Bioorganic & Medicinal Chemistry 19 (2011) 693–701

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Benzopyrazine derivatives: A novel class of growth factor receptor bound protein 7 antagonists

Nigus D. Ambaye, Menachem J. Gunzburg, Reece C. C. Lim, John T. Price, Matthew C. J. Wilce [†], Jacqueline A. Wilce [*,†]

Department of Biochemistry and Molecular Biology, Monash University, Wellington Road, Victoria 3800, Australia

ARTICLE INFO

Article history:
Received 30 July 2010
Revised 8 October 2010
Accepted 12 October 2010
Available online 19 October 2010

Keywords:
Grb7
Adapter protein
SH2 domain
Benzopyrazines
Cancer cells
Similarity search
Virtual screening
ITC
Thermofluor
Melting point shift assay
Cellular growth assay

ABSTRACT

Growth factor receptor bound protein 7 (Grb7) is an adapter protein that functions as a downstream effector of growth factor mediated signal transduction. Over-expression of Grb7 has been implicated in a variety of cancers such as breast, blood, pancreatic, esophageal, and gastric carcinomas. Inhibition of Grb7 has been shown to reduce the migratory and proliferative potential of these cancers, making it an attractive therapeutic target. Starting with a known peptide antagonist, the present work reports the application of a succession of computational ligand design tools comprising a ligand shape based similarity search, molecular docking and a 2D-similarity search to identify small molecular antagonists of the Grb7-SH2 domain from the NCI chemical database. Binding to the Grb7-SH2 domain was then experimentally tested using melting point shift assays and isothermal titration calorimetry. Overall, a total of 11 benzopyrazine based small molecular antagonists were identified with affinity for the Grb7-SH2 domain. Representative compounds tested using ITC were revealed to possess moderate binding affinity in the low micromolar range. Finally, the lead compound (NSC642056) was found to reduce the growth of a Grb7-expressing breast cancer cell line with an IC_{50} of 86 μM. It is expected that the identified antagonists will be useful additions to further explore the function of Grb7 and for the development of inhibitors with therapeutic potential.

1. Introduction

Growth factor Receptor Bound protein 7 (Grb7) is an adapter protein that mediates signals from tyrosine kinases to effect downstream events.[1] Grb7 is of great interest as a therapeutic target in cancers in which it is aberrantly overexpressed with HER2 and impacts on cell proliferation and migration.[2] Grb7 co-overexpression has been implicated in a variety of human cancers such as breast,[3] blood,[4] pancreatic,[5] esophageal[6] and gastric carcinomas.[7] Moreover, the level of Grb7 over-expression is shown to correlate with severity of Barrett's carcinoma[8] and lymphocytic leukemia.[4] In addition to growth factor dependent signaling, Grb7 is also known to associate with focal adhesion kinase[9] and to be a component of integrin-mediated signal transduction which is an important pathway in cancer cell migration.[10] The importance of Grb7 in cancer has been established by a comparative analysis of the characteristics of Grb7 over-expressing and Grb7 knockout cancer cells.[11,12] Finally, Grb7 has been demonstrated to be a targetable protein that can be blocked by synthetic agents resulting in a significant reduction of cancer cell viability.[5] Put together, these investigations make Grb7 a highly attractive target for the design of anti-cancer agents.

Grb7 is a multi-domain protein composed of a proline rich N-terminal domain, a Ras-associating domain, a plekstrin homology (PH) domain, a C-terminal Src homology 2 (SH2) domain and a BPS domain (between the PH and SH2 domains).[13,14] Each domain serves a distinct role that contributes to the overall signaling function of Grb7.[15] The physical association of Grb7 with its activated upstream binding partners is predominantly mediated by the C-terminal SH2 domain of Grb7. The SH2 domain bears a cationic pocket that is the binding site of phosphorylated tyrosines. Specific phosphotyrosines are recognized based upon the sequence surrounding the phosphotyrosine residue.[15] The SH2 domain is also found in a number of Grb7-related and other adaptor proteins guiding their specific interactions with upstream binding partners.[16–18] This step is the most critical event in growth factor dependent Grb7 mediated signal transduction. As such, the SH2 domain forms an essential module for Grb7 mediated oncogenic signaling. Since the Grb7-SH2 domain is readily produced and its crystal structure is available, it provides a clear opportunity for inhibitor development via computational and biochemical studies.[19]

Virtual screening has been developed as a powerful tool for the acceleration of drug discovery, specifically in the early part of lead compound design.[20,21] Ligand based virtual screening, a broad sub-

* Corresponding author. Tel.: + 61 3 9902 9226; fax: +61 3 9902 9500.
E-mail address: jackie.wilce@monash.au (J.A. Wilce).
† Co-senior authors.

0968-0896/$ - see front matter Crown Copyright © 2010 Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2010.10.030

category of virtual screening, is based on the principle that similar structures produce similar effects.[22] Depending on the representation and description of lead structures, several approaches exist that transform such information into a model that can be used to scan compound repositories.[22,23] Most commonly, chemical similarity measures based on two or three-dimensional structure or molecular shape are employed in ligand based virtual screening efforts[24,25] as well as in optimization tools such as quantitative structure–activity relationship,[26,27] molecular docking[28] and pharmacophore mapping.[29,30]

There have been numerous reports of the design and optimization of peptide-derived agents geared towards inhibition of the broad class of growth factor receptor bound adaptor proteins, via their SH2 domains.[31,32] A variety of tactics ranging from phosphorylation[33] macrocyclization[34] conformational restriction[35] and use of amino acid surrogates[36] have been exploited to optimize the binding affinity of the lead peptides. Indeed, this has remarkably resulted in the development of effective inhibitors[37] and contributed to a better understanding of the structural and functional investigations of the SH2 domains. More specifically, Pero et al. have utilized a phage display techniques to discover an 11-residue non-phosphorylated polypeptide (termed G7-18NATE) antagonist of Grb7.[5,38] This peptide does not possess a rigid structure in solution but interacts with the Grb7-SH2 domain with micromolar affinity in a predicted turn-structure at the phosphotyrosine binding site.[19,39] Cell-permeable forms of this peptide have been shown to maintain their binding to Grb7-SH2 and to have anti-proliferative and anti-migratory activity in cancer cell lines.[5,40,41] Together this demonstrates the potential for non-phosphorylated compounds to bind to the Grb7-SH2 domain and to possess anti-cancer activity. Experiments conducted in vitro on G7-18NATE and other phosphorylated peptides demonstrate the potential for peptides to be modified to enhance their affinity for Grb7-SH2.[42,43] However, almost all the identified antagonists to date belong to a single chemical class, that is, peptides, and thus are subject to issues inherent to peptidic structures such as bioavailability, permeability and stability that remain a stumbling block for the application of these molecules in vivo. Most presumably due to these bottlenecks, to date there has not been a single peptide that has progressed to the development of a useful therapeutic agent targeting Grb7 and related growth factor binding proteins, despite extensive efforts.[31] With this in mind, the present work was initiated as a first step in the development of a potent, selective and stable small molecular antagonist of the Grb7 oncoprotein. To the authors' knowledge this is the first time a small molecular antagonist is reported to bind to the Grb7-SH2 domain. This work reports a shape based virtual screening, molecular docking, 2D-similarity search, melting point shift assay and isothermal titration calorimetric study carried out to identify a novel structural class of Grb7 antagonist. We finally report a preliminary functional study carried out with the lead compound that demonstrates its growth inhibitory effects in a breast cancer cell line.

2. Results

2.1. Shape based query using known antagonist of the Grb7-SH2 domain

Given the limited number and lack of structural diversity in the so far reported antagonists of Grb7, a shape based virtual screening was undertaken as a starting point to identify a novel lead. Structural studies of SH2 domains in complex with peptide-based ligands and numerous binding studies of modified peptide ligands have revealed the common features of their interactions.[18] Thus, as a starting point to the filtering of compounds, a shape based query was generated using a tripeptide (*m*-aminobenzoyl-pY(α-Me)pYN-NH_2) shown previously to be the most potent antagonist of growth factor bound adaptor 2 (Grb2) and Grb7 proteins and also the smallest and simplest of the known antagonists.[44] As the experimental conformation of the tripeptide in complex with the Grb7-SH2 domain was not available, a conformational search strategy was pursued to establish possible bioactive conformations. To this end, the MacroModel[45] conformation search[46] engine in Schrodinger[47] was employed to generate stable conformers for interrogation.

In order to select the conformer that may represent the optimum shape for the shaped-based query, a test database of known antagonists of Grb7-SH2 with determined comparative experimental affinities[43] was prepared using the Maestro interface of Schrodinger.[47] Since the lowest energy conformer is not necessarily the bioactive conformer,[46] all of the top five conformers of the query tripeptide were evaluated for their suitability as a shape based query model by considering their ability to rank-order the known antagonists. This assumes that all the known antagonists bind at the same site on the same receptor as the query tripeptide. The program PHASE[48] was employed to carry out a shape based virtual screening experiment.

Table 1 indicates the shape similarity coefficient obtained using the top five stable conformers of the tripeptide as shape based queries against the mini-database of known Grb7 antagonists. The correlation coefficient obtained between the shape based similarity indices of the antagonists and their pIC_{50} values depict the strength of the relationship. The best correlation obtained is 0.93, corresponding to conformer 5. Thus, although the differences in the peptide affinities for their target do not span a wide range, which limits the significance of the correlation, conformer 5 was selected as the basis for the shape based query in the subsequent lead identification procedure.

2.2. Virtual screening of the NCI chemical database

A schematic representation of the virtual screening flowchart used is displayed in Figure 1. The tripeptide conformer 5 was used as a shape based query to screen the National Cancer Institute (NCI) database (which contained 250, 251 structures at the time of this work) [http://dtp.nci.nih.gov/]. This returned a total of 63,300 hits when the similarity score cut-off was set at 0.41. The top hits from the shape based search were further screened using the Glide based molecular docking protocol.[49,50] The top 5000 from the search were docked using Glide's high-throughput virtual screening (Glide HTVS) docking routine. This initial crude docking protocol resulted in 1886 successfully docked hits. All of the 1886 hits were subsequently re-docked using the more accurate Glide standard precision (Glide SP) docking procedure. Inspection of the failed structures revealed that the rejects were either simple or polycyclic hydrocarbons that met the shape based criteria. The Glide SP docking experiment resulted in 1854 of the 1886 hits being successfully re-docked. The top 200 from the Glide SP docking were finally docked with the extra precision Glide (Glide XP) docking routine. The molecular docking protocol was the same throughout the computational experiment. Of the 200 hits identified, 15 potential candidates with favorable Glide SP or Glide XP docking scores, and with good predicted surface complementarity and interactions at the target Grb7-SH2 domain surface upon examination, were selected. The compounds identified and their docking scores are shown in Table 2.

2.3. Predicted binding mode of hit NSC642056

An example of a compound selected by the above criterion, a benzopyrazine numbered NSC642056, is shown in its docked

Table 1
Shape similarity index of known Grb7 antagonists using the top five conformers of the *m*-aminobenzoyl-pY(α-Me)pYN-NH_2 tripeptide as a query

S. No	Known antagonist	pIC_{50}[a]	Shape similarity index[b]				
			Conf. 1	Conf. 2	Conf. 3	Conf. 4	Conf. 5
1	PEpYVNQ-NH_2	5.96	0.54	0.50	0.48	0.39	0.47
2	Ac-pYVNQ-NH_2	5.47	0.51	0.46	0.39	0.43	0.35
3	EpYVNQ-NH_2	5.815	0.51	0.48	0.43	0.45	0.45
4	pYVNQ-NH_2	5.45	0.45	0.48	0.43	0.41	0.38
Correlation coefficient[c]			0.58	0.62	0.63	0.04	0.93

[a] The pIC_{50} values were obtained from the reported IC_{50} value[43] using the equation $pIC_{50} = -\log IC_{50}$.
[b] Conf. 1–5 represent the top five conformers of the query tripeptide used in the virtual screening experiment.
[c] The correlation coefficient indicates the correlation between the pIC_{50} value and the shape similarity value for each conformer.

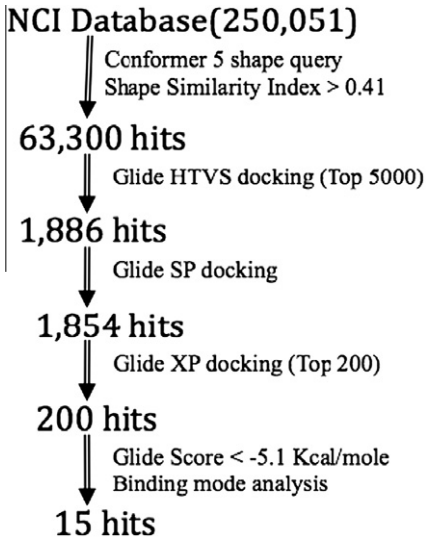

Figure 1. Schematic representation of the virtual screening experiment conducted based on *m*-aminobenzoyl (α-Me)pY-NH_2 tripeptide as shape based query.

geometry in Figure 2. NSC642056 is predicted to form a number of specific binding interactions with the binding site of the Grb7-SH2 domain. The phenyl ring of the benzoyl moiety is appropriately positioned to make favorable hydrophobic/van der Waals interactions with the side chains of Leu481 and Tyr480 within the binding site. A potential hydrophobic/van der Waals bond with the side chains of Val468 and His479 is also apparent with the aromatic system of the hit. In addition, the hit also forms multiple hydrogen bonding interactions with the Grb7-SH2 domain binding site. The carbonyl adjacent to the ester moiety is observed to make two hydrogen bonding interactions with the guanidine group of Arg458. Moreover two other hydrogen bonding contact pairs occur between each of the di-oxo moieties on the anilinobenzene carboxamide with the backbone NHs of Arg462 and Asn461 residues. Finally the carboxamide amino moiety is involved in a hydrogen bonding interactions with the Gln463 of Grb7-SH2 binding site.

2.4. Experimental pre-screening with melting point shift assay

Out of the top 15 potential hits, five were available at the time of this work and kindly provided by the NCI Developmental and Therapeutic program unit of the USA (Drug synthesis and Chemistry branch, http://dtp.nci.nih.gov/) (indicated in Table 2). These were subjected to in vitro binding studies, as supplied by the NCI without further characterization. Experimental pre-screening of these compounds was carried out using a Thermofluor based thermal shift assay.[52,53] This assay is based upon the increased melting temperature of a protein induced by ligand binding. The compounds were dissolved in 1% DMSO/buffer solution and their effect on the melting temperature of the Grb7-SH2 domain was determined. One of out of the five tested compounds, NSC642056, produced a positive melting point shift (Table 3). Interestingly, the finding that hit NSC642056 interacts with the Grb7-SH2 domain correlates with its better shape based similarity coefficient and glide docking score relative to those of the four structures that turned out to be false hits. Indeed, NSC642056 occurs in the top 5 of the priority list while the negative hits were in the bottom 5 of the 15 hits suggested from the combined virtual screening experiment. The change in melting point produced by the hit

Table 2
Final hits identified based on the tripeptide based shape search

Rank	Hit ID.	Shape similarity index	Glide SP docking		Glide XP docking	
			Glide score	Emodel	Glide score	Emodel
1	NSC677186	0.433	−6.348	−79.584	−5.130	−95.586
2	NSC677177	0.463	−6.337	−84.217	−4.790	−62.836
3	NSC677184	0.429	−5.883	−73.463	−4.190	−89.924
4	NSC633047	0.480	−5.792	−68.058	−3.420	−79.200
5	NSC642056*	0.457	−5.749	−71.077	−4.900	−81.612
6	NSC677147	0.437	−5.716	−69.901	−3.200	−77.713
7	NSC677176	0.427	−5.484	−72.001	−3.310	−80.790
8	NSC713163	0.413	−5.368	−44.324	−3.760	−45.198
9	NSC677174	0.427	−5.318	−63.131	−3.560	−74.519
10	NSC100787	0.428	−5.277	−63.580	−4.550	−84.740
11	NSC129843*	0.435	−5.264	−75.089	−3.780	−83.889
12	NSC642057*	0.436	−5.256	−67.214	−3.540	−76.789
13	NSC344032	0.441	−5.252	−64.706	−3.220	−77.365
14	NSC687783*	0.426	−5.241	−65.861	−3.750	−68.422
15	NSC657678*	0.458	−5.223	−61.619	−3.870	−74.176

* Indicates hits that are tested experimentally.

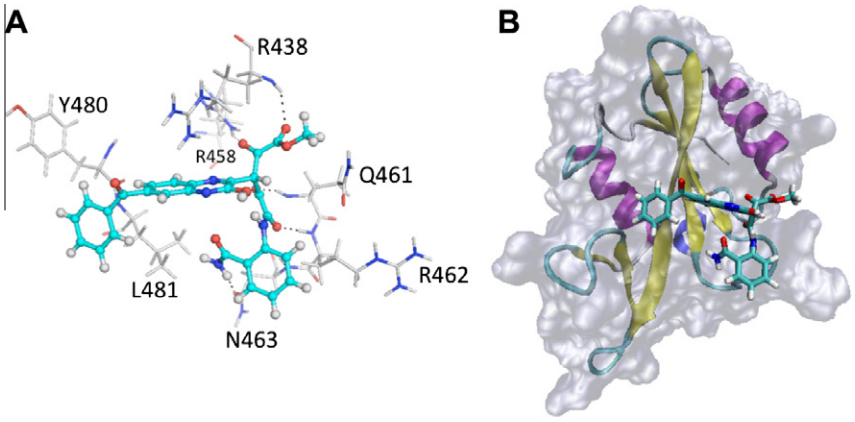

Figure 2. Binding mode of lead NSC642056. (A) The specific interaction between the ligand and Grb7-SH2 domain. The ligand is displayed in ball and stick format while binding site residues are shown in line format. The yellow dotted lines represent hydrogen bonds between the hit and Grb7 binding site; the figure was produced in Maestro of Schrodinger suite. (B) The overall fold of Grb7-SH2 domain depicting its surface, secondary structural elements and the relative positioning of the lead compound. The figure was produced by virtual molecular dynamics (VMD) visualization program.[51]

Table 3
Tanimotto similarity coefficient and melting point shift data for benzopyrazine derivatives

S. No.	Hit code	Tanimotto similarity coefficient (%)	ΔT_m[a] (°C)	
			100 μM	200 μM
1	NSC642040	97	0.45	0.95
2	NSC648605	98	0.05	0.55
3	NSC644743	92	0.95	0.95
4	NSC648607	97	0.45	1.45
5	NSC644745	90	1.30	1.30
6	NSC644750	94	0.45	0.85
7	NSC644770	94	0.45	0.55
8	NSC646820	90	1.20	0.95
9	NSC646823	95	1.00	0.70
10	NSC648589	92	0.90	0.60
A	NSC642056	100	0.60	1.10
B	G7-18NATE	Not applicable	1.35	1.85

[a] The change in melting point (ΔT_m) are obtained from duplicate experiments after the melting point of Grb7-SH2 apo (T_m = 49 °C) was subtracted from that of the hit bound form. NSC642056 was the lead compound while G7-18NATE is the reference polypeptide concurrently run in the melting shift assay.

was 0.6 °C at 100 μM and 1.1 °C at 200 μM which were slightly lower than that produced by a peptide-based (G7-18NATE)[38] Grb7-SH2 inhibitor control (1.35 °C at 100 μM and 1.85 °C at 200 μM). Thus, hit NSC642056 was able to effect stabilization comparable to the known G7-18NATE-based polypeptide inhibitor which suggests that the compound may have potential as an antagonist of Grb7.

2.5. Lead discovery using 2D-similarity searches and binding studies

In an effort to discover further compounds with affinity for Grb7-SH2, structures related to NSC642056 were sought. To this end, analogues of the hit were screened using a two-dimensional (2D) structure based similarity search.[23] The 2D-search was undertaken against the same NCI database using the SMILES[54] description of the lead structure as an input. The use of the lead NSC642056 as a 2D-query returned 22 benzopyrazine hits of which 15 were available at the time of this work from the NCI developmental and Therapeutic program. The 15 compounds were then assessed using the melting point shift assay as described above. The result, shown in Table 3, revealed that out of the 15 structural analogues tested 10 were able to produce a positive melting point shift. The change in melting point induced by the analogues was found to be comparable to that of the lead NSC642056 compound (Fig. 3). The structures of the 10 leads identified in addition to the starting structure are shown in Figure 4. The 10 identified compounds were sub-classified into di-substituted benzopyrazines (NSC644743, NSC646820, NSC646823, and NSC648589), tri-substituted benzopyrazines with benzoyl substituent (NSC642040, NSC648607, and NSC648605) and tri-substituted benzopyrazines with a nitro substituent (NSC644750, NSC644770, and NSC644745). Though these are analogues of NSC642056, the virtual screening criteria adopted in the initial shape based search might have prevented these structures from showing up within the top 15 of the shape based hits.

2.6. Confirmation of binding by ITC

Further characterization of these interactions was undertaken using isothermal titration calorimetry (ITC).[55] ITC measures the amount of heat absorbed or released as one interacting species is titrated into the other. From this, the equilibrium dissociation constant, free energy, enthalpy, entropy and stoichiometry of binding

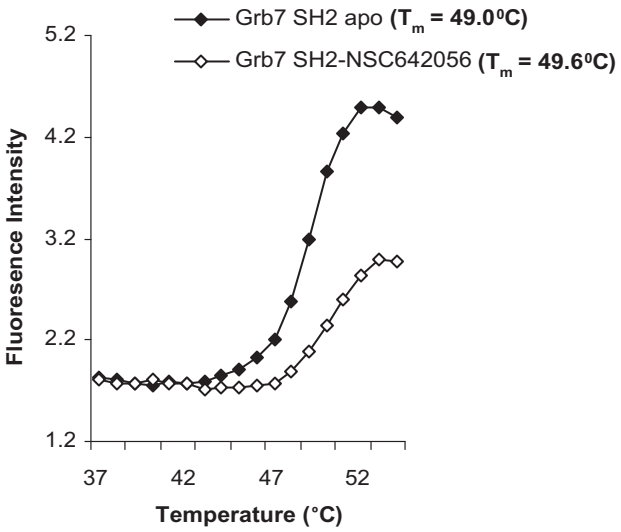

Figure 3. Unfolding curve of Grb7-SH2 domain obtained from thermal melting experiment. The filled square curve corresponds to the apo form of Grb7 (T_m = 49.0 °C) while the open square curve represents that of Grb7–NSC642056 complex (T_m = 49.6 °C).

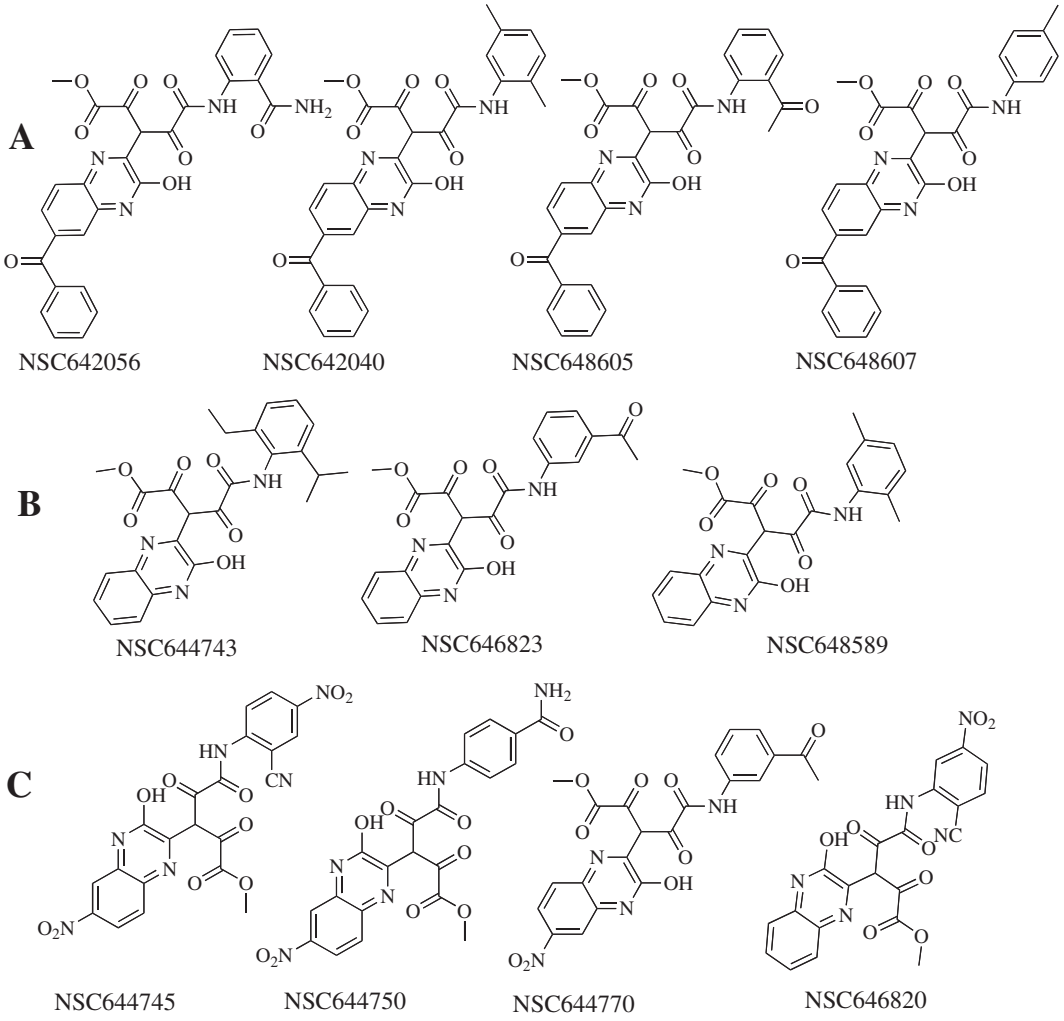

Figure 4. Chemical structures of the benzopyrazine derivatives. (A) Tri-substituted benzopyrazines with benzoyl substituent. (B) Di-substituted benzopyrazine derivatives. (C) Tri-substituted benzopyrazines with a nitro substituent.

can be determined.[56] The binding of NSC642056 as well as 5 representative compounds from three classes of substituted benzopyrazines, identified by the Grb7-SH2 domain melting point assay, were determined by ITC. As shown in Table 4, the equilibrium dissociation constant for hit NSC642056 was found to be 17.15 μM. This not only confirms the binding but demonstrates that it occurs with a moderate affinity. The binding data shows that the compounds interact with favorable enthalpy and unfavorable entropy with a 1:1 binding stoichiometry. In a similar vein, the analogue structures (representative structures from each subclass—NSC642040, NSC644745, NSC644743, NSC646823, and NSC648607)—were subjected to the ITC binding assay. Figure 5 shows the ITC binding isotherm for one of the analogues tested.

Each of the compounds characteristically bound with evolution of heat, that is, through exothermic binding similar to the lead NSC642056 (Table 4). The experimental affinities were found to be of moderate values ranging from K_d = 3.06 μM for NSC646807 to K_d = 17.15 μM for NSC642056 lead. Though the free energy of binding is in a very close range (ca. −7 kcal/mol), deconvolution into its constituents reveals subtle differences in the entropic and enthalpic contributions to binding. In the case of NSC648607 and NSC644743, the enthalpic contribution to binding is greater than the entropic part. However, for leads NSC644745 and NSC646823 the enthalpic/entropic ratio is reversed. The fact that the gross binding affinities remained virtually unaffected despite the differences in the relative contribution of the constituent parts suggests

Table 4
Thermodynamic binding parameters for the benzopyrazine derivatives

Lead	NSC642040	NSC644745	NSC646823	NSC644743	NSC6468607	NSC642056
Stoichiometry	1.11 ± 0.03	1.14 ± 0.03	1.04 ± 0.11	1.06 ± 0.03	1.04 ± 0.02	1.09 ± 0.04
K_b (10^5 M^{-1})	2.06 ± 0.42	1.58 ± 0.35	1.41 ± 0.10	1.85 ± 0.46	3.48 ± 0.87	0.60 ± 0.10
ΔH (Kcal/mol)	−2.94 ± 0.11	−1.33 ± 0.05	−0.74 ± 0.09	−6.11 ± 0.32	−5.54 ± 0.19	−10.23 ± 0.51
ΔS (cal/mol/K)	14.35	19.3	21.1	3.61	6.89	−12.5
$T\Delta S$ (Kcal/mol)	4.28	5.75	6.29	1.08	2.05	−3.73
ΔG (Kcal/mol)	−7.21 ± 0.11	−7.08 ± 0.05	−7.03 ± 0.09	−−7.19 ± 0.32	−7.56 ± 0.19	−6.50 ± 0.51
K_d (μM)	5.06 ± 1.03	6.57 ± 1.20	7.12 ± 0.50	5.76 ± 1.40	3.06 ± 0.77	17.15 ± 2.97

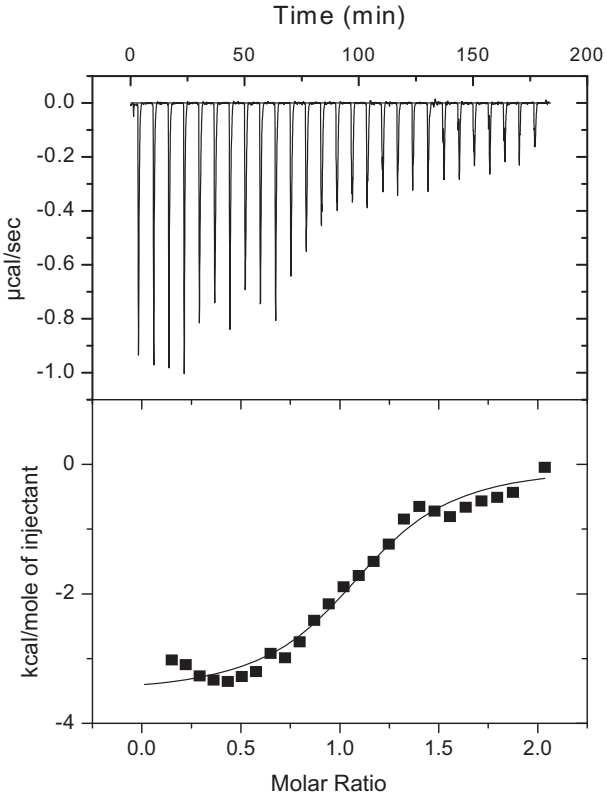

Figure 5. ITC thermogram of a benzopyrazine derivative, NSC642040. The upper panel show raw data obtained from 10 µl injections of compound at 25 °C. The lower panel display plots of integrated total energy exchanged (as kcal/mol of injected compound) as a function of molar ratio of the compound to the Grb7-SH2 domain (squares). Solid lines indicate the fit to a single-site saturation model.

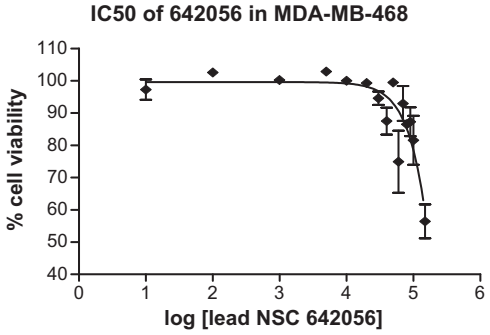

Figure 6. Cellular growth inhibition curve obtained for MDA-MB-468 breast cancer cells treated with the lead NSC642056. The IC_{50} was found to be 86 µM.

that entropy–enthalpy compensation phenomenon taking place. This is commonly observed in ligand–receptor binding studies.[57]

2.7. Cellular growth inhibitory assay

Finally, a cellular growth assay was conducted using the MDA-MB-468 breast cancer cell line to determine whether the lead compound would display activity in an in vivo system. To this effect, the impact of compound NSC642056 on these Grb7-expressing cancer cells was investigated. The result, as shown in Figure 6, shows that the lead compound impacts on the growth of the cells in a concentration dependent manner. The in vivo IC_{50} determined for the cell growth inhibition is found to be 86.0 µM. This is compared to the in vitro K_d of 17.3 µM, and thus suggests a correlation between the inhibitor binding affinity for the Grb7 target and the concentration required for cell growth inhibition. Future investigations, using both cells in which Grb7 is knocked out using RNAi and knocked back in, will allow us to identify the extent to which the inhibitor action is indeed occurring through binding and inhibition of Grb7.

3. Discussion

Few antagonists of Grb7 have previously been identified. Those reported include the polypeptide G7-18NATE[5,38,39] identified through phage display and found to be selective for Grb7 with minimal activity against related proteins. Other short peptides are also reported to bind to the Grb7-SH2 domain in vitro.[43] Being peptidic in nature, however, the potential for their in vivo use is restricted until major issues such as the limited bioavailability, in vivo instability, cell impermeability and high synthetic costs of peptides are resolved.[58] Here we identify compounds of the benzopyrazine class that have an equivalent affinity for the Grb7-SH2 domain as the previously reported peptides (i.e., moderate to low micromolar binding affinity), and suggest that these may offer an alternative scaffold for the development of potent inhibitors of Grb7. Similarly to the in vivo activity of the peptide-based Grb7 inhibitor, G7-18NATE, inhibition of the growth of cancer cells is observed.[40] Future experiments will be required to demonstrate whether the mode of action is specifically via Grb7 inhibition, or also/alternatively through other pathways. The current studies do not establish whether these compounds possess intrinsic specificity for the Grb7-SH2 domain. Being predicted to bind to the peptide recognition surface of the Grb7-SH2 domain, however, we propose that they would be amenable to modification to improve their affinity, pharmacological properties and potentially to confer specificity for Grb7. The compounds would not be the target of proteases, which is a major obstacle for the use of peptides as drugs, and their relatively smaller size is predicted to improve their bioavailability to their intracellular target.

Benzopyrazines from the NCI database were identified in the current study owing to their ability to adopt a structurally similar shape to a known peptide-based inhibitor of the Grb7-SH2 domain. The present work effectively exploited the molecular similarity principle that shape based similarity can reflect similarity in potential binding interactions.[59,60] The shape based virtual screening experiment was able to identify moderate affinity lead candidates, as confirmed by a Thermofluor binding assay. The identification of a substituted benzopyrazine (NSC642056) prompted a 2D-similarity search, resulting in the identification of further benzopyrazines with potential for binding to the Grb7-SH2 domain. Binding to the target was confirmed experimentally using both melting point assays and ITC. Overall, the computational experiments attest to the usefulness of the application of a succession of virtual screening tools in lead compound identification.

The Thermofluor binding assay proved an effective experimental method to confirm ligand binding. It is a rapid and inexpensive assay that does not utilize very much material. Thermofluor exploits the concept of protein stabilization upon ligand binding. Only a one degree change in the melting temperature of the protein was observed upon interaction with the lead compounds. This relatively low value is typical for ligands binding with micromolar affinity and may represent the limit of detection of the method.[54,55,61] Similar thermal shift values were measured for a known peptide inhibitor of the Grb7-SH2 domain that binds with micromolar affinity. ITC was used to confirm the interactions, measure the affinity and also provided insight into the entropic versus enthalpic contributions to binding. Though the identified benzo-

pyrazines bound with only micromolar binding these small molecule antagonists could be useful for the experimental and clinical investigations of Grb7 and related proteins.

In conclusion, starting with a known peptide antagonist, the present work reported the successful application of a succession of computational ligand design tools comprising ligand shape based similarity search, molecular docking and 2D-similarity search to identify a novel class of small molecular antagonists of Grb7. Experimental binding studies with a melting point shift assay and ITC confirmed the interactions and shed some light on the thermodynamic aspects of the binding interactions. That the lead NSC642056 antagonist reduced the growth of breast cancer cells was confirmed in the MDA-MB-435 cell line. Further work will be required to determine whether the mode of action is, indeed, via Grb7 inhibition. It is expected that the newly identified benzopyrazine scaffolds will be useful additional agents to further explore and accelerate functional studies of Grb7 and potentially lead to the development of novel therapeutics.

4. Experimental section

4.1. Generation of the shape based query

MacroModel's[45] conformation search[46] engine in Schrodinger[47] was employed to generate stable conformers of the *m*-aminobenzoyl-pY(α-Me)pYN-NH_2 peptide, previously shown to be a potent inhibitor of the Grb7-SH2 domain.[43] The optimal potential for liquid simulation force field with a distance dependent dielectric constant was used during the conformation search. Partial charges were computed using the force field. Geometry optimization was effected with maximum iteration of 5000 and a gradient convergent criterion was applied with a threshold of 0.05. The systematic torsional sampling option was employed to generate representative conformations of the tripeptide reference query. Unless specifically mentioned, all other conformation search parameters were kept at their default value.

4.2. Shape based similarity virtual screening

The program PHASE[48] was employed to carry out shape based virtual screening experiments. The atom type used for volume scoring was set to none so as not to bias the search towards a particular hit class particularly for the subsequent database screening. The conformer that produced the best correlation between the shape based similarity index and the biological activities of the peptides in the mini-database was then employed to serve as a query to search the enhanced National Cancer Institute (NCI) database. The average similarity index of the known antagonist was then used as a cut-off during the NCI screening.

4.3. Molecular docking

The program Glide was used to perform Grb7-SH2 domain docking studies for top hits from the shape based search. Glide comprises three docking routines that are designed to suit different stages of chemical database screening.[49,50] The PDB file for the Grb7-SH2 domain was read into Maestro in Schrodinger and the protein preparation wizard used to set up the protein including addition of hydrogen atoms and deletion of water molecules beyond 5 Å from protein atoms. Hydrogen bond assignment was effected with the exhaustive sampling option and the system was subsequently relaxed with geometry minimization. A grid of size $12 \times 12 \times 12$ Å was defined using the peptide binding surface about the phosphotyrosine pocket to define the binding site. All the other docking parameters were kept at their default values.

4.4. Expression and purification of Grb7-SH2 domain

The SH2 domain of Grb7 was used as a receptor for the thermal shift and ITC binding experiments. The pGex2T plasmid containing the Grb7-SH2 insert encoding residues 415–532 of human Grb7 protein was expressed in *Escherichia coli* expression system as a GST fusion protein in BL21(DE3)pLysS host strains.[62,63] Cell induction was achieved with 400 μM isopropyl β-D-1-thiogalactopyranoside. Cell resuspension was carried out in ice-cold phosphate buffer saline with 2 mM EDTA, 0.5% Triton-X 100 and lysis by sonication. The fusion protein was purified by glutathione affinity chromatography (GE Healthcare) and cleaved with thrombin. Purification using cation exchange chromatography was then effected using a HiTrap SP column (GE) after an overnight dialysis into 20 mM HEPES pH 7.4, 20% glycerol, and 1 mM DTT at 4 °C. This was followed by dialysis into 50 mM MES pH 6.6, 100 mM NaCl, and 1 mM DTT and final purification by size exclusion chromatography using a Sephadex 75 XK 16/60 column (GE). Grb7-SH2 domain was concentrated and stored at 4 °C. Its purity was verified using SDS–PAGE. The final concentration was determined with UV–vis spectroscopy.[64]

4.5. Thermofluor based melting point shift assay

The thermofluor based melting point shift assay relies on an increase in melting point of the target protein as an indication of compound binding.[52,53] The dye Sypro Orange is used to detect the melting transitions through its binding to exposed hydrophobic regions of the unfolding protein. The melting point of the Grb7-SH2 domain was determined with the use of Quiagen's quantitative RT-PCR instrument (www.qiagen.com). The wavelength of excitation/emission was set to 470/555 and 470/610 nm. A heating rate of at 1 °C per min was programmed over a 30–80 °C temperature range with a holding time of 60 s. The buffer used was 50 mM MES, 100 mM NaCl, pH 6.6, 1 mM dithiothreitol (DTT). The purified Grb7-SH2 domain was added to each tube to yield a final concentration of 20 μM. The protein concentration was established using UV–vis spectrophotometer at 280 nm and a molar extinction coefficient of 8480 M^{-1} cm^{-1}. All the compounds were dissolved in 1% DMSO/buffer solution and a final concentration of 100 μM and 200 μM was added to the protein. After mixing, 25 μl of the protein/hit solution was transferred to the PCR tubes for the melting experiment. A known peptide inhibitor[38] was also tested as a positive control. Each denaturation experiment was conducted in duplicate and in each case the average melting temperature was taken for analysis. The Rotor-Gene 3000 software was used to set up the experiment and analyze the results.

4.6. Isothermal titration calorimetry

ITC binding experiments were performed on a VP-ITC Microcalorimeter (Microcal, Northampton, MA, USA) at 25 °C.[55,56] The Grb7-SH2 domain was dialyzed against 2 L of 50 mM $NaOOCCH_3$ (pH 6.6), 100 mM NaCl, and 1 mM DTT. The concentration of Grb7-SH2 and hits was determined as described in the melting experiment. The hit and protein solutions were degassed and thermostated at 20 °C for 5 min using the ThermoVac of the VP-ITC Microcalorimeter. The reference power of the experiment was set to 20 μCal/s and the cell contents were stirred continuously at 307 rpm throughout the titrations. The feedback mode gain was set to high with fast and auto equilibrations options applied. The reference cell was filled with MQ water. The hits at concentrations in the range of 500–950 μM were titrated into Grb7-SH2 solutions of 50–85 μM concentrations in 7–10 μl injections with a 3–6 min delay between each injection. A binding isotherm was generated by plotting the heat change evolved per injection against the molar

ratio of the leads to Grb7-SH2 domain receptor. A blank determination in which the compounds were titrated against the buffer was carried out to account for heats of dilution and mixing which was then subtracted from the binding data. The corrected data was then fitted by a single binding site model using a non-linear least squares with the Origin (Microcal Software, Northampton, MA, USA).

4.7. Lead optimization by 2D-similarity search

Two-dimensional structure based similarity[22,23] searches of the National Cancer Institute database (http://dtp.nci.nih.gov/) were conducted using the SMILES[53] description of chemical structure using the lead NSC642056 as input. The SMILES string was generated using the enhanced NCI database browser platform. The maximum number of hits was set to 1000. Hits with Tanimotto similarity coefficients above 90% were obtained for further subsequent experimental investigation.

4.8. Cellular growth inhibitory studies

To determine the effects of the compounds on cellular growth, MDA-MB-468 breast cancer cells were resuspended at 5×10^4 cells/mL and 100 μL of this cell suspension was plated in 96 well plates (BD Biosciences, California, USA). The following day cells were treated with a range of concentration of the compounds, each concentration with triplicates wells. Cells were incubated under standard growth conditions with 10% fetal bovine serum over a 4-day period (0–4 days). Adherent cell cultures were fixed in situ after 4 days by the addition of 25 μl of cold 50% trichloroacetic acid solution (Sigma–Aldrich, Missouri, USA) and incubated for 60 min at 4 °C. The supernatant was then discarded and the plate was washed five times with running distilled water before being dried overnight. Fixed cells were stained by the addition of 100 μl of sulforhodamine B (SRB) sodium salt solution (0.4% (w/v) in 1% acetic acid; Sigma–Aldrich, Missouri, USA) and incubated for 10 min at room temperature. Unbound SRB solution was removed by washing five times with 1% acetic acid after which the plates were air-dried overnight. The following day, the bound stain was solubilized with 100 μl of 10 mM Tris buffer (pH 10.5) and the absorbance of the eluted dye was read on the spectrophotometric BMG PolarStar Optima plate reader (BMG Labtech, Offenburg, Germany) at 540 nm. SRB is a total protein stain, which directly correlates with cell number and is used routinely to measure cell growth rate. The Nonlinear regression analysis of the concentration–response plot was conducted using GraphPad Prism software.

Acknowledgments

This work was funded by an Australian Research Council Project Grant awarded to J.A.W. and D.N.K. and a Monash University Graduate Scholarship awarded to N.D.A. M.C.J.W. and J.T.P. are the recipients of a Senior Research Fellowship and a R.D. Wright Fellowship respectively from the National Health and Medical Research Council of Australia. We thank the Drug Synthesis and Chemistry Branch of National Cancer Institute of the USA for providing the compounds used in the screening and testing. The Victorian partnership for advanced computing (VPAC) is acknowledged for granting us access to the computational facility employed in the study.

References and notes

1. Shen, T. L.; Guan, J. L. *Front Biosci.* **2004**, *9*, 192.
2. Nencioni, A.; Cea, M.; Garuti, A.; Passalacqua, M.; Raffaghello, L.; Soncini, D.; Moran, E.; Zoppoli, G.; Pistoia, V.; Patrone, F.; Ballestrero, A. *PLoS One* **2010**, *5*, e9024.
3. Stein, D.; Wu, J.; Fuqua, S. A.; Roonprapunt, C.; Yajnik, V.; D'Eustachio, P.; Moskow, J. J.; Buchberg, A. M.; Osborne, C. K.; Margolis, B. *EMBO J.* **1994**, *13*, 1331.
4. Haran, M.; Chebatco, S.; Flaishon, L.; Lantner, F.; Harpaz, N.; Valinsky, L.; Berrebi, A.; Shachar, I. *Leukemia* **2004**, *18*, 1948.
5. Tanaka, S.; Pero, S. C.; Taguchi, K.; Shimada, M.; Mori, M.; Krag, D. N.; Arii, S. *J. Natl. Cancer Inst.* **2006**, *98*, 491.
6. Tanaka, S.; Mori, M.; Akiyoshi, T.; Tanaka, Y.; Mafune, K.; Wands, J. R.; Sugimachi, K. *Cancer Res.* **1997**, *57*, 28.
7. Kishi, T.; Sasaki, H.; Akiyama, N.; Ishizuka, T.; Sakamoto, H.; Aizawa, S.; Sugimura, T.; Terada, M. *Biochem. Biophys. Res. Commun.* **1997**, *232*, 5.
8. Walch, A.; Specht, K.; Braselmann, H.; Stein, H.; Siewert, J. R.; Hopt, U.; Höfler, H.; Werner, M. *Int. J. Cancer* **2004**, *112*, 747.
9. Han, D. C.; Shen, T. L.; Guan, J. L. *J. Biol. Chem.* **2000**, *275*, 28911.
10. Shen, T. L.; Guan, J. L. *FEBS Lett.* **2001**, *499*, 176.
11. Itoh, S.; Taketomi, A.; Tanaka, S.; Harimoto, N.; Yamashita, Y.; Aishima, S.; Maeda, T.; Shirabe, K.; Shimada, M.; Maehara, Y. *Mol. Cancer Res.* **2007**, *5*, 667.
12. Bai, T.; Luoh, S. W. *Carcinogenesis* **2008**, *29*, 473.
13. Margolis, B. *Prog. Biophys. Mol. Biol.* **1994**, *62*, 223.
14. Han, D. C.; Shen, T. L.; Guan, J. L. *Oncogene* **2001**, *20*, 6315.
15. Holt, L. J.; Daly, R. J. *Growth Factors* **2005**, *23*, 193.
16. Pero, S. C.; Daly, R. J.; Krag, D. N. *Expert. Rev. Mol. Med.* **2003**, *10*, 1.
17. Bradshaw, J. M.; Waksman, G. *Adv. Protein Chem.* **2002**, *61*, 161.
18. Pawson, T. *Adv. Cancer Res.* **1994**, *64*, 87.
19. Porter, C. J.; Matthews, J. M.; Mackay, J. P.; Pursglove, S. E.; Schmidberger, J. W.; Leedman, P. J.; Pero, S. C.; Krag, D. N.; Wilce, M. C.; Wilce, J. A. *BMC Struct. Biol.* **2007**, *7*, 58.
20. Bleicher, K. H.; Bohm, H. J.; Muller, K.; Alanine, A. I. *Nat. Rev. Drug Disc.* **2003**, *2*, 369.
21. Oprea, T. I.; Matter, H. *Curr. Opin. Chem. Biol.* **2004**, *8*, 349.
22. Martin, Y. C.; Kofron, J. L.; Traphagen, L. M. *J. Med. Chem.* **2002**, *45*, 4350.
23. Nikolova, N.; Jaworska, J. *QSAR Comb. Sci.* **2004**, *22*, 1006.
24. Ebalunode, J. O.; Ouyang, Z.; Liang, J.; Zheng, W. *J. Chem. Inf. Model.* **2008**, *48*, 889.
25. Putta, S.; Lemmen, C.; Beroza, P.; Greene, J. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1230.
26. Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
27. Verli, H.; Albuquerque, M. G.; Bicca de Alencastro, R.; Barreiro, E. J. *Eur. J. Med. Chem.* **2002**, *37*, 219.
28. Goldman, B. B.; Wipke, W. T. *Proteins* **2000**, *38*, 79.
29. Ebalunode, J. O.; Dong, X.; Ouyang, Z.; Liang, J.; Eckenhoff, R. G.; Zheng, W. *Bioorg. Med. Chem.* **2009**, *17*, 5133.
30. Masek, B. B.; Merchant, A.; Matthew, J. B. *J. Med. Chem.* **1993**, *36*, 1230.
31. Burke, T. R. *Int. J. Pept. Res. Ther.* **2006**, *12*, 33.
32. Fretz, H.; Furet, P.; Garcia-Echeverria, C.; Schoepfer, J.; Rahuel, J. *Curr. Pharm. Des.* **2000**, *6*, 1777.
33. Liu, W. Q.; Vidal, M.; Olszowy, C.; Million, E.; Lenoir, C.; Dhôtel, H.; Garbay, C. *J. Med. Chem.* **2004**, *47*, 1223.
34. Shi, Z. D.; Wei, C. Q.; Lee, K.; Liu, H.; Zhang, M.; Araki, T.; Roberts, L. R.; Worthy, K. M.; Fisher, R. J.; Neel, B. G.; Kelley, J. A.; Yang, D.; Burke, T. R., Jr. *J. Med. Chem.* **2004**, *47*, 2166.
35. Oishi, S.; Karki, R. G.; Kang, S. U.; Wang, X.; Worthy, K. M.; Bindu, L. K.; Nicklaus, M. C.; Fisher, R. J.; Burke, T. R., Jr. *J. Med. Chem.* **2005**, *48*, 764.
36. Kang, S. U.; Choi, W. J.; Oishi, S.; Lee, K.; Karki, R. G.; Worthy, K. M.; Bindu, L. K.; Nicklaus, M. C.; Fisher, R. J.; Burke, T. R., Jr. *J. Med. Chem.* **2007**, *50*, 1978.
37. García-Echeverria, C. *Curr. Med. Chem.* **2001**, *8*, 1589.
38. Pero, S. C.; Oligino, L.; Daly, R. J.; Soden, A. L.; Liu, C.; Roller, P. P.; Li, P.; Krag, D. N. *J. Biol. Chem.* **2002**, *277*, 11918.
39. Porter, C. J.; Wilce, J. A. *Biopolymers* **2007**, *88*, 174.
40. Pero, S. C.; Shukla, G. S.; Cookson, M. M.; Flemer, S., Jr.; Krag, D. N. *Br. J. Cancer* **2007**, *96*, 1520.
41. Ambaye, N. D.; Lim, R. C.; Clayton, D. J.; Gunzburg, M. J.; Price, J. T.; Pero, S. C.; Krag, D. N.; Wilce, M. C.; Aguilar, M. I.; Perlmutter, P.; Wilce, J. A. *Biopolymers* **2010**. doi:10.1002/bip.21403.
42. Spuches, A. M.; Argiros, H. J.; Lee, K. H.; Haas, L. L.; Pero, S. C.; Krag, D. N.; Roller, P. P.; Wilcox, D. E.; Lyons, B. A. *J. Mol. Recognit.* **2007**, *20*, 245.
43. Luzy, J. P.; Chen, H.; Gril, B.; Liu, W. Q.; Vidal, M.; Perdereau, D.; Burnol, A. F.; Garbay, C. *J. Biomol. Screen.* **2008**, *13*, 112.
44. Liu, W. Q.; Vidal, M.; Gresh, N.; Roques, B. P.; Garbay, C. *J. Med. Chem.* **1999**, *42*, 3737.
45. Mohamadi, F.; Richard, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.
46. Leach, A. R. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH, Inc., 1991; Vol. 2, pp 1–55.
47. Schrodinger LLC, 120 West 45th Street, New York, New York.
48. Dixon, S. L.; Smondyrev, A. M.; Rao, S. N. *Chem. Biol. Drug Des.* **2006**, *67*, 370.
49. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. *J. Med. Chem.* **2006**, *49*, 6177.
50. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004**, *47*, 1739.
51. Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.
52. Pantoliano, M. W.; Petrella, E. C.; Kwasnoski, J. D.; Lobanov, V. S.; Myslik, J.; Graf, E.; Carver, T.; Asel, E.; Springer, B. A.; Lane, P.; Salemme, F. R. *J. Biomol. Screen.* **2001**, *6*, 429.
53. Matulis, D.; Kranz, J. K.; Salemme, F. R.; Todd, M. J. *Biochemistry* **2005**, *44*, 5258.

54. Weininger, D. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31.
55. Freyer, M. W.; Lewis, E. A. *Methods Cell Biol.* **2008**, *84*, 79.
56. Perozzo, R.; Folkers, G.; Scapozza, L. *J. Recept. Signal Transduct. Res.* **2004**, *24*, 1.
57. Eftink, M. R.; Anusiem, A. C.; Biltonen, R. L. *Biochemistry* **1983**, *22*, 3884.
58. Danho, W.; Swistok, J.; Khan, W.; Chu, X. J.; Cheung, A.; Fry, D.; Sun, H.; Kurylko, G.; Rumennik, L.; Cefalu, J.; Cefalu, G.; Nunn, P. *Adv. Exp. Med. Biol.* **2009**, *611*, 467.
59. Johnson, M. A.; Maggiora, G. M. *Concepts and Applications of Molecular Similarity*; John Wiley & Sons: New York, 1990.
60. Mezey, P. G. *Shape in Chemistry: An Introduction to Molecular Shape and Topology*; VCH: New York, 1993.
61. Cummings, M. D.; Farnum, M. A.; Nelen, M. I. *J. Biomol. Screen.* **2006**, *11*, 854.
62. Janes, P. W.; Lackmann, M.; Church, W. B.; Sanderson, G. M.; Sutherland, R. L.; Daly, R. J. *J. Biol. Chem.* **1997**, *272*, 8490.
63. Porter, C. J.; Wilce, M. C.; Mackay, J. P.; Leedman, P.; Wilce, J. A. *Eur. Biophys. J.* **2005**, *34*, 454.
64. Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. *Protein Sci.* **1995**, *4*, 2411.