Pergamon

SCIENCE DIRECT®

Bioorganic & Medicinal Chemistry Letters 13 (2003) 3705–3709

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Identification of a Novel Class of Inhibitor of Human and *Escherichia coli* Thymidine Phosphorylase by In Silico Screening☆

V. A. McNally, A. Gbaj, K. T. Douglas, I. J. Stratford, M. Jaffar, S. Freeman and R. A. Bryce*

School of Pharmacy and Pharmaceutical Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, UK

Received 23 June 2003; revised 22 July 2003; accepted 7 August 2003

Abstract—Structure-based computational screening of the National Cancer Institute database of anticancer compounds identified novel non-nucleobase-derived inhibitors of human thymidine phosphorylase as candidates for lead optimization. The hierarchical in silico screening strategy predicted potentially strong low molecular weight ligands exhibiting a range of molecular scaffolds. Of the thirteen ligands assayed for activity, all displayed inhibitory activity against *Escherichia coli* thymidine phosphorylase. One compound, hydrazine carboxamide 2-[(1-methyl-2,5-dioxo-4-pentyl-4-imidazolidinyl)methylene], was found to inhibit *E. coli* thymidine phosphorylase with an IC_{50} value of 20 μM and an IC_{50} value of 77 μM against human thymidine phosphorylase. As this hydantoin derivative lacks the undesirable ionic sites of existing tight-binding nucleobase-derived inhibitors, such as 5-chloro-6-[(2-iminopyrrolidin-1-yl)methyl]uracil hydrochloride, it provides an opportunity for the design of potent thymidine phosphorylase inhibitors with improved pharmacokinetic properties.

Thymidine phosphorylase (TP), the enzyme catalysing reversible phosphorolysis of thymidine (**1**, Fig. 1) to α-D-2-deoxyribose-1-phosphate and thymine (**2**), is identical to platelet-derived endothelial cell growth factor (PD-ECGF).[1] TP has been implicated in angiogenesis, the development of new capillary blood vessels from existing capillaries and small venules.[2] Inhibition of TP activity, by means of mutagenesis, antibody and small-molecule inhibitor studies, leads to loss of angiogenic activity.[3,4] As TP is over-expressed in cancerous tissue, the protein may be an attractive cancer chemotherapy target for inhibition of tumour angiogenesis and subsequent tumour growth and metastasis.[5]

Inhibitor design has focused upon nucleoside base analogues, principally of uracil[6,7] and, more recently, of purines.[8,9] The most potent inhibitors of human TP currently known are 5-chloro-6-[(2-iminopyrrolidin-1-yl)methyl]uracil hydrochloride (**3**) described by the Taiho Pharmaceutical Co.[5,6] and 5-bromo-6-[(2-aminoimidazol-1-yl)methyl)uracil hydrochloride (**4**)[7] with apparent IC_{50} values in the nanomolar range. These cationic tight-binding inhibitors have been proposed to be analogues of the oxycarbenium ion-like transition state formed during thymidine phosphorolysis (Fig. 2).[8] Despite the nanomolar potency of these nucleobase-derived inhibitors, their ionic nature leads to poor pharmacokinetic properties.

In the present study, we have used a structure-based approach to screen for alternative human TP inhibitors, in particular to identify new scaffolds with submillimolar IC_{50} values suitable for subsequent lead optimization. This strategy involved a hierarchical computational screening of the National Cancer Institute (NCI) database of anticancer compounds (http://www.dtp.nih.gov),[10] using the computer program DOCK[11] to predict protein–ligand binding mode and affinity. Subsequent assay of inhibitory activity against recombinant *Escherichia coli* TP and human TP was performed to evaluate 'hits' obtained from the virtual screen.

The three-dimensional receptor structure was represented by our previously constructed homology model of human TP[12] based on the *E. coli* crystal structure[13] in the open-cleft conformation, which shares 39% sequence identity and 53% similarity with the human

☆Supplementary data associated with this article can be found at doi: 10.1016/j.bmcl.2003.08.010

*Corresponding author. Tel. +44-161-275-8345; fax: +44-161-275-2481; e-mail: r.a.bryce@man.ac.uk

0960-894X/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2003.08.010

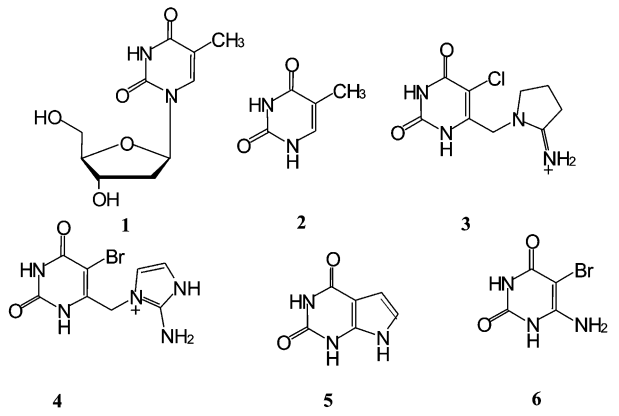

Figure 1. Known TP ligands.

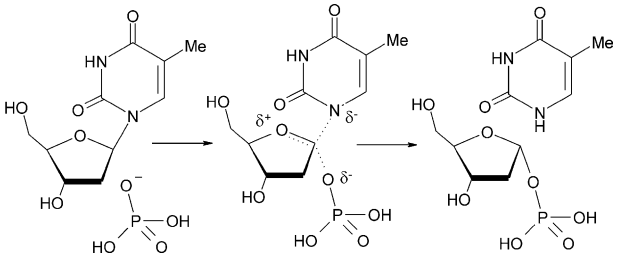

Figure 2. Phosphorolysis of thymidine to thymine catalysed by TP.

enzyme.[12] TP is homodimeric; each 45 kD monomer is composed of a helical domain and a larger α–β domain containing the active site. The active site of human TP, based around residues His85, Leu148, Arg202, Lys221, Ile214 and Val241, was used to generate a 52 sphere negative image for docking calculations. A scoring grid (0.3 Å resolution) was calculated using Gasteiger–Marsili charges.[14] A united-atom AMBER-based potential[15] was employed as a scoring function, with a distance-dependent dielectric of 4r, a non-bond cutoff of 10 Å and a bump overlap of 0.75 Å. 3-D coordinates were assigned to the 250,521 compounds of the NCI database using the program CORINA.[16] These compounds are referred to by their unique NSC identifier using the NCI convention. Visualization of molecular structure was performed using SYBYL.[17]

Prior to docking of the NCI database, we evaluated the ability of DOCK to reproduce the thymine/human TP structure, modelled on the crystallographic binding mode of thymine as observed in the *E. coli* TP X-ray structure (Table 1).[13,18] The crystallographic binding mode of thymine to *E. coli* TP indicates that thymine interacts with a cavity at the top of the active-site cleft, via hydrogen bonds of the heterocyclic ring and hydrophobic interactions of the thymine methyl group. In our human TP model, these hydrogen bonds are with residues Tyr199, Arg202 and Ser217, and hydrophobic interactions with Val208 and Ile214 (Fig. 3). Observation of the inactive human TP mutant Arg202Ser supports the functional significance of this conserved Arg residue in the human enzyme.[3] To evaluate the ability of DOCK to predict TP/ligand binding geometries, we compared the calculated human TP/thymine binding

Table 1. Binding energy, U_{bind} (kcal/mol), predicted from rigid-body docking of known TP ligands, as a function of number of evaluated orientations (n_{orient})

n_{orient}	U_{bind}					
	1	2	3	4	5	6
200	−18.0	−14.1	−21.4	−21.8	−16.5	−16.6
1000	−19.5	−18.0	−25.9	−26.1	−17.7	−17.8
2500	−21.4	−18.5	−26.3	−26.6	−21.2	−20.5
5000	−21.7	−18.4	−25.7	−26.1	−22.2	−21.8
10,000	−22.1	−19.0	−26.6	−26.3	−21.9	−22.2
30,000	−22.3	−18.9	−30.0	−26.3	−22.3	−22.2
50,000	−22.5	−19.0	−30.0	−26.5	−22.3	−22.2

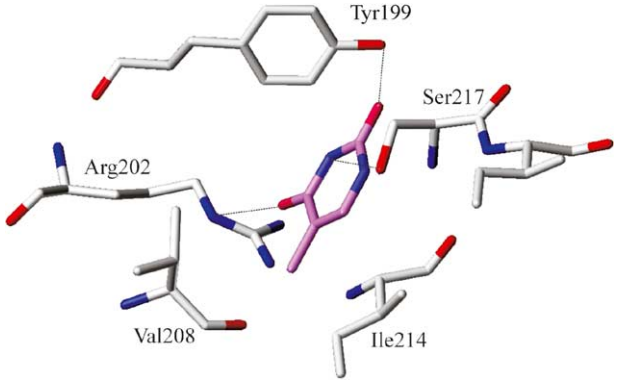

Figure 3. Interactions of thymine (magenta) with active site of TP.

mode with the observed mode exhibited in the *E. coli* crystal structure. We explored the accuracy of the approach as a function of the number of rigid-body orientations evaluated, comparing the root mean square deviation (RMSD) of docked thymine relative to its crystallographic mode. The human TP homology model and the *E. coli* X-ray structure were aligned via superposition of key conserved amino acids. The RMSD converged rapidly from 9.3 Å at 200 orientations to 2.8–3.0 Å from 1000 to 50,000 orientations.

To evaluate the ability of the computational approach to energetically discriminate potent from less potent inhibitors, thymidine and a series of known TP inhibitors (Fig. 1) were also docked as rigid bodies. A range of inhibition data is available for the compounds in Table 1. As **1** is a substrate and we have no evidence that K_m reflects binding strength, there is no quantitative information on how strongly **1** binds to the active site at this stage. Thymine (**2**) has been reported to inhibit mouse liver TP with a K_i of 347±91 mM at 0.15 mM thymidine substrate.[19] Compound **3**, described by Taiho Pharmaceutical Co.,[5] and compound **4**, the imidazolic compound from our laboratories,[7] are tight binding inhibitors, and have been reported with apparent IC_{50} values of 20 and 19 nM, respectively, for *E. coli* TP. The values are apparent[7] because the IC_{50} values determined are limited by the enzyme concentration achievable in the assay (**3** and **4** are stoichiometric inhibitors) and thus represent an upper limit of the true IC_{50}.[7] With human TP, compound **3** has been described

in the literature as a competitive inhibitor with a K_i value of 17 nM.[6] For compound **5**, we determined the IC_{50} value against the *E. coli* enzyme to be 6.3±0.3 μM. For 5-bromo-6-aminouracil (**6**) we determined the IC_{50} value against the human enzyme to be 6.6±0.7 μM (literature value under different conditions is 17 μM[20]); the inhibition of the *E. coli* enzyme by **6** was carefully determined to be linear competitive with a K_i value of 2.6±0.2 μM (literature value under different conditions is 0.80±0.12 μM[9]). We now consider the computational binding energies (Table 1). Binding energies for the ligands converge by 30,000 orientations. Nanomolar inhibitors **3** and **4** were found to have more negative binding energies than micromolar inhibitors **5** and **6** at all levels, by 4.2–7.8 kcal/mol for 30,000 orientations (Table 1). Inhibitors **3** and **4** were of similar calculated interaction energy to the natural substrate **1** and product **2**. Rigid-body docking with 30,000 orientations was employed in subsequent calculations.

Adopting a stepwise in silico screening strategy, an initial filter to select for bioavailability was applied to the NCI database using in-house software, removing ligands of greater than 500 g/mol and absolute charge of greater than 1 e.[21] Application of this filter removed 16% of compounds from the database. A subsequent rigid docking screen of the NCI database highlighted approximately 49,000 compounds with a calculated binding energy for TP more negative than −22 kcal/mol, the predicted affinity of the natural TP substrate, thymidine (**1**, Table 2). A further computational screen was applied to identify from this pool, molecules of similar functional group orientation to known TP ligands (**1**–**6**, Fig. 1), using a modified Tanimoto metric.[22] Combining the top-scoring compounds, 444 ligands were then docked allowing full ligand flexibility. The initial ligand anchor fragment was docked with 30,000 orientations, the subsequent stepwise construction strategy employing a breadth-first search with pruning to 25 configurations per step.[11] The top five docked solutions for each ligand were retained for further analysis. These ligands were then individually scrutinized in the active site for features such as favourable binding interactions, lipophilicity and structural diversity. Of the 40 top-ranked compounds, 13 were available for assay from the NCI repository.

Table 2. Experimental inhibition data for *E. coli* TP for a given concentration of inhibitor, [I] (μM), and flexibly docked interaction energies, U_{bind} (kcal/mol), for NCI compounds identified by in silico screening[a]

Compd	NSC	U_{bind}	[I]	% inhibition
7[b]	632924	−33.2	82	—
8	357688	−33.0	90	28
9	43275	−31.7	106	16
10	81031	−31.4	93	15
11	23145	−31.0	105	23
12	19791	−30.9	89	14
13	122612	−30.8	120	27
14[c]	65043	−30.5	93	84
15	309091	−30.4	130	21
16[d]	659977	−30.3	134	60
17	6487	−30.1	89	6
18	298251	−30.0	100	30
19	3056	−29.9	89	24

[a]Compounds are in rank order according to interaction energy, and are labelled by NSC identifier; see Supplementary Information for chemical identity of all compounds.
[b]An apparent initial increase in velocity was observed.
[c]IC_{50} 19.6±1.5 μM (*E. coli* TP); 77±4.8 μM (human TP).
[d]IC_{50} 94.8±3.6 μM (*E. coli* TP).

Biological evaluation of recombinant human and *E. coli* TP inhibition to determine IC_{50} values of the database screening 'hits' was performed using a continuous spectrophotometric assay.[7,23] This was adapted for a 96-well plate Molecular Devices Spectromax instrument, at a monitoring wavelength of 355 nm with 5-nitro-2′-deoxyuridine (0.13 mM) as the substrate in the presence of the enzyme (45 nM at 25 °C). Percentage inhibition was evaluated at ligand concentrations around 100 μM, although variation was necessitated by issues of solubility. IC_{50} values were determined from the concentration at which a 50% reduction of the initial rate was attained under conditions for which the K_m value for 5-nitro-2′-deoxyuridine was determined as 0.23 mM.[7] The human TP assay was conducted similarly but using a 5-nitro-2′-deoxyuridine substrate concentration of 0.33 mM, and the K_m determined for it was 0.16 mM.

All the compounds identified and available inhibited *E. coli* TP (**7**–**19**, Table 2). The limited amount of material and weak inhibition by many of the compounds prevented IC_{50} determination in a number of cases. In comparison, a control of four randomly selected inhibitor-like compounds exhibited no *E. coli* TP inhibition (see Supplementary Information). Compounds **14** and **16** (Fig. 4) showed the greatest *E. coli* TP inhibition with IC_{50} values of 20 and 95 μM, respectively. A 4-fold apparent decrease in potency is observed for **14** in the human recombinant TP assay, with an IC_{50} of 77 μM. The remaining eleven assayed compounds, as weaker *E. coli* inhibitors, showed no detectable inhibition of human recombinant TP (Table 2). If the predicted inhibitors prove systematically more potent *E. coli* TP inhibitors than human TP ligands when detailed inhibition models have been established, this may indicate some inherent bias in the receptor model stemming from modelling the human enzyme by homology with the microbial form. However, if the inhibition for **14** is assumed to be linear competitive for both *E. coli* and human TP, values for K_i can be calculated, respectively, as 12.5 and 25 mM, using $IC_{50} = K_i(1 + [S]/K_m)$. Ligand **14** is ranked eighth out of the compounds in Table 2 and twenty-sixth from the overall NCI database screen. The binding mode of **14** (Fig. 5) indicates that the 3-methylhydantoin moiety is buried deep within the active-site cleft, close to the phosphate binding pocket. Two polar interactions are made with the phosphate: from the hydantoin ring C_2 oxo, with a H acceptor distance of 2.52 Å; and from the hydantoin ring N_3–H with a H···acceptor distance of 2.93 Å. The semicarbazone sidechain is stabilised by hydrogen bonds from residues on both sides of the moiety (Arg146 and Gln187) as it extends towards the nucleobase-binding cavity (Fig. 5).

Interestingly, compound **11** is a structural analogue of **14** (Fig. 4), based on the hydantoin ring, but without

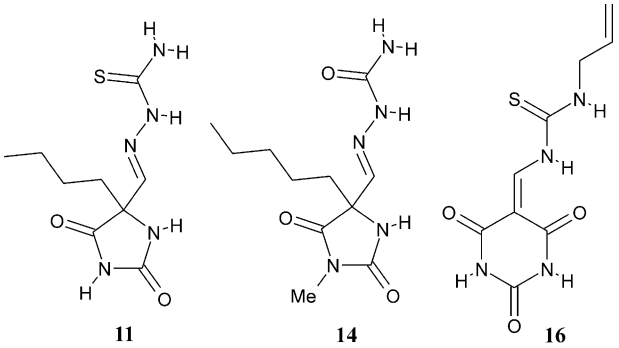

Figure 4. Hit molecules **11**, **14** and **16** from the NCI repository.

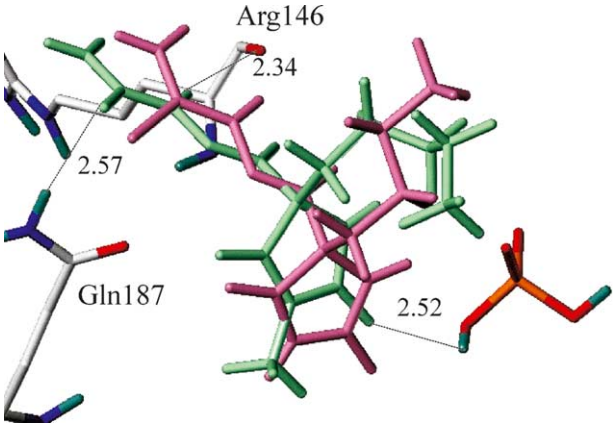

Figure 5. Lowest energy binding mode of **14** (green) with hydrogen bond interactions and the lowest energy binding conformation of **11** (magenta). Distances in Å.

N-methylation of the heterocyclic ring, a reduction in the aliphatic chain to a butyl group, and a thiosemicarbazone instead of semicarbazone pendant group (Fig. 4). Compound **11** is ranked above **14** in the computational screen, possessing a calculated binding affinity of −31 kcal/mol, one half a kcal/mol more favourable than for **14** (Table 2). The small difference in binding energies is reflected by a similarity in predicted binding mode (Fig. 5), with the heterocyclic ring of **11** situated near the phosphate binding pocket and stabilised by polar interactions. His116 and the phosphate interact with the ring oxo groups of **11**, forming hydrogen bonds of H acceptor distance 2.32 and 2.35 Å, respectively. In addition, the ring amino group is stabilised by interaction with Ile143 (2.50 Å). As with the semicarbazone of **14**, the thiosemicarbazone chain of **11** projects into the mid-central active-site cavity, space that the deoxyribose moiety of the thymidine substrate would occupy. The butyl chain is orientated near hydrophobic residues in the phosphate-binding region. Indeed, the pentyl group of **14** stacks more extensively with the side-chain of Ile214, and may account in part for the improved experimental affinity of **14** over **11**.

Ligand **16** exhibits good inhibitory activity against *E. coli* TP. Energetically ranked below **11** and **14**, ligand **16** possesses a six-membered heterocyclic ring similar to thymine (Table 2, Fig. 4). Accordingly, in contrast to **11** and **14**, ligand **16** is predicted to interact with TP at the nucleobase-binding pocket (Fig. 6). This

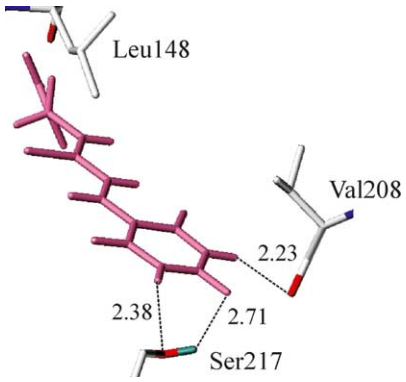

Figure 6. The lowest energy binding mode of **16**. Distances in Å.

ligand therefore binds remote to the phosphate-binding pocket; the distance between ring centroids of **14** and **16** is 9.71 Å. The ring of **16** is stabilised by hydrogen bonding interactions with the hydroxyl sidechain of Ser217, with H acceptor distances of 2.71 and 2.38 Å, and the backbone carbonyl of Val208, with an H acceptor distance of 2.23 Å (Fig. 6). The sidechain of **16** projects into the mid-central active-site cavity, towards the phosphate-binding pocket. The terminal *N*-allyl group is suitably located for a hydrophobic interaction with Leu148.

Although automated docking studies predict a binding mode for **11** and **14** distal to the pyrimidine binding site, the most intuitive TP binding mode for these ligands would involve superposition of the hydantoin ring carbonyl–NH–carbonyl motif upon the equivalent motif in the thymine ring. However, examination of the active site of human TP precludes this mode for **11** and **14**, due to steric clash with Ile214 caused by the bulky aliphatic chain and (thio)semicarbazone substituents at the hydantoin sp^3 carbon. It is important to note that no account is made of protein flexibility in these docking calculations: it is known that TP is capable of domain movement.[13] However, incorporation of large-scale protein flexibility into computational design approaches remains a considerable challenge.

In silico screening of the National Cancer Institute database of anticancer compounds using a three-dimensional structural model of human thymidine phosphorylase has identified a micromolar inhibitor of human TP, ligand **14**, which possesses a novel scaffold unrelated to the molecular framework of the substrate or any previously reported ligands. Docking studies predict binding interactions with TP residues distinct from the pyrimidine binding site. Most previous inhibitors retain the uracil framework, thus maintaining strong binding. In the present study, the ligand structure moves away from the established uracil interactions and it is not surprising that the starting level of inhibition prior to lead optimization is weaker. A number of key ligand features have been proposed as indicative of favourable pharmacokinetic properties:[21] the molecule should possess less than ten hydrogen bond acceptor groups, less than five hydrogen bond donor groups, have a molecular weight of less than 500 g/mol and a

ClogP[24] of less than five. Ligand **14** satisfies all of these criteria, with eight hydrogen bond acceptors and three donor groups, respectively, a molecular weight of 269 g/mol and a ClogP of 0.5. Consequently, this novel TP inhibitor scaffold based on a hydantoin moiety is a lead for the design of potent TP inhibitors with good oral availability.

Acknowledgements

We thank the Ramsay Trust, MRC, BBSRC, the Libyan government and Oxford Biomedica for funding and Dr Phil Edwards for helpful discussions.

References and Notes

1. Furukawa, T.; Yoshimura, A.; Sumizawa, T.; Haraguchi, M.; Akiyama, S.; Fukui, K.; Ishikawa, M.; Yamada, Y. *Nature* **1992**, *356*, 668.
2. Haraguchi, M.; Miyadera, K.; Uemura, K.; Sumizawa, T.; Furukawa, T.; Yamada, Y.; Akiyama, S. *Nature* **1994**, *368*, 198.
3. Miyadera, K.; Sumizawa, T.; Haraguchi, M.; Yoshida, H.; Konstanty, W.; Yamada, Y.; Akiyama, S. *Cancer Res*. **1995**, *55*, 1687.
4. Moghaddam, A.; Zhang, H.-T.; Fan, T.-P. D.; Hu, D. E.; Lees, V. C.; Turley, H.; Fox, S. B.; Gatter, K. C.; Harris, A. L.; Bicknell, R. *Proc. Natl. Acad. Sci. U.S.A*. **1995**, *92*, 998.
5. Matsushita, S.; Nitanda, T.; Furukawa, T.; Sumizawa, T.; Tani, A.; Nishimoto, K.; Akiba, S.; Miyadera, K.; Fukushima, M.; Yamada, Y.; Yoshida, H.; Kanzaki, T.; Akiyama, S. *Cancer Res*. **1999**, *59*, 1911.
6. Fukushima, M.; Suzuki, N.; Emura, T.; Yano, S.; Kazuno, H.; Tada, Y.; Yamada, Y.; Asao, T. *Biochem. Pharmacol*. **2000**, *59*, 1227.
7. Cole, C.; Reigan, P.; Gbaj, A.; Edwards, P. N.; Douglas, K. T.; Stratford, I. J.; Freeman, S.; Jaffar, M. *J. Med. Chem*. **2003**, *46*, 207.
8. Price, M. L. P.; Guida, W. C.; Jackson, T. E.; Nydick, J. A.; Gladstone, P. L.; Juarez, J. C.; Donate, F.; Ternansky, R. J. *Bioorg. Med. Chem. Lett*. **2003**, *13*, 107.
9. Balzarini, J.; Degreve, B.; Esteban-Gamboa, A.; Esnouf, R.; De Clerq, E.; Engelborgs, Y.; Camasara, M. J.; Perez-Perez, M. J. *FEBS Letters* **2000**, *483*, 181.
10. Boyd, M.; Paull, K. D. *Drug Dev.Res*. **1995**, *34*, 91.
11. Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. *J. Comput. Aided Mol. Des*. **2001**, *15*, 411.
12. Cole, C.; Marks, D. S.; Jaffar, M.; Stratford, I. J.; Douglas, K. T.; Freeman, S. *Anti-Cancer Drug Des*. **1999**, *14*, 411.
13. Pugmire, M.; Cook, W.; Jasanoff, A.; Walter, M.; Ealick, S. *J. Mol. Biol*. **1998**, *281*, 285.
14. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.
15. Weiner, P. K.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem*. **1986**, *7*, 230.
16. Gasteiger, J.; Rudolf, C.; Sadowski, J. *Tetrahedron Comput. Methods* **1990**, *3*, 537.
17. *SYBYL 6.8*; Tripos Inc.: St. Louis, MO, USA, 2003.
18. Walter, M.; Cook, W.; Cole, L. *J. Biol. Chem*. **1990**, *265*, 14016.
19. Niedzwicki, J. G.; El Kouni, M. H.; Chu, S. H.; Cha, S. *Biochem. Pharmacol*. **1983**, *32*, 399.
20. Klein, R. S.; Lenzi, M.; Lim, T. H.; Hotchkiss, K. A.; Wilson, P.; Schwartz, E. L. *Biochem. Pharmacol*. **2001**, *62*, 1257.
21. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev*. **1997**, *23*, 3.
22. Willett, P.; Barnard, J. M.; Downs, G. M. *J. Chem. Inf. Comp. Sci*. **1998**, *38*, 983.
23. Wataya, Y.; Santi, D. W. *Anal. Biochem*. **1981**, *112*, 96.
24. Leo, A.; Hoekman, D. *Perspect. Drug. Discov. Des*. **2000**, *18*, 19.