

Development and NMR validation of minimal pharmacophore hypotheses for the generation of fragment libraries enriched in heparanase inhibitors

Rafael Gozalbes · Silvia Mosulén ·
Rodrigo J. Carbajo ·
Antonio Pineda-Lucena

Received: 15 January 2009 / Accepted: 9 April 2009 / Published online: 7 May 2009
© Springer Science+Business Media B.V. 2009

Abstract A combined strategy based on the development of pharmacophore hypotheses and NMR approaches is reported for the identification of novel inhibitors of heparanase, a key enzyme involved in tumor metastasis through the remodeling of the subepithelial and subendothelial basement membranes, resulting in the dissemination of metastatic cancer cells. Several pharmacophore hypotheses were initially developed from the most active heparanase inhibitors known to date and, after their application to a pool of 27 known heparanase inhibitors and a database of 1,120 compounds approved by the FDA, a four-point pharmacophore model was selected as the most predictive. This model was subsequently applied to a database of 686 chemical fragments, and a subset of 100 fragments accomplishing completely or partially the four-point model was selected to perform nuclear magnetic resonance experiments to validate the hypothesis. The experimental studies confirmed the reliability of our pharmacophore model, its applicability to in silico databases in order to reduce the number of compounds to be experimentally screened, and the possibility of generating fragment libraries enriched in heparanase inhibitors.

Keywords Pharmacophore · Fragment-based screening · NMR · Heparanase · Inhibitor

Introduction

Heparanase is an endoglycosidase that acts at the cell-surface and within the extracellular matrix to degrade polymeric heparan sulfate molecules (Fig. 1) into shorter oligosaccharides [1, 2]. Heparanase is involved in several physiological activities such as embryo development, hair growth and wound healing, and it is also implicated in cancer processes such as tumor angiogenesis and metastasis [3, 4]. Increased levels of heparanase have been found in numerous cancer processes at different organs and there is a direct correlation between the overexpression of heparanase and the invasiveness of tumor cells [5]. The implication of heparanase in cancer progression makes it a very attractive target for anti-angiogenic, anti-metastatic and/or anti-inflammatory therapies. The three-dimensional structure of the enzyme has not yet been elucidated, but the active-site residues of human heparanase have been described, and it is widely accepted that there are two essential acidic residues (Glu²²⁵ and Glu³⁴³) involved in the catalytic mechanism, acting as a putative proton donor and a nucleophile, respectively [6].

Despite the increasing interest in finding effective inhibitors against heparanase, and although a number of compounds with some inhibitory effect have been described, only one of them has reached so far the clinical trial phases [7]. Heparanase inhibitors reported until now differ in their chemical nature, origin (from natural products to synthetic compounds found through virtual or experimental screening of chemical libraries) and size (from big heparin-like polymers and oligosaccharide mimetics to small molecules with some “drug-like” features) [8].

Pharmacophore modeling is a useful technique for describing interactions of small molecules with macromolecular targets of therapeutic interest [9–11]. The IUPAC

Rafael Gozalbes and Silvia Mosulén contributed equally to this work.

R. Gozalbes · S. Mosulén · R. J. Carbajo ·
A. Pineda-Lucena (✉)
Structural Biology Laboratory, Department of Medicinal
Chemistry, Centro de Investigación Príncipe Felipe, Avenida
Autopista del Saler 16, 46012 Valencia, Spain
e-mail: apineda@cipf.es

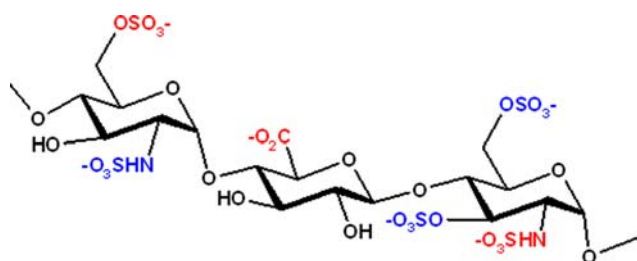


Fig. 1 Minimum heparan sulfate recognition sequence for heparanase. The essential charged groups required for heparanase recognition are highlighted in red, and other charged groups which may improve substrate affinity are shown in blue. Adapted from [8]

defines a pharmacophore as “the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger (or to block) its biological response” [12]. Pharmacophores describe the spatial arrangement of the chemical features necessary for small compounds to interact with the target, and they can be determined either from the complementarities of a ligand interacting with a binding site (structure-based pharmacophores) or, most frequently, from an active compound or the alignment of a set of known active molecules (ligand-based pharmacophores). Pharmacophore representations of ligand-protein interactions can be used to query, in virtual databases, for compounds that have the same or similar distribution of chemical features. Therefore, they provide a way to identify compounds that could potentially interact with the target in the same manner and produce a similar biological activity [11].

The absence of the 3D structure of heparanase precludes the development of structure-based pharmacophores, but the existence of a number of small molecule heparanase inhibitors indicates the potential of ligand-based pharmacophore approaches to identify novel inhibitors against this target. Thus, in the context of the present study, several pharmacophore hypotheses were developed based on the alignment of the most active inhibitors of heparanase. Taking into consideration that the pharmacophore hypotheses were intended to be applied over a fragment library, special attention was paid to the identification of “minimal” pharmacophores, that is, hypotheses where the distances between the different features describing the pharmacophore were relatively short. The different hypotheses were evaluated by *in silico* applying them to a database of structures comprising compounds known to have some heparanase inhibitory activity and a set of compounds approved by the FDA. The quality of the pharmacophore hypotheses was determined by analyzing their ability to retrieve and rank the known inhibitors from the virtual database. Once the selected model was validated, we applied it to a database of chemical fragments,

composed by 686 compounds, and a subset of them, accomplishing totally or partially the pharmacophore model, was selected in order to perform an experimental validation that included the acquisition of a set of NMR and surface plasmon resonance experiments, thus providing a useful way to identify experimentally novel inhibitors against heparanase.

Materials and methods

Protein preparation

A recombinant heparanase construct (details will be described elsewhere) containing the catalytic site was expressed, from a modified version of the pET-15b (Novagen) expression vector containing a TEV cleavage site, in *Escherichia coli* BL21-DE3 codon plus (Stratagene) cells. Cells were grown at 37 °C to an absorbance at 600 nm of 0.6 and induced with 0.8 mM IPTG overnight (15 h) at 25 °C. The expressed protein was purified using metal affinity chromatography under native conditions. Unlabeled, uniformly (^{15}N) and selectively labeled (^{15}N -Glu and ^{15}N -Gly) protein preparations were produced in LB and M9, and NMR samples at different concentrations (5 μM for WaterLOGSY and STD experiments, and 100 μM for HSQC experiments) were prepared. Protein samples selectively labeled in Glu and Gly were used to confirm the interaction of any of the fragments with the catalytic glutamic amino acids (Glu²²⁵ and Glu³⁴³) and/or with the preceding glycine residues (Gly²²³ and Gly³⁴²) and the results were analyzed taking advantage of the in-house backbone assignment obtained for this heparanase construct.

Pharmacophore development

A database of 27 well characterized heparanase inhibitors (Table 1) was constructed *in silico*, in order to generate and validate pharmacophore hypotheses from it. The compounds were extracted from the review of Hammond et al. [8], drawn with MDL Isis Draw 2.5 (www.mdl.com) and exported to the Molecular Operating Environment (MOE) software (Chemical Computing Group, www.chemcomp.com) to develop the pharmacophore hypotheses. The pharmacophore development module in MOE allows the characterization of atoms by their chemical features (hydrogen bond donor or acceptor ability, aromatic character, hydrophobic nature, etc.), the generation of the subsequent pharmacophore hypotheses, and the searching of conformation databases using molecule annotations related to ligand-receptor binding. The plausibility of the theoretical models is measured by the good overlay of active compounds (since they share a common binding mode) and the ability to separate

Table 1 Known heparanase inhibitors

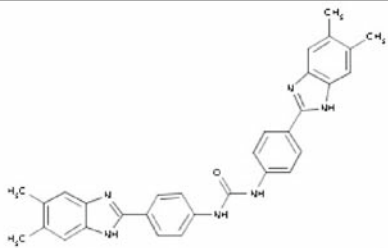
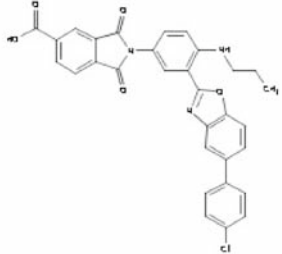
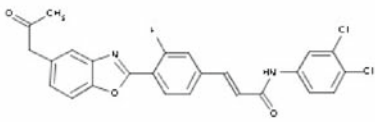
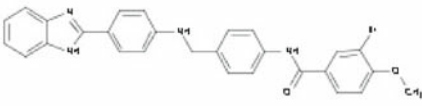
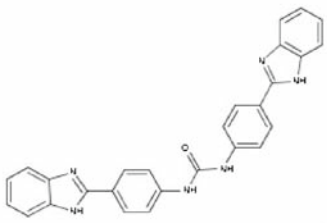
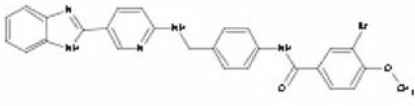
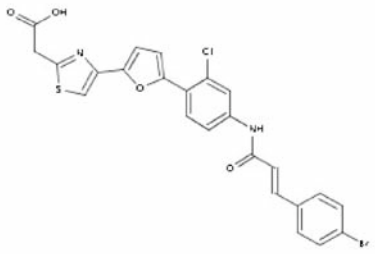
Molecule	IC ₅₀ (μM)	Ranking ^a	Structure
INH-1	0.075	2**	
INH-2	0.20	6*	
INH-3	0.20	17	
INH-4	0.23	4**	
INH-5	0.27	1**	
INH-6	0.29	3**	
INH-7	0.40	11*	

Table 1 continued

INH-8	0.50	5*	
INH-9	0.13 - 19.00	22	
INH-10	1.00	18	
INH-11	1.00 - 10.00	7*	
INH-12	1.50 - 36.00	23	
INH-13	2.50	12*	
INH-14	3.00	8*	
INH-15	5.00 - 45.00	21	
INH-16	8.00	20	

Table 1 continued

INH-17	8.00 - 25.00	16	
INH-18	10.00	15	
INH-19	10.00	14	
INH-20	10.00 - 26.00	N.S.	
INH-21	12.00	19	
INH-22	17.00	N.S.	
INH-23	24.00 - 32.00	10*	
INH-24	25.00	13*	
INH-25	26.00 - 36.00	9*	
INH-26	> 100.00	N.S.	
INH-27	> 100.00	N.S.	

A detailed description of these inhibitors can be found in the review by Hammond et al. [8]. “N.S.” stands for compounds “Not Satisfying” the pharmacophore model

^aA database consisting of the 27 known heparanase inhibitors plus the 1,120 structures from the Prestwick Chemical Library[®] was used to validate the pharmacophore model. The number in the column indicates the order in which the inhibitor was retrieved with respect to the other inhibitors when using the 1,147 chemicals. * Compounds ranked amongst the best 10% of the total database by the pharmacophore model, and ** those ranked amongst the best 2%

actives from inactives. In the context of our work, we developed a number of pharmacophore hypotheses based on the alignment of several subsets derived from the most active heparanase inhibitors (Table 1). Prior to the pharmacophore model development, multiple conformations of each of the 27 known inhibitors were generated with MOE, using the MMFF94x force-field and a 4 kcal/mol energy cutoff, with a maximum of 250 conformations for each compound. Bearing in mind the importance of validating models with unbiased decoys [13], we generated a virtual database consisting of the 27 known heparanase inhibitors plus a collection of 1,120 compounds (Prestwick Chemical Library[®]) that contained 90% marketed drugs (www.prestwickchemical.com), also subjected to the same conformational protocol described previously, for the evaluation of the pharmacophore hypotheses. The different pharmacophore models were evaluated by their accuracy in ranking the known inhibitors in the context of the total database.

Fragment library

A commercial database of 686 chemical fragments, most of them accomplishing the so called “rule of three” (molecular weight < 300 Da, number of hydrogen bond acceptors and hydrogen bond donors ≤ 3 , and $c\text{LogP} \leq 3$) [14] (Fig. 2) and with a relative high chemical diversity as assessed by a Tanimoto calculation using the MACCS keys furnished by MOE (0.30), was purchased from Cerep (www.cerep.com). The same conformational protocol mentioned previously for the 27 known inhibitors was used

for the compounds of the fragment library. Based on the best pharmacophore hypothesis developed, the fragment collection was classified as a function of the number of pharmacophore features satisfied.

Finally, a set of 100 fragments, accomplishing completely or partially the best pharmacophore model and with a high chemical diversity (Tanimoto coefficient of this subset was 0.27), was selected for the NMR experiments.

NMR experiments

All NMR spectra were recorded at 318 K with a Bruker Ultrashield Plus 600 MHz NMR spectrometer equipped with a 5 mm TCI cryogenically cooled probe. A typical NMR sample for WaterLOGSY experiments contained a concentration of 5 μM of protein, 300 μM of ligand (from a DMSO- d_6 stock at 50 mM), 50 mM NaCl, and 25 mM phosphate buffer at pH 6.0. For each sample, a 1D ^1H reference spectrum and a WaterLOGSY spectrum were recorded [15]. 8 K points were used for a sweep width of 9.6 kHz and a total of 1 K scans were accumulated for each WaterLOGSY spectrum. A relaxation delay of 2.5 s was applied with a mixing time of 1.5 s. The water selective pulse in the preparation part was 5 ms, while in the excitation sculpting [16] a 2 ms pulse was used. Gradients of 2 ms and 0.8 ms were applied in the preparation and excitation sculpting parts of the pulse sequence, respectively. An exponential function with a line broadening of 3 Hz was used previous to Fourier transform. Competition STD [17] experiments with suramine were carried out for

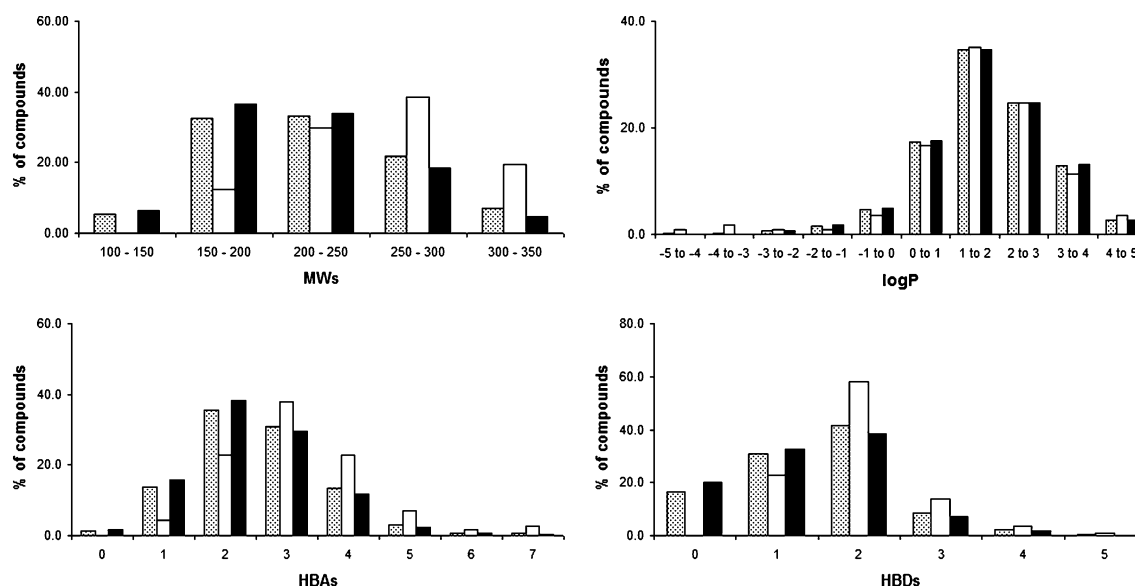


Fig. 2 Percentage of compounds as a function of molecular weights (MW), logP ranges, number of hydrogen bond acceptors (HBA) and donors (HBD). Dotted, white and black bars represent the complete

fragment database, compounds satisfying the four-point pharmacophore hypothesis and those not satisfying the pharmacophore model, respectively

each ligand, with a protein concentration of 5 μM and 500 μM for both ligand and suramine. For the STD experiments a train of Gauss-shaped pulses was applied with a final saturation time of 2.5 s with on/off resonance frequencies of 0.3 and 33.0 ppm. Total transients for STD spectra were 512. Heteronuclear 2D ^{15}N HSQC were acquired with spectral widths of 8 kHz for protons and 1.5 kHz for ^{15}N , accumulating a total of 16 scans. NMR data were processed using the program Topspin (Bruker GmbH, Karlsruhe, Germany).

Surface plasmon resonance experiments

All the surface plasmon resonance experiments were performed at room temperature using a Biacore T100 instrument. Heparanase was diluted with 10 mM sodium acetate buffer (pH 5) to a final concentration of 50 $\mu\text{g mL}^{-1}$ and immobilized on a carboxymethylated (CM5) sensor chip. The running buffer was 10 mM HEPES (pH 7.4), 150 mM NaCl, 50 μM EDTA, 0.005% surfactant P20. Activation of the sensor chip surface was performed with a mixture of N-hydroxysuccinimide (NHS, 115 mg mL^{-1}) and 1-ethyl-3(3-diaminopropyl) carbodiimide hydrochloride (EDC, 750 mg mL^{-1}) for 7 min at 10 $\mu\text{L min}^{-1}$. The amount of heparanase immobilized on the activated surface was 2,400 RU. After immobilization of the protein, a 7 min injection (at 5 $\mu\text{L min}^{-1}$) of 1 M ethanolamine (pH 8.5) was used to quench excess active succinimide ester groups.

Surface plasmon resonance binding experiments were performed in the same running buffer used for heparanase

immobilization plus 2% dimethyl sulfoxide (DMSO) to alleviate solubility problems of the fragments. The flow rate and the contact time were 30 $\mu\text{L min}^{-1}$ and 120 s, respectively. Fragment concentrations ranging from 5 μM to 1 mM were always used in these experiments. The sensor surface was regenerated between experiments removing any formed complex using 10 mM glycine (pH 2.5) for 30 s at 30 $\mu\text{L min}^{-1}$. Analysis of the binding curves was performed with software provided by Biacore.

Results and discussion

Pharmacophore model generation

Since we intended to apply the pharmacophore models to a database of fragments, we considered choosing a “minimal” hypothesis in which the distances between the pharmacophore points were relatively short. Taking this into account, we developed several pharmacophore hypotheses, and after their application to a set of compounds (the 27 known heparanase inhibitors of Table 1 plus the Prestwick Chemical Library) of comparable chemical complexity that guaranteed a computationally robust validation, we selected the model displayed in Fig. 3 as the most reliable and predictive, since most of the known inhibitors were correctly classified (Table 1). The selected model was developed from the alignment of the three most active inhibitors (INH-1, INH-2 and INH-3), and includes two hydrophobic/aromatic positions, a hydrogen donor and a hydrogen

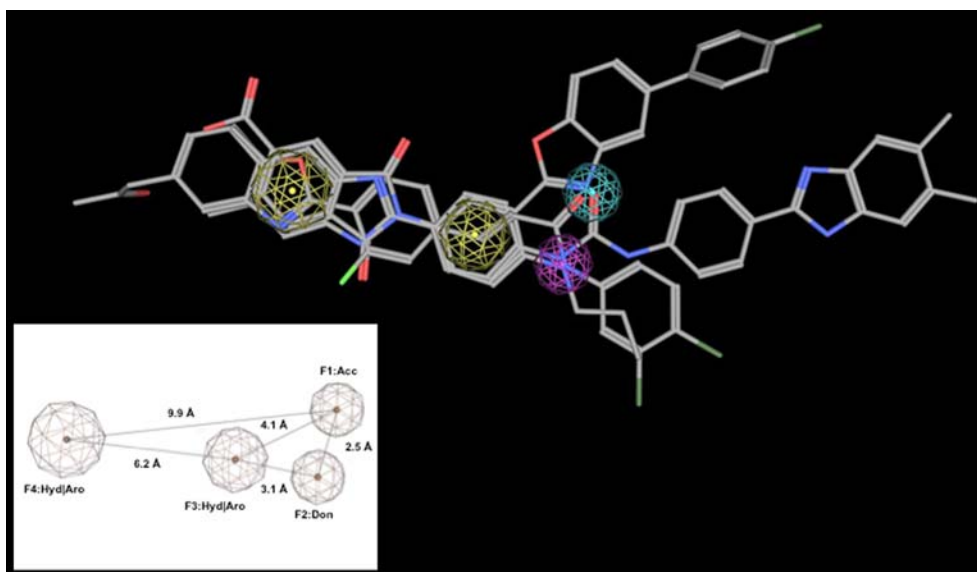


Fig. 3 Pharmacophore model obtained from the alignment of the three most active known heparanase inhibitors (INH-1, INH-2 and INH-3). Pharmacophore features include a hydrogen donor (magenta

sphere), a hydrogen acceptor (blue sphere) and two hydrophobic/aromatic regions (yellow spheres)

acceptor. The scoring function used for evaluating the ability of the different models in retrieving the known inhibitors and ranking them as good candidates was the root of the mean square deviation (rmsd) between the query features and their matching ligand annotation points, as defined by the MOE “Pharmacophore Search” tool.

The selected pharmacophore model was also applied to a commercial library composed by 686 fragments, finding that 114 of them (16.6% of the library) satisfied the four-point pharmacophore query. After evaluating the “rule of three” for these 114 compounds and comparing the results with those obtained for the complete fragment library, it was observed that compounds passing the pharmacophore filter were globally more complex than the rest of the library, thus having higher molecular weights as well as number of hydrogen bond donors and acceptors (Fig. 2). This is an expected result, specially considering that the largest distance between any of the pharmacophore points is nearly 10 Å (Fig. 3), thus making unattainable for the smallest fragments to satisfy this requirement.

Reliability of the pharmacophore model based on the known inhibitors

As shown in Table 1, twenty-three out of 27 compounds known to have heparanase inhibitory activity satisfied the four-point pharmacophore hypothesis, with the model rejecting only four compounds, two of them being the ones showing the lowest inhibitory activity (i.e., highest IC_{50} values, INH-26 and INH-27). A clear trend of activity was also found, with the best predicted compounds being the most active ones. Furthermore, we analyzed together the 27 reported inhibitors and the 1,120 structures from the Prestwick Chemical Library to evaluate the reliability of the model in retrieving the known inhibitors and ranking them as the best candidates. The results showed that 13 out of the 23 known inhibitors passing the pharmacophore filter were ranked amongst the best 10% compounds, and that for the eight inhibitors exhibiting sub-micromolar IC_{50} values (INH-1 to INH-8), four of them were amongst the 2% best scored compounds (including three compounds not used to develop the pharmacophore model), thus proving that, in the hypothetical context of an *in silico* screening campaign, choosing the best scored compounds would provide a higher percentage of hits.

Interestingly, we also applied the four-point pharmacophore model to the heparan sulfate (Fig. 1), the natural substrate of heparanase, and found out that this compound fully satisfied the complete model.

Chemical diversity and chemotype analysis

An interesting issue when applying pharmacophore approaches is the possibility of identifying structures with

the same pharmacophore features but very different chemical structures [11] and this also holds true in our case. As shown in Fig. 4, the pharmacophore model was also correctly satisfied by other inhibitors (e.g., INH-9 and INH-10, the best inhibitors from other chemotypes), despite the fact that the pharmacophore hypothesis was based on the alignment of three compounds from very similar chemical series. Fig. 5 shows the structures of the three fragments that scored best, as well as their superposition with the pharmacophore model, demonstrating that the same observation (i.e., different chemical structures compatible with the same pharmacophore model) could be extrapolated to the fragment library.

A further analysis of the known inhibitors was also carried out based on their particular chemical nature. Thus, Table 1 includes several inhibitors that, at first instance, could be considered to exhibit a similar behavior against heparanase taking into account that all of them are characterized by a long aliphatic chain. Nevertheless, the four-point pharmacophore model was able to adequately predict that INH-9, INH-18 and INH-19 were effective inhibitors of heparanase, since they satisfied the four-point criterion, and INH-26 and INH-27 were in fact “inactive” compounds (or at least they are the less active from the series, their IC_{50} being higher than 100 μ M). However, this analysis also shows the limitations of the model since INH-22, the closest structure to INH-26 and INH-27, did not satisfy the model, despite the fact of showing an improvement of one log of activity when compared to the other two compounds. It seems evident that the slight chemical difference between these three compounds is not appreciated by the model.

NMR validation of the pharmacophore model

To experimentally validate the pharmacophore model, we carried out a number of NMR experiments over a collection of 100 fragments (Table 2) accomplishing totally (50 compounds out of the 114 fragments passing the 4-point pharmacophore features, i.e., a 44% of that group) or partially (25 compounds satisfying 3-points, and other 25 fragments covering only 2-points) the pharmacophore model. The compounds of each subset were selected based on their chemical diversity with the exception of the group fulfilling completely the pharmacophore model that included the 18 best compounds scored by the pharmacophore model, plus 32 chemically diverse compounds accomplishing totally the model. The compounds were identified by their MACCS keys with MOE, and the Tanimoto algorithm was used for clustering them.

Initially, the validation consisted in the acquisition of a set of WaterLOGSY experiments over that subset of fragments. This sensitive method allows an easy identification of molecules binding to the target protein in the μ M-range,

Fig. 4 Structures (*top*) of the three most active heparanase inhibitors (from left to right, INH-1, INH-9 and INH-10) from different chemical subfamilies and superposition with the pharmacophore model (*bottom*)



and does not require neither labeled nor large amounts of biomolecule. The results of these studies showed that the higher the number of pharmacophore criteria fulfilled by the fragments, the higher the percentage of hits obtained. Thus, 66% of the compounds that satisfied all the pharmacophore features showed binding activity to heparanase in the WaterLOGSY studies, representing a 6 and 30% enrichment when compared to compounds accomplishing three or only two pharmacophore features, respectively.

These results show the potential of this approach in fragment-based hit generation where it could be useful to start testing a reduced number of fragments and then to proceed with a second phase of hit explosion by HTS. Also, this strategy has additional advantages for those researchers working in the academic environment where the possibility of managing big collections of compounds is somewhat limited.

NMR-based hit generation

Although the WaterLOGSY experiment is extremely useful in the identification of compounds interacting with a protein target, it has the disadvantage of not easily providing information about the binding site. Also, the signal intensity in the WaterLOGSY experiment greatly depends on the orientation and solvent accessibility [18], so it is very difficult to draw conclusions about affinity based on this information. Thus, and in attempt to identify novel inhibitors against heparanase, the subset of 100 fragments (Table 2) was

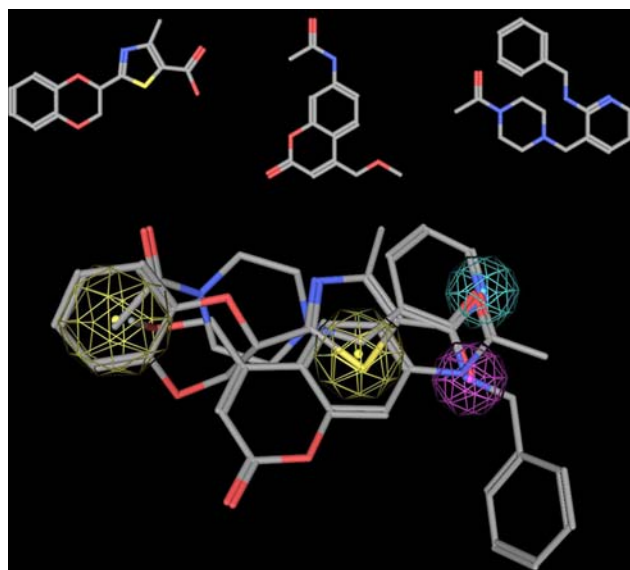


Fig. 5 Structures (*top*) of the three best scored compounds (from left to right, FRAG-8, FRAG-28 and FRAG-3) from the fragment library and superposition with the four-point pharmacophore model (*bottom*). As described in the text, the pharmacophore model is compatible with different chemical structures

subjected to saturation transfer difference (STD) and competition STD experiments (Fig. 6). The purpose of this exercise was first to reduce the number of false positives (WaterLOGSY and STD experiments on their own are error-prone and, very often, produce false positives due to changes in the pH, aggregation or non-specific binding [19]) by

Table 2 Subset of 100 fragments selected for the NMR validation of the pharmacophore model

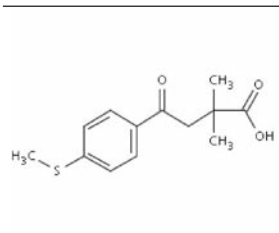
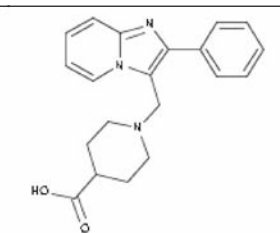
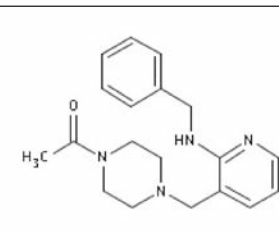
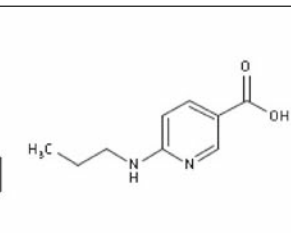
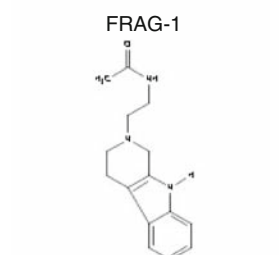
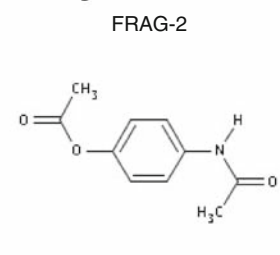
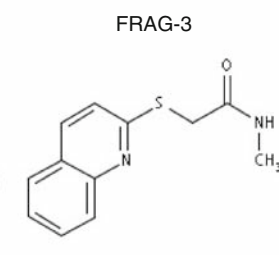
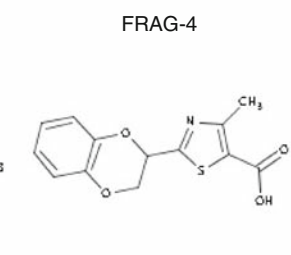
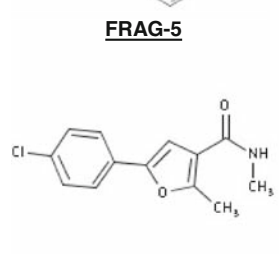
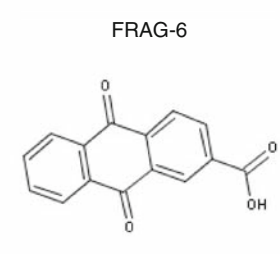
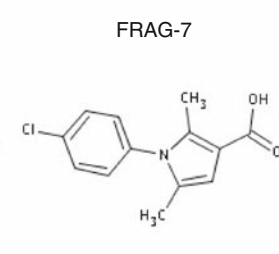
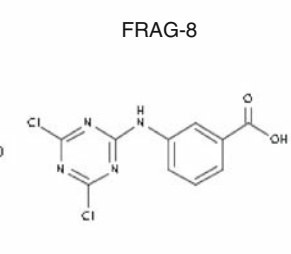
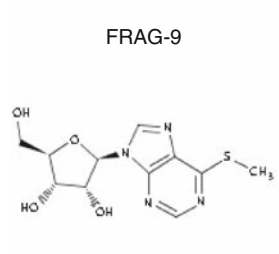
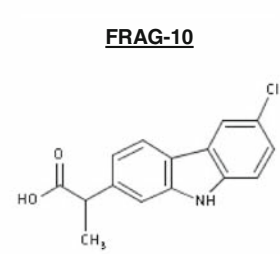
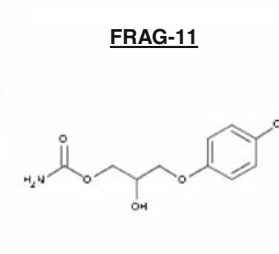
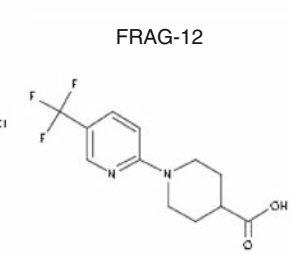
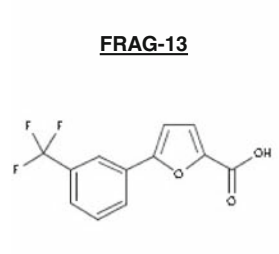
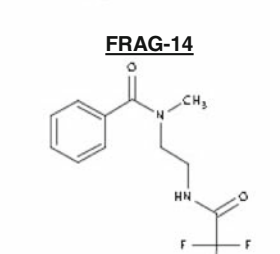
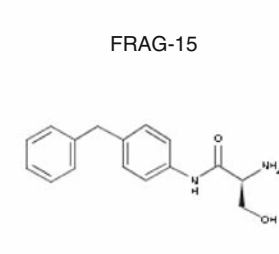
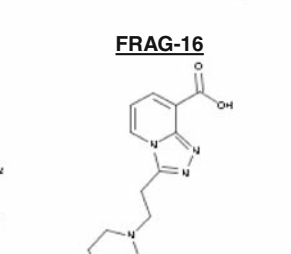
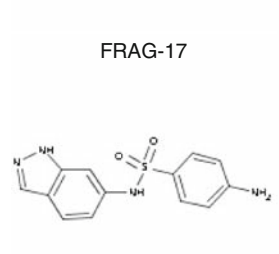
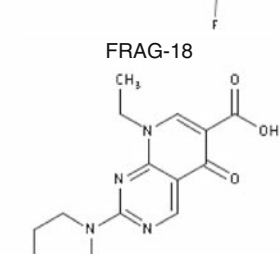
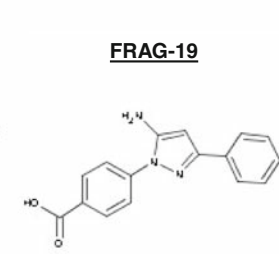
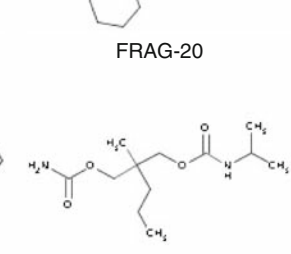
 FRAG-1	 FRAG-2	 FRAG-3	 FRAG-4
 FRAG-5	 FRAG-6	 FRAG-7	 FRAG-8
 FRAG-9	 FRAG-10	 FRAG-11	 FRAG-12
 FRAG-13	 FRAG-14	 FRAG-15	 FRAG-16
 FRAG-17	 FRAG-18	 FRAG-19	 FRAG-20
 FRAG-21	 FRAG-22	 FRAG-23	 FRAG-24

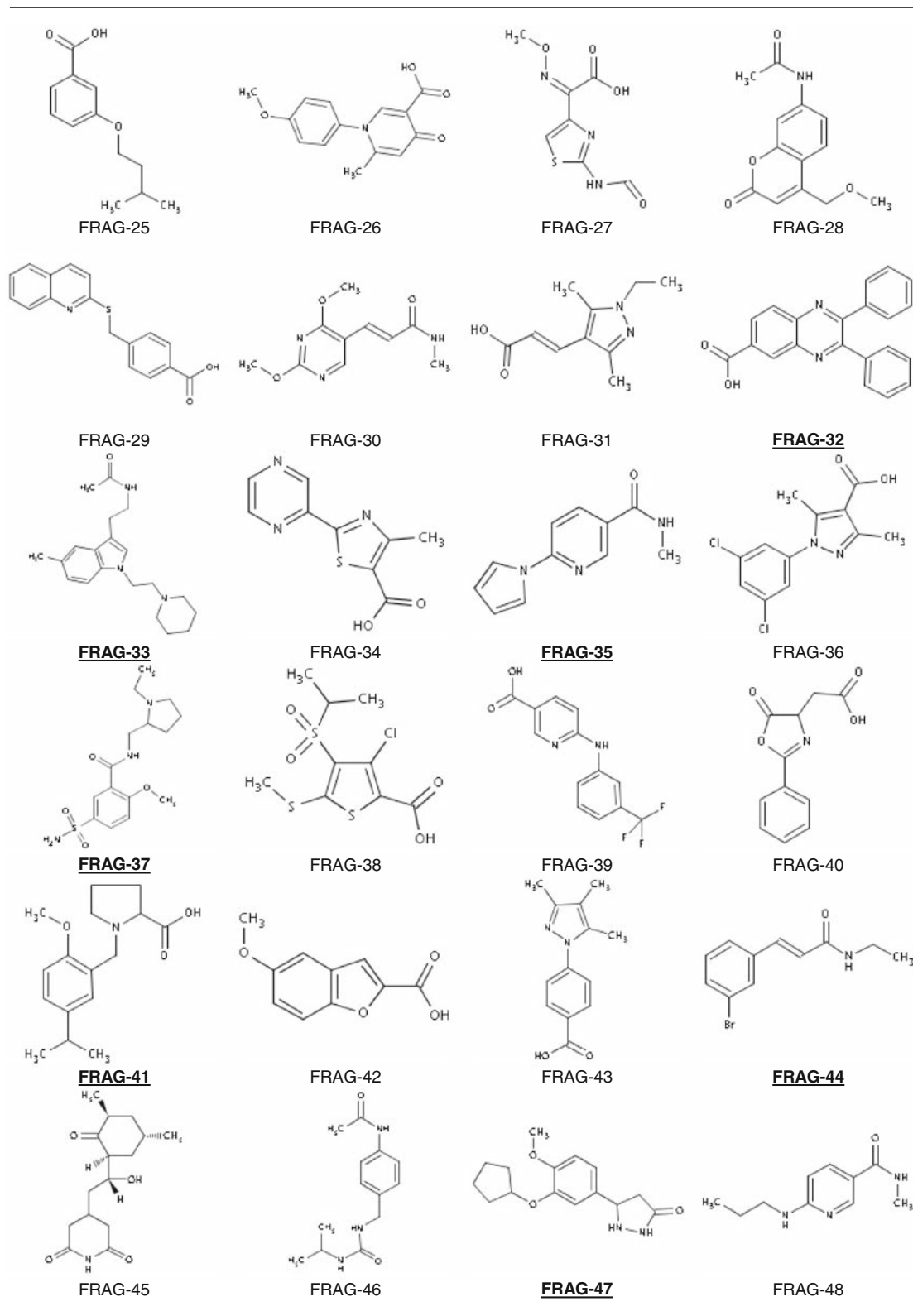
Table 2 continued

Table 2 continued

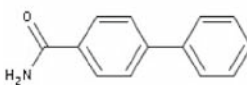
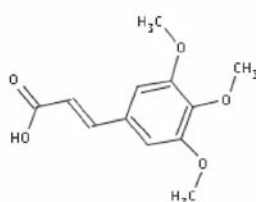
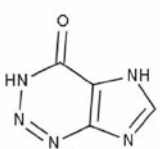
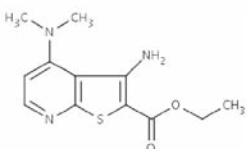
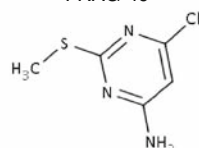
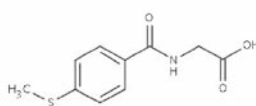
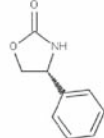
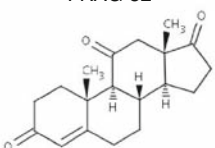
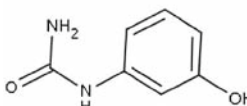
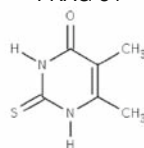
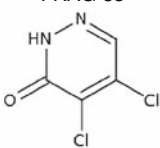
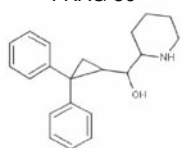
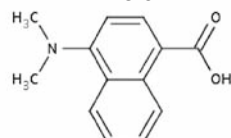
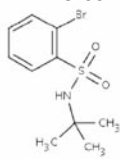
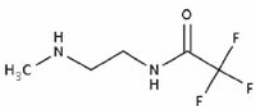
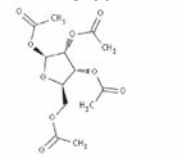
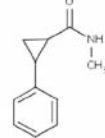
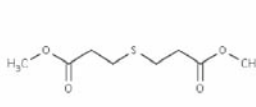
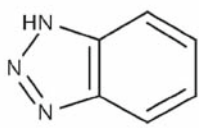
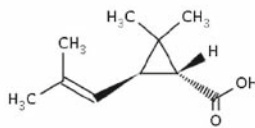
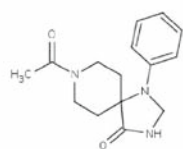
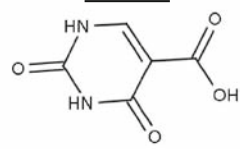
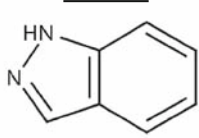
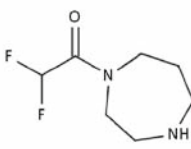
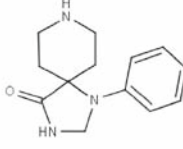
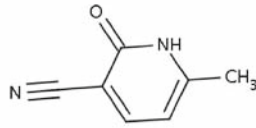
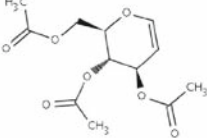
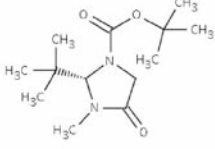
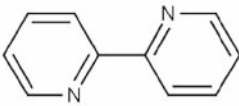
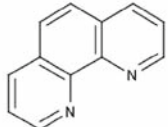
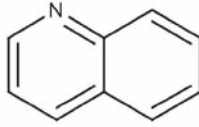
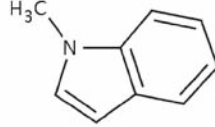
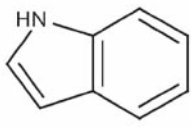
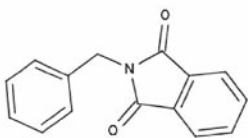
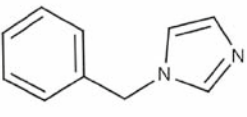
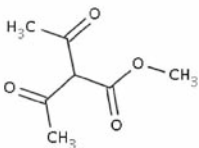
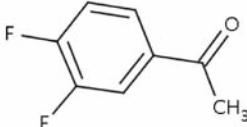
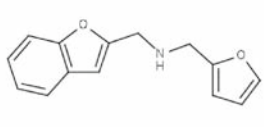
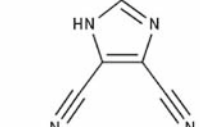
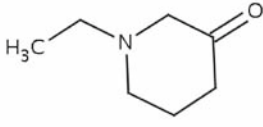
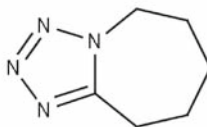
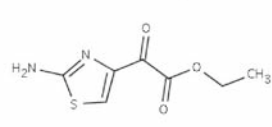
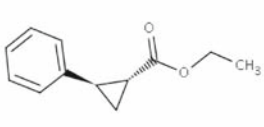
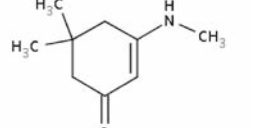
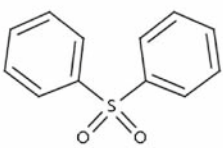
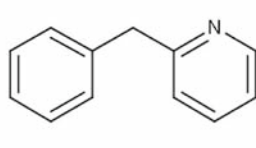
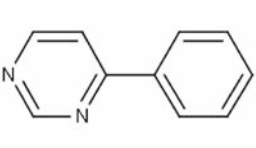
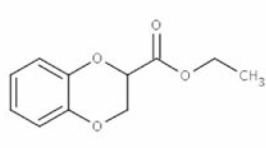
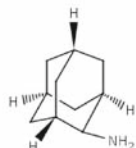
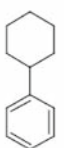
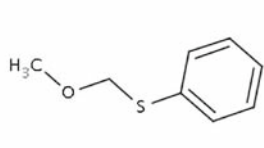
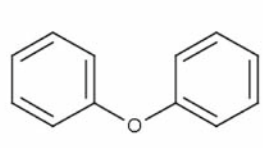
			
FRAG-49	FRAG-50	FRAG-51	FRAG-52
			
FRAG-53	FRAG-54	FRAG-55	FRAG-56
			
FRAG-57	FRAG-58	FRAG-59	FRAG-60
			
FRAG-61	FRAG-62	FRAG-63	FRAG-64
			
FRAG-65	FRAG-66	FRAG-67	FRAG-68
			
FRAG-69	FRAG-70	FRAG-71	FRAG-72
			
FRAG-73	FRAG-74	FRAG-75	FRAG-76
			
FRAG-77	FRAG-78	FRAG-79	FRAG-80
			
FRAG-81	FRAG-82	FRAG-83	FRAG-84

Table 2 continued

			
FRAG-85	<u>FRAG-86</u>	FRAG-87	FRAG-88
			
FRAG-89	FRAG-90	FRAG-91	FRAG-92
			
FRAG-93	FRAG-94	FRAG-95	FRAG-96
			
FRAG-97	FRAG-98	FRAG-99	FRAG-100

Fragments fulfilling 4 (FRAG-1 to FRAG-50), 3 (FRAG-51 to FRAG-75) and 2 (FRAG-76 to FRAG-100) features of the pharmacophore model. Marked in *bold* and *underlined* are those fragments of comparable chemical complexity found to be interacting with heparanase as judged by both WaterLOGSY and competition STD experiments

assessing the interactions by two independent methods [20], and second to obtain information about the binding site of the interacting compounds via competition experiments [21]. The competition STD experiments were acquired in the presence of suramine, a known ligand of heparanase ($K_D = 0.5 \mu\text{M}$). As it could be expected, when the interaction between the heparanase and the fragments was assessed by two independent methods (WaterLOGSY and STD experiments), we found a decrease in the hit rate for the whole subset. However, the comparison of fragments in the same size range (MW and topological distance [22]) showed that there is still a strong correlation between the fulfillment of the pharmacophore model and the hit rate. Thus, it was found that 30% of the fragments completely satisfying the pharmacophore model showed binding activity to heparanase, representing a 9% and 20% enrichment when comparing to compounds accomplishing three or two pharmacophore features, respectively (Table 2, bold and underlined). In some cases, and just for fragments we found to be interacting with heparanase as assessed both by the WaterLOGSY and the competition STD experiments, HSQC experiments were acquired. The HSQC-based chemical shift mapping

experiments proved that all the binders were indeed interacting with the catalytic site of heparanase (data not shown).

Finally, surface plasmon resonance studies were performed to measure the K_D of the interactions between the most promising compounds and heparanase. Fragments for these experiments were selected based on their overall profile in terms of the response in the WaterLOGSY and competition STD experiments. Most of them show dissociation constants in the mM range (1–5 mM) and binding efficiencies ($\text{BEI} = \text{p}K_D/\text{MW}$ (kDa)) equivalent to those reported [23] for fragments of that size. With only one exception, FRAG-86 (Table 2), the most promising compounds belonged to the subset of fragments completely fulfilling the 4-point pharmacophore model. This compound, in particular, is the largest (MW: 263.7 Da, heavy atoms: 17 atoms, diameter: 10) of the subset of fragments satisfying two features of the pharmacophore model. In fact, when looking in detail, this fragment is fulfilling almost three out of the four pharmacophore features. Using this approach, we have identified several fragments that, based on the biophysical experiments carried out, could potentially be good starting points for future hit-to-lead optimizations.

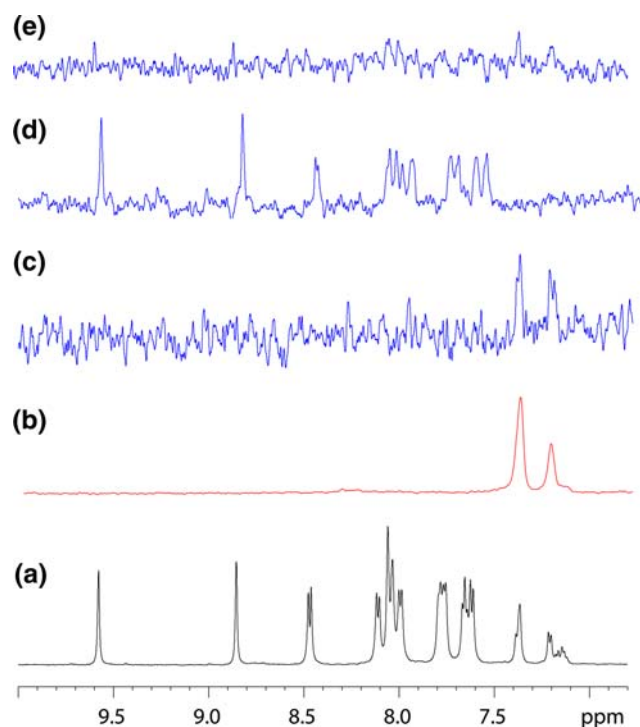


Fig. 6 Example of the set of NMR protein-ligand interaction experiments carried out for each fragment, showing a detail of the aromatic region of the spectrum. **a** ¹H spectrum of a mixture of suramine (known inhibitor) and FRAG-47 (500 μ M each); **b** Water-LOGSY spectrum of FRAG-47 (5 μ M heparanase, 300 μ M fragment) showing positive interaction; a non-interacting compound would show the same signals but inverted, whereas interacting ligands, in this experimental setting, are characterized by positive signals or resonances less negative than a reference spectrum without protein (not shown); **c** STD experiment of FRAG-47 (5 μ M heparanase, 500 μ M fragment) confirming the interaction between this fragment and heparanase (the absence of signals, in this case, would indicate no binding); **d** STD spectrum of the known inhibitor suramine (5 μ M heparanase, 500 μ M suramine), proving its binding with heparanase; **e** STD competition experiment between FRAG-47 and suramine (5 μ M heparanase, 500 μ M fragment, 500 μ M suramine): a decrease in the intensity of the resonances from the fragment and suramine compared to **c** and **d** indicates that FRAG-47 and suramine are competing for the same binding site

Conclusions

In conclusion, we have shown that the development of “minimal” pharmacophore models could be a useful tool for further reducing the number of compounds to be tested in fragment-based screening without compromising the success and chemical diversity associated to these strategies. Although rarely used in the context of fragment discovery approaches [24], pharmacophore models offer an attractive opportunity for academic groups interested in the field but limited by reduced economic budgets and the possibility of managing big chemical libraries. In this particular case, we have proved that hit enrichment can be obtained and that novel inhibitors can be sought and

identified for a very relevant target, involved in cell invasion and metastasis, using a four-point pharmacophore model. Until now, and despite the interest in finding effective inhibitors against this relevant therapeutic target, only few chemotypes have been explored, most of them mimicking the chemical features of its natural substrate, the heparan sulfate. Experimental validation, using a set of NMR experiments, of the identified hits has been carried out and confirmation of the interaction with the catalytic site has been performed. Furthermore, the initial hits have been characterized in terms of binding efficiencies and they showed comparable values to others reported in the bibliography [23]. Taken together, the results open a new perspective for finding novel chemical scaffolds while reducing the number of fragments to be screened against therapeutic targets.

Acknowledgments The authors wish to thank the Spanish Ministerio de Ciencia e Innovación (MCINN, SAF2008-01845), the Fundación de Investigación Médica Mutua Madrileña and the Centro de Investigación Príncipe Felipe for economic support. We also acknowledge the financial support provided by the *Access to Research Infrastructures activity* in the 6th FP of the EC (Contract #RII3-026145, EU-NMR), and the CERM for providing technical support. S.M. and R.J.C are recipients of a FPI predoctoral fellowship and a contract of the Ramón y Cajal program from the MCINN, respectively.

References

1. Nasser NJ (2008) Heparanase involvement in physiology and disease. *Cell Mol Life Sci* 65(11):1706–1715. doi:10.1007/s0018-008-7584-6
2. Levy-Adam F, Feld S, Suss-Toby E et al (2008) Heparanase facilitates cell adhesion and spreading by clustering of cell surface heparan sulfate proteoglycans. *PLoS One* 3(6):e2319. doi:10.1371/journal.pone.0002319
3. Ilan N, Elkin M, Vlodavsky I (2006) Regulation, function and clinical significance of heparanase in cancer metastasis and angiogenesis. *Int J Biochem Cell Biol* 38(12):2018–2039. doi:10.1016/j.biocel.2006.06.004
4. Vlodavsky I, Ilan N, Naggi A et al (2007) Heparanase: structure, biological functions, and inhibition by heparin-derived mimetics of heparan sulfate. *Curr Pharm Des* 13(20):2057–2073. doi:10.2174/138161207781039742
5. Vlodavsky I, Friedmann Y (2001) Molecular properties and involvement of heparanase in cancer metastasis and angiogenesis. *J Clin Invest* 108(3):341–347
6. Hulett MD, Hornby JR, Ohms SJ et al (2000) Identification of active-site residues of the pro-metastatic endoglycosidase heparanase. *Biochemistry* 39(51):15659–15667. doi:10.1021/bi002080p
7. Khachigian LM, Parish CR (2004) Phosphomannopentaose sulfate (PI-88): heparan sulfate mimetic with clinical potential in multiple vascular pathologies. *Cardiovasc Drug Rev* 22(1):1–6
8. Hammond E, Bytheway I, Ferro V (2006) Heparanase as a target for anticancer therapeutics: new developments and future prospects. In: Delehedde M, Lortat-Jacob H (eds) *New developments in therapeutic glycomics*. Research Signpost, Kerala, India
9. Güner OF (2000) *Pharmacophore perception, development, and use in drug design*. International University Line, La Jolla, California

10. Güner OF (2002) History and evolution of the pharmacophore concept in computer-aided drug design. *Curr Top Med Chem* 2(12):1321–1332. doi:[10.2174/1568026023392940](https://doi.org/10.2174/1568026023392940)
11. Wolber G, Seidel T, Bendix F et al (2008) Molecule-pharmacophore superpositioning and pattern matching in computational drug design. *Drug Discov Today* 13(1–2):23–29. doi:[10.1016/j.drudis.2007.09.007](https://doi.org/10.1016/j.drudis.2007.09.007)
12. Wermuth CG, Ganellin CR, Lindberg P et al (1998) Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1997). *Annu Rep Med Chem* 33:385–395. doi:[10.1016/S0065-7743\(08\)61101-X](https://doi.org/10.1016/S0065-7743(08)61101-X)
13. Huang N, Shoichet BK, Irwin JJ (2006) Benchmarking sets for molecular docking. *J Med Chem* 49(23):6789–6801. doi:[10.1021/jm0608356](https://doi.org/10.1021/jm0608356)
14. Congreve M, Carr R, Murray C et al (2003) A ‘rule of three’ for fragment-based lead discovery? *Drug Discov Today* 8(19):876–877. doi:[10.1016/S1359-6446\(03\)02831-9](https://doi.org/10.1016/S1359-6446(03)02831-9)
15. Dalvit C, Fogliatto G, Stewart A et al (2001) WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability. *J Biomol NMR* 21(4):349–359. doi:[10.1023/A:1013302231549](https://doi.org/10.1023/A:1013302231549)
16. Hwang T-L, Shaka AJ (1995) Water suppression that works: excitation sculpting using arbitrary waveforms and pulse field gradients. *J Magn Reson A* 112:275–279. doi:[10.1006/jmra.1995.1047](https://doi.org/10.1006/jmra.1995.1047)
17. Meyer B, Peters T (2003) NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew Chem Int Ed Engl* 42(8):864–890. doi:[10.1002/anie.200390233](https://doi.org/10.1002/anie.200390233)
18. Ludwig C, Michiels PJA, Wu X et al (2008) SALMON: solvent accessibility, ligand binding, and mapping of ligand orientation by NMR spectroscopy. *J Med Chem* 51(1):1–3. doi:[10.1021/jm701020f](https://doi.org/10.1021/jm701020f)
19. Stockman BJ, Kothe M, Kohls D et al (2009) Identification of allosteric PIF-pocket ligands for PDK1 using NMR-based fragment screening and ^1H - ^{15}N TROSY experiments. *Chem Biol Drug Des* 73(2):179–188. doi:[10.1111/j.1747-0285.2008.00768.x](https://doi.org/10.1111/j.1747-0285.2008.00768.x)
20. Klages J, Coles M, Kessler H (2007) NMR-based screening: a powerful tool in fragment-based drug discovery. *Analyst* 132(7):693–705. doi:[10.1039/b709658p](https://doi.org/10.1039/b709658p)
21. Siegal G, Ab E, Schultz J (2007) Integration of fragment screening and library design. *Drug Discov Today* 12(23–24):1032–1039. doi:[10.1016/j.drudis.2007.08.005](https://doi.org/10.1016/j.drudis.2007.08.005)
22. Petitjean M (1992) Applications of the radius-diameter diagram to the classification of topological and geometrical shapes of chemical compounds. *J Chem Inf Comput Sci* 32(4):331–337. doi:[10.1021/ci00008a012](https://doi.org/10.1021/ci00008a012)
23. Dalvit C (2007) Ligand- and substrate-based ^{19}F NMR screening: principles and applications to drug discovery. *Prog Nucl Magn Reson Spectrosc* 51:243–271. doi:[10.1016/j.pnmrs.2007.07.002](https://doi.org/10.1016/j.pnmrs.2007.07.002)
24. Ji H, Stanton BZ, Igarashi J et al (2008) Minimal pharmacophoric elements and fragment hopping, an approach directed at molecular diversity and isozyme selectivity. Design of selective neuronal nitric oxide synthase inhibitors. *J Am Chem Soc* 130(12):3900–3914. doi:[10.1021/ja0772041](https://doi.org/10.1021/ja0772041)