

Optimization of Azoles as Anti-Human Immunodeficiency Virus Agents Guided by Free-Energy Calculations

Jacob G. Zeevaart,[†] Ligong Wang,^{†,‡} Vinay V. Thakur,[†] Cheryl S. Leung,[†]
Julian Tirado-Rives,[†] Christopher M. Bailey,[‡] Robert A. Domaoal,[‡]
Karen S. Anderson,[‡] and William L. Jorgensen^{*,†}

Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, and

Department of Pharmacology, Yale University School of Medicine,
New Haven, Connecticut 06520-8066

Received March 14, 2008; E-mail: william.jorgensen@yale.edu

Abstract: Efficient optimization of an inactive 2-aniliny-5-benzoxadiazole core has been guided by free energy perturbation (FEP) calculations to provide potent non-nucleoside inhibitors of human immunodeficiency virus (HIV) reverse transcriptase (NNRTIs). An FEP "chlorine scan" was performed to identify the most promising sites for substitution of aryl hydrogens. This yielded NNRTIs **8** and **10** with activities (EC_{50}) of 820 and 310 nM for protection of human T-cells from infection by wild-type HIV-1. FEP calculations for additional substituent modifications and change of the core heterocycle readily led to oxazoles **28** and **29**, which were confirmed as highly potent anti-HIV agents with activities in the 10–20 nM range. The designed compounds were also monitored for possession of desirable pharmacological properties by use of additional computational tools. Overall, the trends predicted by the FEP calculations were well borne out by the assay results. FEP-guided lead optimization is confirmed as a valuable tool for molecular design including drug discovery; chlorine scans are particularly attractive since they are both straightforward to perform and highly informative.

Introduction

In our pursuit of new anti-human immunodeficiency virus (HIV) agents, we have focused on non-nucleoside inhibitors of HIV reverse transcriptase (NNRTIs). Three drugs in this class arose in the 1990s (nevirapine, delavirdine, and efavirenz) and in January 2008 a new NNRTI, known as etravirine or TMC125, was approved.¹ NNRTIs are true inhibitors, which bind to an allosteric pocket in the vicinity of RT's polymerase active site.² A well-known problem with anti-HIV chemotherapy is the rapid mutation of the virus to yield drug-resistant strains. The occurrence of undesirable side effects is also often problematic and the onset may be immediate or arise after extended treatment. In either event, the availability of alternative therapeutic options is essential. Thus, continual anti-HIV drug discovery efforts are required to yield new drugs with alternative pharmacological characteristics as well as activity against the current and future spectrum of variants.¹

Simultaneously, we have sought improved computational methods of general utility to streamline the discovery of therapeutic agents that are both potent and have auspicious pharmacological properties. The basic goal is to minimize the number of compounds

that have to be synthesized and assayed to yield a drug candidate. Our approach features creation and evaluation of virtual libraries, estimation of pharmacological properties, and lead optimization guided by free-energy perturbation (FEP) calculations to assess relative protein–ligand binding affinities.³ Potent, structurally diverse anti-HIV agents have been discovered.³

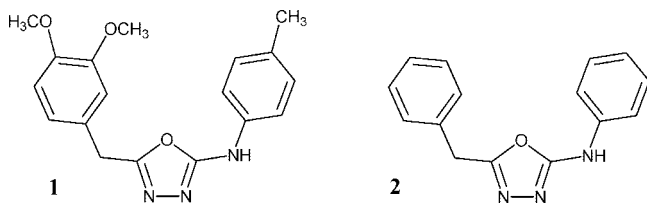
In a recent report, leads were sought by docking a library of commercially available compounds into the NNRTI binding site.⁴ Though known NNRTIs were retrieved well, purchase and assaying of representative, top-scoring compounds from the library failed to yield any active anti-HIV agents. Persisting, the highest-ranked library compound **1** was pursued computationally to seek constructive modifications. Specifically, the substituents were removed to yield the anilinybenzoxadiazole core, **2**. A set of small substituents was reintroduced in place of each hydrogen; scoring with the BOMB program and FEP results led to synthesis and assaying of several polychloro analogues with EC_{50} values as low as 310 nM in an HIV-

[†] Department of Chemistry.[‡] Department of Pharmacology.

- (1) For reviews, see (a) 2007 *AIDS Epidemic Update*; UNAIDS: Geneva, 2007. (b) Flexner, C. *Nat. Rev. Drug Discovery* **2007**, *6*, 959–966. (c) De Clerq, E. *Nat. Rev. Drug Discovery* **2007**, *6*, 1001–1018. (2) (a) Kohlstaedt, L. A.; Wang, J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A. *Science* **1992**, *256*, 1783–1790. (b) Smerdon, S. J.; Jäger, J.; Wang, J.; Kohlstaedt, L. A.; Chirino, A. J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3911–3915.

- (3) (a) Jorgensen, W. L.; Ruiz-Caro, J.; Tirado-Rives, J.; Basavapathruni, A.; Anderson, K. S.; Hamilton, A. D. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 663–667. (b) Ruiz-Caro, J.; Basavapathruni, A.; Kim, J. T.; Wang, L.; Bailey, C. M.; Anderson, K. S.; Hamilton, A. D.; Jorgensen, W. L. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 668–671. (c) Thakur, V. V.; Kim, J. T.; Hamilton, A. D.; Bailey, C. M.; Domaoal, R. A.; Wang, L.; Anderson, K. S.; Jorgensen, W. L. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5664–5667. (d) Kim, J. T.; Hamilton, A. D.; Bailey, C. M.; Domaoal, R. A.; Wang, L.; Anderson, K. S.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2006**, *128*, 15372–15373. (4) Barreiro, G.; Guimarães, C. R. W.; Tubert-Brohman, I.; Lyons, T. M.; Tirado-Rives, J.; Jorgensen, W. L. *J. Chem. Inf. Model.* **2007**, *47*, 2416–2428.

infected T-cell assay.⁵ The present report documents the FEP-guided developments that have now led to analogues of **2** with potencies in the 10–20 nM range. The essentially exhaustive FEP-guided lead optimization can serve as a model for future applications.



Computational Details

FEP calculations were carried out in the context of Monte Carlo (MC) statistical mechanics simulations to predict relative free energies of binding. The MC/FEP calculations are performed to interconvert two ligands unbound in water and bound to the protein; standard protocols were followed.^{3a} Briefly, initial structures were generated with the molecule growing program BOMB starting from the PDB file 1s9e;⁶ the ligand was removed and replaced by cores such as ammonia or **2** that are used by BOMB to grow the desired analogues in the binding site.⁵ A reduced model of the protein was utilized that consisted of the ca. 175 amino acid residues closest to the NNRTI binding site; a few remote side chains were neutralized so that there was no net charge for the protein. The MC/FEP calculations are executed with MCPRO,⁷ which also adds 1250 and 2000 water molecules in 25-Å caps for the complexes and unbound ligands, respectively. The energetics for the systems are described classically with the OPLS-AA force field for the protein, OPLS/CM1A for the ligands, and TIP4P for water molecules.⁸ For the MC simulations, all degrees of freedom were sampled for the ligand, while the TIP4P water molecules only translated and rotated, as usual; bond angles and dihedral angles for protein side chains were also sampled, while the backbone was kept fixed after conjugate-gradient relaxation.

The present study included some methodological testing of FEP protocols. The calculations used either double-wide (DW) sampling⁹ with 14 windows or simple overlap (OS) sampling with 11 windows.¹⁰ A window refers to a MC simulation at one point along the mutation coordinate λ , which interconverts two ligands as λ goes from 0 to 1; two free-energy changes are computed at each window, corresponding to a forward and backward increment. The spacing between windows, $\Delta\lambda$, is 0.1 except with 14 windows the spacing is 0.05 for $\lambda = 0-0.2$ and $0.8-1$, which addresses the fact that the free energy often changes most rapidly in these regions. Each window for the unbound ligands in water consisted of at least 10×10^6 (10M) configurations of MC equilibration followed by

30M configurations of averaging. For the bound calculations, the equilibration period was again at least 10M configurations followed by 20M configurations of averaging for the DW simulations and 10M for the OS alternative. In a recent comprehensive comparison for perturbations between substituted benzenes in water, the accuracy and precision from the 14-window DW and 11-window OS protocols were similar.^{10b} Thus, the intention here was to expand the testing to protein–ligand cases and to explore the possibility of halving the averaging period for the bound OS calculations (the most time-consuming step). If successful, the overall time for the FEP calculations would also be cut roughly in half, shortening the total time from about 2 weeks (for 14-DW) to 1 week (for 11-OS) to obtain a computed change in free energy of binding between two ligands by use of a 2.4 GHz Pentium processor. With commitment of 13 processors, one each for the 11 bound windows and 2 for the unbound calculations, a predicted change in binding affinity could then be obtained in 1 day.

All MC simulations were run at 298 K. The reported uncertainties ($\pm 1\sigma$) for the free energy changes were obtained from the fluctuation in separate averages over batches of 2M configurations.¹¹ Equation 1 is used, where m is the number of batches, θ_i is the average of property θ for the i th batch, and $\langle\theta\rangle$ is the overall average for θ .

$$\sigma^2 = \sum_i^m (\theta_i - \langle\theta\rangle)^2 / m(m-1) \quad (1)$$

Experimental Details

The synthesized compounds were primarily oxadiazole and oxazole derivatives. 1,3,4-Oxadiazole-2-amines were prepared either via cyclization of phenylacetic hydrazides and phenyl isocyanide dichlorides, as described previously,⁵ or via cyclization of the hydrazidecarboxamide by heating with POCl₃. Representative examples are provided in Schemes 1 and 2.

1,3,4-Oxadiazole-2,5-diamines were prepared in a similar manner via a hydrazine-1,2-dicarboxamide intermediate starting from the substituted phenylisocyanates (Scheme 3).¹²

Finally, oxazoles were synthesized as shown in Scheme 4 by minor modification of Froyen's procedure.¹³ Substituted 2-azidoacetophenones, which are readily obtained from corresponding arylacetic acids, were converted in situ to the iminophosphoranes. Following condensation with various arylisothiocyanates to yield β -keto carbodiimides, the desired oxazoles emerged in good yield. Additional synthetic and spectral details are provided in Supporting Information.

For the biology, the primary assay determined activities against the wild-type IIIB strain of HIV-1¹⁴ by use of MT-2 human T-cells¹⁵ at a multiplicity of infection (MOI) of 0.1; EC₅₀ values are obtained as the dose required to achieve 50% protection of the infected cells by the MTT colorimetric method. CC₅₀ for inhibition of MT-2 cell growth by 50% is obtained simultaneously.¹⁶ Analogous assays were performed with variant strains of the virus

(5) Barreiro, G.; Kim, J. T.; Guimarães, C. R. W.; Bailey, C. M.; Domaaal, R. A.; Wang, L.; Anderson, K. S.; Jorgensen, W. L. *J. Med. Chem.* **2007**, *50*, 5324–5329.

(6) Himmel, D. M.; Das, K.; Clark, A. D., Jr.; Hughes, S. H.; Benjahad, A.; Oumouch, S.; Guillemont, J.; Coupa, S.; Poncelet, A.; Csoka, I.; Meyer, C.; Andries, K.; Nguyen, C. H.; Grierson, D. S.; Arnold, E. *J. Med. Chem.* **2005**, *48*, 7582–7591.

(7) Jorgensen, W. L.; Tirado-Rives, J. *J. Comput. Chem.* **2005**, *26*, 1689–1700.

(8) (a) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236. (b) Jorgensen, W. L.; Tirado-Rives, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6665–6670. (c) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

(9) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050–3054.

(10) For recent reviews, see (a) Chipot, C.; Pohorille, A. In *Free Energy Calculations: Theory and Applications in Chemistry and Biology*; Chipot, C., Pohorille, A., Eds.; Springer Series in Chemical Physics, Vol. 86; Springer-Verlag: Berlin, 2007; pp 33–75. (b) Jorgensen, W. L.; Thomas, L. T. *J. Chem. Theory Comput.* **2008**, *4*, 869–876.

(11) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Clarendon: Oxford, 1987.

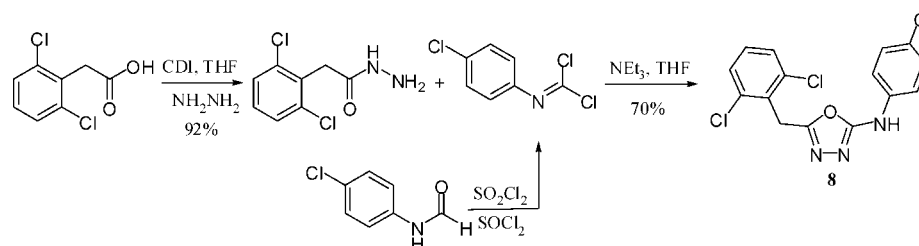
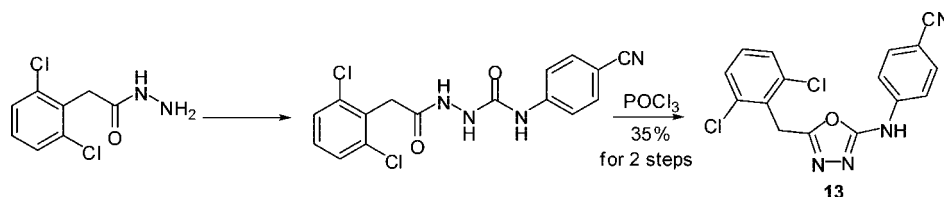
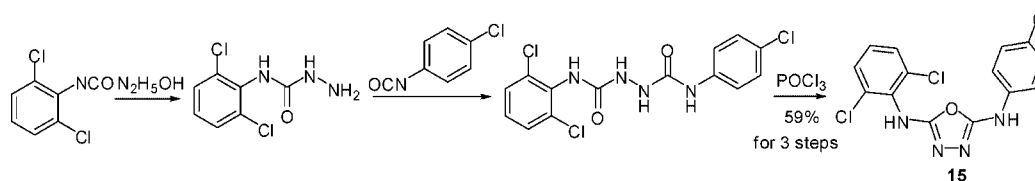
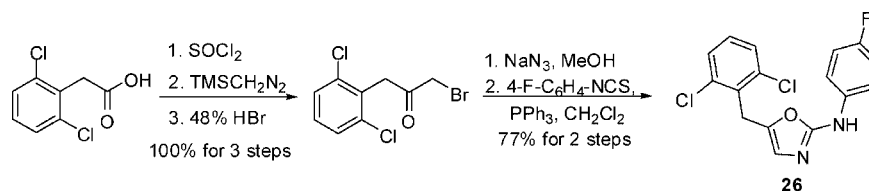
(12) Gehlen, H.; Moeckel, K. *Liebigs Ann. Chem.* **1965**, 685, 176–180.

(13) Froyen, P. *Phosphorus, Sulfur Silicon Relat. Elements* **1991**, *60*, 81–84.

(14) (a) Popovic, M.; Read-Connole, E.; Gallo, R. C. *Lancet* **1984**, *2*, 1472–1473. (b) Popovic, M.; Sarngadharan, M. G.; Read, E.; Gallo, R. C. *Science* **1984**, *224*, 497–500. (c) Ratner, L.; et al. *Nature* **1985**, *313*, 277–284.

(15) (a) Haertle, T.; Carrera, C. J.; Wasson, D. B.; Sowers, L. C.; Richmann, D. D.; Carson, D. A. *J. Biol. Chem.* **1988**, *263*, 5870–5875. (b) Harada, S.; Koyanagi, Y.; Yamamoto, N. *Science* **1985**, *229*, 563–566.

(16) (a) Lin, T. S.; Luo, M. Z.; Liu, M. C.; Pai, S. B.; Deutschman, G. E.; Cheng, Y. C. *Biochem. Pharmacol.* **1994**, *47*, 171–174. (b) Ray, A. S.; Yang, Z.; Chu, C. K.; Anderson, K. S. *Antimicrob. Agents Chemother.* **2002**, *46*, 887–891.

Scheme 1. Synthesis of 1,3,4-Oxadiazole-2-amines**Scheme 2.** Alternative Preparation of 1,3,4-Oxadiazole-2-amines**Scheme 3.** Synthesis of 1,3,4-Oxadiazole-2,5-diamines**Scheme 4.** Synthesis of 2,5-Oxazole Derivatives

that encode the Tyr181Cys (Y181C)¹⁷ and Lys103Asn/Y181C (K103N/Y181C)¹⁸ mutant forms of HIV-RT.

Results

Chlorine Scan. Given the false positive **1**, a chlorine scan was performed on the core **2** to seek auspicious sites for substitution.⁵ Thus, FEP calculations were performed for the interconversion of each hydrogen in the phenyl rings of **2** with a chlorine. The predicted structure of the core bound to RT is illustrated in Figure 1; it was built by use of BOMB and then subjected to conjugate gradient optimization by MCPRO. Confidence in the computed structure comes from past experience with complexes for other NNRTIs, including numerous crystal structures.^{2–6} Key features are the placement of the benzyl ring in the π -box, formed by Tyr181, Tyr188, Phe227, and Trp229, and the hydrogen bond between the oxygen of Lys101 and the amino group of **2**. The 10 complexes for the monochloro derivatives of **2** were also built with BOMB and each was converted to **2** in the FEP calculations. The FEP-DW results are listed in Table 1 with $\Delta\Delta G_b$ representing the predicted change in free energy of binding for the chlorine to

hydrogen conversion. A positive value indicates that chlorine is preferred over hydrogen. The structure in Table 1 also corresponds to the structure in Figure 1; in the FEP calculations C2 and C6 are not equivalent as they do not interconvert during the MC simulations.

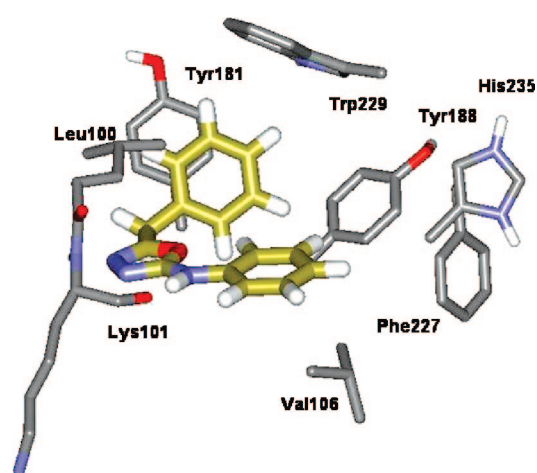
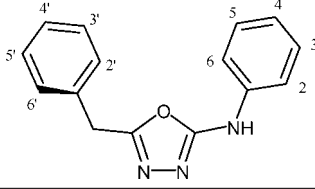


Figure 1. Some key residues in the NNRTI binding site with the core **2** bound, as created by BOMB and optimized by MCPRO. The full model contains ca. 175 protein residues and the ligand.

(17) Richman, D.; Shih, C.-K.; Lowy, I.; Rose, J.; Prodanovic, P.; Goff, S.; Griffin, J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 11241–11245.

(18) Nunberg, J. H.; Schleif, W. A.; Boots, E. J.; O'Brien, J. A.; Quintero, J. C.; Hoffman, A. M.; Emini, E. A.; Goldman, M. E. *J. Virol.* **1991**, *65*, 4887–4892.

Table 1. FEP-DW Results for Replacements of Chlorine by Hydrogen


Cl → H	ΔG_{bound} (kcal/mol)	ΔG_{un} (kcal/mol)	$\Delta\Delta G_b^a$ (kcal/mol)	σ^b
C2	2.71	0.51	2.21	0.10
C3	1.33	-2.00	3.33	0.11
C4	6.62	2.42	4.20	0.09
C5	-0.54	-1.71	1.17	0.10
C6	-6.86	-4.26	-2.60	0.12
C2'	4.86	1.43	3.43	0.08
C3'	-4.01	0.69	-4.70	0.11
C4'	1.01	1.81	-0.80	0.09
C5'	1.71	0.80	0.91	0.08
C6'	4.89	1.12	3.77	0.08

^a Computed change in free energy of binding from FEP-DW calculations. ^b Statistical uncertainties computed from eq 1.

The FEP-DW results indicate that the most promising places for chlorines are at C3, C4, C2', and C6'. The most desirable situation is when equivalent positions, for example, C2' and C6', are both favorable. If only one is favorable, then a penalty of $RT \ln 2$ (0.6 kcal/mol) can be expected for loss of one rotameric state upon binding. This is not an issue at a symmetrical position, such as C4 and C4'. Thus, the message from the FEP-DW results is that the placement of chlorines at C4, C2', and C6' should be particularly favorable and substitution at C3 (C5) would be the next most favorable.

The calculations were repeated with the FEP-OS methodology, as summarized in Table 2. The results agree well with those in Table 1, so this time-saving alternative is promising, and the message for the most favorable substitutions remains the same. The largest discrepancy occurs for C5, with the FEP-DW prediction being slightly favorable while the FEP-OS result is slightly unfavorable. With the FEP-OS calculations, the perturbations were also carried out in the gas phase. This allows computation of the change in free energy of hydration as $\Delta G_{\text{un}} - \Delta G_{\text{gas}}$ in Table 2. Each monochloro analogue is found to be more hydrophobic than **2** by 1–2 kcal/mol; in both the gas-phase and aqueous MC simulations, the average dipole moments are typically about 1 D less for the monochloro analogues than the parent. The accord for equivalent positions now reflects interconversion by rotation in the unbound aqueous and gas-phase simulations. It should be noted that the results in Tables 1 and 2 are far from obvious if one just displays the computed structures for the 10 monochloro analogues of **2** bound to RT. All of the structures appear reasonable with the possible exception of the C6 analogue, which seems to have the chlorine and benzyl ring crowded together. One might then execute conjugate gradient optimizations on the 10 complexes and the core **2**. With a dielectric constant of 2, the resultant protein–ligand interaction energies are -46 (C5), -47 (C3'), -49 (C6), -50 (C3), -51 (C5'), -52 (C2' and C4'), -54 (C2 and C6'), and -55 (C4) kcal/mol for the monochlorides and -50 kcal/mol for the unsubstituted core. The correlation of these results with those in Tables 1 and 2 is weak with the exception of the optimism for chlorination at C4 and C6'.

A final potential concern for design is addressed in Table 3. Perhaps the predictions would change as the core is progressively chlorinated. Table 3 records the results for substitution in the phenyl

ring for the 2',6'-dichloro analogue of **2**. In this case, adding a third chlorine at C4 should still be highly favorable and adding a chlorine at C3 (C5) also remains promising. A chlorine at C2 (C6) is not likely to be beneficial.

The related experimental data for the chlorine scan are provided by compounds **2–10** in Table 4. The results for **2–5**, **8**, and **10** were reported previously.⁵ Perfect correlation between the whole-cell assay results and the computed changes in free energies of binding is not expected. Direct correlation with the FEP results would require an unavailable binding assay yielding K_d values. However, good correlations have been observed consistently for NNRTI series in assays for inhibition of reverse transcription and in cell-based anti-HIV analyses.¹⁹ For lead optimization, the critical element is that the computations need to be qualitatively correct in providing guidance toward enhanced activity. The message from the FEP results was that it should be most beneficial to pursue chlorination at C4, C2', C6', and possibly C3.

The false positive **1**, the core **2**, and **3–5**, which do not have the 2',6'-dichloro substitution, are all inactive (Table 4). However, the results for the series **6–8** nicely conform to the pattern in Tables 1–3; the 2,2',6'-trichloride **6** is still inactive, while moving the 2-chloro substituent to C3 and C4 does yield active compounds with EC_{50} values of 4.3 and 0.82 μM . The accord is upheld for the tetrachlorides **9** and **10**, where again addition of chlorine at C3 is more beneficial than at C2. Thus, the recommendations from the FEP results in Tables 1–3 for optimal addition of three and four chlorines directly lead to NNRTIs **8** and **10** with activities of 820 and 310 nM. Compound **11** with a methyl group at C3 was also prepared and found to be less potent than the chlorine analogue **10**.

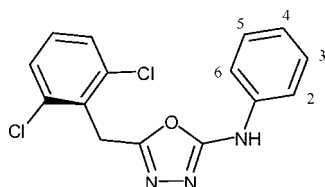
Why a chlorine scan? Addition of a single non-hydrogen atom is desirable since it minimizes potential steric conflicts. The other obvious possibilities are then a fluorine or methyl scan. The methyl scan is a good idea since the electronic character is different, electron-donating instead of electron-withdrawing. Consequently, there will be cases in which methyl is more activating than chlorine, and it is generally desirable to perform both FEP scans. If predicted free energies of binding for both series are obtained relative to the unsubstituted core, then ideal sites for introduction of both groups would be revealed. However, if only one scan is to be performed, chlorine has the technical advantage of being the most straightforward since the methyl to hydrogen perturbation requires the creation of three dummy atoms. Other considerations that can favor chlorine are that aryl chlorides are less prone to metabolism than methyl analogues and, synthetically, chlorine can provide a handle for coupling reactions. The problem with fluorine is that it is smaller; computed solvent-accessible surface areas with the OPLS-AA force field are 266, 275, 290, and 298 \AA^2 for benzene, fluorobenzene, chlorobenzene, and toluene. Thus, a chlorine and a methyl group have similar size, while fluorine is more like a hydrogen in size and its successful introduction does not guarantee that a chlorine or methyl group could be accommodated.

Optimization of the C4 Substituent. At this point, further refinements were probed by FEP calculations, starting with the substituent at C4. Analogous calculations had been carried out for the NNRTI series in Scheme 5 to optimize the substituent X.^{3a–c} The present case is reminiscent of the thiazole-containing example, though the presence of the dimethylallyloxy substituent at C5 in

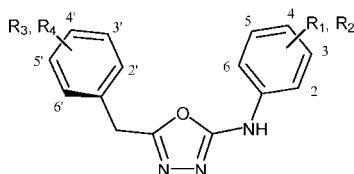
(19) Rizzo, R. C.; Udier-Blagovic, M.; Wang, D. P.; Watkins, E. K.; Kroeger Smith, M. B.; Smith, R. H., Jr.; Tirado-Rives, J.; Jorgensen, W. L. *J. Med. Chem.* **2002**, *45*, 2970–2987.

Table 2. Overlap Sampling Results for Replacements of Chlorine by Hydrogen for Core 2

Cl → H	ΔG_{bound} (kcal/mol)	ΔG_{un} (kcal/mol)	ΔG_{gas} (kcal/mol)	$\Delta \Delta G_{\text{aq}}^a$ (kcal/mol)	$\Delta \Delta G_b^b$ (kcal/mol)	σ^c
C2	3.37	0.62	2.57	−1.95	2.75	0.13
C3	1.06	−1.94	−0.57	−1.37	3.00	0.14
C4	6.47	2.20	3.53	−1.33	4.27	0.11
C5	−2.04	−1.47	−0.50	−0.97	−0.57	0.13
C6	−6.67	−4.26	−1.88	−2.38	−2.41	0.14
C2′	3.98	1.01	2.05	−1.04	2.97	0.11
C3′	−3.80	0.67	1.44	−0.77	−4.47	0.13
C4′	0.15	1.94	2.75	−0.81	−1.79	0.12
C5′	1.73	0.67	1.62	−0.95	1.06	0.10
C6′	5.08	0.88	1.71	−0.83	4.20	0.09

^a Computed change in free energy of hydration from FEP-OS calculations. ^b Computed change in free energy of binding from FEP-OS calculations.^c Statistical uncertainties computed from eq 1.**Table 3.** FEP-DW Results for Replacements of Chlorine by Hydrogen

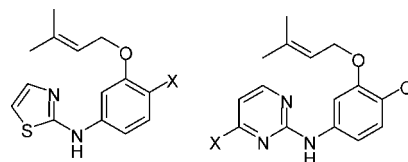
Cl → H	ΔG_{bound} (kcal/mol)	ΔG_{un} (kcal/mol)	$\Delta \Delta G_b^a$ (kcal/mol)	σ^b
C2	2.51	0.95	1.57	0.11
C3	1.42	−1.38	2.81	0.10
C4	6.88	2.38	4.50	0.08
C5	1.15	−1.23	2.38	0.13
C6	−8.77	−2.60	−6.17	0.13

^a Computed change in free energy of binding from FEP-DW calculations. ^b Statistical uncertainties computed from eq 1.**Table 4.** Oxadiazoles: Anti-HIV-1 Activity (EC_{50}) and Cytotoxicity (CC_{50})

compd	R ₁	R ₂	R ₃	R ₄	EC_{50}^a (μ M)	CC_{50}^b (μ M)
1	4-CH ₃	H	3'-OCH ₃	4'-OCH ₃	NA	61
2	H	H	H	H	NA	>100
3	4-Cl	H	H	H	NA	31
4	4-Cl	H	4'-Cl	H	NA	>100
5	4-Cl	H	5'-Cl	H	>10	32
6	2-Cl	H	2'-Cl	6'-Cl	NA	>100
7	3-Cl	H	2'-Cl	6'-Cl	4.3	71
8	4-Cl	H	2'-Cl	6'-Cl	0.82	20
9	4-Cl	2-Cl	2'-Cl	6'-Cl	2.4	45
10	4-Cl	3-Cl	2'-Cl	6'-Cl	0.31	>100
11	4-Cl	3-CH ₃	2'-Cl	6'-Cl	1.3	>100
12	4-CH ₂ OCH ₃	H	2'-Cl	6'-Cl	4.3	100
13	4-CN	H	2'-Cl	6'-Cl	0.13	40
14	4-CN	H	2'-F	6'-F	0.23	90
d4T					1.4	>100
nevirapine					0.11	>10
efavirenz					0.002	>0.1

^a For 50% protection in MT-2 cells; antiviral curves used triplicate samples at each concentration. NA for $EC_{50} > CC_{50}$. ^b For 50% inhibition of MT-2 cell growth; toxicity curves also used triplicate samples.

place of the benzyl group in the oxadiazole series could affect the preferences for X. The FEP results are recorded in Table 5. In

Scheme 5. Thiazole- and Pyrimidine-Containing NNRTIs³**Table 5.** FEP-DW Results for Optimization of the 4-Substituent X

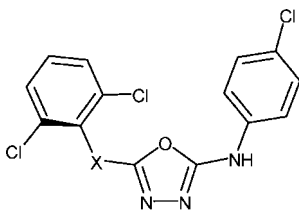
X →	Y	ΔG_{bound} (kcal/mol)	ΔG_{un} (kcal/mol)	$\Delta \Delta G_b^a$ (kcal/mol)	σ^b
CN	Cl	1.53	0.34	1.19	0.13
Cl	F	1.22	−0.54	1.75	0.05
F	H	5.56	3.27	2.29	0.05
Cl	H	6.99	2.23	4.76	0.06
CH ₃	H	5.49	3.88	1.61	0.12
CH ₃ CH ₂	CH ₃	−3.27	−1.93	−1.34	0.13
CH ₃ CH ₂	OCH ₃	−6.46	−4.91	−1.55	0.16
CF ₃	CH ₃	−10.01	−10.63	0.62	0.08
CH ₃ OCH ₂	CH ₃ CH ₂	0.79	−0.63	1.41	0.17

^a Computed change in free energy of binding from FEP-DW calculations. ^b Statistical uncertainties computed from eq 1.

short, the only substituent that is predicted to be preferred over chlorine is a cyano group, and the relative free energies of binding (in kilocalories per mole) are H (0.0) > Et (−0.3) > Me (−1.6) > methoxymethyl (−1.7) > OMe (−1.8) > CF₃ (−2.2) > F (−2.3) > Cl (−4.0) > CN (−5.2). As shown in Table 4, the methoxymethyl (MOM) and cyano derivatives **12** and **13** were prepared and bracket the chloro analogue **8** in activity, as predicted. The activity has progressed to 130 nM with **13** from 820 nM for **8**. For comparison, it is noted in Table 4 that the nucleoside d4T is a 1.4 μ M inhibitor in the MT-2 assay, while nevirapine and efavirenz are 110-nM and 2-nM NNRTIs.

The 2',6'-difluoro compound **14** was also prepared and is somewhat less potent at 230 nM than **13**. Our interest in the difluoro analogue arose from prediction of relatively poor solubility for **13** by QikProp.²⁰ The predicted intrinsic aqueous solubility increases from 10^{−5.8} M for **13** to 10^{−5.3} M for **14**. As noted previously, analysis of 1700 oral drugs with QikProp shows that 90% have predicted solubilities above 10^{−5.7} M.^{3c}

(20) Jorgensen, W. L. *QikProp*, v 3.0; Schrödinger LLC: New York, 2006.

Table 6. Summary of FEP-DW Results for Optimization of the Linker X


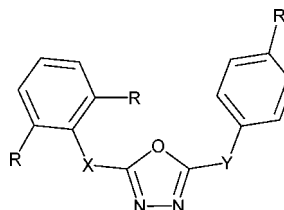
X	$\Delta\Delta G_b$ (kcal/mol)	σ
CH ₂	0	0
(R)-CHCH ₃	0.50	0.21
(S)-CHCH ₃	2.80	0.18
NH	1.20	0.21
NCH ₃	-1.60	0.29
O	1.37	0.16
S	-1.40	0.19

On the other hand, the predicted octanol/water partition coefficients, $QP \log P$, of 3.5 and 3.1 for **13** and **14** are both well below the 90% limit of 5.0.^{3c}

The preference for a cyano group at C4 was also found for the thiazole and pyrimidines series in Scheme 5, with typical reductions of a factor of 3–10 in EC₅₀ values over the corresponding chlorides.^{3a–c} As discussed previously,^{3b} other NNRTIs also feature a chlorine or cyano group in a similar position including efavirenz, TMC125, 8-chloro-TIBO, and UC781; however, the activity data are variable in these series with no consistent preference. For our series, we attribute the preference for cyano and chloro over the alternatives above to (a) their strengthening of the hydrogen bond with Lys101 by acidification of the hydrogen of the para-amino group and (b) constructive interaction of the C–Cl and C–CN dipoles with the electric field at the binding site. Concerning (a), MP2/6–311+G(d,p) results for hydrogen-bonded complexes of a water molecule with para-substituted phenols (*p*-HOPhY) show striking hydrogen-bond strengthening in the order Y = NO₂ > CN > CF₃ > Cl > F > H > OH > CH₃ > NH₂.²¹ For the inhibitors, the binding to the amide carbonyl of Lys101 is expected to be more affected by this substituent change than the binding to a water molecule in the unbound state. Concerning (b), His235 was included in Figure 1 to illustrate its proximity to C4 of the inhibitor. Upon conjugate-gradient minimization of the complex of RT with **13**, the distance between the cyano nitrogen and C_α of His235 is only 3.3 Å. Consequently, it is highly favorable to have a δ⁺–δ[−] dipole as provided by the 4-chloro or cyano substituent directed at His235.

Optimization of the Benzylic Linker. Attention next turned to the benzylic CH₂ group, designated X in the structure in Table 6. FEP calculations were performed starting from the trichloride **8** for X being CH₂, NH, O, S, NCH₃, and (R)- and (S)-CHCH₃. Multiple FEP cycles were considered and the final results for predicted free energies of binding relative to **8** are summarized in Table 6. In brief, no modification was predicted to be particularly constructive in comparison to the results from the previous substituent scans. Some related compounds were synthesized and assayed, as summarized in Table 7.

Conversion to the amino analogue was computed to be unfavorable by 1.2 kcal/mol, which was confirmed by the observed inactivity of **15** compared to **8** and weak activity of

Table 7. Benzylic Analogues: Anti-HIV-1 Activity (EC₅₀) and Cytotoxicity (CC₅₀)


compd	R	R'	X	Y	EC ₅₀ ^a (μM)	CC ₅₀ ^a (μM)
8	Cl	Cl	CH ₂	NH	0.82	20
15	Cl	Cl	NH	NH	> 100	> 100
16	Cl	Cl	CH ₂	CH ₂	18	43
13	Cl	CN	CH ₂	NH	0.13	40
17	Cl	CN	NH	NH	9.8	> 100
18	Cl	CN	NCH ₃	NH	0.27	> 100
19^b	Cl	CN	CHCH ₃	NH	1.20	54
20^b	F	CN	CHCH ₃	NH	0.13	8.5

^a See footnotes in Table 4. ^b Racemate.

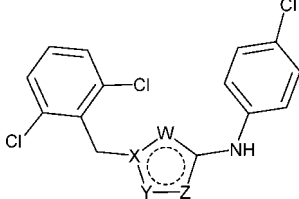
17 compared to **13**. Replacement of the pro-*R* hydrogen with a methyl group is predicted to be 2.3 kcal/mol more favorable than the pro-*S* substitution but still 0.5 kcal/mol less favorable than for the methylene version, **8**. Indeed, **19**, though racemic, loses a factor of 9 in potency compared to **13**, while curiously the difluoro analogue **20** is a little more potent than **14** (Table 4) and nearly 10-fold more potent than **19**. The FEP calculations were run only for the 2',6'-dichlorides, so there may be some differences in conformational details that are responsible for the varying effects of the addition of the benzylic methyl group.

The principal hope was for the methylamino analogue, which was predicted to be 2.1 kcal/mol better bound than the (R)-CHCH₃ isostere and 1.6 kcal/mol better bound than the methylene analogue. However, the one related compound that was prepared, **18**, turned out to be a factor of 2 less active than its methylene analogue, **13**. In view of this and the results in Table 6, enthusiasm could not be mustered for pursuing the oxa and thia (X = O and S) analogues.

As a curiosity, **16** was also prepared to obtain an estimate for the importance of the putative hydrogen bond between the amino group of **8** and Lys101; the answer is that there is a 22-fold loss in activity for deletion of the hydrogen bond. In summary, modifications were considered for the methylene linker between the oxadiazole ring and the 2,6-dihalophenyl group, but none provided significant gains in potency.

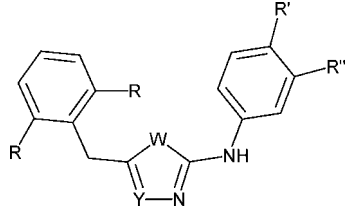
Optimization of the Heterocycle. The final target was the heterocycle. The oxadiazole arose from the docking study,⁴ but there was no reason to assume that it was the optimal five-membered ring heterocycle. Thus, MC/FEP calculations were performed to convert the thiophene analogue of **8** to 10 other common possibilities. The structures were all created by BOMB, and thiophene was taken as the reference because it has the maximum number of atoms; the FEP calculations are set up such that atoms are deleted, if appropriate, in progressing from λ = 0 to 1. The results are summarized in Table 8, and the order of predicted increasing activity is for the 1,3,4-thiadiazole, thiazole, imidazole, furan, isoxazole, thiophene, 1,2,3-triazole, isothiazole, pyrazole, 1,3,4-oxadiazole, and oxazole. Remarkably, only the oxazole analogue was predicted to be more potent than the oxadiazole. It is expected that the predicted range of free energies of binding, ΔΔG_b, in Table 8 is exaggerated. This may arise from the lack of sampling of the protein backbone in the MC simulations, which can be expected to disfavor the

(21) Jorgensen, W. L.; Jensen, K. P.; Alexandrova, A. A. *J. Chem. Theory Comput.* **2007**, 3, 1987–1992.

Table 8. FEP-DW Results for Optimization of the Heterocycle


W	X	Y	Z	ΔG_{bound} (kcal/mol)	ΔG_{un} (kcal/mol)	$\Delta\Delta G_{\text{b}}^a$ (kcal/mol)	σ^b
S	C	CH	CH	0	0	0	0
S	C	N	N	2.23	-2.68	4.91	0.47
S	C	CH	N	-7.18	-11.76	4.58	0.32
CH	N	CH	N	-19.18	-20.57	1.39	0.47
O	C	CH	CH	-7.55	-8.24	0.69	0.22
CH	C	O	N	-22.10	-22.52	0.42	0.25
CH	N	N	N	-12.21	-12.06	-0.15	0.43
CH	C	S	N	-16.80	-16.16	-0.64	0.52
CH	C	NH	N	-29.37	-28.08	-1.29	0.29
O	C	N	N	-24.92	-21.45	-3.47	0.29
O	C	CH	N	-29.65	-23.61	-6.04	0.22

^a Computed change in free energy of binding from FEP-DW calculations. ^b Statistical uncertainties computed from eq 1.

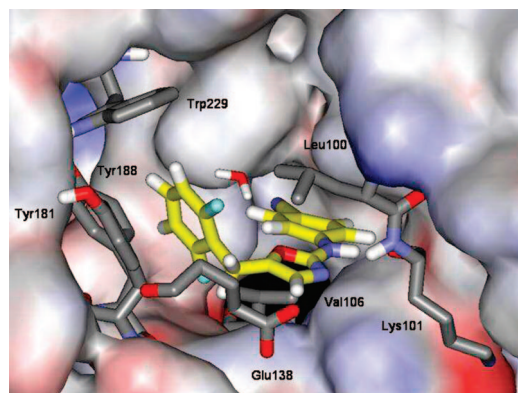
Table 9. Heterocycle Analogues: Anti-HIV-1 Activity (EC_{50}) and Cytotoxicity (CC_{50})


compd	R	R'	R''	W	Y	EC_{50}^a (μM)	CC_{50}^a (μM)
21	Cl	Cl	H	S	N	NA	>100
22	Cl	Cl	H	S	CH	3.1	29
23	Cl	Cl	Cl	S	CH	1.4	9.8
24	Cl	CN	H	S	CH	0.43	>100
25	Cl	Et	H	O	CH	2.8	>100
26	Cl	F	H	O	CH	0.43	12
27	Cl	Cl	H	O	CH	0.11	2.8
28	Cl	CN	H	O	CH	0.022	2.1
29	F	CN	H	O	CH	0.013	7.4

^a See footnotes in Table 4. NA for $\text{EC}_{50} > \text{CC}_{50}$.

larger, sulfur-containing heterocycles. The changes in heterocycles are also accompanied by significant changes in the electrostatic interactions with the ligands; this would benefit from inclusion of polarization effects, which are not explicitly represented by fixed charged force fields such as OPLS.

For some testing of the predictions, the thiadiazole **21** was prepared and indeed found to be inactive (Table 9). The thiazole **22** was also predicted to be inactive. It too was prepared and, though less active than **8**, it did yield an EC_{50} value of 3.1 μM and the trichloro version **23** came in at 1.4 μM . However, the synthetic focus naturally turned to oxazoles, and it was gratifying that **27** emerged as a 110-nM NNRTI, showing a nice boost over its oxadiazole counterpart **8** at 820 nM. Modification of the 4-chloro substituent of **27** to cyano in **28** then provided the expected additional increment to 22 nM. And, finally, the 2,6-difluorobenzoxazole **29** was prepared and found to be a very potent NNRTI at 13 nM. The predicted properties for **29** from QikProp are also all in the acceptable ranges for a potential oral drug;^{3c} the molecular weight is 311, and the predicted aqueous solubility, log *P*, and Caco-2 cell permeability are

**Figure 2.** Image constructed from the last configuration of an MC/FEP calculation for **29** bound to HIV-RT. Carbon atoms of **29** are in yellow. Fragments of some residues in the binding site are shown along with one of the 1250 water molecules that were included.

$10^{-5.6}$ M, 3.6, and 1054 nm/s. The compound has two predicted primary metabolites arising from oxidation at the benzylic methylene group and epoxidation of the oxazole.

An illustration of the computed structure for **29** bound to RT is provided in Figure 2. The preference for the oxazole over the oxadiazole can be attributed to improved protein–ligand electrostatic interactions. The oxazole nitrogen is more basic than N3 in the oxadiazole, which should improve the interaction with the backbone NH of Lys101. In addition, the proximity of N4 of the oxadiazole and the side-chain carboxylate of Glu138 is expected to be unfavorable. It is also noted that, in the MC simulations, one or two water molecules are often present in the binding site situated in the region between the two phenyl rings of the inhibitors. One is seen in the snapshot in Figure 2; the water molecules are typically forming π -type hydrogen bonds with aryl rings of the inhibitor and protein, and when more than one is present, they are also hydrogen-bonded to each other. On the other hand, a water molecule is not found hydrogen-bonded to the nitrogen of any C4-cyano groups in the bound inhibitors; the key stabilizing interaction in this case is the charge–dipole interaction with His235 (Figure 1).

Returning to Table 9, it may be noted that the difluoro (**29**) over dichloro (**28**) preference here repeats the pattern for **19** and **20** but not for **13** and **14**, again suggesting some conformational subtleties. In addition, to confirm further the structure–activity predictions for the substituent at C4, the ethyl and fluoro derivatives **25** and **26** were also synthesized. The observed activity order of $\text{CN} > \text{Cl} > \text{F} > \text{Et}$ for **25**–**28** exactly mirrors the predictions in Table 5. As a last thrust, FEP calculations were performed for possible replacement of the oxazole C4 hydrogen of **28** or **29** by F, Et, Me, CF_3 , and CH_2OH . The calculations were performed for **29**, and the five analogues were predicted to be less well bound than **29** by 0.8, 1.5, 1.8, 2.2, and 3.9 kcal/mol, respectively. This is another example where expectations based on visual inspection of modeled structures, as in Figure 2, are ambiguous. The qualitative FEP result was confirmed experimentally for the C4-methyl derivative of **28**, which was found to be 7-fold less potent than **28**. Another option that had been considered was replacement of the aminophenyl ring in **8** with aminopyridinyl. The phenyl to 2-pyridinyl perturbation was unfavorable by 0.9 kcal/mol, while 4-pyridinyl and 3-pyridinyl were favorable, but by only 0.4 and 0.9 kcal/

Table 10. Activity (EC_{50}) and Cytotoxicity (CC_{50}) for Wild-Type and Variant Strains of HIV-1^a

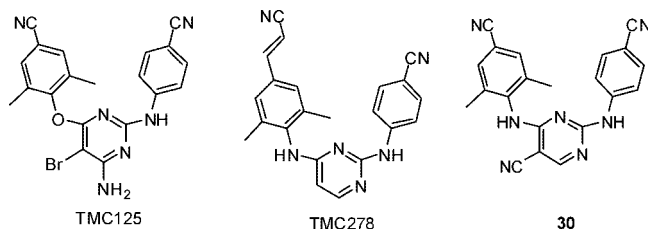
compd	WT		Y181C		K103N/Y181C	
	EC_{50} (μ M)	CC_{50} (μ M)	EC_{50} (μ M)	CC_{50} (μ M)	EC_{50} (μ M)	CC_{50} (μ M)
13	0.13	40	21	72	24	>100
27	0.11	2.8	NA	6	NA	15
28	0.022	2.1	NA	4	NA	5
29	0.013	7.4	NA	8	ND	ND
d4T	1.4	>100	0.68	>100	0.28	>100
efavirenz	0.002	>0.1	0.013	>0.1	0.050	>0.1
TMC125	0.001	>0.1	0.014	ND	0.005	ND

^a See footnotes in Table 4. NA for $EC_{50} > CC_{50}$; ND = not determined.

mol before correction for the RT ln 2 rotamer loss. So, these possibilities were also abandoned.

Viral Mutations

In pursuing optimization for the core **2**, potential performance against mutant strains of HIV-1 was considered. Resistance typically arises from protein mutations that (a) introduce unfavorable interactions with the inhibitor, (b) reduce favorable interactions with the inhibitor, or (c) stabilize the unliganded (apo) form of the protein. In response to (a) and (b), it is viewed as desirable to keep the inhibitors relatively small and to incorporate several rotatable bonds. This is also expressed in the wiggle and jiggle idea that has been claimed as responsible for the excellent resistance profiles for the NNRTIs TMC125 (etravirine) and TMC278 (rilpivirine).²³ These molecules and the progeny of **2** all feature four rotatable bonds connecting the three rings.

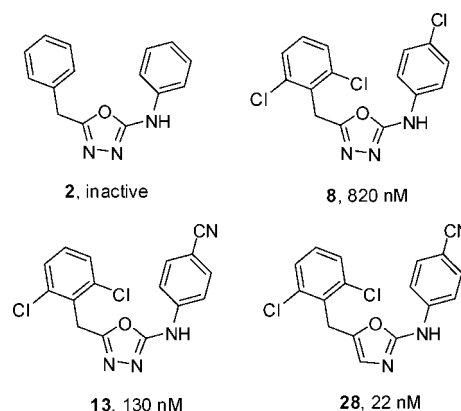


The Y181C and K103N/Y181C mutants have been particularly problematic for NNRTI therapy,¹ so several of the more potent NNRTIs from the present work were tested against corresponding strains of HIV in the infected T-cell assay. The results are summarized in Table 10 and include reference data for d4T, efavirenz, and TMC125. The findings for efavirenz and TMC125 are consistent with prior reports using similar assays.²⁴ Here, efavirenz loses factors of 6.5 and 25 in potency to the Y181C and K103N/Y181C variants and has a 2-nM EC_{50} for wild-type HIV-1, while 2- and 40-fold losses and a 1-nM EC_{50} for the wild type were previously reported.²⁴ For TMC125, the prior EC_{50} values are 1.4, 7, and 4.3 nM for WT, Y181C, and the double mutant, while our results are 1, 14, and 5 nM. However, in spite of the similar expected flexibility for TMC125 and the oxadiazole and

oxazoles in Table 10, the new compounds exhibit no significant activity against the two mutant strains of the virus. In fact, when one looks closely at the data for TMC125 analogues, one sees that the results for the mutant panel vary substantially for small substituent changes in compounds with no change to the core structure or number of rotatable bonds.²⁴ For example, compound **30** is a 0.5 nM NNRTI toward wild-type HIV-1 but loses 150-fold to the Y181C variant.²⁴ Thus, some flexibility is undoubtedly needed for resilience, but there are additional subtleties that affect the resistance profiles. Further analyses and design efforts are in progress on these issues.

Conclusion

MC/FEP calculations have been used to guide efficient optimization of the inactive core **2** to yield **28** and **29** as 10–20 nM inhibitors of HIV reverse transcriptase. Key milestones were provided by the chlorine scan, which proposed **8**, optimization of the C4 substituent to yield **13**, and prediction of the preference for the oxazole **28**. Additional modifications for the benzylic linker, replacement of the aminophenyl ring, and C4-substitution of the oxazole were also considered and were deemed less promising. The study illustrates the benefits for inhibitor design of closely coupled computer modeling, synthetic organic chemistry, and biological assaying. Future work will focus on application of the technology to other cores and simultaneous optimization for wild-type and mutant viral targets.



Acknowledgment. Gratitude is expressed to the National Institutes of Health (AI44616, GM32136, GM49551) for support. Receipt of the following reagents through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH, is also greatly appreciated: MT-2 cells, catalog no. 237, and nevirapine-resistant HIV-1 (N119), catalog no. 1392, from Dr. Douglas Richman; HTLV-III_B/H9, no. 398, from Dr. Robert Gallo; and HIV-1_{IIIB} (A17 variant) from Dr. Emilio Emini.

Supporting Information Available: Synthetic details and NMR and HRMS spectral data for compounds in Tables 4, 6, and 8 except **1–5**, **8**, **10**, and **11**, for which the data were previously reported;⁵ also full citations for refs 14c, 23a, and 24. These materials are available free of charge via the Internet at <http://pubs.acs.org>.

JA8019214

- (22) Wang, D.-P.; Rizzo, R. C.; Tirado-Rives, J.; Jorgensen, W. L. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2799–2802.
 (23) (a) Das, K.; et al. *J. Med. Chem.* **2004**, *47*, 2550–2560. (b) Das, K.; Bauman, J. D.; Clark, A. D., Jr.; Frenkel, Y. V.; Lewi, P. J.; Shatkin, A. J.; Hughes, S. H.; Arnold, E. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 1466–1471.
 (24) Ludovici, D. W.; et al. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2235–2239.