Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry 16 (2008) 1784–1795
Bioorganic & Medicinal Chemistry

Virtual screening approaches for the identification of non-lipid autotaxin inhibitors

Abby L. Parrill,* Uniqua Echols, Tran Nguyen, Truc-Chi T. Pham, Adrienne Hoeglund and Daniel L. Baker*

Department of Chemistry, The University of Memphis, Memphis, TN 38152, USA

Received 17 September 2007; revised 4 November 2007; accepted 5 November 2007
Available online 12 November 2007

Abstract—Autotaxin (ATX, NPP-2) catalyzes the conversion of lysophosphatidyl choline (LPC) to lysophosphatidic acid (LPA), a mitogenic cell survival factor that stimulates cell motility. The high expression of both ATX and receptors for LPA in numerous tumor cell types has produced substantial interest in exploring ATX as an anticancer chemotherapeutic target. ATX inhibitors reported to date are analogs of LPA, a phospholipid, and are more hydrophobic than is typical of orally bioavailable drugs. This study applied both structure-based and ligand-based virtual screening techniques with hit rates of 20% and 37%, respectively, to identify a promising set of non-lipid, drug-like ATX inhibitors. Structure-based virtual screening necessitated development of a homology model of the ATX catalytic domain due to the lack of structural information on any mammalian NPP family member. This model provided insight into the interactions necessary for ATX inhibition, and produced a suitably diverse training set for the development and application of binary QSAR models for virtual screening. The most efficacious compound identified in this study was able to completely inhibit ATX-catalyzed hydrolysis of 1 μM FS-3 (a synthetic, fluorescent LPC analog) at a 10 μM concentration.

1. Introduction

Autotaxin (ATX, NPP-2) is a lysophospholipase D enzyme[1,2] initially discovered as a motility factor in the conditioned media of A2058 melanoma cells.[3] Since its discovery, increased ATX expression has been identified in renal carcinoma,[4] metastatic breast cancer,[5–7] thyroid carcinoma,[8] Hodgkin lymphoma[9] and invasive glioblastoma multiforme.[10,11] Conversely, downregulation of ATX expression occurs in response to the tumor suppressor, CST6.[12] The pathophysiology of ATX in cancer survival, proliferation, metastasis and angiogenesis is linked to the biology of lysophosphatidic acid (LPA),[2,13,14] the product of ATX-catalyzed hydrolysis of lysophosphatidyl choline (LPC). These numerous reports uniformly support investigation of ATX inhibition as a novel cancer chemotherapeutic strategy.

The relevance of ATX to cancer progression has stimulated a number of reports of ATX inhibitors. L-histidine was the first reported ATX inhibitor.[15] However, millimolar concentrations of histidine were required, and the inhibitory effect was counteracted by the inclusion of submillimolar concentrations of zinc sulfate, suggesting an inhibition mechanism involving interaction with the two native active site metal ions. A series of subsequent publications have described additional ATX inhibitors (Fig. 1).[16–21] These inhibitors include the products of ATX-catalyzed hydrolysis of LPC and sphingosyl phosphorylcholine (SPC), LPA (1) and S1P (2), respectively.[16] Inhibition of ATX by LPA and S1P suggests that product feedback inhibition may contribute to regulation of ATX function in vivo. Previously reported ATX inhibitors share several common structural features, including a phosphate, thiophosphate or phosphonate headgroup attached either with or without a linker to an alkyl chain, which can vary in overall length and can be either saturated or unsaturated. These compounds both lack substantial structural diversity and fail to meet Lipinski's empirical rules that characterize 90% of orally bioavailable drugs.[22,23] Identification of novel non-lipid structural classes capable of inhibiting ATX is of substantial value both to demonstrate that ATX

Keywords: Autotaxin; Binary QSAR; Lead identification; Metalloenzyme inhibition.
* Corresponding authors. Tel.: +1 901 678 2638 (A.L.P.), +1 901 678 4178 (D.L.B.); fax: +1 901 678 3447.; e-mail addresses: aparrill@memphis.edu; dlbaker@memphis.edu

0968-0896/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2007.11.018

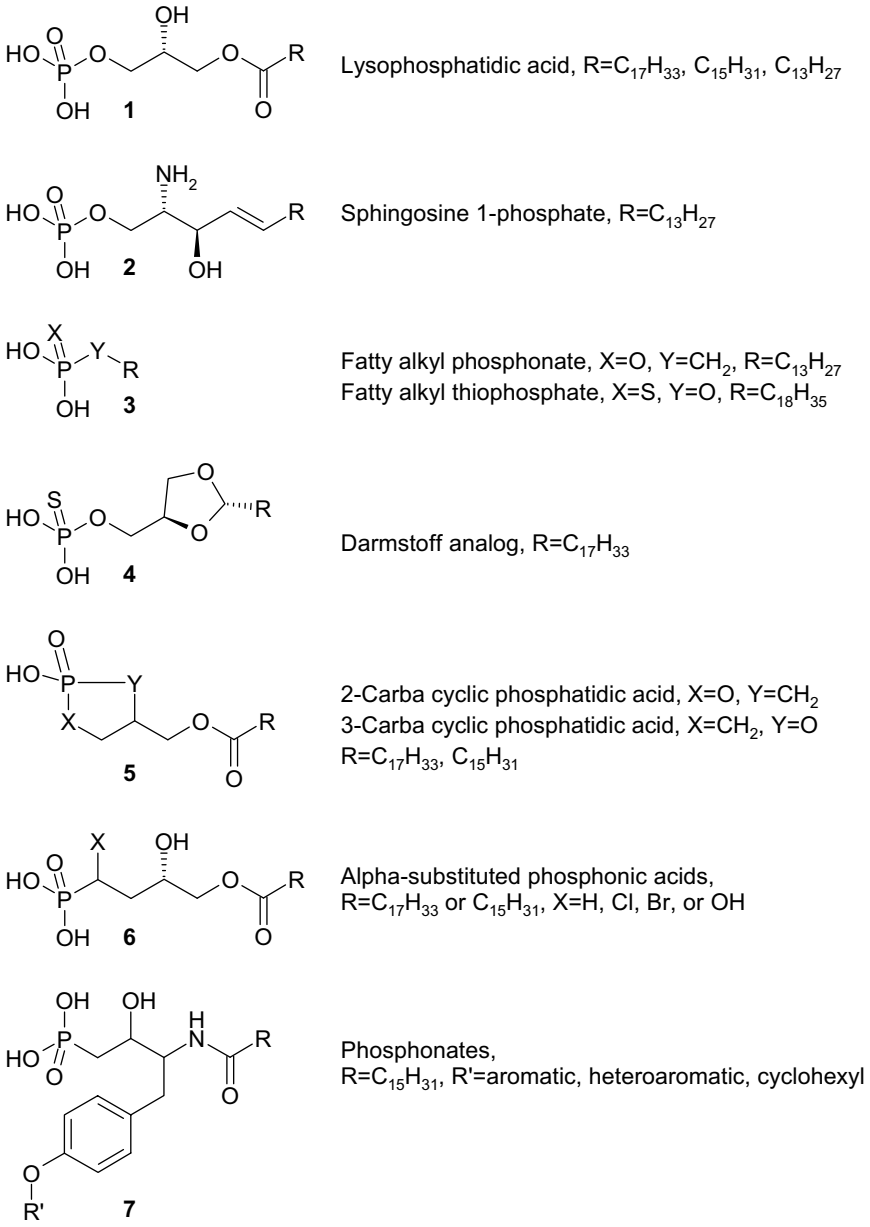

Figure 1. Published ATX inhibitor structures.[16–21]

is a 'druggable' target as well as to stimulate optimization of leads with drug-like physicochemical characteristics.

The lack of diversity in reported ATX inhibitors is due, in part, to the lack of a characterized three-dimensional structure of the enzyme. The ATX sequence of over 860 amino acids is divided into several domains, including a central catalytic domain composed of about 400 amino acids.[24,25] ATX is a member of the nucleotide pyrophosphatase/phosphodiesterase (NPP) family, as well as the alkaline phosphatase superfamily.[26] Crystallographic structures of several alkaline phosphatase superfamily members have been available for decades. These crystal structures show remarkable structural conservation in a small core surrounding the catalytic site,[27–30] but unfortunately show completely different structural characteristics outside this conserved core. Sequence homology of the alkaline phosphatases with ATX does not exceed 14% and is therefore insufficient for generation of a high quality homology model in any region outside the approximately 100-amino acid structurally conserved core. The recent report of a crystal structure of a bacterial NPP enzyme[31] with 30% identity to the ATX catalytic core domain enabled the development of a structural model of the ATX catalytic domain that may prove useful in structure-based drug design. Although a significant improvement, such a homology model must be applied cautiously as involvement of the C-terminal nuclease-like domain in substrate recognition has been suggested[25] from studies of NPP family domain-swapping chimeras.

This work applies an integrated approach utilizing several computational strategies for the identification of new classes of ATX inhibitors with drug-like physicochemical properties. These strategies include database screening guided by either docked complexes of candidate structures into a homology modeled catalytic core domain structure or a binary QSAR model based on a training set of known ATX substrates and inhibitors. Candidate inhibitors were screened for their ability to inhibit ATX-catalyzed hydrolysis of the synthetic fluorescent substrate, FS-3, in the micromolar concentration range. Lead compounds in several structural classes were identified using preliminary docking studies, including diphenyl diazerenes, thioureas, isoindoles, and xanthenes (Fig. 2). A binary QSAR model developed using these compounds and inhibitors reported in the literature as a training set showed a remarkable 71% accuracy when predicting structures with a calculated probability of >90% to show at least 20% inhibition at a 10 μM concentration. Although these compounds show modest potencies, they provide promising evidence that ATX is a 'druggable' target for the development of orally bioavailable therapeutics and validate our virtual screening approaches. Future studies will focus on refining the binary QSAR model with a more restrictive classification of active compounds in order to identify more potent leads.

2. Results

2.1. Model

A model of the ATX catalytic domain, residues 162–539 from the sequence of human ATX, was developed by homology to *Xac* NPP. All but 4 (1%) of the 317 non-proline and non-glycine residues occupy regions of Ramachandran space categorized as core secondary structures (232 residues, 73%), allowed regions (75 residues, 24%) or generously allowed regions (six residues, 2%). This model was used in docking studies with two LPC species, 12:0 and 14:0, shown previously[1] to be optimal ATX substrates. Unrestrained docking studies failed to identify a pose placing the phosphorus atom of either substrate within a reasonable distance of T210, the catalytic residue responsible for initiating hydrolysis by forming a covalent bond to the phosphorus atom. This result is consistent with the report that substrate recognition elements reside outside the catalytic domain.[25] Poses consistent with the role of LPC as an ATX substrate were identified when an anionic atom was required to occur within a 3.6 Å sphere centered 4.6 Å from both metal ions and 2.8 Å from the oxygen atom of T210. The phosphorus atom still preferred a location 6–7 Å from the oxygen atom of T210. Figure 3 shows the docked positions of LPC 12:0 and LPC 14:0.

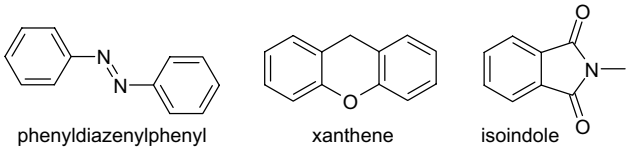

Figure 2. Substructures used to identify binary QSAR external validation set.

2.2. Model-guided compound selection

The ATX catalytic domain homology model was used as a docking target to identify promising candidate inhibitors for experimental screening. Docking candidates were selected from ChemBridge using the online search tool, www.hit2lead.com. Both similarity and substructure searches were performed. Initial searches focused on anionic groups such as phosphate, sulfate, and carboxylate as well as bioisosteres of these anionic groups including phosphonamide, sulfonamide, and oxocarboxylate. Compounds showing better than 50% ATX inhibition at 10 μM by enzyme activity assay (Fig. 4) were used as similarity targets in subsequent searches. Ninety-five structures (Supplementary Fig. 1) were selected for screening out of 500 compounds docked against the ATX catalytic domain homology model. Experimental screening results shown in Table 1 demonstrate that 19 of the 94 structures tested inhibited 20% or more of the ATX-catalyzed FS-3 hydrolysis at 10 μM. The selection process was therefore 20% accurate in selecting compounds with some degree of ATX inhibition.

Docking studies with the ATX catalytic domain homology model were able to provide insight into activity differences of very similar structures. In particular, H2L 5564949 and H2L 5564676 are constitutional isomers that differ only in the location of a carboxylic acid functional group on the terminal aromatic ring (Fig. 5). Although their chemical structures are very similar, their biological activities are quite distinct. H2L 5564949 inhibited ATX activity by 51% at 10 μM, whereas H2L 5564676 had no effect at the same concentration. H2L 5564949 displays a number of favorable interactions in the docked complex (Fig. 6A). These interactions include a cation–π interaction between the carboxylate functional group and R284, an ion-pairing interaction between the carboxylate functional group and R285, and metal ligating interactions with the isoindole and diazerine functional groups. In contrast, H2L 5564676 failed to dock into the ATX catalytic domain model, suggesting that insufficient space was available for this molecule. Relocation of the carboxylate group of H2L 5564949 from the meta to the para position (Fig. 6B) suggests that a very close, energetically unfavorable interaction would occur with the backbone of R284. These models are consistent with the difference in the ability of H2L 5564949 (meta isomer) and H2L 5564676 (para isomer) to inhibit ATX.

Docking studies additionally provided insight into unforeseen binding modes. H2L 7839888 was selected for in silico evaluation on the basis of the oxocarboxylate functional group, which was expected to interact favorably with the active site metal ions. Figure 7 shows that the thiourea functional group of H2L 7839888 is involved in ligating one of the active site metal ions. This

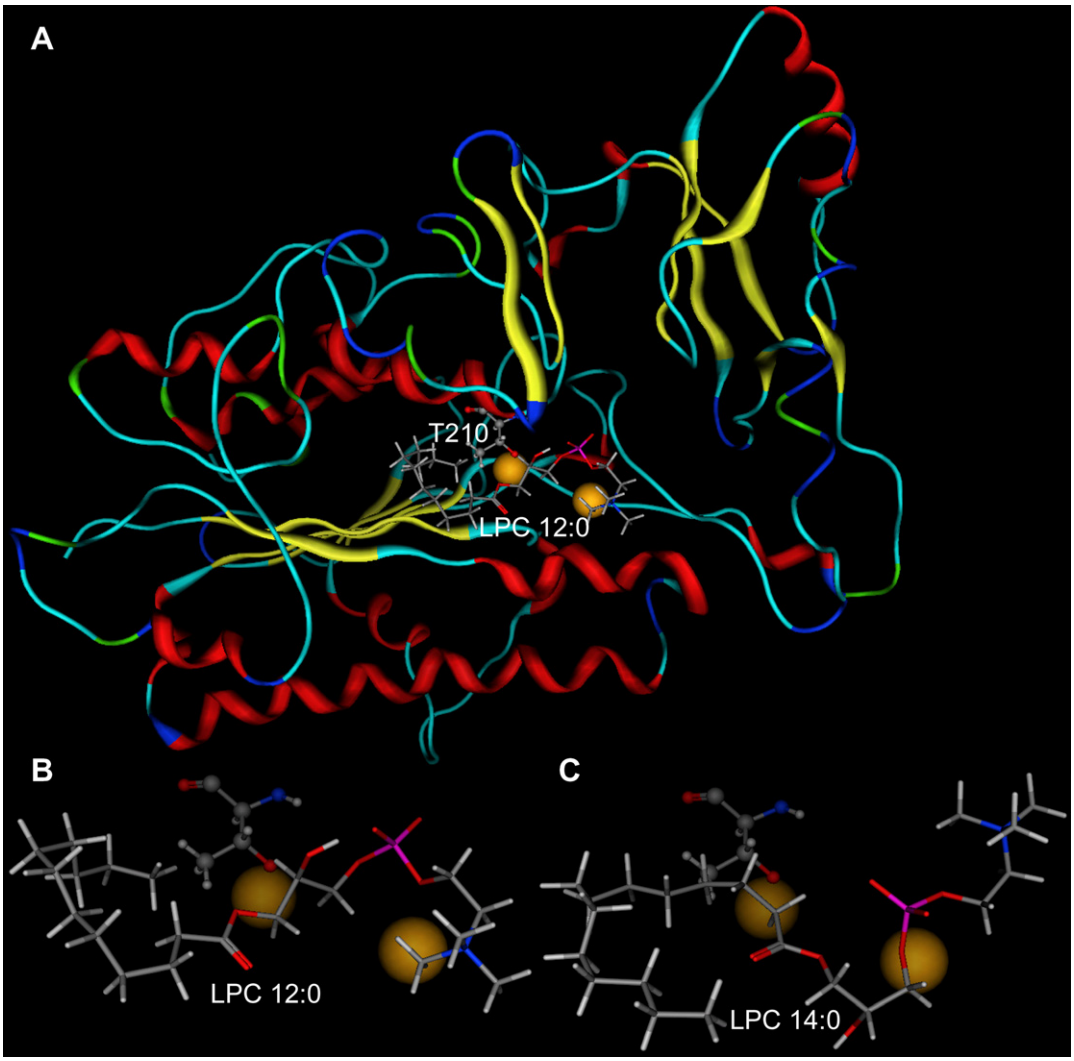

Figure 3. ATX complexes with LPC 12:0 and LPC 14:0. Metal ions are shown as spheres, LPC species are shown as stick models, and T210 is shown as a ball and stick model. All panels are shown from the same perspective. (A) ATX (ribbons) complex with LPC 12:0. (B) Close up of LPC 12:0 position relative to T210 and metal ions. (C) Close up of LPC 14:0 position relative to T210 and metal ions.

interaction and the electrostatic interactions of the carboxylate functional group with R284 and K249 are likely responsible for the 59% inhibition observed.

2.3. Binary QSAR model

Two four-component binary QSAR models were developed using different structural representations of the training set. Model 1 utilized the structural representation downloaded directly from www.hit2lead.com, which lacked hydrogen atoms and biologically relevant ionization states. This model relies on calculated descriptors, such as the number of acidic functional groups, to compensate for these inaccuracies. The advantage to this structural representation is that new candidate structures can be downloaded quickly and activity probabilities can be calculated without delays for structure processing and correction. Model 2 utilized structures after hydrogen addition, correction of ionization states to match the assay conditions, and geometry optimization. Thus predictions on new compounds require these additional processing steps. Models 1 and 2 provided a total accuracy of 76% and 77%, respectively, compared to a chance accuracy of 62% on the training set. The training set was biased toward inactive compounds (68 out of 97), and thus showed a better accuracy predicting inactive compounds (91% by both models) than active compounds (41% and 45% by Models 1 and 2, respectively). Nevertheless, these accuracies were considerably better than the chance accuracies of 80% on inactives and 20% on actives. In order to validate the binary QSAR model on an external test set, 828 structures were selected for activity predictions based on substructure searches in the ChemBridge online search facility with diphenyldiazenylphenyl, xanthene, and isoindole substructures (Fig. 2). Predictions were also made on 40 candidate structures selected by similarity to H2L 7839888, containing a thiourea group not well-parameterized in the MMFF94 forcefield. Both binary QSAR models were used to generate a probability that each structure would be in the active group. External validation sets of 54 and 19 structures with activity probability predictions greater than 0.6 were selected by

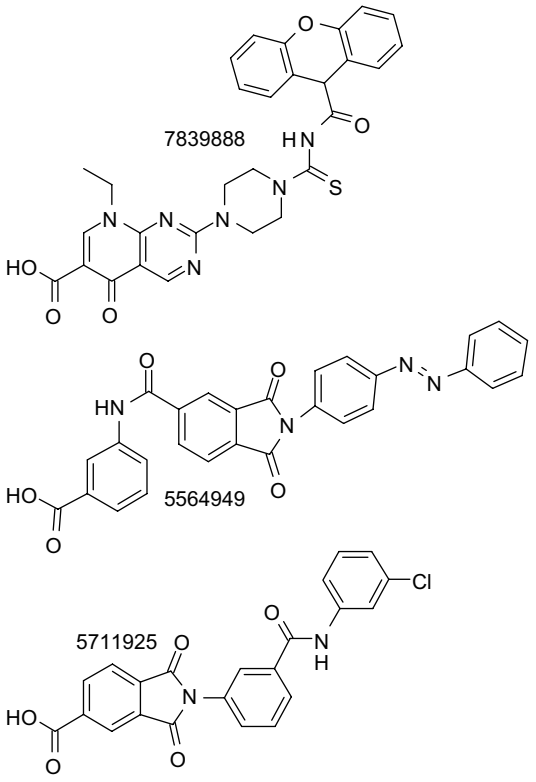

Figure 4. Compounds selected based on docking studies that showed greater than 50% inhibition of ATX-catalyzed FS-3 hydrolysis at 10 μM.

Models 1 and 2 (Supplementary Figs. 3 and 4), respectively, for experimental screening. Tables 2 and 3 show the prediction probabilities and experimental screening results for both validation sets. The validation set for Model 1 produced 20 active compounds out of the 54 tested. The validation set for Model 2 produced 7 compounds out of 19 tested. Both models were therefore 37% accurate on the external validation sets, an accuracy substantially greater than the 20% obtained using docking results to guide compound selection. Notably, Model 1 offers four substantial advantages as a tool for guiding experimental screening efforts aimed at identifying ATX inhibitors. First, it identified a greater number of actives out of the same set of candidates. This indicates that Model 2 was incorrectly classifying a greater number of active compounds as inactives, and thus would have a lower accuracy on inactive compounds. Second, Model 1 classified the most efficacious compound, H2L 7905958 (Fig. 8), as active, whereas Model 2 classified this compound as inactive. Thus Model 2 may be limited to identifying new compounds within the efficacy range represented in the training set. Third, Model 1 can be applied more rapidly as there is no need to add explicit hydrogen atoms or assign appropriate ionization states. This indicates that the descriptors utilized in the model reflect the aspects of these structures that are relevant to ATX inhibition. Finally, Model 1 showed a substantially improved accuracy of 71% on the subset of compounds with a predicted activity probability greater than 0.9 (Fig. 9).

3. Discussion

ATX is an attractive target for the development of anticancer chemotherapeutic compounds due to its established elevated expression in multiple tumor cell types,[7,8,10,11,32–34] correlated elevated expression of receptors for its reaction product, LPA,[11,35–40] and the established role of LPA in cell survival, mitogenesis, and migration. Nevertheless, efforts to identify drug-like inhibitors of ATX are hindered by several factors. These include lack of three-dimensional structural information of the entire ATX enzyme, lack of drug-like inhibitors with limited flexibility that define the required features and geometry required for ATX inhibition, as well as the relatively low structural diversity among previously published inhibitors.

Lack of three-dimensional structural information about the structure of ATX makes the application of docking as a structure-based virtual screening tool very challenging. Two recent examples of structure-based virtual screening utilized crystallographic structures of integrin αvβ3[41] and glucosamine-6P synthase[42] as docking targets and achieved hit rates of 14% and 23%, respectively. In this context, our 20% hit rate using structure-based virtual screening against a homology model of the catalytic domain, approximately one-third of the total sequence, is quite respectable.

The lack of structurally rigid inhibitors prevents definition of a single pharmacophore, or geometry of chemical features responsible for biological activity. A pharmacophore is an essential tool for ligand-based screening. A recent comparison of virtual screening methods to identify biogenic amine G protein-coupled receptor antagonists demonstrated that hit rates as high as 100% could be obtained using ligand-based screening.[43] However, all pharmacophores evaluated in this study were developed using at least nine drug-like molecules with limited flexibility. A set of rigid inhibitors that could define the ATX pharmacophore was simply not available at the start of this study.

The low structural diversity among the previously reported ATX inhibitors is problematic for the development of QSAR models, which are best applied to predict activity within the structural diversity represented in the training set. QSAR models are desirable screening tools, however, as a binary QSAR accuracy of 98%[44] and an 80% hit rate using a conventional QSAR model as a screening tool[45] have recently been reported. The lack of diversity among known ATX inhibitors was solved by first applying structure-based virtual screening to identify additional structures, both active and inactive, in an effort to provide a more heterogeneous training set. The resultant binary QSAR models were then applied in virtual screening applications with greater confidence. A validation set of 54 structures was selected using a binary QSAR model. Experimental screening confirmed that 37% of these structures inhibited at least 20% of the ATX-catalyzed FS-3 hydrolysis at 10 μM, comparable to the 41% accuracy on actives observed on the training set. The accuracy may have

Table 1. Screening results for compounds selected based on docking studies

Hit2Lead ID #	% ATX activity at 10 μM	Hit2Lead ID #	% ATX activity at 10 μM	Hit2Lead ID #	% ATX activity at 10 μM
7839888	**41 ± 9**	7886845	89 ± 4	6511389	NI
5564949	**45 ± 10**	5186721	89 ± 5	6636019	NI
5711925	**48 ± 5**	7930936	90 ± 9	6639593	NI
7084279	**55 ± 2**	5303472	91 ± 3	6819710	NI
7349910	**56 ± 7**	7982730	91 ± 7	6834600	NI
5536728	**58 ± 9**	5114733	NI	6861828	NI
7953880	**67 ± 7**	5129846	NI	6863549	NI
5772007	**70 ± 4**	5150931	NI	6893809	NI
5233626	**71 ± 13**	5173584	NI	7037059	NI
5161561	**71 ± 5**	5185417	NI	7510305	NI
5731577	**71 ± 6**	5213868	NI	7527305	NI
5766971	**71 ± 6**	5213875	NI	7687795	NI
6378071	**74 ± 4**	5233516	NI	7692052	NI
6637525	**74 ± 4**	5243698	NI	7697730	NI
7685625	**76 ± 10**	5252613	NI	7723878	NI
7906002	**78 ± 12**	5271796	NI	7784508	NI
5276712	**79 ± 4**	5322223	NI	7887795	NI
5579103	**79 ± 7**	5374577	NI	7906001	NI
9010457	**80 ± 8**	5482541	NI	7909638	NI
5186695	82 ± 6	5545681	NI	7927537	NI
7905969	83 ± 9	5564676	NI	7928422	NI
5271345	85 ± 18	5584804	NI	7933221	NI
7988917	85 ± 14	5731983	NI	7934179	NI
6501049	85 ± 4	5734259	NI	7945953	NI
5266219	86 ± 5	5748750	NI	7958761	NI
5636998	86 ± 5	5764004	NI	7961791	NI
7778075	86 ± 4	5791967	NI	7967798	NI
5733393	87 ± 6	5846151	NI	7973958	NI
6665417	88 ± 9	6035829	NI	7981128	NI
7932784	88 ± 10	6085556	NI	9009993	NI
5197025	88 ± 9	6095403	NI		
7978373	89 ± 13	6140430	NI		

NI: No inhibition. Bold type indicates compounds reducing the ATX-catalyzed response by at least 20%.

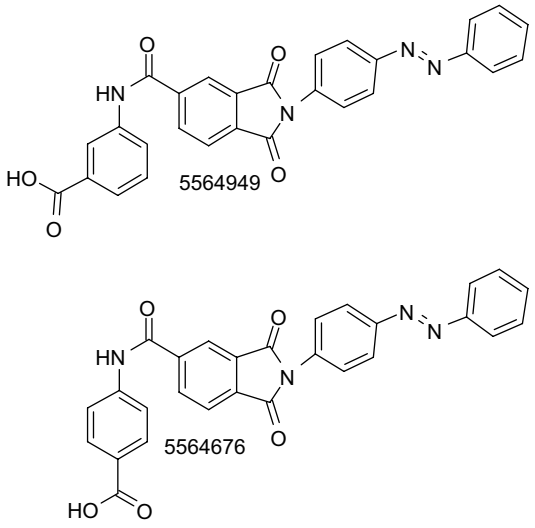

Figure 5. Structurally similar constitutional isomers with drastically different biological activities.

been improved by removing limits on the number of principal components used in building the model, however, a training set with a relatively small number of actives warranted limiting the number of principal components.

The primary goal of this study was to demonstrate that ATX is a 'druggable' target. Previously reported ATX inhibitors[16–21] are all analogs of the catalytic product, LPA, and fail to meet most metrics applied to determine drug-likeness. The most common of these metrics are Lipinski's rules.[46] These rules are based on the observation that poor biological distribution occurs when solubility and permeability are reduced due to excessive molecular weight ($\geqslant$500), large numbers of hydrogen bond donors or acceptors ($\geqslant$5 or $\geqslant$10, respectively), and calculated log P ($\geqslant$5). Several extensions to Lipinski's rules have appeared in the literature. Veber et al., found that increasing flexibility as reflected in an increasing number of rotatable bonds reduced oral bioavailability.[47] Nine compounds were identified in the current study as reducing the ATX-catalyzed FS-3 hydrolysis by 50% or more at 10 μM. Two of these compounds (H2L 5711925 and H2L 7921385) comply with all of Lipinski's rules and have fewer than 10 rotatable bonds, thus also meeting the extended flexibility metric. Five additional compounds (H2L 7839888, H2L 5186522, H2L 5761473, H2L 5538444 and H2L 7905958) exceed only the molecular weight limit by amounts ranging from seven (H2L 7905958) to 148 (H2L 5186522). One compound (H2L 5564949) exceeds only the calculated log P criterion. The final compound (H2L 5210574) exceeds two criteria, calculated log P

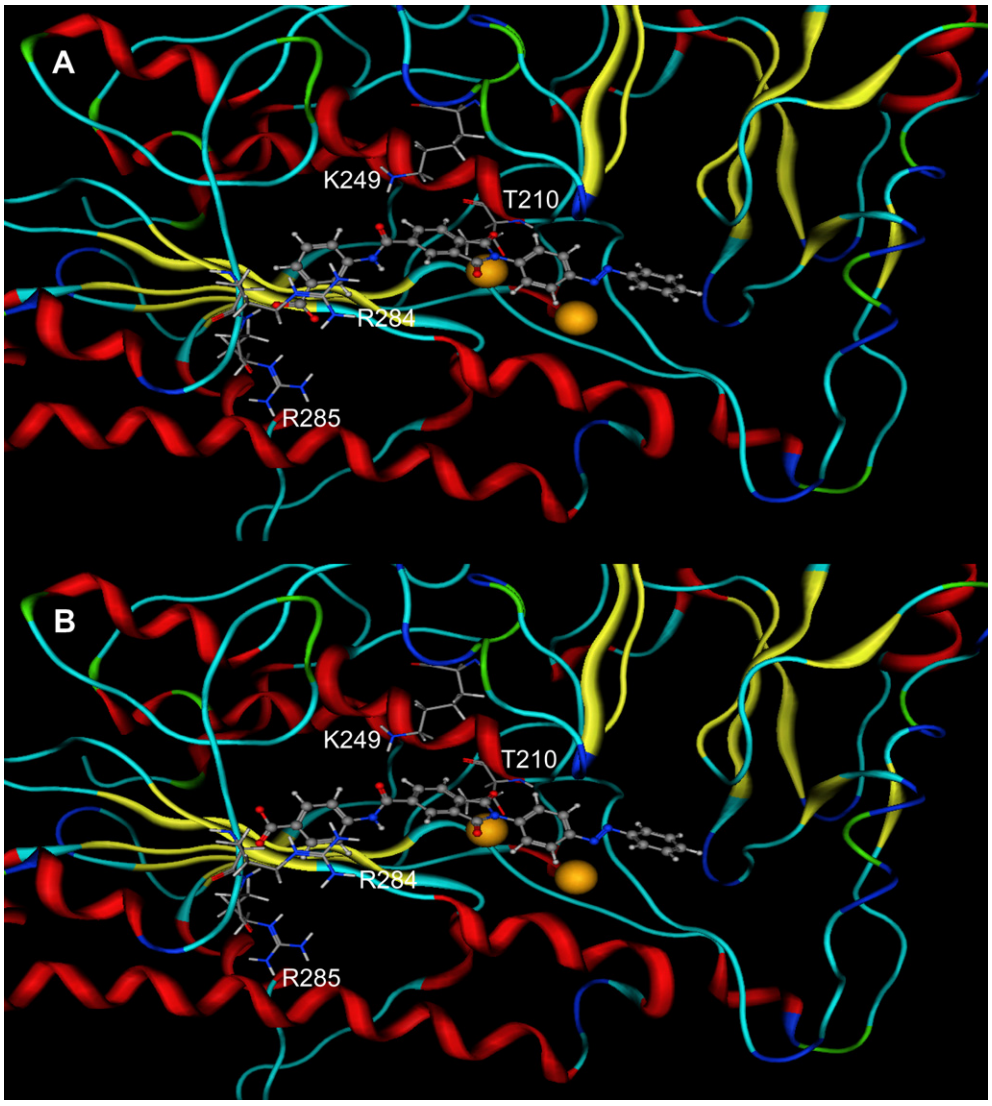

Figure 6. ATX homology model (ribbon) provides structural insight into activity difference of closely related constitutional isomers. Residues of interest are shown as stick models. (A) Docked position of Hit2Lead ID 5564949 (ball and stick model) displays cation–π interaction with R284, ion-pairing interaction with R285, and metal ligating interactions with diazerene and isoindole functional groups. (B) Hit2Lead ID 5564676 (ball and stick model), generated from docked position of Hit2Lead ID 5564949, displays steric clash with R284 backbone.

and molecular weight. None of the compounds violated the recommended limits on rotatable bonds, hydrogen bond acceptors, or hydrogen bond donors. Thus the goal of demonstrating that ATX is a 'druggable' target has been achieved.

4. Conclusions

A set of structurally diverse and drug-like inhibitors of ATX catalytic function have been identified. Nine of these are expected to exhibit IC_{50} values less than 10 μM based on the observed reduction of ATX-catalyzed FS-3 hydrolysis by 50% or more at this concentration. Several tools were developed to guide virtual screening efforts. The first tool is a model of the ATX catalytic domain developed by homology to the recently reported crystal structure of a bacterial NPP family member. Twenty percent of the compounds selected using structure-based virtual screening with this model inhibited ATX-catalyzed FS-3 hydrolysis. The second tool is a binary QSAR model that can be rapidly applied to structural representations based only on heavy atoms, without requiring addition of hydrogen atoms or assignment of ionization states. The binary QSAR model was 37% accurate in the selection of active compounds for screening. The structures identified with these tools can guide the refinement of the binary QSAR model using a lower threshold for activity in order to guide future screening efforts toward increasingly potent compounds.

5. Experimental

5.1. ATX model development

A model of human ATX was developed by homology to the Xac nucleotide pyrophosphatase/phosphodiesterase (NPP) crystal structure (entry 2GSU[31] in the

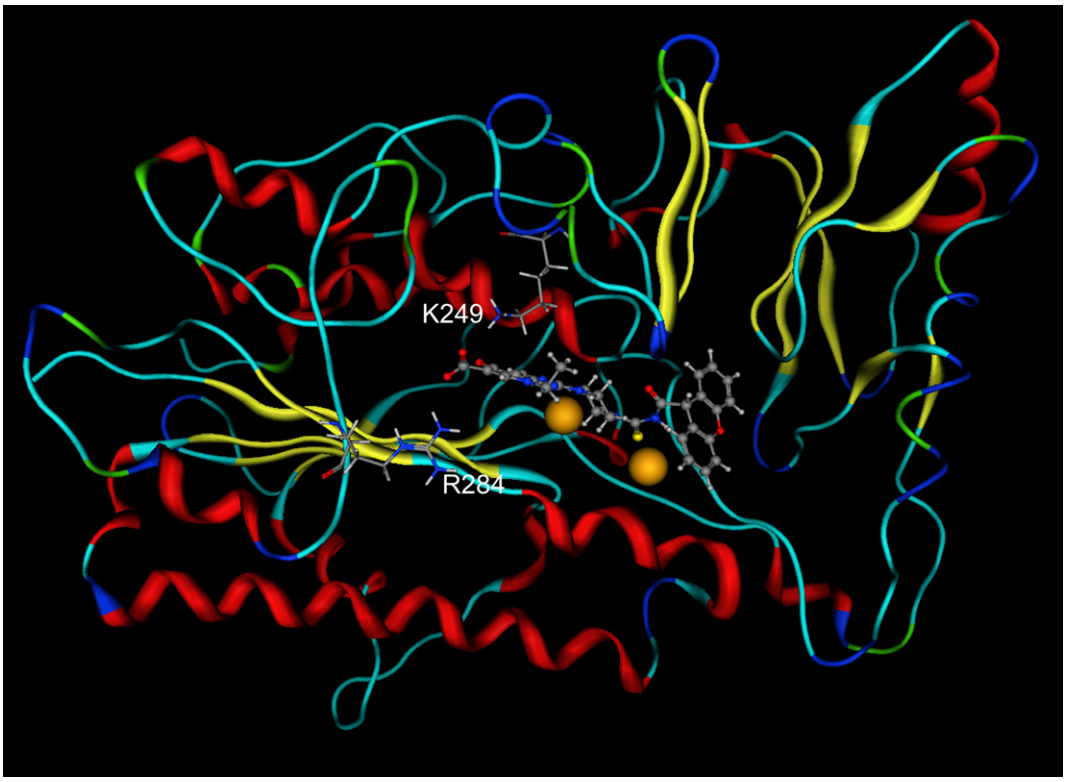

Figure 7. ATX homology model (ribbon) provides structural insight into unforeseen binding modes. Residues of interest are shown as stick models. Docked position of Hit2Lead ID 7839888 (ball and stick model) displays ionic interactions with R284 and K249, and a metal ligating interaction via the thiourea functional group.

Table 2. Validation of binary QSAR Model 1

Hit2Lead ID #	Inhibitory probability (*p*)	% ATX activity at 10 μM	Hit2Lead ID #	Inhibitory probability (*p*)	% ATX Activity at 10 μM
7905962	**1.0000**	**62 ± 2**	5562974	0.8159	NI
5135442	**0.9836**	**71 ± 6**	**5161216**	**0.8138**	**59 ± 5**
5161502	**0.9434**	**58 ± 8**	5560737	0.8125	93 ± 9
7970685	0.9433	NI	5128758	0.8024	NI
5186522	**0.9147**	**49 ± 10**	**5538444**	**0.7936**	**31 ± 4**
5476934	**0.9058**	**53 ± 9**	5161338	0.7932	NI
5232155	0.9036	NI	**7921385**	**0.7845**	**21 ± 6**
5211047	**0.8973**	**67 ± 5**	**5161768**	**0.7793**	**67 ± 4**
5761473	**0.8945**	**46 ± 6**	5751093	0.7772	83 ± 9
5847961	**0.8945**	**65 ± 5**	5187104	0.7692	91 ± 6
5161576	**0.8901**	**59 ± 5**	5773585	0.7666	86 ± 9
5561207	0.8872	NI	**5232935**	**0.7611**	**70 ± 2**
5161663	0.8834	NI	5541077	0.7463	NI
5135449	0.8727	NI	**6041006**	**0.7326**	**58 ± 5**
5161337	0.8706	NI	5162333	0.7234	91 ± 6
5210546	**0.8631**	**71 ± 4**	5161411	0.7213	87 ± 5
5186517	0.8569	NI	**7905958**	**0.7012**	**0 ± 4**
5581731	0.8533	NI	5233620	0.7003	91 ± 4
5161334	0.8523	NI	5161340	0.6985	NI
5161319	0.8489	NI	5251588	0.695	NI
5161545	0.8457	NI	5143009	0.6888	NI
5186564	0.8432	90 ± 4	5116845	0.6756	85 ± 4
5175107	0.8342	NI	5836563	0.6616	91 ± 4
5210569	0.832	90 ± 5	**6764570**	**0.6531**	**62 ± 5**
6051020	**0.8232**	**80 ± 5**	**5210574**	**0.6523**	**31 ± 5**
5532563	0.8224	90 ± 4	5538674	0.6501	81 ± 5
5323948	0.8179	87 ± 6	5161774	0.6401	93 ± 5

NI: No inhibition. Bold type indicates compounds reducing the ATX-catalyzed response by at least 20%.

Table 3. Validation of binary QSAR Model 2

Hit2Lead ID #	Inhibitory probability (*p*)	% ATX activity at 10 μM	Hit2Lead ID #	Inhibitory probability (*p*)	% ATX activity at 10 μM
5161411	1.0000	87 ± 5	5116845	0.9861	85 ± 4
5161502	**1.0000**	**58 ± 8**	**5135442**	**0.9761**	**71 ± 6**
5211047	**1.0000**	**67 ± 5**	5186564	0.8418	90 ± 4
5541077	1.0000	NI	**5761473**	**0.7584**	**46 ± 6**
5543060	1.0000	NI	5161338	0.6536	97 ± 4
5560737	1.0000	93 ± 9	**5210574**	**0.6529**	**31 ± 5**
5187104	0.9998	91 ± 6	5210569	0.6276	90 ± 5
5175107	0.9994	NI	5751093	0.6257	83 ± 9
5135449	0.9874	NI	**5210546**	**0.6147**	**71 ± 4**
5186522	**0.9871**	**49 ± 10**			

NI: No inhibition. Bold type indicates compounds reducing the ATX-catalyzed response by at least 20%.

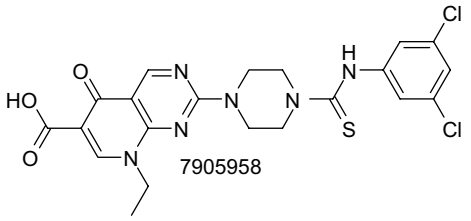

Figure 8. Structure of H2L 7905958, the most efficacious compound identified in the current study.

Protein Data Bank[48]). The sequence of human ATX (GenBank entry Q13822)[49] was aligned with the sequence of Xac NPP using MOE[50] and amino acids 1–161 and 540–863 of ATX were deleted as they did not correspond to amino acids contained in the Xac NPP structure. The ATX model was then generated by the homology model function in MOE, and divalent metal cations were transferred to corresponding locations in ATX from the Xac NPP crystal structure. Ionization states of histidine residues near divalent metal cations were adjusted to zero charge, with the deprotonated imidazole nitrogen nearest the cation. The ionization state of T210 was adjusted to −1, consistent with the pH profile demonstrated for Xac NPP.[31]

5.2. ATX docking studies

Complexes of ATX with small molecules including substrates, inhibitors, and candidate inhibitors were generated by docking using MOE.[50] Each small molecule structure was constructed in the ionization state expected at pH 7. Functional groups with pK_a values near seven, such as phosphate monoesters, were deprotonated due to the expected anionic stabilization provided by the active site metal ions. Residues with atoms within 6.5 Å of the active site metal ions were defined as the placement site. Small molecules were placed into ATX using the αPMI placement method, which matches the small molecule principal moments of inertia to alpha spheres distributed in closely packed locations within the docking site. Small molecule poses were scored using the affinityDG function, which approximates the enthalpic contribution to binding-free energy. This combination of placement and scoring methods was selected based on comparison of results on two test systems (unpublished data). The αPMI placement and affinityDG energy function were found to reproduce the crystallographic position of rosiglitazone in PPARγ as well as our previous docked position of sphingosine 1-phosphate (S1P) within the validated model of the $S1P_1$ receptor[51] generated by Autodock 3.0.[52] MOE was used for docking studies with ATX due to the lack of metal ion parameters in Autodock 3.0. Docking simulations of small molecules with long alkyl

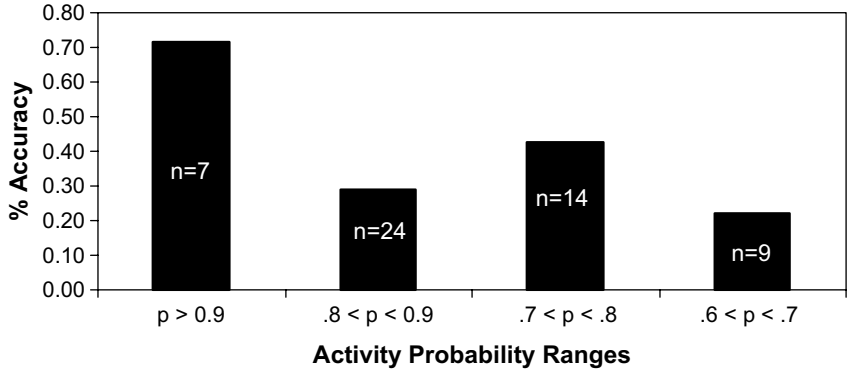

Figure 9. Higher accuracy observed when activity probability (p) is greater than 0.9. Number of training set compounds in each probability range (binary QSAR model 1) is indicated in each bar.

chains additionally were performed multiple times due to the stochastic nature of the conformational sampling procedure and included an increased number of samples per conformation (50) and an increased number of poses (500) to improve sampling. Conformations for these highly flexible molecules were also precomputed in the absence of the receptor using the stochastic conformational search algorithm and all resulting conformations were rigidly docked into the receptor. Substrate docking studies were performed in the presence and absence of a restraint requiring an anionic atom to occur inside a 3.6 Å sphere centered 4.6 Å from both metal ions and 2.8 Å from the oxygen atom of T210. In all cases, the pose with the lowest score after elimination of ligand poses with strained conformations was selected for analysis and discussion.

5.3. Binary QSAR model development

Two binary QSAR models were constructed using 97 chemical structures identified in the current study using the structure-based virtual screening approach (Supplementary Fig. 1) or with reported ATX inhibition data in the literature (Supplementary Fig. 2).[16–21] These models differ in the representation of the small molecule inhibitor. Model 1 was constructed using structures downloaded from www.hit2lead.com without further processing. These structures do not have explicit hydrogen atoms nor do they have explicit charges on ionizable groups. Model 2 was constructed using structures corrected after download to include explicit hydrogens and the ionization state expected under the assay conditions (pH 8) with subsequent geometry optimization with the MMFF94[53] force field. Relatively few highly potent structures were available in the training set. Compounds were therefore classified active if at least 20% inhibition was observed at a concentration of 10 μM. Several hundred structural descriptors were calculated for each structure (both representations), all available in the current version of the MOE software with the exception of those dependent on the orientation of the molecules in the coordinate system and those based on semi-empirical quantum mechanical methods. These individual descriptors were reduced to a set of orthogonal principal components. The binary models were fitted using four principal components.

5.4. Candidate inhibitor selection

Candidate inhibitors were selected for docking using the online search tool www.hit2lead.com (ChemBridge). Search criteria included an anionic group or metal ligating group (phosphate, phosphonamide, sulfate, sulfonamide, diazerenyl, carboxylate, sulfhydryl), and sufficient additional structural details (such as an additional aromatic ring, molecular weight restriction between 400 and 500, log P range between 2 and −2) to restrict the hitlist to 350 structures or fewer. Hits were selected for purchase and experimental assays based on their interaction with active site metal ions, ability to fill a substantial portion of the region surrounding the metal ions, observation of additional complementary interactions with ATX, or to expand the structure–activity relationship around experimentally confirmed hits.

Candidate inhibitors were selected for validation of the binary QSAR models using the online search tool www.hit2lead.com (ChemBridge). Search criteria included a metal ligating group represented in the training set (diazerenyl, thiourea, isoindole, xanthene). No further pruning of hitlists was performed as binary QSAR predictions for hundreds of compounds can be done in less than a minute. Hits were selected for purchase and experimental validation if the prediction probability of ATX inhibition was 0.6 or greater.

5.5. Cell culture

MDA-MB-435 cells were cultured at 37 °C and 5% CO_2 in Dulbecco's modified Eagle medium (DMEM) (MediaTech, Herndon, VA) containing 5% fetal bovine serum (Hyclone, Logan, UT), 100 U/ml penicillin, 100 μg/ml streptomycin (Hyclone, Logan, UT), and 292 μg/ml L-glutamine (Hyclone, Logan, UT). Cells were grown to ~80% confluence at which time the cells were washed twice with sterile phosphate-buffered saline prior to the addition of serum-free DMEM containing L-glutamine. Conditioned medium was collected after 24–30 h, supplemented with 20% ethylene glycol, and was clarified by centrifugation at 3000*g* and 4 °C for 10 min. The media was concentrated ~10 fold and buffer exchanged into Tris (50 mM, pH 7.4) containing 20% ethylene glycol using an Amicon 8050 cell fitted with a PM30 filter (Millipore, Billerica, MA). Aliquots of concentrated conditioned media were stored at 4 °C until needed.

5.6. ATX inhibition assay

ATX inhibition was assayed using FS-3 (Echelon Biosciences, Inc., Salt Lake City, UT, USA) as a substrate and ~10 times concentrated conditioned serum-free medium (CCM) from MDA MB-435 cells as the source of ATX. Assays were performed in 96-well plates with CCM comprising one-third of the total volume and final FS-3 and charcoal-stripped fatty acid-free BSA concentrations of 1 and 30 μM in assay buffer (1 mM each $CaCl_2$ and $MgCl_2$, 5 mM KCl, 140 mM NaCl, 50 mM Tris, pH 8.0). Fluorescence was read at 5 minute intervals by a Synergy2 system (BioTek, Winooski, VT) with excitation and emission wavelengths of 485 and 538 nm, respectively. Results are shown at 1 h, at which point all fluorescence changes as a function of time were linear. Fluorescence readings were normalized to vehicle control after subtraction of fluorescence in the absence of CCM. Data are shown as means ± SD of at least three wells.

Acknowledgments

Project SEED Stipends for U.E. and T.N. were provided by the American Chemical Society Project SEED Endowment, Buckman Laboratories, Cargill Research Center, Schering-Plough HealthCare Products, and Valero Memphis. Thanks to the Chemical Computing Group for the MOE software.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2007.11.018.

References and notes

1. Tokumura, A.; Majima, E.; Kariya, Y.; Tominaga, K.; Kogure, K.; Yasuda, K.; Fukuzawa, K. *J. Biol. Chem.* **2002**, *277*, 39436–39442.
2. Umezu-Goto, M.; Kishi, Y.; Taira, A.; Hama, K.; Dohmae, N.; Takio, K.; Yamori, T.; Mills, G. B.; Inoue, K.; Aoki, J., et al. *J. Cell Biol.* **2002**, *158*, 227–233.
3. Stracke, M. L.; Krutzsch, H. C.; Unsworth, E. J.; Arestad, A.; Cioce, V.; Schiffmann, E.; Liotta, L. A. *J. Biol. Chem.* **1992**, *267*, 2524–2529.
4. Stassar, M. J.; Devitt, G.; Brosius, M.; Rinnab, L.; Prang, J.; Schradin, T.; Simon, J.; Petersen, S.; Kopp-Schneider, A.; Zoller, M. *Br J. Cancer* **2001**, *85*, 1372–1382.
5. Debies, M. T.; Welch, D. R. *J. Mammary Gland Biol. Neoplasia* **2001**, *6*, 441–451.
6. Euer, N.; Schwirzke, M.; Evtimova, V.; Burtscher, H.; Jarsch, M.; Tarin, D.; Weidle, U. H. *Anticancer Res.* **2002**, *22*, 733–740.
7. Yang, S. Y.; Lee, J.; Park, C. G.; Kim, S.; Hong, S.; Chung, H. C.; Min, S. K.; Han, J. W.; Lee, H. W.; Lee, H. Y. *Clin. Exp. Metastasis* **2002**, *19*, 603–608.
8. Kehlen, A.; Englert, N.; Seifert, A.; Klonisch, T.; Dralle, H.; Langner, J.; Hoang-Vu, C. *Int J. Cancer* **2004**, *109*, 833–838.
9. Baumforth, K. R.; Flavell, J. R.; Reynolds, G. M.; Davies, G.; Pettit, T. R.; Wei, W.; Morgan, S.; Stankovic, T.; Kishi, Y.; Arai, H., et al. *Blood* **2005**, *106*, 2138–2146.
10. Hoelzinger, D. B.; Mariani, L.; Weis, J.; Woyke, T.; Berens, T. J.; McDonough, W. S.; Sloan, A.; Coons, S. W.; Berens, M. E. *Neoplasia* **2005**, *7*, 7–16.
11. Kishi, Y.; Okudaira, S.; Tanaka, M.; Hama, K.; Shida, D.; Kitayama, J.; Yamori, T.; Aoki, J.; Fujimaki, T.; Arai, H. *J. Biol. Chem.* **2006**, *281*, 17492–17500.
12. Song, J.; Jie, C.; Polk, P.; Shridhar, R.; Clair, T.; Zhang, J.; Yin, L.; Keppler, D. *Biochem. Biophys. Res. Commun.* **2006**, *340*, 175–182.
13. Clair, T.; Aoki, J.; Koh, E.; Bandle, R. W.; Nam, S. W.; Ptaszynska, M. M.; Mills, G. B.; Schiffmann, E.; Liotta, L. A.; Stracke, M. L. *Cancer Res.* **2003**, *63*, 5446–5453.
14. Hama, K.; Aoki, J.; Fukaya, M.; Kishi, Y.; Sakai, T.; Suzuki, R.; Ohta, H.; Yamori, T.; Watanabe, M.; Chun, J., et al. *J. Biol. Chem.* **2004**, *279*, 17634–17639.
15. Clair, T.; Koh, E.; Ptaszynska, M.; Bandle, R. W.; Liotta, L. A.; Schiffmann, E.; Stracke, M. L. *Lipids Health Dis.* **2005**, *4*, 5.
16. van Meeteren, L. A.; Ruurs, P.; Christodoulou, E.; Goding, J. W.; Takakusa, H.; Kikuchi, K.; Perrakis, A.; Nagano, T.; Moolenaar, W. H. *J. Biol. Chem.* **2005**, *280*, 21155–21161.
17. Gududuru, V.; Zeng, K.; Tsukahara, R.; Makarova, N.; Fujiwara, Y.; Pigg, K. R.; Baker, D. L.; Tigyi, G.; Miller, D. D. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 451–456.
18. Jiang, G.; Xu, Y.; Fujiwara, Y.; Tsukahara, T.; Tsukahara, R.; Gajewiak, J.; Tigyi, G.; Prestwich, G. D. *Chem. Med. Chem.* **2007**, *2*, 679–690.
19. Cui, P.; Tomsig, J. L.; McCalmont, W. F.; Lee, S.; Becker, C. J.; Lynch, K. R.; Macdonald, T. L. *Bioorg Med. Chem. Lett.* **2007**, *17*, 1634–1640.
20. Baker, D. L.; Fujiwara, Y.; Pigg, K. R.; Tsukahara, R.; Kobayashi, S.; Murofushi, H.; Uchiyama, A.; Murakami-Murofushi, K.; Koh, E.; Bandle, R. W., et al. *J. Biol. Chem.* **2006**, *281*, 22786–22793.
21. Durgam, G. G.; Virag, T.; Walker, M. D.; Tsukahara, R.; Yasuda, S.; Liliom, K.; van Meeteren, L. A.; Moolenaar, W. H.; Wilke, N.; Siess, W., et al. *J. Med. Chem.* **2005**, *48*, 4919–4930.
22. Keller, T. H.; Pichota, A.; Yin, Z. *Curr. Opin. Chem. Biol.* **2006**, *10*, 357–361.
23. Lipinski, C. A. *Adv. Drug Delivery Rev.* **1997**, *23*, 3–25.
24. Gijsbers, R.; Ceulemans, H.; Bollen, M. *Biochem. J.* **2003**, *371*, 321–330.
25. Cimpean, A.; Stefan, C.; Gijsbers, R.; Stalmans, W.; Bollen, M. *Biochem. J.* **2004**, *381*, 71–77.
26. Gijsbers, R.; Ceulemans, H.; Stalmans, W.; Bollen, M. *J. Biol. Chem.* **2001**, *276*, 1361–1368.
27. Bond, C. S.; Clements, P. R.; Ashby, S. J.; Collyer, C. A.; Harrop, S. J.; Hopwood, J. J.; Guss, J. M. *Structure* **1997**, *5*, 277–289.
28. Galperin, M. Y.; Jedrzejas, M. J. *Prot. Struct. Funct. Genetics* **2001**, *45*, 318–324.
29. Kim, E. E.; Wyckoff, H. W. *J. Mol. Biol.* **1991**, *218*, 449–464.
30. Lukatela, G.; Krauss, N.; Theis, K.; Selmer, T.; Gieselmann, V.; von Figura, K.; Saenger, W. *Biochemistry* **1998**, *37*, 3654–3664.
31. Zalatan, J. G.; Fenn, T. D.; Brunger, A. T.; Herschlag, D. *Biochemistry* **2006**, *45*, 9788–9803.
32. Kawagoe, H.; Stracke, M. L.; Nakamura, H.; Sano, K. *Cancer Res.* **1997**, *57*, 2516–2521.
33. Zhang, G.; Zhao, Z.; Xu, S.; Ni, L.; Wang, X. *Chin. Med. J. (Engl.)* **1999**, *112*, 330–332.
34. Black, E. J.; Clair, T.; Delrow, J.; Neiman, P.; Gillespie, D. A. *Oncogene* **2004**, *23*, 2357–2366.
35. Shida, D.; Watanabe, T.; Aoki, J.; Hama, K.; Kitayama, J.; Sonoda, H.; Kishi, Y.; Yamaguchi, H.; Sasaki, S.; Sako, A., et al. *Lab. Invest.* **2004**, *84*, 1352–1362.
36. Yun, C. C.; Sun, H.; Wang, D.; Rusovici, R.; Castleberry, A.; Hall, R. A.; Shim, H. *Am. J. Physiol. Cell Physiol.* **2005**, *289*, C2–C11.
37. Kitayama, J.; Shida, D.; Sako, A.; Ishikawa, M.; Hama, K.; Aoki, J.; Arai, H.; Nagawa, H. *Breast Cancer Res.* **2004**, *6*, R640–R646.
38. Yamada, T.; Sato, K.; Komachi, M.; Malchinkhuu, E.; Tobo, M.; Kimura, T.; Kuwabara, A.; Yanagita, Y.; Ikeya, T.; Tanahashi, Y., et al. *J. Biol. Chem.* **2004**, *279*, 6595–6605.
39. Boucharaba, A.; Serre, C. M.; Gres, S.; Saulnier-Blache, J. S.; Bordet, J. C.; Guglielmi, J.; Clezardin, P.; Peyruchaud, O. *J. Clin Invest.* **2004**, *114*, 1714–1725.
40. Shida, D.; Kitayama, J.; Yamaguchi, H.; Hama, K.; Aoki, J.; Arai, H.; Yamashita, H.; Mori, K.; Sako, A.; Konishi, T., et al. *Exp. Cell Res.* **2004**, *301*, 168–178.
41. Zhou, Y.; Peng, H.; Ji, Q.; Qi, J.; Zhu, Z.; Yang, C. *Bioorg Med. Chem. Lett.* **2006**, *16*, 5878–5882.
42. Floquet, N.; Richez, C.; Durand, P.; Maigret, B.; Badet, B.; Badet-Denisot, M. A. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1966–1970.
43. Evers, A.; Hessler, G.; Matter, H.; Klabunde, T. *J. Med. Chem.* **2005**, *48*, 5448–5465.
44. Ijjaali, I.; Petitet, F.; Dubus, E.; Barberan, O.; Michel, A. *Bioorg. Med. Chem.* **2007**, *15*, 4256–4264.
45. Zhang, S.; Wei, L.; Bastow, K.; Zheng, W.; Brossi, A.; Lee, K. H.; Tropsha, A. *J. Comput. Aided Mol. Des.* **2007**, *21*, 97–112.
46. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3–26.
47. Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. *J. Med. Chem.* **2002**, *45*(12), 2615–2623.

48. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.
49. Murata, J.; Lee, H. Y.; Clair, T.; Krutzsch, H. C.; Arestad, A. A.; Sobel, M. E.; Liotta, L. A.; Stracke, M. L. *J. Biol. Chem.* **1994**, *269*, 30479–30484.
50. MOE. In 2006.08 ed. Montreal: Chemical Computing Group; 2006.
51. Fujiwara, Y.; Osborne, D. A.; Walker, M. D.; Wang, D. A.; Bautista, D. A.; Liliom, K.; Van Brocklyn, J. R.; Parrill, A. L.; Tigyi, G. *J. Biol. Chem.* **2007**, *282*, 2374–2385.
52. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
53. Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490–519.