# Theoretical Studies Suggest a New Antifolate as a More Potent Inhibitor of Thymidylate Synthase

**Tai-Sung Lee and Peter A. Kollman***

*Contribution from the Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0446*

*Received July 19, 1999. Revised Manuscript Received January 28, 2000*

**Abstract:** A set of free energy calculations were performed for selected antifolate compounds as inhibitors of thymidylate synthase (TS). These calculations reproduced the nonadditive behavior of different substitutions on selected compounds described in the experiments by Jones et al[1]. (*J. Med. Chem.* **1996**, *39* (4), 904−917). Molecular dynamics (MD) simulations showed that the nonadditive behavior was due to the interference between the way different substituents interacting with key protein side chains. Pictorial representation of free energy change (PROFEC) and free energy calculations suggested that a compound not previously considered would bind more tightly to TS than the best binding known compounds in this series of propargyl inhibitors and, thus, could be a promising candidate in anticancer drug design.

## I. Introduction

Thymidylate synthase (TS) catalyzes the conversion of dUMP and 5,10-$CH_2$-$H_4$ folate into dTMP. It is the only known path for dTMP synthesis, and dTMP synthesis is believed to be one of the rate-limiting steps in DNA replication. Hence inhibitors of this enzyme can be used to block dTMP synthesis and can be of possible chemotherapeutic use in cancer.[2]

Much effort has been devoted to design and test inhibitors of TS. 10-Propargyl-5,8-dideazafolic acid has been shown to be an effective inhibitor that is clinically active, but renally toxic.[3−7] More soluble and nontoxic modified compounds have been developed.[8−10] In particular, polyglutamatic acid compounds were designed to have tighter binding and better pharmacological properties.[11−17]

However, the poor transport property of polyglutamatic acid compounds in the cell limits their potential as TS-inhibiting drugs.[18] Therefore, the development of new inhibitors with better transport properties is essential. Jones et al. designed and synthesized a total of 31 lipophilic quinazoline inhibitors of TS.[1] These compounds all lack the bulky glutamate residues which are likely responsible for poor transport properties. Most of these compounds have the same *N*-((3,4-dihydro-2-methyl-6-quinazolinyl)methyl)-*N*-prop-2-ynylaniline structure (shown in the upper part of Figure 1). The aniline was substituted with simple lipophilic substituents at either position 3 or 4, or both. Those compounds were tested for their inhibition of *Escherichia coli* TS and human TS and also for their inhibition of the growth in tissue culture of a murine leukemia, a human leukemia, and a thymidine kinase-deficient human adenocarcinoma. Their inhibition results suggested the following: (i) A 3-substituent such as $CF_3$, iodo, or ethynyl enhances binding by up to 1 order of magnitude and in the case of $CF_3$ was proven to fill a nearby pocket in the enzyme. (ii) A simple strongly electron withdrawing substituent in the 4-position enhances binding by 2 orders of magnitude. (iii) Attempts to combine the enhancements of i and ii in the same molecule were generally unsuccessful.

Some of those compounds provided almost as much binding affinity as the tight binding polyglutamatic acid inhibitors.

(1) Jones, T. R.; Varney, M. D.; Webber, S. E.; Lewis, K. K.; Marzoni, G. P.; Palmer, C. L.; Kathardekar, V.; Welsh, K. M.; Webber, S.; Matthews, D. A.; Appelt, K.; Smith, W. W.; Janson, C. A.; Villafranca, J. E.; Bacquet, R. J.; Howland, E. F.; Booth, C. L.; Herrmann, S. M.; Ward, R. W.; White, J.; Moomaw, E. W.; Bartlett, C. A.; Morse, C. A. *J. Med. Chem.* **1996**, *39* (4), 904−917.

(2) Santi, D. V.; Danenberg, P. V. In *Chemistry and Biochemistry of Folates. Vol. 1: Folates and Pterins*; Blakley, R. L., Benkovic, S. J., Eds.; John Wiley & Sons: New York, 1984; pp 345−398.

(3) Berman, E. M.; Werbel, L. M. *J. Med. Chem.*, **1991**, *34*, 479−485.

(4) Jones, T. R.; Calvert, A. H.; Jackman, A. L.; Brown, S. J.; Jones, M.; Harrap, K. *Eur. J. Cancer* **1981**, *17*, 11−19.

(5) Jones, T. R.; Calvert, A. H.; Jackman, A. L.; Eakin, M. A.; Smithers, M. J.; Betteridge, R. F.; Newell, D. R.; Hayter, A. J.; Stocker, A.; Harland, S. J.; Davies, L.; Harrap, K. R. *J. Med. Chem.* **1985**, *28*, 1468−1476.

(6) Calvert, A. H.; Alison, D. L.; Harland, S. J.; Robinson, B. A.; Jackman, A. L.; Jones, T. R.; Newell, D. R.; Siddik, Z. H.; Wiltshaw, E.; McElwain, T. J.; Smith, I. E.; Harrap, K. R. *J. Clin. Oncol.* **1986**, *4*, 1245−1252.

(7) Jodrell, D. I.; Newell, D. R.; Morgan, S. E.; Clinton, S.; Bensted, J. P. M.; Hughes, L. R.; Calvert, A. H. *Br. J. Cancer* **1991**, *64*, 833−838.

(8) Jones, T. R.; Thornton, T. J.; Flinn, A.; Jackman, A. L.; Newell, D. R.; Calvert, A. H. *J. Med. Chem.* **1989**, *32*, 847−852.

(9) Hughes, L. R.; Jackman, A. L.; Oldfield, J.; Smith, R. C.; Burrows, K. D.; Marsham, P. R.; Bishop, J. A. M.; Jones, T. R.; O'Connor, B. M.; Calvert, A. H. *J. Med. Chem.* **1990**, *33*, 3060−3067.

(10) Jackman, A. L.; Newell, D. R.; Gibson, W.; Jodrell, D. I.; Taylor, G. A.; Bishop, J. A.; Hughes, L. R.; Calvert, A. H. *Biochem. Pharmacol.* **1991**, *42*, 1885−1895.

(11) Moran, R. G.; Colman, P. D.; Rosowsky, A.; Forsch, R. A.; Chan, K. K. *Mol. Pharmacol.* **1985**, *27*, 156−166.

(12) Sikora, E.; Jackman, A. L.; Newell, D. R.; Calvert, A. H. *Biochem. Pharmacol.* **1988**, *37*, 4047−4054.

(13) Jackman, A. L.; Taylor, G. A.; O'Connor, B. M.; Bishop, J. A.; Moran, R. G.; Calvert, A. H. *Cancer Res.* **1990**, *50*, 5212−5218.

(14) Jackman, A. L.; Taylor, G. A.; Gibson, W.; Kimbell, R.; Brown, M.; Calvert, A. H.; Judson, I. R.; Hughes, L. R. *Cancer Res.* **1991**, *51*, 5579−5586.

(15) Clarke, S.; Ward, J.; Planting, A.; Spiers, J.; Smith, R.; Verweij, J. *Proc. Am. Assoc. Cancer Res.* **1992**, *33*, 406.

(16) Hagan, R. L.; Duch, D. S.; Smith, G. K.; Hanlon, M. H.; Shane, B.; Freisheim, J. H.; Hynes, J. B. *Biochem. Pharmacol.* **1991**, *41*, 781−787.

(17) Duch, D. S.; Banks, S.; Dev, I. K.; Dickerson, S. H.; Ferone, R.; Heath, L. S.; Humphreys, J.; Knick, V.; Pendergast, W.; Singer, S.; Smith, G. K.; Waters, K.; Wilson, H. R. *Cancer Res.* **1993**, *53*, 810−818.

(18) Albrecht, A. M.; Biedler, J. L. In *Folate Antagonists as Therapeutic Agents*; Sirotnak, F. M., Burchall, J. J., Ensminger, W. B., Montgomery, J. A., Eds.; Academic Press: San Diego, 1984; Vol. 1, pp 317−353.

10.1021/ja9925554 CCC: $19.00 
Published on Web 04/25/2000

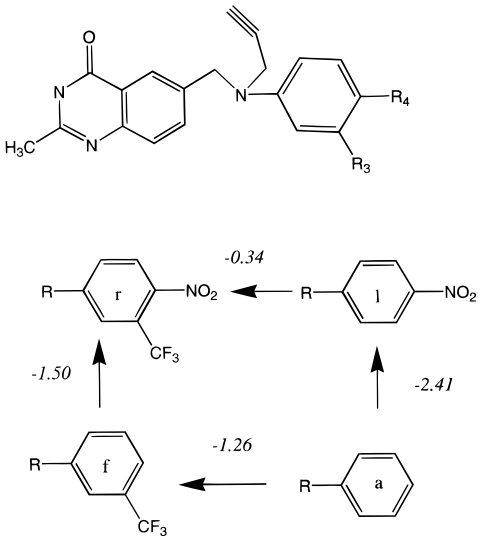



**Figure 1.** Four compounds designed by Jones et al.[1] The upper is the common structure, and the lower shows different substitutions for these compounds. Each compound is named as the letter appearing on the aniline ring.

Unfortunately this series of compounds generally have low water solubility, and they probably do not penetrate cells particularly well. Hence those compounds are still not good drug candidates, and modifications to this type of compound are needed.

Understanding the interaction of TS with its inhibitors at the atomic level will be helpful in the development of better inhibitors. Molecular dynamics (MD) simulations have been used to understand the interaction between proteins and ligands. MD-based free energy methods have been successfully applied to determine the relative binding free energies of ligands to hosts in many molecular systems.[19] In addition, the pictorial representation of free energy change (PROFEC) method[20] allows one to average over molecular dynamics trajectories and generate contour maps that can suggest how binding free energy might be changed upon modification of the ligand. It has been shown how PROFEC is often able to reproduce general trends in binding free energies.[20,21]

In this paper we have performed molecular dynamics simulations and free energy calculations on the compounds with 3-trifluoromethyl- and 4-nitro-substituted compounds designed by Jones et al.[1] (Figure 1). Those compounds show strong nonadditive results: The 3-trifluoromethyl substitution alone gives 1.26 kcal/mol binding enhancement while the 4-nitro-substituted compound is bound more tightly than the unsubstituted one by 2.41 kcal/mol. However, the binding enhancement for the 3-trifluoromethyl-4-nitro substitution is only 2.75 kcal/mol. The relative binding free energies from our free energy calculations agree well with the experimental results. The MD results also rationalize the nonadditive behavior of these 3-trifluoromethyl- and 4-nitro-substituted compounds. The PROFEC calculations predicted that 5- and 6-fluoro substitutions could enhance binding. The free energy calculations for those 5- and 6-fluorine-substituted compounds confirmed the PROFEC prediction. The results also predict that the 4-nitro-5,6-difluoro-substituted compound would lead to near additive effects from 4-nitro and 5,6-difluoro substitutions. Thus, this compound would bind more tightly than any other compounds considered to date.

## II. Methods

The simulations described here were performed with the molecular dynamics simulation package AMBER 4.1[22,23] using the all-atom force field of Cornell et al.[24] The charges, van der Waals parameters, equilibrium bond lengths, bond angles, and dihedrals for standard residues were taken from the AMBER database. Some parameters for the inhibitors we studied here, however, have not yet been developed. Those parameters are obtained in the following way. The atomic charges of the inhibitors are obtained by fitting the electrostatic potentials with the RESP method.[25] The fitting procedure is as follows: The electrostatic potential of compound **a** (Figure 1) was obtained by a single-point ab initio quantum mechanical calculation with the 6-31G* basis set on the 3-21G* optimized geometry using the Gaussian 94 package.[26] The charges of the chemical equivalent atoms, e.g., H in $CH_3$, were restrained to be the same. After the charges of compound **a** were obtained, the same fitting procedure was performed for each inhibitor, with the charges of the fused ring and the propargyl group constrained to the same values as in compound **a**. In this way, only those atoms of the aniline ring needed to be mutated during free energy calculations. For the fused ring, other parameters are assigned to be the same as similar atom types in the force field database. For the propargyl group, a model molecule, i.e., propyne $CH_3C{\equiv}CH$, was chosen and the second derivatives of total energy with respect to different geometric parameters, i.e., bond lengths, bond angles, and dihedrals, were calculated at the MP2/6-31+G(d) level with the full optimized geometry. The force constant for each geometrical parameter was obtained from the corresponding second derivatives of the total energy. The equilibrium geometrical parameters are obtained from the X-ray crystal structure of the complex of *E. coli* TS and compound **f** in Figure 1. For the substituted aniline ring, the parameters for interactions involving $NO_2$ and $CF_3$ were missing and were obtained using the model molecules, nitrobenzene and trifluoromethylbenzene, by the same procedure. The force constant of the dihedral angle between the nitro group and the benzene plane was chosen so that the experimental value of the rotation barrier of the nitro group (2.9 kcal/mol)[27] is reproduced. A complete listing of force constants and equilibrium geometric parameters for all inhibitors are listed in the Supporting Information. More details for the simulation methods are described in the following subsections.

(19) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395−2417.
(20) Radmer, R.; Kollman, P. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 215−227.
(21) Wang, L.; Veenstra, D.; Radmer, R.; Kollman, P. *Proteins: Struct. Funct. Genet.* **1998**, *32*, 438−458.
(22) Pearlman, D.; Case, D.; Caldwell, J.; Ross, W.; Cheatham, T.; Debolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1−41.
(23) Pearlman, D.; Case, D.; Caldwell, J.; Ross, W.; Cheatham, T.; Ferguson, D.; Seibel, G. M.; Singh, U. C.; Kollman, P. *Amber 4.1*; University of California, San Francisco: San Francisco, 1995.
(24) Cornell, W.; Cieplak, P.; Bayly, C.; Gould, I.; Merz, K.; Ferguson, D.; Spellmeyer, D.; Fox, T.; Caldwell, J.; Kollman, P. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(25) Bayly, C.; Cieplak, P.; Cornell, W.; Kollman, P. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(26) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.;; Pople, J. A. *Gaussian 94*, revision E.2; Gaussian, Inc.: Pittsburgh, 1995.
(27) Hehre, W. J.; Radom, L.; Pople, J. A. *J. Am. Chem Soc.* **1972**, *94*, 1496−1504.

**A. Molecular Dynamics (MD).** We began with the X-ray crystal structure of the complex of *E. coli* TS and compound **f** in Figure 1. The structures for other inhibitors are mutated from compound **f** using the LEaP module in the AMBER package. The MD simulations used the following protocol. Unless specified, all simulations were performed with a time step of 2 fs with the SHAKE algorithm.[28]

For inhibitors in water, a 10-Å shell of water molecules was added around the inhibitor molecule. A 10-Å nonbonded cutoff was used with periodic boundary conditions. The temperature was kept at 300 K, and the pressure was kept at 1 atm. A 60-ps equilibration was performed before the production phase of the simulation.

For protein−inhibitor complexes, a 200-step conjugate minimization was performed to remove bad contacts between the protein molecule and the inhibitor. Water molecules were added within 20 Å from the C4 atom of the aniline ring of the inhibitor. Only residues within 14 Å from the C4 atom of the aniline ring of the inhibitor were allowed to move, and a 14-Å nonbonded cutoff was used. A 2000-step conjugate minimization within this 14-Å region was performed, followed by a 200-ps equilibration. Counterions were added to the protein−inhibitor complexes before simulations. We found that the electrostatic interactions due to residues Asp81 and Glu82 had a strong influence on the simulations. The likely reason is that these two residues are very close to the inhibitor and they are also very flexible. Since those two residues are exposed to the solvent, it is reasonable to add counterions near them. The following procedure was used to add the counterions near Asp81 and Glu82: Asp81 and Glu82 were first protonated, and the carboxyl hydrogens were replaced by $Na^+$ ions, followed by a 200-step conjugate energy minimization to move the $Na^+$ ions into lower energy positions. During all simulations, these two $Na^+$ ions were restrained to their positions by a 20 kcal/Å$^2$ harmonic potential. This restraint procedure is used to reduce the fluctuations in the electrostatic free energies.

The structures after equilibration were used as initial structures for subsequential simulations.

**B. Free Energy Calculations.** In this study, the thermodynamic integration method (TI)[19,20] was used for the free energy calculations. In the TI method, for the transformation of one thermodynamic state into another, a coupling parameter $\lambda$ is introduced and the Hamiltonians of the two states are defined as $H_0(\lambda = 0)$ and $H_1(\lambda = 1)$. The total free energy change for the transformation is

$$\Delta G = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda \tag{1}$$

where $\langle \partial H(\lambda)/\partial \lambda \rangle_\lambda$ is an ensemble average at $\lambda$. In practice, the integral is calculated by the trapezoidal integration method, in which a number of evenly spaced windows with different $\lambda$ values ranging from 0 to 1 are chosen and, at each window $\langle \partial H(\lambda)/\partial \lambda \rangle_\lambda$ is calculated by averaging over molecular dynamics trajectories. If assuming that the kinetic energy contribution can be neglected, $\Delta G$ can be approximated by

$$\Delta G = \sum_0^1 \left\langle \frac{\partial V(\lambda)}{\partial \lambda} \right\rangle_{\lambda_i} \Delta\lambda \tag{2}$$

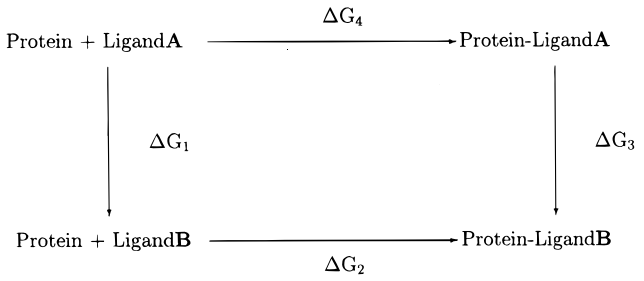
where $\lambda_i$ is the $\lambda$ value of the $i$th window and $\Delta\lambda$ is the interval between successive windows. $V(\lambda)$ is the potential function that describes the atomic interactions in the system. It has the following form in the force field of Cornell et al:[24]

$$V = \sum_{\text{bonds}} K_r(r - r_{\text{eq}})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(\mathrm{n}\phi - \gamma)] + \sum_{i<j}\left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right] + \sum_{i<j} \frac{q_i q_j}{\epsilon R_{ij}} \tag{3}$$

where those terms represent the bond, bond angle, dihedral, van der Waals, and electrostatic interactions, respectively. $A_{ij}$ and $B_{ij}$ are calculated on the basis of the following equations:

$$A_{ij} = \epsilon_{ij}(R_{ij}^*)^{12}$$
$$B_{ij} = 2\epsilon_{ij}(R_{ij}^*)^{6} \tag{4}$$
$$\epsilon_{ij} = (\epsilon_i \epsilon_j)^{1/2}$$
$$R_{ij}^* = R_i^* + R_j^* \tag{5}$$

The relative binding free energy of two different but similar ligands can be determined via the thermodynamic cycle:[29]




Because free energy is a state function, the total free energy change for the whole cycle, $\Delta G_1 + \Delta G_2 - \Delta G_3 - \Delta G_4$, should be zero. Hence the relative binding free energy, $\Delta G_4 - \Delta G_2$, is equal to $\Delta G_1 - \Delta G_3$. Rather than calculate the horizontal processes of the thermodynamic cycle at great computational expense, the less expensive vertical processes can be calculated and used to determine the relative binding free energy: the parameters that describe ligand **A** are perturbed into the parameters that describe ligand **B** in both the enzyme and solution to get $\Delta G_1$ and $\Delta G_3$.

The starting position and velocity for each atom in each free energy calculation were taken from the corresponding final structure of MD equilibration. The protocol for the free energy calculations is the same as the protocol used in the MD simulations. For inhibitors in water, the thermal integration (TI) method with 101 windows was used. For each window, there are 1 ps of equilibration and 1 ps of data collection, i.e., 100 ps of data collection for the whole simulation. Both forward ($\lambda = 0$ to $\lambda = 1$) and backward ($\lambda = 1$ to $\lambda = 0$) simulations were performed for each mutation. The error bars were estimated by the differences between results from the forward and backward simulations. For protein−inhibitor complexes, the simulation procedures are the same except simulations of different lengths of data collection time were performed to ensure the convergence of results.

**C. Pictorial Representation of Free Energy Change (PROFEC).** The basic idea of the PROFEC method is to generate a contour map around the ligand, which can indicate

(28) Ryckaert, J.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, *23*, 327−341.

(29) Straatsma, T.; McCammon, J. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407−435.

**Table 1.** Calculated Relative Binding Free Energies for Four Mutations: Compound **a** to Compound **f**, Compound **f** to Compound **r**, Compound **a** to Compound **l**, and Compound **l** to Compound **r**[a]

| | **a → f** | | | **f → r** | | | **a → l** | | | **l → r** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| condition | $\Delta G_w$ | $\Delta G_p$ | $\Delta\Delta G$ | $\Delta G_w$ | $\Delta G_p$ | $\Delta\Delta G$ | $\Delta G_w$ | $\Delta G_p$ | $\Delta\Delta G$ | $\Delta G_w$ | $\Delta G_p$ | $\Delta\Delta G$ |
| 25 ps | | | | | | | | | | | | |
| backward | | 9.24 | | | 7.63 | | | 19.50 | | | −1.37 | |
| | | 9.32 | | | 7.91 | | | | | | | |
| forward | | 9.84 | | | 8.18 | | | 17.96 | | | −2.81 | |
| | | 10.64 | | | 5.02 | | | | | | | |
| average | | 9.54 | −2.05 ± 0.35 | | 7.90 | −2.65 ± 0.32 | | 18.73 | −3.18 ± 0.79 | | −2.09 | −2.92 ± 0.83 |
| | | 9.98 | *−1.61 ± 0.71* | | 6.46 | *−4.08 ± 1.48* | | | | | | |
| 100 ps | | | | | | | | | | | | |
| backward | 11.54 | 9.84 | | 10.24 | 7.22 | | 21.93 | 18.55 | | 1.11 | 0.95 | |
| | | 9.96 | | | 7.89 | | | | | | | |
| forward | 11.64 | 9.50 | | 10.41 | 8.04 | | 21.89 | 18.85 | | 0.74 | −0.59 | |
| | | 12.83 | | | 12.18 | | | | | | | |
| average | 11.59 | 9.67 | −1.92 ± 0.22 | 10.33 | 7.63 | −2.92 ± 0.45 | 21.91 | 18.70 | −3.21 ± 0.17 | 0.92 | 0.18 | −0.64 ± 0.88 |
| | | 11.39 | *−0.20 ± 1.48* | | 10.03 | *−0.52 ± 2.18* | | | | | | |
| 200 ps | | | | | | | | | | | | |
| backward | | | | 10.59 | 7.93 | | | 18.84 | | 0.71 | −0.02 | |
| | | | | | 9.10 | | | | | | | |
| forward | | | | 10.51 | 7.83 | | | 18.65 | | 0.94 | 0.01 | |
| | | | | | 11.44 | | | | | | | |
| average | | | | 10.55 | 7.88 | −2.67 ± 0.09 | | 18.75 | −3.17 ± 0.11 | 0.83 | −0.01 | −0.83 ± 0.13 |
| | | | | | 10.27 | *−0.28 ± 1.21* | | | | | | |
| 300 ps | | | | | | | | | | | | |
| backward | | | | | | | | 18.92 | | | | |
| forward | | | | | | | | 18.66 | | | | |
| average | | | | | | | | 18.78 | −3.12 ± 0.15 | | | |
| exptl value | | −1.26 | | | −1.50 | | | −2.41 | | | −0.34 | |

[a] $\Delta G_w$ is the free energy change of ligand mutation in water while $\Delta G_p$ is the free energy change of ligand mutation in protein. $\Delta\Delta G$ is the relative binding free energy. The $\Delta G_w$ with the longest simulation time was used in the $\Delta\Delta G$ calculation for each mutation. The numbers in italics are the results from simulations without counterions. Units are kcal/mol.

how the binding affinity of the ligand will change when additional particles are added near the ligand. The contour map is obtained by introducing test particles on the grid points in the vicinity of the ligand and evaluating the insertion free energies of the test particles on the grid points from the molecular dynamics simulation. Two contour maps for the inhibitors in water and the protein−inhibitor complexes need to be generated, and their difference map is used to indicate the binding affinity change for introducing particles around the ligand of interest. A detailed discussion of the PROFEC method can be found in ref 20. Here, we only briefly describe the method and list the computational procedures.

The position and orientation of the grid is defined using one atom of the ligand of interest to indicate the origin of the grid, a second atom to define the *x* axis, and a third atom to define the *xy* plane. The *z* axis is generated by the cross product of the unit vectors of *x* and *y* axes. In this study, the C4 atom of the aniline ring of the inhibitor was used as the origin of the grid. The C3 and C5 atoms attached to this C4 atom were used to define the *x* axis and the *xy* plane, respectively. During calculations, the origin was shifted to the center of aniline ring to get a clearer picture around the aniline ring. The free energy cost of adding a Lennard-Jones particle at a specific grid point is calculated by

$$\Delta G(i,j,k) = -RT \ln\langle \exp(-\Delta V(i,j,k)/RT)\rangle_0 \qquad (6)$$

where *i*, *j*, and *k* are the grid indices. $\Delta G(i,j,k)$ is the change of free energy resulting from the addition of the Lennard-Jones particle. $\Delta V(i,j,k)$ is the van der Waals interaction energy between the particle and the surrounding atoms, which include all atoms other than the aniline ring. $\langle\ \rangle$ denotes the ensemble average. The parameters of the Lennard-Jones particle are $R_{\text{probe}} = 1.75$ Å and $\epsilon_{\text{probe}} = 0.061$ kcal/mol, which corresponds to a fluorine atom. The electrostatic contributions for introducing particles at the grid points can be examined by calculating the derivative of free energy with respect to the charge at each grid point. The electrostatic contribution to the free energy is given by

$$\left[\frac{dG(i,j,k)}{dq}\right]_{LJ(i,j,k)} = \frac{\langle \Phi(i,j,k)\exp(-\Delta V(i,j,k)/RT)\rangle_0}{\langle \exp(-\Delta V(i,j,k)/RT)\rangle_0} \qquad (7)$$

where *q* is the charge of the introducing particle and $\Phi(i,j,k)$ is the electrostatic potential at each grid point. LJ(*i*,*j*,*k*) indicates that the derivative is calculated assuming that a Lennard-Jones particle is added at the point (*i*,*j*,*k*). In practice, the contour map generated by insertion of Lennard-Jones particles can be colored according to the free energy derivatives of the electrostatic contributions. This can be used to suggest how a ligand charge distribution should be changed to improve the binding affinity.

Simulations (500 ps) were performed for compound **a** and compound **f** both in water and in protein. Trajectories were recorded every 0.1 ps, and there are 5000 structures for each complex. The PROFEC module of the AMBER package was used for the contour map calculations based on those trajectories. The protocol used in MD simulation was the same as mentioned in section II.A. The grid spacing along each direction is 0.30 Å. The numbers of grid points for the *x*, *y*, and *z* directions are 25, 25, and 5, respectively.

All MD and free energy simulations were performed using the parallel mode (8 CPUs) of the AMBER program on the SGI cluster at the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-

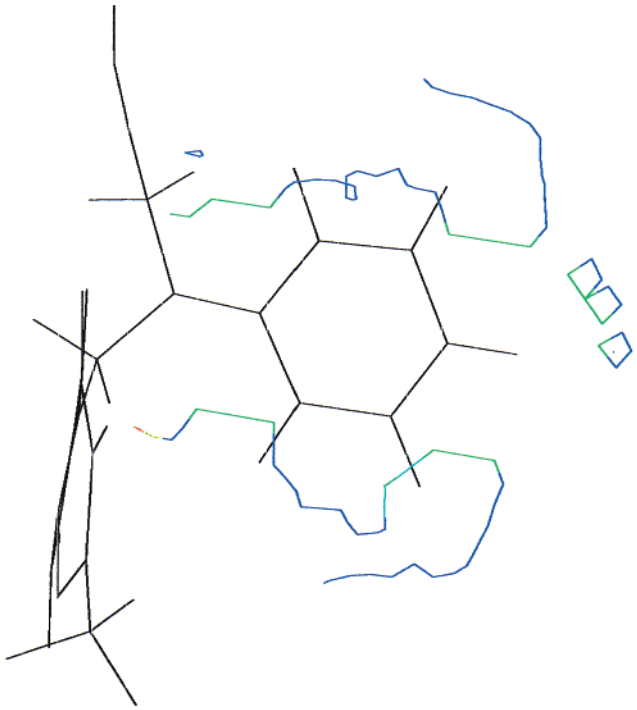


**Figure 2.** The PROFEC contour map of zero van der Waals interaction for compound **a**.

Champaign. PROFEC calculations were performed on a Digital 4100 workstation with 4 × 533 MHz CPUs and 512 MB of memory.

## III. Results

**A. Free Energy Calculation Results.** Table 1 lists the results of free energy calculations for four mutations, compound **a** to compound **f**, compound **f** to compound **r**, compound **a** to compound **l**, and compound **l** to compound **r**. The errors of free energy calculations were estimated by the difference of the forward result and the backward result. The free energy changes in water are very stable with error bars less than 0.2 kcal/mol for 100-ps simulations. For protein−ligand complexes the errors are, however, larger and fluctuating for different ligands. Fortunately, the free energy changes of protein−ligand complexes converge with respect to the simulation time. The results agree with experimental results within 1 kcal/mol. The nonadditive behavior of 3- and 4-substitution is also clearly shown. The 3-trifluoromethyl substitution (compound **f**) alone gives 1.92 kcal/mol (experimental: 1.26) binding enhancement while the 4-nitro-substituted compound (compound **l**) is bound tighter than the unsubstituted one by 3.12 kcal/mol (experimental: 2.41). However, the binding enhancement for the 3-trifluoromethyl substitution on the 4-nitro-substituted compound (compound **l** to compound **r**) is only 0.83 kcal/mol (experimental: 0.34). With 4-nitro substitution, the binding enhancement for the 3-trifluoromethyl substitution decreased 1.09 kcal/mol, from 1.92 to 0.83 kcal/mol (experimental: 0.92, from 1.26 to 0.34 kcal/mol).

**B. Pictorial Representation of Free Energy Change (PROFEC) Results.** We have performed PROFEC calculations for compound **a** and compound **f**. The PROFEC contour maps of zero van der Waals interaction are shown in Figures 2 (compound **a**) and 3 (compound **f**). To be clear, only the slice of contour map through the aniline ring is shown. Since all substituent groups of the aniline ring lay in the aniline ring plane, this slice of contour map is sufficient for adding particles. The coloring in those maps indicates the sign of charge that the introducing particle should have. Blue, green, and red indicate a negative charge, a neutral charge, and a positive charge, respectively. For compound **a**, the contour map suggests that a large and nonpositively charged substituent at the 3- or 4-position or a small and negatively charged substituent at the 5- or 6-position could enhance the binding affinity. For compound **f**, which has a large substituent at the 3-position, the contour map changes significantly and suggests no binding affinity enhancement at the 5- or 6-position and the 4-position substituent should be smaller and negatively charged.

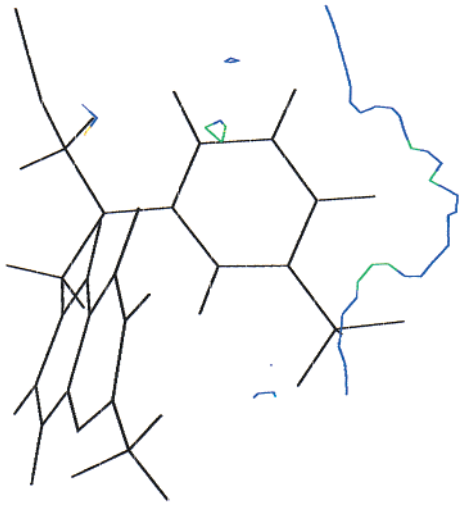


**Figure 3.** The PROFEC contour map of zero van der Waals interaction for compound **f**.

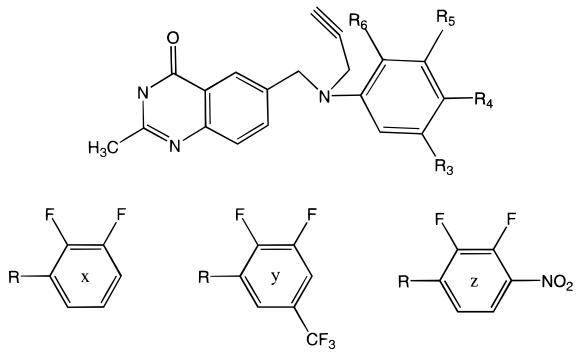



**Figure 4.** Compounds designed according to PROFEC suggestions. The upper is the common structure, and the lower shows different substitutions for these compounds. Each compound is named as the letter appearing on the aniline ring.

These results agree with the free energy results. We designed three compounds, compound **x**, compound **y**, and compound **z**, shown in Figure 4, and performed free energy calculations for those compounds to support the PROFEC suggestions quantitatively.

**C. Free Energy Results for PROFEC Predicted Inhibitors.** Table 2 lists the results of free energy calculations for four mutations, compound **a** to compound **x**, compound **f** to compound **y**, compound **l** to compound **z**, and compound **x** to compound **y**. The errors of those calculations are similar to the free energy calculations listed in Table 1. The 5,6-difluoro substitution (compound **x**) alone gives 2.98 kcal/mol binding enhancement, which supports the suggestion from the PROFEC results of compound **a**. This enhancement is at the same level (2.38 kcal/mol) for the 4-nitro-substituted compound (compound **l** to compound **z**). However, this effect is much less (0.26 kcal/mol) on the 3-trifluoromethyl-substituted compound (compound **f** to compound **y**). Similarly, the binding enhancement for the

**Table 2.** Calculated Relative Binding Free Energies for Four Mutations: Compound **a** to Compound **x**, Compound **f** to Compound **y**, Compound **l** to Compound **z**, and Compound **x** to Compound **y**[a]

| | **a → x** | | | **f → y** | | | **l → z** | | | **x → y** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| condition | $\Delta G_w$ | $\Delta G_p$ | $\Delta\Delta G$ | $\Delta G_w$ | $\Delta G_p$ | $\Delta\Delta G$ | $\Delta G_w$ | $\Delta G_p$ | $\Delta G$ | $\Delta G_w$ | $\Delta G_p$ | $\Delta\Delta G$ |
| 25 ps | | | | | | | | | | | | |
| backward | | | | | 8.35 | | | −5.39 | | | 4.27 | |
| forward | | | | | 8.91 | | | −5.51 | | | 4.94 | |
| average | | | | | 8.63 | 0.02 ± 0.33 | | −5.45 | −2.50 ± 0.25 | | 4.60 | 0.77 ± 0.34 |
| 100 ps | | | | | | | | | | | | |
| backward | 16.18 | 13.42 | | 8.56 | 8.83 | | −2.75 | −5.45 | | 3.84 | 4.29 | |
| forward | 15.88 | 12.64 | | 8.66 | 8.91 | | −3.13 | −5.13 | | 3.83 | 5.02 | |
| average | 16.03 | 13.03 | −3.00 ± 0.54 | 8.61 | 8.87 | 0.26 ± 0.09 | −2.94 | −5.29 | −2.35 ± 0.35 | 3.83 | 4.65 | 0.82 ± 0.37 |
| 200 ps | | | | | | | | | | | | |
| backward | | 13.26 | | | | | −3.25 | −5.49 | | | 4.10 | |
| forward | | 12.83 | | | | | −2.90 | −5.29 | | | 5.08 | |
| average | | 13.04 | −2.98 ± 0.35 | | | | −3.07 | −5.39 | −2.44 ± 0.29 | | 4.59 | 0.76 ± 0.49 |
| 300 ps | | | | | | | | | | | | |
| backward | | | | | | | | −5.30 | | | | |
| forward | | | | | | | | −5.35 | | | | |
| average | | | | | | | | −5.32 | −2.38 ± 0.21 | | | |

[a] $\Delta G_w$ is the free energy change of ligand mutation in water while $\Delta G_p$ is the free energy change of ligand mutation in protein. $\Delta\Delta G$ is the relative binding free energy. The $\Delta G_w$ with the longest simulation time was used in the $\Delta\Delta G$ calculation for each mutation. Units are kcal/mol.

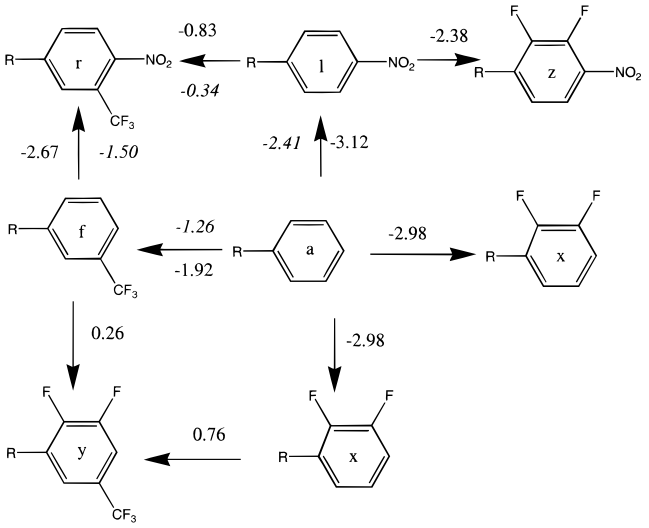



**Figure 5.** The free energy calculation results for all mutations. The numbers in italics are the relative binding free energies calculated from experimental $K_i$ values. Other numbers are the relative binding free energies from the simulations. Units are kcal/mol.

3-trifluoromethyl substitution decreases from 1.26 kcal/mol favorable to 0.76 kcal/mol unfavorable when the 5- and 6-positions are substituted by fluorine atoms (compound **y** to compound **x**). The decrease is 2.02 kcal/mol.

The first three mutations only involve two fluorine substitutions, and one may be confused by the significantly different free energy changes for those mutations in water. The reason is the following: We included all intragroup interactions in the free energy calculations. To avoid the artificial contributions from intramolecular interactions, one needs to compare $\Delta\Delta G$ for different mutations, not $\Delta G$ for any single mutation. We have performed 100-ps free energy calculations in vacuum for three mutations, **a** to **x**, **f** to **y**, and **l** to **z**. The results are 14.86 ± 0.13, 7.36 ± 0.34, and −3.75 ± 0.14 kcal/mol, which lead to the solvation free energy changes of 1.17, 1.25, and 0.68 kcal/mol for mutations **a** to **x**, **f** to **y**, and **l** to **z**, respectively. Since aromatic fluorine should be more hydrophobic but also more polar than aromatic hydrogen, small changes in solvation free energies for these mutations seems reasonable. The fact that all three solvation free energy changes are within 0.6 kcal/mol suggests that there is little nonadditive effect in adding ortho difluoro in this system.

Our results are summarized in Figure 5 and suggest that combined 4-nitro and 5,6-difluoro substitution will give about 5.5 kcal/mol binding enhancement over the parent compound, which corresponds to a ∼10000-fold change in $K_i$. To support these results, 300-ps free energy simulations were performed for the mutation **a** to **l** and **l** to **z** (Tables 1 and 2). The results are similar to the 100- and 200-ps simulations, suggesting that the results should be reliable.

## IV. Discussion

There are three kinds of substitutions in our simulations, i.e., 3-trifluoromethyl, 4-nitro, and 5,6-difluoro substitutions. The free energy calculation results show that the effects of 3-trifluoromethyl and 4-nitro substitutions are nonadditive. The effects of 3-trifluoromethyl and 5,6-difluoro substitutions are also nonadditive while the effects of 4-nitro and 5,6-difluoro substitutions are additive. To explain the nonadditive behavior of relative binding free energies for different substitutions, we examined each substituent and its neighbor residues from MD trajectories as follows:

We first performed MD simulations for all inhibitors. A 500-ps simulation was performed for each protein−inhibitor complex. Trajectories were recorded every 0.1 ps, and there are 5000 structures for each complex. The protocol used in the simulations was the same as mentioned in the section II.A. The MD trajectories were analyzed by the CARNAL module of the AMBER 4.1 package.[22,23] To give a reference point for this analysis, we also present stereoviews of the final structures of TS bound compound **f** and compound **z** in Figures 6 and Figure 7.

For 3-trifluoromethyl substitution, the substituent has been proposed to interact with a hydrophobic pocket formed by residues Leu172 and Val262, the edge of Trp80, and the backbone carbonyl of Ile79 (all residue numbers refer to the *E. coli* enzyme).[30] The binding enhancement of 3-position substitution could be due to van der Waals interactions between the substituent and the hydrophobic pocket. To examine this interaction, the distances between each fluorine atom of the

(30) Matthews, D. A.; Appelt, K.; Oatley, S. J.; Xuong, N. H. *J. Mol. Biol.* **1990**, *214*, 923−936.

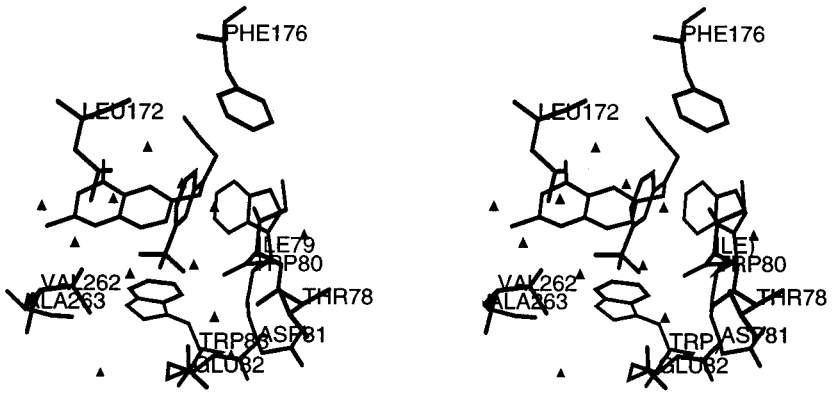



**Figure 6.** Stereoview of TS bound compound **f**. Only residues within 7 Å from the ligand are shown. Hydrogens are not shown for clarity. Black triangles are water molecules.

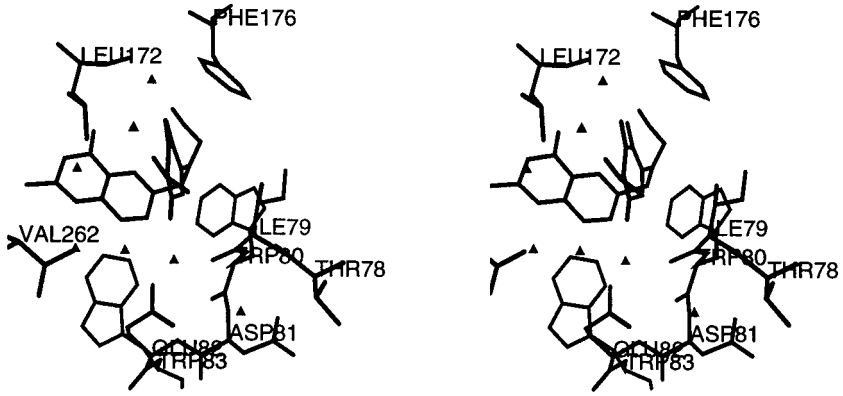



**Figure 7.** Stereoview of TS bound compound **z**. Only residues within 7 Å from the ligand are shown. Hydrogens are not shown for clarity. Black triangles are water molecules.

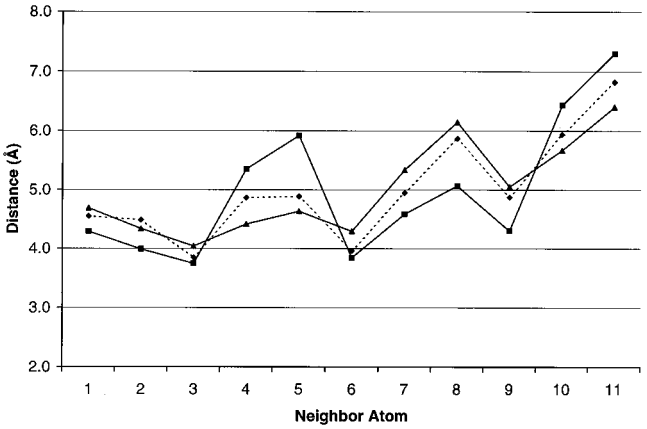



**Figure 8.** The average distances between neighbor atoms and the closest fluorine atom of the 3-trifluoro group. Three compounds were measured: compound **f** (diamond), compound **r** (square), and compound **y** (triangle). The neighbor atoms used are numbered as follows: 1, C$\gamma$ of Leu172; 2, C$\delta$1 of Leu172; 3, C$\delta$2 of Leu172; 4, C$\beta$ of Val262; 5, C$\gamma$1 of Val262; 6, C$\alpha$ of Ile79; 7, C$\beta$ of Ile79; 8, C$\gamma$1 of Ile79; 9, C$\gamma$2 of Ile79; 10, carbonyl C of Trp80; and 11, carbonyl O of Trp80.

3-trifluoromethyl substituent and those atoms forming the hydrophobic pocket were measured for each MD structure. Since the positions of fluorine atoms of the 3-trifluoromethyl substituent can be permuted by C−C bond rotation, we defined the distance between a neighbor atom and the 3-trifluoromethyl substituent as the distance between this neighbor atom and the closest fluorine atom. The neighbor atoms used in the measurement are C$\gamma$, C$\delta$1, and C$\delta$2 of Leu172, C$\beta$ and C$\gamma$1 of Val262, C$\alpha$, C$\beta$, C$\gamma$1, and C$\gamma$2 of Ile79, and the carbonyl group (C and O) of Trp80. The average values are shown in Figure 8. It is clear that C$\delta$2 of Leu172, and C$\alpha$ and C$\gamma$2 of Ile79, are the nearest three atoms to the 3-trifluoromethyl group. The changes of those average distances from compound **f** to compound **y** are less than 0.4 Å, suggesting that the 5,6-difluoro substitution does not affect the 3-trifluoromethyl substitution in a significant way. However, the 4-nitro substitution causes changes of average distances more than 0.5 Å for C$\gamma$1 of Val262, and C$\gamma$1 and C$\gamma$2 of Ile79 (from compound **f** to compound **l**). This result implies that the environment of the 3-trifluoromethyl substituent changes when the 4-nitro substitution is added. To quantitatively measure the change of the local environment of the 3-trifluoromethyl substituent due to different substitutions, we also calculated the root-mean-square deviation (RMSD) of the average distances in Figure 8 compared to compound **f**. The result is 0.29 Å for compound **y** and 0.50 Å for compound **l**.

For 5,6-difluoro substitution, we have performed a similar analysis. The two fluorine atoms of the 5,6-difluoro substituent are not interchangeable, hence the distances to neighbor atoms for each atom were measured. The neighbor atoms used in the measurement are C$\gamma$, C$\delta$1, and C$\delta$2 of Leu172, C$\alpha$, C$\beta$, C$\gamma$1, and C$\gamma$2 of Ile79, and C$\gamma$, C$\delta$1, C$\delta$2, C$\epsilon$1, C$\epsilon$2, and C$\zeta$ of Phe176. The average values are shown in Figure 9. Although the average distances differ from each other for compounds measured here (**x**, **y**, and **z**), the values of compound **x** have a pattern similar to the values of compound **z**. The positions of the two fluorine atoms relative to Phe176 are almost the same for compound **x** and compound **z**, while the positions of the fluorine atoms for compound **y** are quite different. This suggests that the environment of the 5,6-difluoro substituent changes when the 3-trifluoromethyl substitution is added, while the 4-nitro substitution causes less change. The RMSD of the

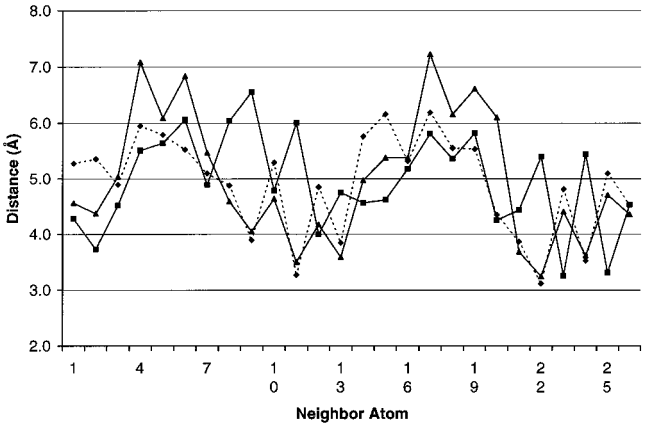



**Figure 9.** The average distances between 5- and 6-position fluorine atoms and neighbor atoms. Three compounds were measured: compound **x** (diamond), compound **y** (square), and compound **z** (triangle). The numbering of the *x* axis is as follows: Numbers 1−13 are for the distances between the 5-position fluorine atom and neighbor atoms; numbers 14−26 are for the distances between the 6-position fluorine atom and neighbor atoms. The neighbor atoms used are numbered as follows: 1, C$\gamma$ of Leu172, 2, C$\delta$1 of Leu172; 3, C$\delta$2 of Leu172; 4, C$\alpha$ of Ile79; 5, C$\beta$ of Ile79; 6, C$\gamma$1 of Ile79; and 7, C$\gamma$2 of Ile79; 8, C$\gamma$ of Phe176; 9, C$\delta$1 of Phe176; 10, C$\delta$2 of Phe176; 11, C$\epsilon$1 of Phe176; 12, C$\epsilon$2 of Phe176; and 13, C$\zeta$ of Phe176; 14, C$\gamma$ of Leu172; 15, C$\delta$1 of Leu172; 16, C$\delta$2 of Leu172; 17, C$\alpha$ of Ile79; 18, C$\beta$ of Ile79; 19, C$\gamma$1 of Ile79; and 20, C$\gamma$2 of Ile79; 21, C$\gamma$ of Phe176; 22, C$\delta$1 of Phe176; 23, C$\delta$2 of Phe176; 24, C$\epsilon$1 of Phe176; 25, C$\epsilon$2 of Phe176; and 26, C$\zeta$ of Phe176.

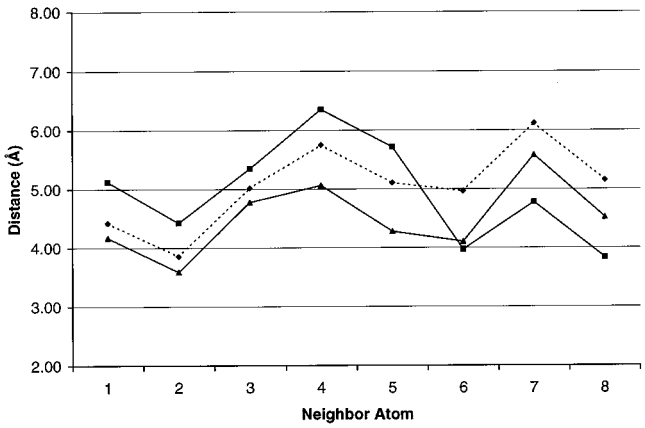



**Figure 10.** The average distances between neighbor atoms and the closest oxygen atom of the 4-nitro group. Three compounds were measured: compound **l** (diamond), compound **r** (square), and compound **z** (triangle). The neighbor atoms used are numbered as follows: 1, C$\gamma$ of Leu172; 2, C$\delta$1 of Leu172; 3, C$\delta$2 of Leu172; 4, carbonyl C of Leu172; 5, carbonyl O of Leu172; 6, C$\epsilon$1 of Phe176; 7, C$\epsilon$1 of Phe176; and 8, C$\zeta$ of Phe176.

average distances compared to compound **x** are 1.25 and 0.71 Å for compound **y** and compound **z**, respectively.

For 4-nitro substitution, there are only two residues, Leu172 and Phe176, within 5 Å from the closest oxygen of the nitro group. The neighbor atoms used in the measurement are C$\gamma$, C$\delta$1, C$\delta$2, carbonyl C and O of Leu172, and C$\epsilon$1, C$\epsilon$1, and C$\zeta$ of Phe176. The average values are shown in Figure 10. From those results, it is not very clear which substituent, 5,6-difluoro or 3-trifluoromethyl, will interfere with the environment of the 4-nitro substituent more. Hence the RMSD of the average distances compared to compound **l** were calculated. The result is 0.88 Å for compound **r** and 0.59 Å for compound **z**. It suggests that the environment of the 4-nitro substituent changes more when the 3-trifluoromethyl substitution is added.

We can make a short summary of the above analysis: 3-trifluoromethyl substitution is perturbed more by 4-nitro substitution than by 5,6-difluoro substitution; 5,6-difluoro substitution is perturbed more by 3-trifluoromethyl substitution than by 4-nitro substitution; and 4-nitro substitution is perturbed more by 3-trifluoromethyl substitution than by 5,6-difluoro substitution. This analysis qualitatively rationalizes the free energy calculation results we mentioned above: only the effects of combined 4-nitro substitution and 5,6-difluoro substitution are additive.

The error bars shown in Table 1 and Table 2 are also an indication of the relatively larger environment changes for mutations **l** to **r** and **x** to **y**. These two mutations need longer simulation time for converged results, probably because of slower equilibration due to larger environment changes.

Compared to experimental results, free energy calculations of four mutations, **a** to **f**, **f** to **r, a** to **l**, and **l** to **r**, show systematic errors of 0.5−1.2 kcal/mol. This suggests a systematic defect in the simulations. One approximation we made in the simulations was the position constraints for the counterions near Asp81 and Glu82. One can test this approximation by applying techniques to treat electrostatic interaction with no cutoff, e.g., the particle mesh Ewald method (PME),[31,32] using periodic boundary conditions, as well as performing a longer simulation. These kinds of simulations, however, are still very expensive for the sizes of systems we have studied in this work. We also have performed simulations without counterions for the mutations **a** to **f** and **f** to **r**, shown in Table 1. The error bars for those simulations are much larger, which is probably due to the fluctuation of electrostatic interactions from residues Asp81 and Glu82. It is true that, in some cases, counterions are not so critical to achieve lower error bars in the free energy calculations. This will be the case if there are no charged residues near the inhibitor, or the nearby charged residues are not solvent exposed.

Our PROFEC analysis and free energy calculation results for the mutations **a** to **x**, **f** to **y**, **l** to **z**, and **x** to **y** suggest that PROFEC is a useful way to identify regions around the ligand where substituents could be added. However, one must be aware that the PROFEC method may not be effective in suggesting multiple substitutions in a single step because of interference between substituents. A new PROFEC calculation should be carried out after each substitution to get more reliable suggestions.

## V. Conclusion

In conclusion, free energy calculation results and an analysis from MD structures suggest how substituents at different positions in an aromatic ring might interfere with the way other substituents interact with the protein, leading to nonadditive effects in binding enhancements. For the compounds we modeled in this paper, the interference between 4-nitro substitution and 5,6-difluoro substitution is the smallest. We predict compound **z**, which has these two substitutions, to be a good inhibitor of TS. Our results also support the use of PROFEC with free energy calculations as an efficient and accurate way to predict modifications in drug design. Using PROFEC, we suggested that the 5- and 6-position substituents are able to provide extra binding enhancement, as shown subsequently in our free energy calculations for compound **x**, **y**, and **z**.

(31) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(32) Essmann, U.; Perera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

Substitutions at these positions could be a new direction for TS inhibitors. Since these substitutions enhance binding but are likely to be less soluble, one should add additional substitutions at positions not buried in the binding site to increase water solubility. For example, the 4-position is exposed to the solvent and one might be able to add an ester group with a long polar tail to increase water solubility while achieving a binding enhancement similar to that found with a nitro group (Figure 7). Nonetheless, the improved potency of binding from compound **z**, which we predict to bind better than the strongest binding inhibitor in this series so far tested, compound **r**, could allow larger water solubility to be achieved while retaining high potency and thus have potential as a promising new anticancer TS inhibitor.

**Acknowledgment.** T.-S. L. is an NIH postdoctoral fellow under the training grant F32-GM19410-01. We are glad to acknowledge the use of the UCSF Computer Graphics Lab, supported by RR-1081, T. Ferrin, P.I. P.A.K. is pleased to acknowledge research support from the NIH (GM-29072). We thank Dr. David A. Matthews of Agouron Pharmaceuticals, Inc., San Diego, CA, for providing the X-ray crystal structure for compound **f**. The NSF is acknowledged for providing computational resources at the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign. T.-S. L. also thanks Dr. Lu Wang and Mr. Wei Wang for helpful discussions.

**Supporting Information Available:** A complete listing of force constants and equilibrium geometric parameters and the assignment of atom types for all inhibitors (PDF). This material is available via the Internet at http://pubs.acs.org.

JA9925554