens the ligand signals. The basis of the competition binding experiment was to add a material with established high affinity for the factor Xa binding site on the premise that this new high affinity and specific ligand should displace the screening hit. Such displacement would be signalled by a sharpening of the 1D ^{1}H spectrum, or a decrease in intensity of the saturation transfer difference spectrum of the small molecule. The high affinity ligand is *N*α-tosylglycyl-3-amidinophenylalanine methyl ester hydrochloride (TPAM, or Pefabloc, **9**, K_D = 2 μM).[27] This material has been used as a basis for inhibitor design and the X-ray crystal structure of the related trypsin complex of this ligand shows that the benzamidine group is in the S1 pocket and the tolylsulfonamide is in the S4 pocket.[28]



TPAM (**9**)

Addition of 0.1 mg TPAM had no significant effect on either the linewidths or the intensity of the saturation transfer difference spectra of **6** with factor Xa. The linewidths of the solutions containing ligands **5**, **7** and **8** were not significantly affected by TPAM, but the saturation transfer difference spectra show clearly that **5** and **7** are displaced and **8** is displaced somewhat more weakly (Fig. 5). Thus these experiments show that ligands **5** and **7** (and probably **8**) bind at the same site as TPAM. These experiments also demonstrate that the saturation transfer difference method is a very effective indicator signal of binding, and is preferred to simple linewidth measurements.

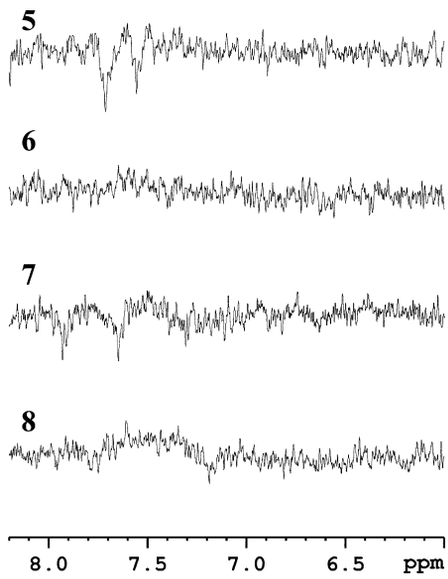


Fig. 5 Results of competition binding experiments with TPAM. The traces are the difference between saturation transfer difference spectra acquired with and without the competition ligand. The spectra correspond to the appearance in solution of free ligand that was displaced from the protein binding site. The noise in this data is high, but clear negative signals can be seen from the aromatic signals of ligands **5** and **7** at 7.7/7.55 ppm and 7.85/7.65 ppm respectively. The phenyl ring protons of **8** are less distinct, but recognisable at 7.8 and 7.2 ppm.

Factor Xa binding sites

The two strongly basic fragments benzamidine (**5**) and 2-guanylbenzimidazole (**6**) appeared as the strongest hits and the follow up studies showed that they were the highest affinity ligands for factor Xa. This is to be expected because of the known high affinity of the S1 pocket for the amidine fragment. Hence the screening phase successfully identified the key moiety for good factor Xa ligands. Benzamidine was also clearly displaced by TPAM indicating that benzamidine is definitely bound at the protein active site. The failure to observe a displacement of 2-guanylbenzimidazole (K_D 30 μM) by TPAM (K_D 2 μM) may simply be due to the high affinity of 2-guanylbenzimidazole.

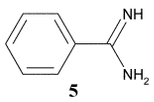
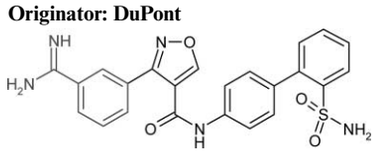
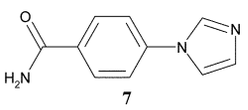
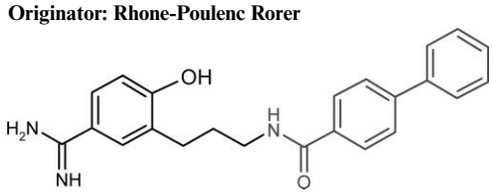
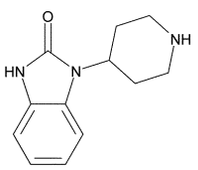
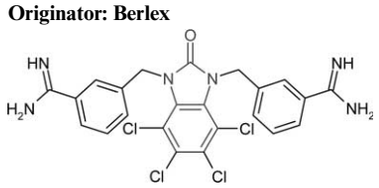
4-(2-Keto-1-benzimidazolinyl)piperidine (**8**) had the next highest affinity of factor Xa (K_D 110 μM) and was weakly displaced by TPAM. 4-(1*H*-Imidazol-1-yl)benzenecarboxamide (**7**) was the weakest affinity ligand that was followed up, and this ligand was clearly displaced by TPAM. It seems likely that **7** and **8** are good ligands for the S4 pocket.

Some ligands (*e.g.* pyridine-4-carboximidamide) were expected to have some affinity for factor Xa, but were not picked up in the primary screen. The binding of aminoisoquinoline was also found to be weak although this moiety is known from X-ray crystallography studies to bind to the S4 pocket of the binding site and is present in another class of factor Xa inhibitors, the weakly basic factor Xa inhibitors, *e.g.* **3**.

From low affinity fragments to lead generation

Having identified three competitive low affinity, low molecular weight fragments (compounds **5**, **7**, and **8**), the next step is to hypothesise the site of binding so library explosion around the

Table 2 Appearance of hit fragments in known factor Xa ligands

Low affinity hit	Pocket	Known lead
 **5**	**S1/S4**	**Originator: DuPont** 
 **7**	**S4**	**Originator: Rhone-Poulenc Rorer** 
 **8**	**Core**	**Originator: Berlex** 

fragments can be efficiently guided. At this stage, however, instead of starting a chemistry program on these fragments, we took the approach of examining whether those fragments, or fragments isosteric to them, are known to be present in factor Xa inhibitors. The results are presented in Table 2.

Benzamidine (**5**) is a common fragment binding at the S1 pocket in serine protease inhibitors in general but also at the S4 pocket in factor Xa inhibitors in particular. Therefore, not surprisingly, many factor Xa inhibitors could be identified containing a benzamidine moiety. Three examples are given in Table 2, one of them having a bis-benzamidine, thus providing evidence of the ability of this fragment to bind to both S1 and S4 pockets. Interestingly, few substituted benzamidines were present in our compound screening collection. High-throughput screening (using a chromogenic substrate to measure anti-factor Xa activity) gave percentages of inhibition between 9%–25%@10^{-5} M for all those benzamidine analogues. Consequently, this fragment was completely overlooked, even though it is one of the most common fragments for the S1 pocket in serine protease inhibitors. This is a revealing example of the complementary role that structure-based NMR screening can play alongside conventional high-throughput screening.

Compound **7** represents, to the best of our knowledge, a novel active fragment in factor Xa inhibitors. However, isosteric fragments to compound **7** can be indeed found in known factor Xa inhibitors binding to the S4 pocket. An example is shown in Table 2, where the imidazole moiety in compound **7** is replaced by a phenyl ring in the Rhone-Poulenc Rorer compound.

Finally, a substructure search of compound **8** did not lead to any factor Xa hits in our database, which again reflects the potential novelty of the fragments that can be identified by NMR screening. However, although the substructure *per se* was not found in known factor Xa inhibitors, the core was indeed present in the central scaffold of a bis-benzamidine series of factor Xa inhibitors from Berlex. Therefore, compound **8** will be assigned to bind in the core part of the binding cavity, with the carbonyl group potentially making a hydrogen bond with the backbone of Gly-216.

Conclusion

This work demonstrates that a single spectroscopy technique may be used successfully to study all of the important events related to ligand–receptor interactions (detection of binding, quantitation of K_D, and validation of specificity), and each step can be studied independently by following a logical flow of protocols. The use of a single experimental technique allowed the work to be performed with a small project team and was economical with materials. The procedure can be made even more efficient by performing the screening in the presence of a medium affinity ligand, as has recently been proposed by several workers.[29,30] Such an approach rolls all of the steps together, and can also give a simultaneous measure of K_D.

Experimental

General

Human factor Xa was purchased from Haemotological Technologies Inc, Vermont, USA, as a powder lyophilised from solutions buffered with ammonium carbonate. TPAM was purchased from Kordia Life Sciences, Leiden, the Netherlands. Stock solutions of the small molecule ligands were made at concentrations of 100 mM or 250 mM in DMSO-d_6 and stored at −20 °C. Reference NMR spectra of the individual compounds were obtained in the appropriate buffer system by adding stock solutions such that the final compound concentrations were 500 μM in 0.4 ml. The samples prepared for screening against factor Xa contained three ligands at 500 μM each, 35 μM factor Xa, 100 mM deutero-TRIS, 5% deuteroglycerol at pD 8.2. The ligands were studied as mixtures of three in order to shorten the time the protein spends in solution, reducing the possibility of protein degradation, aggregation and solubility issues.

NMR screening

NMR screening was carried out on a Bruker DRX 400 MHz spectrometer equipped with a Bruker Efficient Sample Transport (BEST) and operating under ICONNMR. This equipment consists of a Gilson liquid handler interfaced to a 3 mm flow probe for high throughput NMR data collection. 1D solvent suppressed spectra were obtained with presaturation of the residual water peak, with a relaxation delay of 1.5 s and 512 scans. The saturation transfer difference spectra[31] were obtained using a continuous wave saturation of 1.6 s at 0.8 ppm. The reference spectrum was obtained by saturating at 30 ppm. The solvent suppression was achieved using WATERGATE.[32] The number of scans was 1024, giving a total

experimental time for both spectra of 1 hour 15 minutes per sample. Acquisition data were processed using XWINNMR version 2.5.

Binding affinities

Hits identified from the primary screen were diluted to 1 mM with deuterated 100 mM TRIS buffer, pD 8.2, for use as stock solutions in the next phase (giving a DMSO-d_6 content of between 0.4 and 1% in D_2O). Protein solutions were prepared in this same buffer, with protein concentration determined by UV, and using an extinction coefficient of $E^{1\%}_{1\text{ cm, }280\text{ nm}} = 11.6$. Binding constants were measured from the 1D ^{1}H NMR spectra of a series of solutions at varying ligand/protein ratios in 5 mm NMR tubes. Stock solutions were 100 μM factor Xa and 1 mM ligand. The starting point consisted of 500 μl of the protein solution plus 100 μl ligand. After recording the spectrum a further 100 μl ligand solution was added and the spectrum was recorded again. This sequence was repeated five times so that data from solutions containing the following compositions were collected – $[P]_0$, μM, $[L]_0$, μM; 83.3, 167; 71.4, 286; 62.5, 375; 55.6, 444; 50.0, 500; corresponding to ligand/protein ratios ($[L]_0/[P]_0$) of 2 : 1, 4 : 1, 6 : 1, 8 : 1 and 10 : 1. A reference NMR spectrum of each ligand at a concentration of 500 μM was also collected. NMR spectra were collected on a Bruker DRX 400 MHz spectrometer. Both 1D ^{1}H and 1D solvent suppressed spectra were obtained, with an average of 1000 scans, leading to an experiment time of about one hour per data point. The ^{1}H linewidths were measured by fitting the experimental spectral line to a simulated pattern using the MestRe-C programme.[33] Fitting a simulated binding curve to the experimental data (estimation of K_D) was performed with a home written spread sheet. This is available from the authors upon request.

Competition binding

The materials used for quantitative K_D work were carried forward to the competition binding experiments. These solutions consisted of 39 μM factor Xa plus 500 μM ligand in approximately 1.0 mL of buffered D_2O. The integrity of the samples was checked by recording a conventional 1D ^{1}H spectrum. 0.1 mg TPAM was added to each tube (190 μM) and the 1D ^{1}H spectrum and saturation transfer difference spectrum were recorded again. The on resonance saturation was set at 0.917 ppm and off resonance was set at 34.19 ppm, with a saturation duration of 1.6 s. WATERGATE was used to suppress the water signal.

References

1 B. J. Stockman and C. Dalvit, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2002, **41**, 187–231.
2 R. Carr and H. Jhoti, *Drug Discovery Today*, 2002, **7**, 522–527.
3 S. W. Muchmore and P. J. Hajduk, *Curr. Opin. Drug Discovery Dev.*, 2003, **6**, 544–549.
4 J. Fejzo, C. A. Lepre, J. W. Peng, G. W. Bemis, W. Guy, M. A. Murko and J. M. Moore, *Chem. Biol.*, 1999, **6**, 755–769.
5 M. J. P. van Dongen, J. Uppenberg, S. Svensson, T. Lundbäck, T. Åkerud, M. Wikström and J. Schultz, *J. Am. Chem. Soc.*, 2002, **124**, 11874–11880.
6 G. W. Bemis and M. A. Murko, *J. Med. Chem.*, 1996, **39**, 2887–2893.
7 E. W. Davie, K. Fujikawa and W. Kisiel, *Biochemistry*, 1991, **30**, 10363–10370.
8 W. R. Ewing, H. W. Pauls and A. P. Spada, *Drugs Future*, 1999, **24**, 771–787.
9 A. E. P. Adang and J. B. M. Rewinkel, *Drugs Future*, 2000, **25**, 369–383.
10 K. Padmanabhan, K. P. Padmanabhan, A. Tulinsky, C. H. Park, W. Bode, R. Huber, D. T. Blankenship, A. D. Cardin and W. Kisiel, *J. Mol. Biol.*, 1993, **232**, 947–966.
11 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.
12 S. Maignan, J.-P. Guilloteau, Y. M. Choi-Sledeski, M. R. Becker, W. R. Ewing, H. W. Pauls, A. P. Spada and V. Mikol, *J. Med. Chem.*, 2003, **46**, 685–690.
13 Z. Lin and M. E. Johnson, *FEBS Lett.*, 1995, **370**, 1–5.
14 S. Katakura, T. Nagahara, T. Hara and M. Iwamoto, *Biochem. Biophys. Res. Commun.*, 1993, **197**, 965–972.
15 M. L. Quan and R. R. Wexler, *Curr. Top. Med. Chem.*, 2001, **1**, 137–149.
16 Y. M. Choi-Sledeski, M. R. Becker, D. M. Green, R. Davis, W. R. Ewing, H. J. Mason, C. Ly, A. Spada, G. Liang, D. Cheney, J. Barton, V. Chu, K. Brown, D. Colussi, R. Bentley, R. Leadley, C. Dunwiddie and H. W. Pauls, *Bioorg. Med. Chem. Lett.*, 1999, **9**, 2539–2544.
17 Y. K. Yee, A. L. Tebbe, J. H. Linebarger, D. W. Beight, T. J. Craft, D. Gifford-Moore, T. Goodson, Jr., D. K. Herron, V. J. Klimkowski, J. A. Kyle, J. S. Sawyer, G. F. Smith, J. M. Tinsley, R. D. Towner, L. Weir and M. R. Wiley, *J. Med. Chem.*, 2000, **43**, 873–882.
18 Molecular Design Limited (MDL), 14600 Catalina St., San Leandro, CA 94577.
19 Our in-house data base contains *c.* 15000 drugs reported primarily in *Drugs of the Future*, *Drug Data Report*, MartinDale, Pharmaprojects, *USP Dictionary of USAN*, and *International Drug Names*.
20 B. Meyer and T. Peters, *Angew. Chem., Int. Ed.*, 2003, **42**, 864–890.
21 K. A. Connors, *Binding Constants: The Measurement of Molecular Complex Stability*, John Wiley and Sons, New York, 1987.
22 L. Fielding, *Curr. Top. Med. Chem.*, 2003, **3**, 39–53.
23 Y. Takeuchi, H. J. C. Yeh, K. L. Kirk and L. A. Cohen, *J. Org. Chem.*, 1978, **43**, 3565–3570.
24 J. J. Fischer and O. Jardetzky, *J. Am. Chem. Soc.*, 1965, **87**, 3237–3244.
25 S. H. Smallcombe, B. Ault and J. H. Richards, *J. Am. Chem. Soc.*, 1972, **94**, 4585–4590.
26 C. Coibion and P. Laszlo, *Nouv. J. Chim.*, 1978, **2**, 309–316.
27 J. Stürzebecher, U. Stürzebecher, H. Vieweg, G. Wagner, J. Hauptmann and F. Markwardt, *Thrombosis Res.*, 1989, **54**, 245–252.
28 B. Gabriel, M. T. Stubbs, A. Bergner, J. Hauptmann, W. Bode, J. Stürzebecher and L. Moroder, *J. Med. Chem.*, 1998, **41**, 4240–4250.
29 W. Jahnke, P. Floersheim, C. Ostermeier, X. Zhang, R. Hemmig, K. Hurth and D. P. Uzunov, *Angew. Chem., Int. Ed.*, 2002, **41**, 3420–3423.
30 A. H. Siriwardena, F. Tian, S. Noble and J. H. Prestegard, *Angew. Chem., Int. Ed.*, 2002, **41**, 3454–3457.
31 M. Mayer and B. Meyer, *Angew. Chem., Int. Ed.*, 1999, **38**, 1784–1788.
32 V. Sklenar, M. Piotto, R. Leppik and V. Saudek, *J. Magn. Reson., Ser. A*, 1993, **102**, 241–245.
33 http://qobrue.usc.es.