Available online at www.sciencedirect.com
SCIENCE DIRECT
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 16 (2006) 663–667

Computer-aided design of non-nucleoside inhibitors of HIV-1 reverse transcriptase

William L. Jorgensen,[a,*] Juliana Ruiz-Caro,[a] Julian Tirado-Rives,[a] Aravind Basavapathruni,[b] Karen S. Anderson[b] and Andrew D. Hamilton[a]

[a]*Department of Chemistry, Yale University, New Haven, CT 06520-8107, USA*
[b]*Department of Pharmacology, Yale University School of Medicine, New Haven, CT 06520-8066, USA*

Received 23 September 2005; accepted 11 October 2005
Available online 2 November 2005

Abstract—Design principles are delineated for non-nucleoside inhibitors for HIV-1 reverse transcriptase (NNRTIs). Simultaneous optimization of binding affinity for wild-type RT, tolerance for viral mutations, and physical properties is pursued. Automated lead generation with the growing program *BOMB*, Monte Carlo simulations with free-energy perturbation theory for lead optimization, and property analysis with *QikProp* are featured. An initial 30 μM lead has been optimized rapidly to the 10 nM level.

The AIDS crisis continues with ca. 40 million people infected by HIV-1, and 3 million deaths in 2004.[1] The therapeutic situation is challenged by rapid mutation of the virus to yield resistant strains. Anti-HIV drugs are normally given in combination, typically two nucleoside inhibitors of HIV-1 reverse transcriptase (NRTIs) and either a HIV protease inhibitor (PI) or a non-nucleoside RT inhibitor (NNRTI).[2] Long-term solutions include a vaccine; however, for those already infected, the current drugs do not systemically eliminate the virus, so they must be taken chronically. Many of the drugs require high dosages, which lead to compliance difficulties and costs that are only manageable in affluent nations. Patients often migrate among combinations, and long-term use of NRTIs and PIs can yield morphologic and metabolic complications including wasting and lipodystrophy.[2] The needs for new drugs and points of attack on HIV are profound. New drugs are expected to show enhanced activity against at least parts of the current spectrum of mutants.

Our own efforts have been aimed at design of new NNRTIs. Three compounds in this class, nevirapine, delavirdine, and efavirenz, were approved by the FDA during 1996–1998. The first two succumb to many single point mutations in RT, and delavirdine is further debilitated by a high pill burden (400 mg tid). Efavirenz has an improved resistance profile and dosing (600 mg qd), though resistance still arises from common mutations including K103N and Y188L, and neurological side effects including dizziness and nightmares are frequent. Several NNRTIs with further improved resistance profiles such as DPC083, HBY097, and S-1153 (capravirine) stalled in clinical trials, while TMC125 (etravirine) continues on in spite of high dosage (900 mg bid).[3]

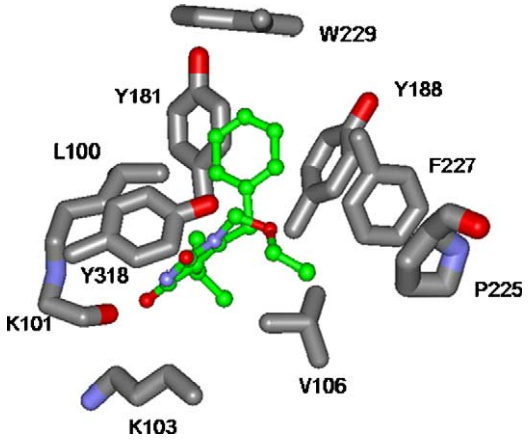


Figure 1. Key residues forming the NNRTI binding site from the 1rt1 X-ray structure for the emivirine–RT complex (Ref. 6). Y318 is in front.

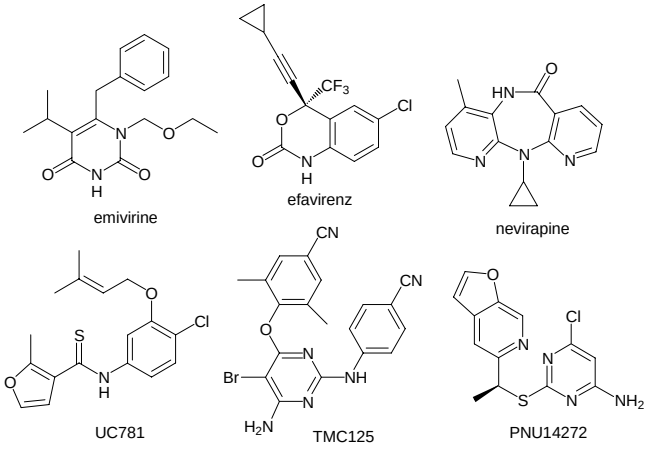
Keywords: NNRTI; Structure-based drug design; Anti-HIV drugs; Computer-aided drug design.
* Corresponding author. Tel.: +1 203 432 6278; fax: +1 203 432 6299; e-mail: william.jorgensen@yale.edu

0960-894X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2005.10.038

Thus, there is much room for improvement, and we have pursued computer-aided design of new NNRTIs with simultaneous goals of enhanced performance against common RT mutants, high bioavailability, and facile synthesis. The approach features lead generation with the growing program *BOMB*, property predictions with *QikProp*, and assistance in lead optimization with Monte Carlo simulations using free-energy perturbation theory (MC/FEP).

Design considerations

WT structure. Upon binding in an allosteric pocket, NNRTIs cause a change in the rate-limiting step in DNA polymerization catalyzed by HIV-1 RT.[4] The binding site has been well characterized by X-ray crystallography.[5] As illustrated in Figure 1 for emivirine,[6] key features for NNRTI design have emphasized placement of an unsaturated group in the arene pocket formed by Y181, Y188, W229, and F227, and a hydrogen-bond donor to interact with the K101 oxygen atom. Many options have been successfully explored for the unsaturated group including benzyl (emivirine and loviride), phenoxyl (TMC125), thiophenoxyl (capravirine), ethynyl- and vinylcyclopropyl (efavirenz, DPC961, and DPC083), and dimethylallyl (TIBO and UC781) derivatives. For the hydrogen-bond donor, amide and thioamide analogs are the rule; notable exceptions are nevirapine and capravirine, which have a water molecule bridging between an N-acceptor site and the K101 oxygen. Relative to the NH donor, the unsaturated group can be viewed as being positioned straight back (emivirine, efavirenz, and TIBOs) or approaching from the left toward Y181 (trovirdine, PNU14272, and TMC125) or from the right (UC10 and UC781).



Resistance. Common single mutations of RT that confer resistance to NNRTIs include L100I, K101E, K103N, V106A, Y181C, Y188C/L/H, G190A/S, P225H, and F227C/L. Not surprisingly, the affected residues line the NNRTI binding site (Fig. 1). Important lessons from NNRTI research to date are that some internal flexibility and small size are desirable to allow the NNRTI to adjust to the mutations; a perfect fit for WT RT with a rigid inhibitor is not desirable.[7,8] Relatively rigid, first-generation NNRTIs such as nevirapine, delavirdine, and TIBO are rendered ineffective by most of the above mutations, especially Y181C, Y188L, and K103N. A principal benefit of the second-generation efavirenz is that it holds up well to Y181C.[3,9] Though efavirenz is rigid, the comparatively small ethynylcyclopropyl group makes limited contact with Y181; it is directed toward Y188, and Y188L causes a >100-fold loss of activity.[8,9] Rigidity around the hydrogen-bond donor NH group is also undesirable for NNRTIs including efavirenz, TIBO, and emivirine making them vulnerable to the L100I mutation.[7] Much progress has been made with more nimble inhibitors such as capravirine,[3] UC781,[10] TMC125,[8] and rilpivirine.[11] The latter three feature adjustable dihedral angles for the unsaturated group and at the hydrogen-bond donor site. Activity losses below a factor of 10 are obtained for most single mutations except for Y188L with UC781, TMC125 performs better still, and rilpivirine has below 10 nM EC_{50}s against WT HIV-1 and 23 single and double mutants including K103N with Y181C or L100I.[11]

Properties. The NNRTI binding site is hydrophobic, and NNRTI potency generally increases with hydrophobicity. This has led to bioavailability issues often associated with poor aqueous solubility. Attempts at oral formulation of UC781 and TMC120 were abandoned and use as topical microbicides is being pursued. In general, there is little correlation between oral dosage and potency for NNRTIs, for example, nevirapine is far less potent than TMC125, but it is given in much lower dosage (Table 1). When predicted solubility falls below $10^{-5.5}$ M, it is likely that oral formulation will be challenging.[12]

Computer-aided design

Lead generation. The The *BOMB* program was used to construct thousands of NNRTI leads. *BOMB* builds ligands into a binding site by adding user-selected substituents to a core.[13] Up to four substituents are added to the core and/or linked together. The *BOMB* libraries

Table 1. Dosage, potency, and solubility for selected NNRTIs

NNRTI	Dosage (mg)[a]	EC_{50} (nM)[b]	QP log S[c]
DPC083	100–200	1	−5.4
Nevirapine	400	90	−3.6
Efavirenz	600	1	−5.0
Emivirine	1000	4	−3.8
Calanolide A	1200	53	−5.4
Delavirdine	1200	16	−5.6
UC781		9	−5.8
TMC120		0.9	−6.3
TMC125	1800	1.5	−6.5
Rilpivirine		0.5	−6.4

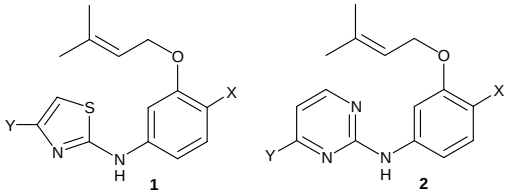
[a] Standard daily approved dosages or in phase II/III clinical trials.
[b] Typical values for protection of MT or CEM cells from HIV-1 infection.
[c] Predicted aqueous solubility from *QikProp* v 2.3; S in mol/l.

contain more than 100 cores and 600 substituents, which are common fragments in drugs. An extensive conformational search is performed for each ligand, each conformer is optimally positioned, and the lowest-energy one is scored and output. The structure optimization is performed with the OPLS-AA force field, and the scoring functions have been trained on SAR data for >300 protein–ligand complexes. Prime considerations are the protein–ligand potential energy (E_{PL}), and desolvation and surface-area burial for the ligand. The output from *BOMB* includes a spreadsheet with a row for each ligand containing its predicted activities, protein–ligand energetic and structural results, and *QikProp* results for predicted properties including solubility and cell permeabilities.[12,14]

Given the above design considerations, focused virtual libraries were generated in two motifs, U–Het–NH–Ph and Het–NH–Ph–U, where U is an unsaturated, hydrophobic group and Het is an aromatic heterocycle. Synthesis was envisaged via metal coupling or aromatic substitution chemistry. The former motif has been more commonly pursued, for example, TMC125, so initial attention was directed at the latter option. The ligands were built with NH_3 as the core, positioned to hydrogen bond with the K101 carbonyl group. The nearest 159 RT residues from the 1s9e crystal structure were included.[8] Het covered 61 five- and six-membered heterocycles, and 47 alternatives for U were placed *meta* to the amino group; *ortho* and *para* have synthetic and steric liabilities. Conjugate-gradient optimizations were carried out for all complexes with well-scoring candidates. Consensus was sought for low score ($\propto$ log activity), low E_{PL}, and acceptable solubility ($\log S > -5$). Some resultant top choices for Het and U are listed in Table 2.

Lead optimization. The 100 lead compounds suggested in Table 2 all warrant pursuit. High potency was also expected to require addition of one or two small substituents. In order to narrow the possibilities for the heterocycle and a *para* substituent in the phenyl ring, MC/FEP simulations were carried out. For these studies U was taken to be ODMA in view of its rank in Table 2, precedent with UC781, and envisioned synthesis via allylation. The FEP calculations yield relative free energies of binding; standard protocols were followed.[7,9] Briefly, the structures from *BOMB* were used, and 1250 and 2000 TIP4P water molecules were added in 25 Å caps for the complexes and unbound ligands, respectively. Each FEP utilized 20 free-energy increments and typically covered 20×10^6 (20 M) configurations for equilibration and 40 M configurations for averaging.

For the heterocycle, the FEPs converted Het = phenyl to 2-, 3-, and 4-pyridinyl, 2-pyrimidinyl, and pyrazinyl, and Het = 3-pyrrolyl to 2-, and 3-furanyl, 2-oxazolyl, 2-thiazolyl, and 4-1,2,3-triazolyl. The last choice has the best E_{PL}, but a score of −4.8 owing to a substantial desolvation penalty. The results are listed in Table 3 along with those for the nine options that were considered for the *para* substituent, X. The reported uncertainties ($\pm 1\sigma$) were obtained by separate averages over batches of 1 M configurations; from round-robin cycles, the true σ's are closer to ca. 1.0 and 0.3 kcal/mol for the Het and X results. Thus, it is anticipated that the most potent NNRTIs from these options will have Het = 2-pyrimidinyl, 2-thiazolyl, 4-1,2,3-triazolyl or 2-oxazoyl, and X = CN or Cl. Our initial synthetic efforts were then directed toward the thiazole and pyrimidine series, **1** and **2**. Oxazoles are being pursued, though their synthesis is more challenging than for the thiazoles.



Activity results

The accompanying paper details the synthesis and assaying against WT HIV-1 for ca. 50 compounds.[15] These efforts have confirmed the predicted trends in Table 3. Other options for U and some substituents Y for Het have also been explored. A subset of results focusing on **1** and **2** is summarized in Table 4. Compound **1a** and the corresponding ethyl and phenyl ethers were

Table 2. Top Het and U choices for Het–NH–*m*-Ph–U leads

Het[a]	Scr[a]	U[b]	Scr[b]	QP log *S*[c]
3-Pyrazolyl	−6.3	$CH_2COOCH(CH_3)_2$	−5.2	−3.2
3-Pyridin-2-onyl	−7.2	2-Thienylmethoxy	−6.0	−3.6
3-Pyrrolyl	−7.2	*N*-Me-3-pyrrolylmethyl	−6.0	−3.6
2-Oxazolyl	−5.8	2-Me-*c*-pent-1-enylmethoxy	−6.7	−2.8
4-Oxazolyl	−6.0	*O*-3,3-Dimethylallyl (ODMA)	−6.0	−3.2
2-Pyrazinyl	−5.1	2-Furanylmethoxy	−5.8	−3.5
2-Pyrimidinyl	−5.0	2-Me-*c*-pent-1-enylmethyl	−6.0	−3.7
2-Thiazolyl	−5.8	3-Furanylmethoxy	−5.9	−3.5
2-Pyridinyl	−5.7	3-Pyrrolylmethyl	−5.6	−4.1
2-Furanyl	−6.9	Phenoxy	−5.7	−3.8

[a] For Het–NH–*m*-PhODMA, Het ordered by E_{PL}; *BOMB* score in column 2.
[b] For 2-thiazolyl-NH-*m*-PhU, U ordered by E_{PL}; *BOMB* score in column 4.
[c] *QikProp* predicted aqueous solubility for Het–NH–*m*-PhODMA.

Table 3. FEP results for Het–NH–*p*-X-PhODMA optimization

Het	$\Delta\Delta G$[a]	X	$\Delta\Delta G$[b]
Phenyl	0.0	Pr	8.9 ± 0.3
3-Pyridinyl	−0.8 ± 0.2	*i*-Pr	2.6 ± 0.2
4-Pyridinyl	−3.3 ± 0.2	OMe	0.7 ± 0.2
Pyrazinyl	−4.2 ± 0.4	H	0.0
2-Pyridinyl	−4.8 ± 0.3	Et	−1.2 ± 0.2
2-Pyrimidinyl	−11.4 ± 0.5	F	−2.4 ± 0.0
3-Pyrrolyl	0.0	CH_3	−2.8 ± 0.1
3-Furanyl	−3.6 ± 0.3	CN	−4.2 ± 0.1
2-Furanyl	−6.7 ± 0.2	Cl	−4.4 ± 0.1
2-Thiazolyl	−9.6 ± 0.4		
4-1,2,3-(1*H*)Triazolyl	−10.8 ± 0.4		
2-Oxazolyl	−13.1 ± 0.3		

[a] $\Delta\Delta G$ in kcal/mol relative to phenyl or 3-pyrrolyl.
[b] $\Delta\Delta G$ in kcal/mol relative to H for Het = 2-thiazolyl.

Table 4. Anti-HIV-1 activity (EC_{50}) and cytotoxicity (CC_{50}), μM[a]

Compound	Het	X	Y	EC_{50}	CC_{50}
1a	2-Thiazolyl	H	H	10	23
1b	2-Thiazolyl	Pr	H	23	23
1c	2-Thiazolyl	*i*-Pr	H	NA	31
1d	2-Thiazolyl	Et	H	5.0	31
1e	2-Thiazolyl	Me	H	3.0	17
1f	2-Thiazolyl	OMe	H	3.8	42
1g	2-Thiazolyl	Cl	H	0.30	26
1h	2-Thiazolyl	CN	H	0.21	0.5
2a	2-Pyrimidinyl	H	H	30	>100
2e	2-Pyrimidinyl	Me	H	2.8	78
2g	2-Pyrimidinyl	Cl	H	0.20	2.5
2h	2-Pyrimidinyl	CN	H	NA	<0.2
2i	2-Pyrimidinyl	Cl	Me	NA	<0.2
2j	2-Pyrimidinyl	Cl	CF_3	0.065	2.5
2k	2-Pyrimidinyl	Cl	NH_2	0.075	0.5
2l	2-Pyrimidinyl	Cl	SMe	0.018	2.8
2m	2-Pyrimidinyl	Cl	OMe	0.010	9.0
3e	2-Pyridinyl	Me	H	NA	35
3g	2-Pyridinyl	Cl	H	3.2	15
3h	2-Pyridinyl	CN	H	NA	0.29
4a	Pyrazinyl	H	H	NA	23
4g	Pyrazinyl	Cl	H	NA	15
d4T				3 ± 1	>100
Nevirapine				0.12	>10
UC781				0.002	>100
TMC125				<0.002	>1

[a] EC_{50} is the dose to achieve 50% protection of MT-2 cells from HIV-1 cytopathogenicity using the MTT method. NA for $EC_{50} > CC_{50}$.

the first compounds to be prepared; the latter two did not show activity below 100 μM, while the DMA ether has an EC_{50} of 10 μM. Optimization of X in **1** did follow the pattern of Table 3 closely with gains for X = Et and Me, and great improvement for Cl (0.3 μM) and CN (0.2 μM). The parent **2a** was also found to be active at the 30 μM level and a 150-fold boost is obtained for the Cl analog, **2g**. However, generally, the X = CN derivatives such as **2h** showed enhanced cytotoxicity, which over-shadowed antiviral activity. Variation of Het also supported the lower predicted activity for the pyridinyl and pyrazinyl compounds (**3** and **4**). Further evaluations with *BOMB* indicated benefits for addition

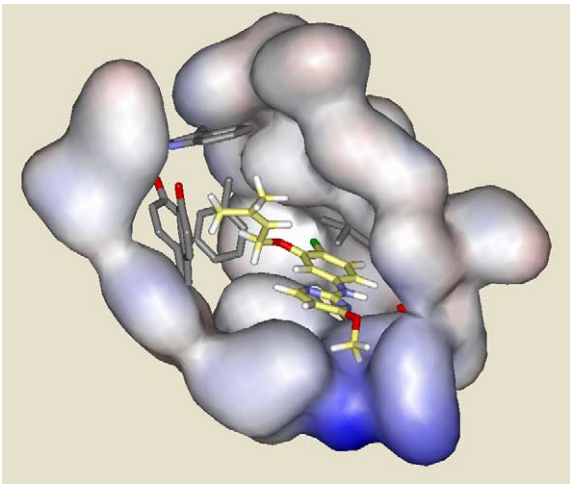


Figure 2. Computed structure for **2m** bound to RT; Y181, Y188, F227, W229 on the left; L100 and the K101 C=O on the right. The methoxy group of **2m** can be accommodated either to the right, as shown, or left.

of some small groups for Y. Indeed, introduction of an amino group, as in TMC125, gave a factor of 2–3 improvement in EC_{50} for **2k** over **2g**, but greater cytotoxicity. However, the methoxy derivative **2m** was readily found as a 10 nM NNRTI with a 900-fold safety margin. The predicted structure for its complex with WT RT is shown in Figure 2.

In summary, the present computational approach has been effective in identifying a 30 μM lead compound that could be rapidly progressed to a 10 nM NNRTI.

Acknowledgment

Gratitude is expressed to the National Institutes of Health (AI44616, GM32136, GM35208, and GM49551) for support.

References and notes

1. AIDS Epidemic Update: December 2004: UNAIDS: Geneva, 2005.
2. *HIV Medicine 2005*; Hoffmann, C., Kamps, B. S., Eds.; Flying Publisher: Paris, 2005.
3. Pauwels, R. *Curr. Opin. Pharmacol.* **2004**, *4*, 437.
4. Spence, R. A.; Kati, W. M.; Anderson, K. S.; Johnson, K. A. *Science* **1995**, *267*, 988.
5. Turner, B. G.; Summers, M. F. *J. Mol. Biol.* **1999**, *285*, 1.
6. Hopkins, A. L.; Ren, J.; Esnouf, R. M.; Wilcox, B. E.; Jones, E. Y.; Ross, C.; Miyasaka, T.; Walker, R. T.; Tanaka, H.; Stammers, D. K.; Stuart, D. I. *J. Med. Chem.* **1996**, *39*, 1589.
7. Wang, D.-P.; Rizzo, R. C.; Tirado-Rives, J.; Jorgensen, W. L. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2799.
8. Das, K.; Clark, A. D., Jr.; Lewi, P. J.; Heeres, J.; de Jonge, M. R.; Koymans, L. M. H.; Vinkers, H. M.; Daeyaert, F.; Ludovici, D. W.; Kukla, M. J.; De Corte, B.; Kavash, R. W.; Ho, C. Y.; Ye, H.; Lichtenstein, M. A.; Andries, K.; Pauwels, R.; de Béthune, M.-P.; Boyer, P. L.; Clark, P.;

Hughes, S. H.; Janssen, P. A. J.; Arnold, E. *J. Med. Chem.* **2004**, *47*, 2550.
9. Rizzo, R. C.; Wang, D.-P.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2000**, *122*, 12898.
10. Esnouf, R. M.; Stuart, D. I.; De Clerq, E.; Schwartz, E.; Balzarini, J. *Biochem. Biophys. Res. Commun.* **1997**, *234*, 458.
11. Janssen, P. A. J.; Lewi, P. J.; Arnold, E.; Daeyaert, F.; de Jonge, M.; Heeres, J.; Koymans, L.; Vinkers, M.; Guillemont, J.; Pasquier, E.; Kukla, M.; Ludovici, D.; Andries, K.; de Béthune, M.-P.; Pauwels, R.; Das, K.; Clark, A. D., Jr.; Frenkel, Y. V.; Hughes, S. H.; Medaer, B.; de Knaep, F.; Bohets, H.; De Clerck, F.; Lampo, A.; Williams, P.; Stoffels, P. *J. Med. Chem.* **2005**, *48*, 1901.
12. Jorgensen, W. L.; Duffy, E. M. *Adv. Drug Deliv. Rev.* **2002**, *54*, 355.
13. Jorgensen, W. L. *BOMB, v 2.5*; Yale University: New Haven, CT, 2004.
14. Jorgensen, W. L. *QikProp, v 2.3*; Schrödinger LLC: New York, 2005.
15. 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.* **2005**, *15,* doi:10.1016/j.bmcl.2005.10.037.