Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters 17 (2007) 722–726
Bioorganic & Medicinal Chemistry Letters

Pharmacophore identification of KSP inhibitors

Fei Liu, Qi-Dong You* and Ya-Dong Chen

Department of Medicinal Chemistry, China Pharmaceutical University, 24 Tongjiaxiang, Nanjing 210009, PR China

Received 22 August 2006; revised 24 October 2006; accepted 26 October 2006
Available online 1 November 2006

Abstract—A three-dimensional pharmacophore model was developed based on 25 currently available KSP (kinesin spindle protein) inhibitors in Catalyst software package. The best pharmacophore hypothesis (Hypo1), consisting of four chemical features (one hydrogen-bond acceptor, one hydrogen-bond donor, one aromatic ring, and one hydrophobic group), has a correlation coefficient of 0.965. The results of our study provide a valuable tool in designing new leads with desired biological activity by virtual screening.

Antimitotic agents are a major class of cytotoxic drugs in the treatment of cancer. The targets of antimitotic agents are microtubules, which are composed of α- and β-tubulins.[1] By interfering with the tubulins' polymerization and depolymerization, antimitotic agents inhibit mitotic spindle and arrest dividing cells in metaphase. However, although the Vinca alkaloids, taxol, epothilones, and other antimitotic agents have achieved great success in the treatment of cancer, these drugs have undesired mechanism-based side effects (e.g., neurotoxicity).[2] This is because that microtubules are also involved in many other cellular processes such as maintenance of organelles, cell shape, cell motility, synaptic vesicles, and intracellular transport. So it is necessary to propose new strategies so as to target microtubule-associated proteins or mitotic checkpoint proteins to inhibit mitotic spindle and cell division.[3]

Members of the kinesin superfamily play important roles in cargo transport, spindle and chromosome movement, and regulation of microtubule dynamics.[4] KSP (kinesin spindle protein, also known as *Hs* Eg5) is a plus-end-directed motor of the BimC kinesin subfamily which is responsible for the formation of the bipolar spindle.[5] KSP plays an important role in the early stage of mitosis and mediates centrosome separation. Inhibition of KSP leads to a stable mitotic block with monoastral microtubule arrays.[6] The function of KSP provides a novel route for the manipulation of the cell cycle and the induction of apoptosis.

Keywords: Pharmacophore; Mitotic kinesin; KSP inhibitors.
* Corresponding author. Tel./fax: +86 25 83271351; e-mail: youqd@cpu.edu.cn

0960-894X/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2006.10.083

In 1999, Mayer et al. found the first small molecule, monastrol (compound **20**), which targeted KSP and led to mitotic arrest.[7] In recent years, monastrol analogues and several kinds of other KSP inhibitors have been reported.[8–19] This makes it possible to make a ligand-based design for KSP inhibitors. In a ligand-based molecule design, identification of pharmacophore is one of the most important steps. In this paper, we identified the pharmacophore of KSP inhibitors for the first time.

Computational methodology. The study was performed using the Catalyst software package (version 4.11, Accelrys Inc., San Diego, CA) on a SGI Origin 3800 workstation. Chemical-feature-based pharmacophore hypotheses can be generated automatically using the HypoGen algorithm within Catalyst, provided that structure–activity relationship data of a well-balanced set of compounds are available. Two assumptions must be made about the data: (1) all compounds used in the training set have to bind to the same receptor in roughly the same fashion and (2) compounds having more binding interactions with the receptor are more active than those with fewer.

Training set and test set selection and conformational models. The constructed pharmacophore model can be as good as the information data input. Nevertheless, to achieve such a quality, some must-obey rules should be respected in a three-dimensional quantitative structure–activity relationship (3D-QSAR) generation using Catalyst. Thus, for instance, the set must be widely populated (at least 16 items) by structurally diverse representatives covering an activity range of at least four orders of magnitude. The most active compounds

should inevitably be included in the training set, and all biologically relevant data should be obtained by homogeneous procedures.

The training set consists of 25 compounds (No. 1–25 Fig. 1) and is selected to generate HypoGen hypotheses by considering structural diversity and wide coverage of activity range in terms of IC_{50} ranging from 0.5 nM to 145 μM (Table 2). All structures in the training set were built and minimized to the closest local minimum based on a modified CHARMM force field within Confirm module.

Conformational models of the training set compounds were generated using a Monte Carlo-like algorithm together with poling.[20,21] Catalyst provides two types of conformational analysis: Fast and Best. In this case, the Best option was used, specifying 250 as the maximum number of conformers with a constraint of 20 kcal/mol energy. All other parameters used were default.

Generation of pharmacophore hypotheses with HypoGen. Taking into account the chemical nature of the compounds considered in this work, the following four features were selected to form the essential information in the hypothesis generation process: hydrogen-bond acceptor (HBA), hydrogen-bond donor (HBD), hydrophobic group (Hp), and ring aromatic (Ar). The uncertain factor for each compound represents the ratio range of uncertainty in the activity value based on the expected statistical straggling of biological data collection. Here this factor was defined as 3.

Pharmacophores were then computed using HypoGen module implemented in Catalyst software package and the top 10 scoring hypotheses were exported.

Validation of HypoGen hypotheses. Catalyst produced 10 hypotheses (Hypo1–Hypo10). Hypo1, which consisted of four features: one hydrogen-bond acceptor (HBA), one hydrogen-bond donor (HBD), one hydrophobic group (Hp), and one aromatic ring (Ar), is the best pharmacophore hypothesis in this study (Fig. 2). This is characterized by the highest cost difference, lower error cost, lowest root mean square (rms) divergence, and best correlation coefficient (Table 1). The null cost of the 10 top-scored hypotheses was equal to 219.236, the fixed

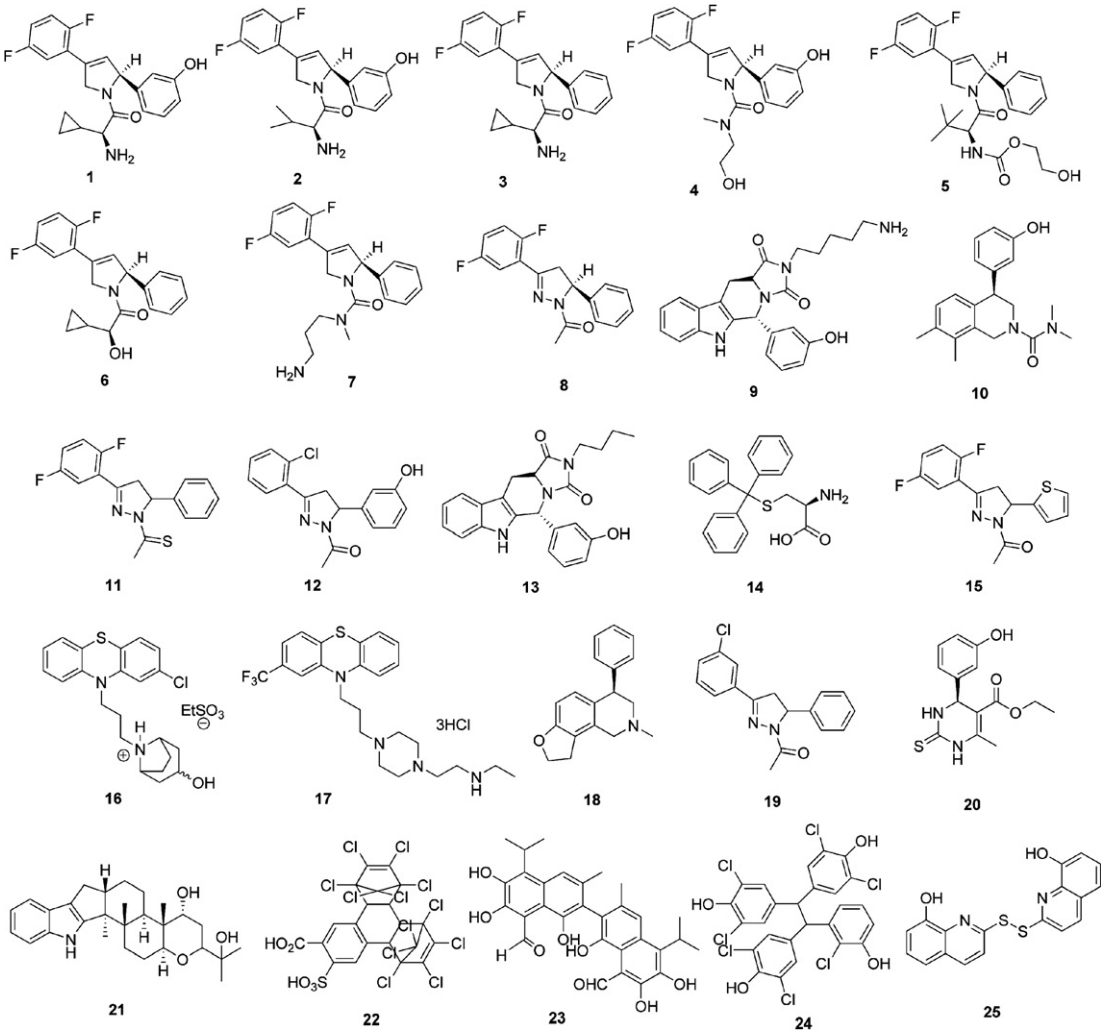

Figure 1. 2D chemical structures of the 28 molecules forming the training set used to obtain HypoGen pharmacophore hypotheses.

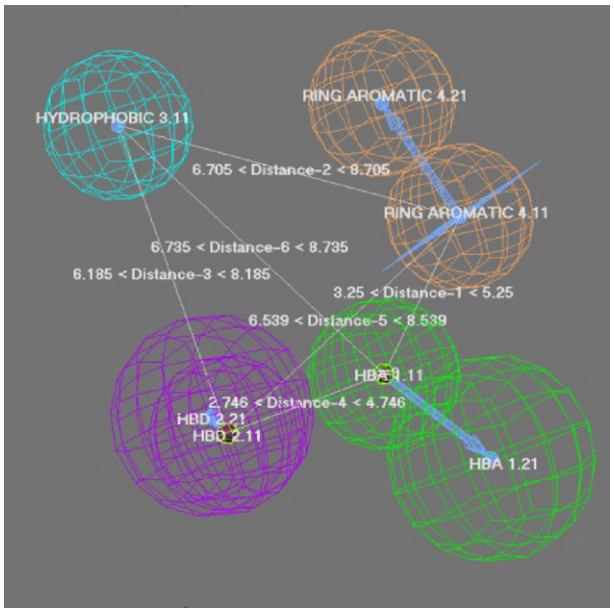

Figure 2. Top-scoring HypoGen pharmacophore Hypo1. Features are color-coded as follows: aromatic ring, orange; hydrogen-bond acceptor, green; hydrophobic, blue; hydrogen-bond donor, violet.

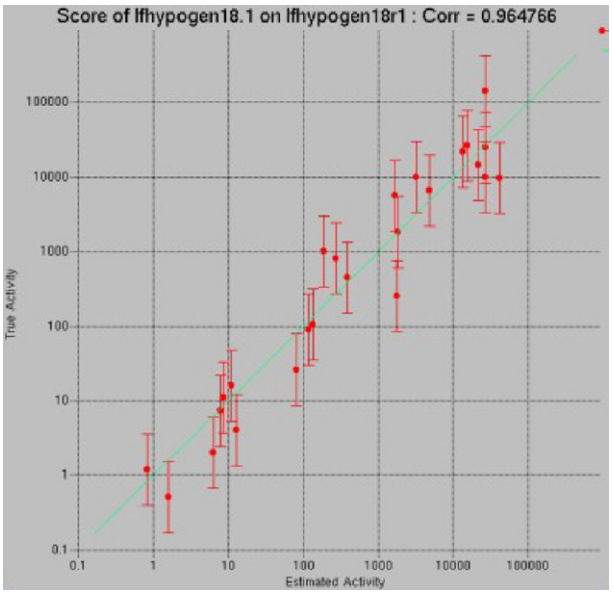

Figure 3. The regression of actual versus predicted activities by the Hypo1 hypothesis for the training set inhibitors onto a linear relationship.

Table 1. Information of statistical significance and predictive power presented in cost values for top 10 hypotheses[a]

Hypothesis	Features[a]	Total cost	Δcost	rms	Correlation (r)
1	ADHR	112.192	107.044	0.873	0.965
2	ADHR	125.067	94.169	1.347	0.913
3	ADHR	126.844	92.392	1.441	0.899
4	ADHR	126.912	92.324	1.436	0.900
5	ADHR	127.096	92.140	1.375	0.910
6	ADHR	128.002	91.234	1.396	0.907
7	ADHR	129.643	89.593	1.512	0.888
8	ADHR	130.329	88.907	1.536	0.884
9	ADDH	130.461	88.775	1.525	0.886
10	AHRR	130.867	88.369	1.546	0.883

[a] Abbreviations used for features: A, hydrogen-bond acceptor; D, hydrogen-bond donor; H, hydrophobic group; R, aromatic ring.

cost value was 100.529, and the configuration cost was 15.314. As the total cost of Hypo1 was equal to 112.192, the large difference between null and total hypothesis cost, (Δcost)107.044, coupled with a high correlation coefficient, (r) 0.965 (Fig. 3), and a reasonable root mean square (rms) deviation 0.873 ensures that a true correlation will very likely be estimated by the model.

Besides this cost analysis, another validation method to characterize the quality of hypothesis is represented by its capacity for correct activity prediction. The difference between estimated activity values and experimental activity values is represented as error (ratio between the estimated and experimental activity), with a negative sign if the actual activity is higher than the estimated. As we can see from Table 2, most of the IC_{50} values were predicted correctly.

Table 2. Experimental biological data and estimated IC_{50} values of the training set molecules based on the pharmacophore model Hypo1

Compound	Experimental IC_{50} (nM)	Estimated IC_{50} (nM)	Error	Reference
1	0.5	1.5	+3	14
2	1.2	0.79	−1.5	14
3	2	6	+3	13
4	4	12	+3	14
5	7.4	7.5	+1	13
6	11	8.2	−1.3	13
7	16	11	−1.5	13
8	26	77	+3	12
9	90	110	+1.3	10
10	104	130	+1.2	17
11	250	1700	+7	12
12	450	370	−1.2	10
13	800	270	−3	12
14	1000	180	−5.6	9
15	1800	1800	+1	12
16	5600	1600	−3.4	12
17	6600	4700	−1.4	9
18	9700	40,000	+4.2	9
19	9800	3100	−3.2	17
20	10,000	26,000	+2.6	12
21	15,000	21,000	+1.5	12
22	22,000	13,000	−1.7	9
23	25,000	26,000	+1.1	9
24	27,000	15,000	−1.8	9
25	145,000	26,000	−5.6	9

Figure 4 depicts the mapping of Hypo 1 onto a highly active compound (compound **3**). As we can see from this figure, compound **3** fits all features of the pharmacophore model Hypo1 very well. The phenyl group of this molecule overlaps with the aromatic ring (Ar) feature of Hypo1, and the 2,4-difluorophenyl group serves as a hydrophobic group (Hp). The hydrogen-bond acceptor

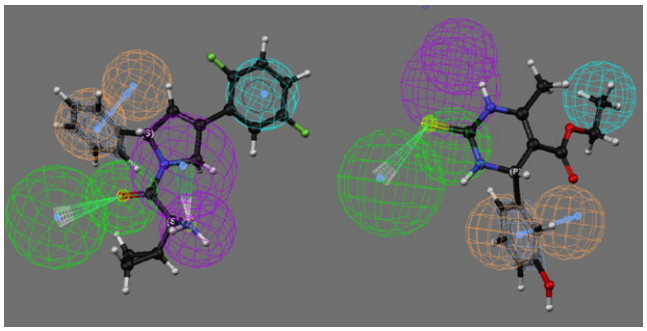

Figure 4. Mapping of compound **3** (left) and compound **20** (monastrol, right) onto Hypo1. Features are color-coded as follows: ring aromatic, orange; hydrogen-bond acceptor, green; hydrophobic, blue; hydrogen-bond donor, violet.

(HBA) is located on the carbonyl group and the hydrogen-bond donor (HBD) is located on the neighboring primary amino group.

Figure 4 also shows, as an example, the mapping on Hypo1 of the training set compound **20** (monastrol). In this figure, we can see that 3-hydroxy phenyl group overlaps with the aromatic ring (Ar) feature of Hypo1. The ethyl group serves as the hydrophobic group (Hp) and the thiocarbonyl group serves as the hydrogen-bond acceptor (HBA), whereas the hydrogen-bond donor (HBD) mapping is missing.

The validity and the predictive character of Hypo1 were further assessed by using the test set molecules. The structural data for the test set are shown in Figure 5. All molecules in the test set were built and minimized as well as used in conformational analysis like all molecules in the training set. In this test set analysis, out of 15 compounds, 14 compounds had the error values of less

Table 3. Experimental biological data and estimated IC_{50} values of the test set molecules based on the pharmacophore model Hypo1

Compound	Experimental IC_{50} (nM)	Estimated IC_{50} (nM)	Error	Reference
26	1.3	0.45	−2.9	14
27	5.2	13	+2.5	13
28	12	150	+12.5	19
29	50	390	+7.8	13
30	84	560	+6.7	12
31	280	190	−1.5	19
32	306	190	−1.6	17
33	808	470	−1.7	17
34	1800	1600	−1.1	12
35	2270	240	−9.5	17
36	3100	1100	−2.8	12
37	3600	890	−4	12
38	3600	3300	−1.1	12
39	8230	6300	−1.3	17
40	25,200	29,000	+1.2	17

than 10 (Table 3), representing a not more than one order difference between estimated and actual activity. The Hypo1 also showed a good correlation with the test set ($r = 0.886$).

In addition, the structure of compound **3**-KSP-ADP complex was analyzed using the Ligplot4.22 program to identify some specific contacts between atoms of ligand and receptor (Fig. 6).[22] And so we could learn how the chemical features in the pharmacophore acted with the KSP receptor. The X-ray structure of compound **3**-KSP-ADP suggested that the difluorophenyl group and the phenyl group are properly positioned into two pockets of the receptor, which means that the Hydrophobic 3.11 and Ring Aromatic 4.11 are positioned in the two pockets. Besides, the amino group and the carbonyl group of compound **3**, which serve as HBA 2.11 and HBD 1.11, respectively, were placed

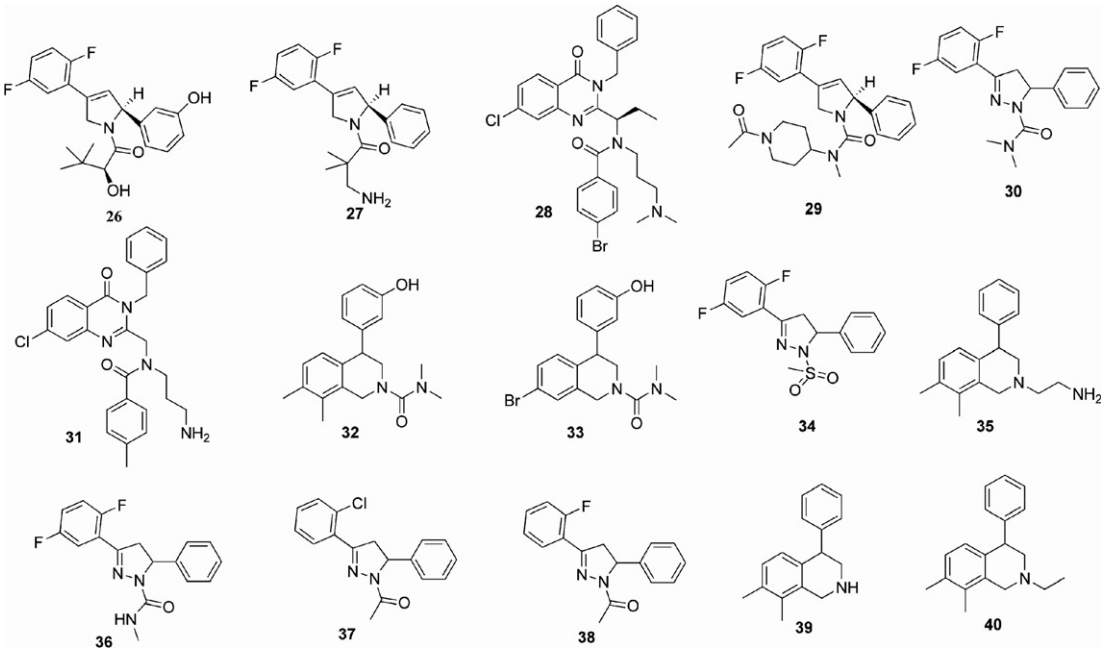

Figure 5. 2D chemical structures of the 15 molecules forming the test set.

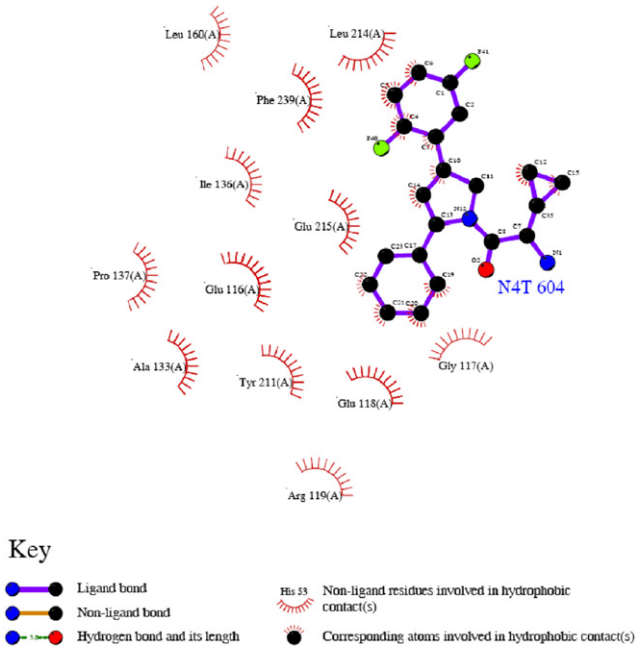

Figure 6. The structure of compound **3**-KSP-ATP complex in Ligplot 4.22 program.

out of the pocket to minimize the energy cost of desolvation.

In summary, a three-dimensional pharmacophore model was developed based on 25 KSP inhibitors by a ligand-based computational approach. This pharmacophore hypothesis consists of one hydrogen-bond acceptor, one hydrogen-bond donor, one aromatic ring, and one hydrophobic group, and has a correlation coefficient of 0.965. Besides, this hypothesis is further validated by using an external test set of 15 compounds. The most active compounds (e.g., compound **3**) fit very well with this top scoring hypothesis. Thus, our pharmacophore model should be helpful in identifying novel lead compounds and providing a valuable tool in designing new KSP inhibitors.

References and notes

1. Gunderson, G. G. *Nat. Rev. Mol. Cell Boil.* **2002**, *3*, 296.
2. Rowinski, E. K.; Chaudhry, V.; Cornblath, D. R.; Donehower, R. C. *J. Natl. Cancer Inst.* **1993**, *15*, 107.
3. Wood, K. W.; Cornell, W. D.; Jackson, J. R. *Curr. Opin. Pharmacol.* **2001**, *1*, 370.
4. Wittmann, T.; Hyman, A.; Desai, A. *Nat. Cell Biol.* **2001**, *3*, 28.
5. Lawrence, C. J.; Dawe, R. K.; Christie, K. R.; Cleveland, D. W.; Dawson, S. C.; Endow, S. A.; Goldstein, L. S. B.; Goodson, H. V.; Hirokawa, N.; Howard, J.; Malmberg, R. L.; McIntosh, J. R.; Miki, H.; Mitchison, T. J.; Okada, Y.; Reddy, A. S. N.; Saxton, W. M.; Schliwa, M.; Scholey, J. M.; Vale, R. D.; Walzak, C. E.; Wordeman, L. *J. Cell Biol.* **2004**, *167*, 19.
6. Sawin, K. E.; Mitchison, T. J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 4289.
7. Mayer, T. U.; Kapoor, T. M.; Haggarty, S. J.; King, R. W.; Schreiber, S. L.; Mitchison, T. J. *Science* **1999**, *289*, 971.
8. Nakazawa, J.; Yajima, J.; Usui, T.; Ueki, M.; Takatsuki, A.; Imoto, M.; Toyoshima, Y.; Osada, H. *Chem. Biol.* **2003**, *10*, 131.
9. DeBonis, S.; Skoufias, D. A.; Lebeau, L.; Lopez, R.; Robin, G.; Margolis, R. L.; Wade, R. H.; Kozielski, F. *Mol. Cancer Ther.* **2004**, *3*, 1079.
10. Hotha, S.; Yarrow, J. C.; Yang, J. G.; Garrett, S.; Renduchintala, K. V.; Mayer, T. U.; Kapoor, T. M. *Angew. Chem.* **2003**, *115*, 2481; . *Angew. Chem. Int. Ed.* **2003**, *42*, 2379.
11. Sakowicz, R.; Finer, J. T.; Beraud, C.; Crompton, A.; Lewis, E.; Fritsch, A.; Lee, Y.; Mak, J.; Moody, R.; Turincio, R.; Chabala, J. C.; Gonzales, P.; Roth, S.; Weitman, S.; Wood, K. W. *Cancer Res.* **2004**, *64*, 3276.
12. Cox, C. D.; Breslin, M. J.; Mariano, B. J.; Coleman, P. J.; Buser, C. A.; Walsh, E. S.; Hamilton, K.; Huber, H. E.; Kohl, N. E.; Torrent, M.; Yan, Y.; Kuo, L. C.; Hartman, G. D. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 2041.
13. Fraley, M. E.; Garbaccio, R. M.; Arrington, K. L.; Hoffman, W. F.; Tasber, E. S.; Coleman, P. J.; Buser, C. A.; Walsh, E. S.; Hamilton, K.; Fernandes, C.; Schaber, M. D.; Lobell, R. B.; Tao, W.; South, V. J.; Yan, Y.; Kuo, L. C.; Prueksaritanont, T.; Shu, C.; Torrent, M.; Heimbrook, D. C.; Kohl, N. E.; Huber, H. E.; Hartman, G. D. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1775.
14. Garbaccio, R. M.; Fraley, M. E.; Tasber, E. S.; Olson, C. M.; Hoffman, W. F.; Arrington, K. L.; Torrent, M.; Buser, C. A.; Walsh, E. S.; Hamilton, K.; Schaber, M. D.; Fernandes, C.; Lobell, R. B.; Tao, W.; South, V. J.; Yan, Y.; Kuo, L. C.; Prueksaritanont, T.; Slaughter, D. E.; Shu, C.; Heimbrook, D. C.; Kohl, N. E.; Huber, H. E.; Hartman, G. D. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1780.
15. Cox, C. D.; Torrent, M.; Breslin, M. J.; Mariano, B. J.; Whitman, D. B.; Coleman, P. J.; Buser, C. A.; Walsh, E. S.; Hamilton, K.; Schaber, M. D.; Lobell, R. B.; Tao, W.; South, V. J.; Kohl, N. E.; Yan, Y.; Kuo, L. C.; Prueksaritanont, T.; Slaughter, D. E.; Li, C.; Mahan, E.; Lu, B.; Hartman, G. D. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3175.
16. Kim, K. S.; Lu, S.; Cornelius, L. A.; Lombardo, L. J.; Borzilleri, R. M.; Schroeder, G. M.; Sheng, C.; Rovnyak, G.; Crews, D.; Schmidt, R. J.; Williams, D. K.; Bhide, R. S.; Traeger, S. C.; McDonnell, P. A.; Mueller, L.; Sheriff, S.; Newitt, J. A.; Pudzianowski, A. T.; Yang, Z.; Wild, R.; Lee, F. Y.; Batorsky, R.; Ryder, J. S.; Ortega-Nanos, M.; Shen, H.; Gottardis, M.; Roussell, D. L. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3937.
17. Tarby, C. M.; Kaltenbach, R. F., 3rd; Huynh, T.; Pudzianowski, A.; Shen, H.; Ortega-Nanos, M.; Sheriff, S.; Newitt, J. A.; McDonnell, P. A.; Burford, N.; Fairchild, C. R.; Vaccaro, W.; Chen, Z.; Borzilleri, R. M.; Naglich, J.; Lombardo, L. J.; Gottardis, M.; Trainor, G. L.; Roussell, D. L. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2095.
18. Duhl, D. M.; Renhowe, P. A. *Curr. Opin. Drug. Discov. Devel.* **2005**, *8*, 431.
19. Bergnes, G.; Ha, E.; Feng, B.; Yao, B.; Smith, W. W.; Tochimoto T.; Lewis, E. R.; Lee, Y. Y.; Moody, R.; Turincio, R. A.; Finer, J. T.; Wood, K. W.; Sakowicz, R.; Crompton, A. M.; Chabala, J. C.; Morgans, Jr. D. J.; Sigal, N. H.; Sabry, J. H. American Association for Cancer Research (AACR), April, **2002**.
20. Smellie, A.; Kahn, S. D.; Teig, S. J. *Chem. Inf. Comput. Sci.* **1995**, *35*, 285.
21. Smellie, A.; Kahn, S. D.; Teig, S. J. *Chem. Inf. Comput. Sci.* **1995**, *35*, 295.
22. PDB deposition number: 2FL2.