Pergamon

SCIENCE DIRECT®

Bioorganic & Medicinal Chemistry Letters 13 (2003) 3337–3340

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Activity Predictions for Efavirenz Analogues with the K103N Mutant of HIV Reverse Transcriptase

Marina Udier-Blagović, Edward K. Watkins,[†] Julian Tirado-Rives and William L. Jorgensen*

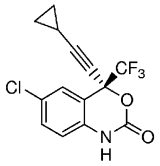
Department of Chemistry, Yale University, New Haven, CT 06520-8107, USA

Received 23 January 2003; accepted 12 May 2003

Abstract—Monte Carlo-extended linear response (MC/ELR) calculations are used to examine the binding of efavirenz analogues with the K103N mutant of HIV-1 reverse transcriptase (HIVRT). A regression equation previously reported for the wild type (WT) enzyme is shown to predict 47 experimental activities for the K103N mutant with a $q^2 = 0.55$ and avg error of only 0.46 kcal/mol. Further analysis identifies the key features for binding to the K103N mutant: ligand flexibility, burial of hydrophobic surface area, and protein–ligand van der Waals interactions.

The emergence of mutant forms of HIV-1 undermines the successful treatment of HIV infection. For patients failing non-nucleoside reverse transcriptase inhibitor (NNRTI) therapy, the lysine to asparagine modification for residue 103 of the RT p66 subunit is a pan-class resistance mutation.[1,2] The activities of all three FDA-approved NNRTIs, nevirapine, delavirdine, and efavirenz (Sustiva), are diminished by factors of 40–200 by the K103N mutation. Though efavirenz has the most favorable overall resistance profile, 90% of patients who have rebounded after efavirenz-inclusive HAART possess the K103N mutation.[3]

Efavirenz



It is important to understand and circumvent the effects of K103N in order to develop improved drug therapies. NNRTIs bind near the polymerase active site and cause inhibition by allosteric modifications.[1] In a recent study, computer simulations of HIVRT–ligand complexes were used to develop a regression equation that well reproduced WT anti-HIV activities for more than 200 NNRTIs representing eight diverse structural classes.[4] The current study (a) examines the generality of the WT regression equation for prediction of activities against the K103N mutant with a series of 47 efavirenz analogues, and (b) identifies controlling factors for binding with the K103N mutant.

Binding affinities were computed using a modification of the linear response method proposed by Åqvist et al.[5] In their approach, the free energy of binding (ΔG_b) was cast simply as a linear combination of two physically motivated descriptors:

$$\Delta G_b = \alpha\langle\Delta E_{vdW}\rangle + \beta\langle\Delta E_{Coul}\rangle \quad (1)$$

In eq 1, α and β are constants and ⟨⟩ represents an ensemble average of the difference in van der Waals (Lennard–Jones) and Coulombic interaction energies in the bound and unbound states.

This methodology has been generalized by our group to include other potentially important descriptors.[6,7] The resultant extended linear response (ELR) approach corresponds to eq 2 where the c_n are optimised

*Corresponding author. Tel.: +1-203-432-6278; fax: +1-203-432-6299; e-mail: william.jorgensen@yale.edu
[†]Present address: Schrödinger, Inc., 120 West Forty-Fifth Street, New York 10036-4041, USA.

0960-894X/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0960-894X(03)00681-4

$$\Delta G_b = \sum_n c_n \xi_n + c_0 \quad (2)$$

coefficients and the ξ_n are values of descriptors that are averaged during MC or molecular dynamics simulations. This allows consideration of additional binding-relevant terms such as changes in internal energy, solvent-accessible surface areas (SASA) and numbers of hydrogen bonds.[6,7] LR approaches rely on experimental data to train the scoring functions, eqs 1 and 2, though no additional experimental data are required for prediction of binding free energies of new systems. Only simulations for the bound complexes and unbound inhibitors are needed to obtain values for the descriptors and binding predictions. In the present case, the application of a regression equation obtained for WT HIVRT is evaluated for a mutant protein. Successful demonstration of such transferability would be auspicious for development of a general ELR expression that could be applied in the absence of specific training.

In the earlier study, eq 3 was optimized to reproduce experimental anti-HIV activities for over 200 diverse NNRTI's, and it performed well in the prediction of activities for test compounds in new chemotypes ($q^2 = 0.5$–0.6).[4]

$$\begin{aligned}\Delta G_{\text{calcd}} \approx \text{RT ln}\text{IC}_{50} &= 0.15\langle EXX_{\text{LJ}}\rangle - 0.22\langle \Delta HB_{\text{tot}}\rangle \\ &- 0.56\langle \textit{water-bridges}\rangle + 0.24\langle \#\textit{rotor}\rangle \\ &+ 0.0061\langle \Delta FOSA\rangle + 0.015\langle \Delta WPSA\rangle \\ &+ 0.14\langle \Delta \textit{dipole}\rangle - 0.73\langle \textit{DtoProPi}\rangle + 0.036\langle \Delta \textit{eintra}\rangle \\ &+ 1.2\,(H_{\text{corr}}) + 0.94(U_{\text{corr}}) - 1.3 \end{aligned} \quad (3)$$

The terms which emerged as significant are all reasonable: the protein–ligand Lennard–Jones interaction energy (EXX_{LJ}), the change in the total number of hydrogen bonds for the inhibitor upon binding (ΔHB_{total}), the number of bridging water molecules that mediate hydrogen bonding between ligand and protein (*water-bridges*), a count of the number of rotatable bonds in the ligand (*#rotor*), the changes in hydrophobic and weakly polar (halogens, P, and S) SASA, the change in dipole moment for the inhibitor, the number of hydrogen bonds donated by the ligand to aryl groups in the protein (*DtoProPi*), the change in intramolecular strain energy for the ligand, and offsets for non-efavirenz-like cores (H_{corr} and U_{corr}, which are both zero here).

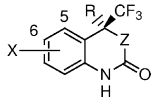
To test eq 3 for the K103N mutation, the 47 efavirenz analogues with experimental data listed in Table 1 were treated. A complication is that these data are IC_{90} values from an assay that correlates inhibition of viral replication with mRNA levels, while eq 3 was derived from IC_{50} data for the more common cell protection assay. However, we find that available experimental RT *ln* IC_{90} and RT *ln* IC_{50} data for 85 efavirenz analogues with WT HIVRT correlate well ($r^2 = 0.76$), though the RT *ln* IC_{90} values are lower by an average of 2.1 kcal/mol. The details for the MC/ELR calculations are the same as before.[4] The model for the unbound state consisted of the inhibitor centered in a 22-Å radius droplet containing 1485 TIP4P water molecules.[12] The bound state included the inhibitor, the 123 nearest protein residues and a 22-Å cap with 851 TIP4P water molecules. Two crystal structures have been reported for efavirenz bound to K103N HIVRT.[13] The more recent 1ikv structure shows little difference for the protein from virtually all structures of WT HIVRT with inhibitors including efavirenz.[13b] However, the earlier 1fko structure[13a] shows significant torsional changes for several binding-site residues, especially Y181, which has the phenol rotated to the uncommon 'down' position, roughly perpendicular to the normal 'up' arrangement as in 1ikv. The MC/ELR calculations were executed for all 47 complexes starting from both crystal structures.

The cores, 1,4-dihydro-2*H*-3,1-benzoxazin-2-one and 3,4-dihydro-1*H*-quinazolin-2-one, were positioned in the binding pocket to overlay well with efavirenz. All analogues were created with *GenMol*,[14] which adds the substituents, generates multiple conformers of the ligand, optimizes each conformer's dihedral angles and its position in the rigid binding site, and selects the energetically most favorable complexes. The unbound inhibitors were generated analogously with *GenMol*. To relax the complexes, 50 steps of conjugate-gradient minimization using a distance-dependent dielectric of $4r$ were performed. The OPLS-AA force field[15] was used throughout except the partial atomic charges for the inhibitors were obtained from AM1 calculations using the CM1 procedure[16] and scaled by a factor of 1.08 to reflect enhanced polarization in condensed phases. All MC simulations were performed at the experimental temperature of 37 °C using the *MCPRO* program.[17]

The observed activities for the K103N mutant are well correlated by regression eq 3 using either crystal structure, particularly since the range of the experimental data is not large (3.3 kcal/mol); $q^2 = 0.55$ for 1fko (Fig. 1) and $q^2 = 0.45$ for 1ikv. The average error for the 1fko-based predictions is only 0.46 kcal/mol, but this does not include any adjustment for the IC_{90} versus IC_{50} issue. The average errors are affected by such offsets, and it is best to focus on the correlation coefficients.

The reasonable q^2 values imply that the important factors for binding with the WT enzyme are also important for binding with the K103N mutant. We actually favor the 1ikv structure as more representative of K103N–ligand complexes. We have also performed MC/ELR calculations for the 47 efavirenz analogues with WT HIVRT. Application of eq 3 to predict the WT activities and the activities for the K103N mutant uniformly yields the experimental pattern that the activities for the mutant are lower with the 1ikv structure, while they are incorrectly greater with use of the 1fko structure. Results of more rigorous free-energy perturbation calculations on the effects of the K103N mutation also

Table 1. Efavirenz analogues



No.	Z	R	X	ΔG_{exp}[a]	ΔG_{calcd}[b]	No.	Z	R	X	ΔG_{exp}[a]	ΔG_{calcd}[b]
B01[c]	O	CC-c-Pr	6-Cl	−10.29	−10.50	**U10**[3]	NH	CC-c-Pr	6-OCH_3	−10.93	−10.09
B02[9]	O	$OCH_2CHCHCH_3$ (*Z*)	6-Cl	−9.83	−10.31	**U11**[3]	NH	CC-c-Pr	5-F,6-Cl	−11.42	−11.36
B03[9]	O	$OCH_2CHCHCH_3$ (*E*)	6-Cl	−9.50	−10.16	**U12**[3]	NH	CC-c-Pr	5,6-diCl	−11.42	−12.03
B04[9]	O	$OCH_2CHC(CH_3)_2$	6-Cl	−10.47	−11.02	**U13**[3]	NH	$CCCH(CH_3)_2$	6-Cl	−11.29	−11.19
B05[9]	O	OCH_2CCCH_3	6-Cl	−9.66	−9.78	**U14**[3]	NH	$CCCH(CH_3)_2$	5-Cl,6-F	−11.53	−11.40
B06[9]	O	OCH_2CHCH_2	5,6-diF	−8.65	−9.30	**U15**[3]	NH	$CCCH(CH_3)_2$	6-F	−11.15	−11.05
B07[9]	O	OCH_2CHCCl_2	5,6-diF	−9.99	−11.35	**U16**[3]	NH	$CCCH(CH_3)_2$	6-OCH_3	−10.72	−10.00
BH1[9]	O	CC-2-pyridyl	6-Cl	−8.88	−9.96	**U17**[3]	NH	$CCCH_2CH_3$	5,6-diF	−11.57	−10.69
BH2[9]	O	CC-3-pyridyl	6-Cl	−10.13	−9.62	**U18**[3]	NH	$CCCH_2CH_3$	6-Cl	−11.19	−10.90
BH3[9]	O	CC-2-furanyl	6-Cl	−10.10	−10.20	**U19**[3]	NH	$CCCH_2CH_3$	6-F	−10.56	−10.47
BH4[9]	O	CC-3-furanyl	6-Cl	−9.64	−9.62	**U20**[3]	NH	CC-Ph	5,6-diF	−9.96	−9.29
BH5[9]	O	CC-3-pyridyl	6-F	−9.63	−8.59	**U21**[3]	NH	CC-Ph	6-Cl	−9.80	−9.73
BH6[9]	O	CC-3-furanyl	6-F	−9.56	−9.21	**U22**[3]	NH	CC-Ph	6-F	−9.64	−9.47
BH7[9]	O	CC-3-pyridyl	5,6-diF	−10.68	−9.12	**U23**[3]	NH	CC-Ph	6-OCH_3	−9.72	−9.09
BH8[9]	O	CC-3-furanyl	5,6-diF	−9.73	−9.06	**U24**[11]	NH	CC-c-Pr	5-OCH_3,6-Cl	−11.95	−11.85
U01[8]	NH	CHCH-c-Pr (*E*)	5,6-diF	−10.90	−11.32	**U25**[11]	NH	CC-Ph	5-OCH_3,6-Cl	−10.29	−9.91
U02[8]	NH	CHCH-c-Pr (*E*)	6-Cl	−10.74	−11.48	**U26**[11]	NH	CC-c-Pr	5-OH,6-Cl	−11.15	−11.46
U03[3]	NH	CC-c-Pr	6-Cl	−11.35	−11.11	**UH1**[3]	NH	CC-2-pyridyl	5,6-diF	−10.31	−10.07
U04[3]	NH	CC-c-Pr	5,6-diF	−11.29	−10.75	**UH2**[d]	NH	CC-2-pyridyl	6-Cl	−10.11	−10.44
U05[3]	NH	$CCCH(CH_3)_2$	5,6-diF	−11.42	−10.99	**UH3**[3]	NH	CC-2-pyridyl	5-Cl,6-F	−10.56	−9.83
U06[3]	NH	CC-c-Pr	5-F	−10.43	−10.34	**UH4**[3]	NH	CC-2-pyridyl	6-F	−9.39	−9.40
U07[3]	NH	CC-c-Pr	5-Cl	−10.64	−11.06	**UH5**[3]	NH	CC-2-pyridyl	6-OCH_3	−9.61	−9.72
U08[3]	NH	CC-c-Pr	5-Cl,6-F	−11.42	−11.55	**UH6**[11]	NH	CC-3-pyridyl	5-OCH_3,6-Cl	−10.12	−11.44
U09[3]	NH	CC-c-Pr	6-F	−10.81	−10.61						

[a]Experimental free energies $\Delta G_{exp} = RT \ln (IC_{90})$ in kcal/mol at 37 °C. Literature IC_{90} values are divided by two for assays performed on racemic mixtures.[8]
[b]Computed free energies (kcal/mol) from ELR, eq 3 using coordinates from crystal structure 1fko.
[c]Efavirenz, IC_{90} average from refs 9 and 10.
[d]IC_{90} average from refs 3 and 11.

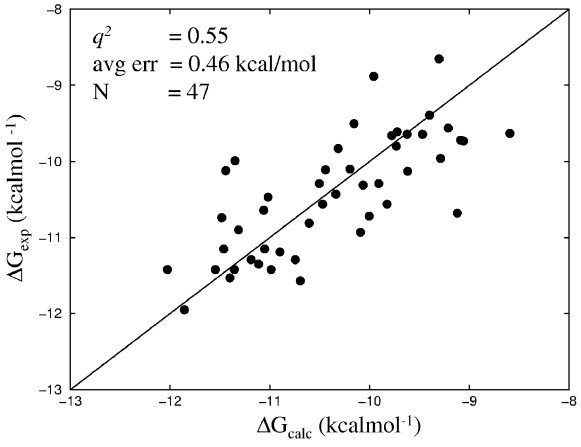


Figure 1. Calculated free energies (ΔG_{calc}) from eq 3 versus experimental activities (ΔG_{exp}) for efavirenz analogues with K103N HIVRT.

yield the same pattern favoring the conventional structure for the K103N complexes.[18]

In general, some reduced binding to the mutant can be attributed to poorer van der Waals interactions (less favorable EXX_{LJ}) between the side chain of N103 with the ligands than for K103, as illustrated in Figure 2 for the complexes with efavirenz. Furthermore, the mutation also leads to unfavorable electrostatic interactions between an amide hydrogen of N103 and the hydrogen on nitrogen in the benzoxazin-2-one ring of efavirenz (Fig. 2). The rarely observed K103Q and K103T mutations induced upon treatment with carboxanilide analogue UC-42 and BHAP U-90152 are also likely to diminish the potency of NNRTIs through a combination of poorer van der Waals and electrostatic interactions.[19]

When eq 3 is optimized to yield the best ELR fit to the K103N experimental data, only three descriptors remain that have high statistical significance, as specified in eq 4 for 1ikv and eq 5 for 1fko.

$$\Delta G_{calc} = 0.10\langle EXX_{LJ}\rangle + 0.0059\langle \Delta FOSA\rangle + 0.79 \times (rotor) - 8.2 \quad (4)$$

$$\Delta G_{calc} = 0.12\langle EXX_{LJ}\rangle + 0.0076\langle \Delta FOSA\rangle + 0.80 \times (rotor) - 6.2 \quad (5)$$

The correlation coefficients r^2 for eqs 4 and 5 are 0.61 and 0.67, with average unsigned errors of only 0.38 and 0.35 kcal/mol. The quantitative agreement between the coefficients in eqs 4 and 5 and their corresponding values in the original model, eq 3, is excellent, confirming the robustness of the model.

All three descriptors in eqs 4 and 5 and their signs are physically intuitive. Favorable protein–ligand Lennard–

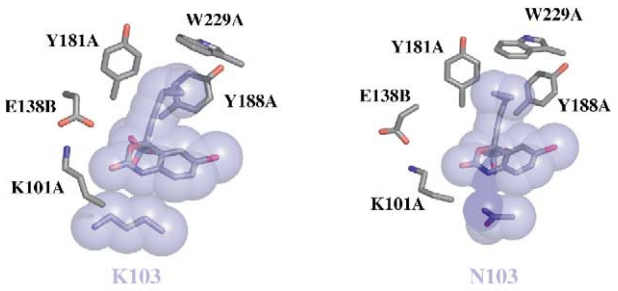


Figure 2. Typical computed structures for efavirenz bound to WT HIVRT (left) and to the K103N mutant (right).

Jones interactions indicate a good steric fit and enhance binding. The *#rotor* is expected to provide a measure of the entropy loss due to conformational restrictions for the ligands upon binding. Finally, SASA is always reduced upon binding, and burial of the hydrophobic component is favorable. Analysis of the components also permits specific insights for pairs of inhibitors. For example, with K103N, the activity of **B05** is less than for **B04**, mostly influenced by EXX_{LJ}. The butynyl-ether of **B05** has fewer favorable van der Waals interactions with residues P95, Y188, and W229 than the branched dimethyl-allyl-ether of **B04**. Similarly, the activity of **B03** is poorer than for **B01**, as influenced by *#rotor*. The butenyloxy group of **B03** is more freely rotatable than the cyclopropylethynyl group of **B01**, so **B03** pays a greater entropic penalty upon binding. Finally, the activity of **U21** is lower than for **U02**, mostly influenced by $\Delta FOSA$. The phenylethynyl group of **U21** leads to less burial of hydrophobic SASA than the cyclopropylethenyl group of **U02**.

An ELR scoring function developed previously for wild type HIV-1 RT was found to predict well activities for efavirenz analogues with the clinically significant K103N mutant. The methodology allows identification of important features for enhancing activity with the mutant enzyme in readily understandable terms, namely, optimization of protein-ligand van der Waals interactions, reduction of ligand flexibility, and increased burial of hydrophobic surface area.

Acknowledgements

The authors thank the NIH (AI44616) for financial support.

References and Notes

1. Bacheler, L. T. *Drug Resist. Updates* **1999**, *2*, 56.
2. De Clercq, E. *Biochim. Biophys. Acta* **2002**, *1587*, 258.
3. Corbett, J. W.; Ko, S. S.; Rodgers, J. D.; Gearhart, L. A.; Magnus, N. A.; Bacheler, L. T.; Diamond, S.; Jeffrey, S.; Klabe, R. M.; Cordova, B. C.; Garber, S.; Logue, K.; Trainor, G. L.; Anderson, P. S.; Erickson-Viitanen, S. K. *J. Med. Chem.* **2000**, *43*, 2019.
4. Rizzo, R. C.; Udier Blagoviæ, M.; Wang, D.-P.; Watkins, E. K.; Smith, M. B. K.; Smith, R. H.; Tirado-Rives, J.; Jorgensen, W. L. *J. Med. Chem.* **2002**, *45*, 2970.
5. Åqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Eng.* **1994**, *7*, 385.
6. Rizzo, R. C.; Tirado-Rives, J.; Jorgensen, W. L. *J. Med. Chem.* **2001**, *44*, 145.
7. Pierce, A. C.; Jorgensen, W. L. *J. Med. Chem.* **2001**, *44*, 1043.
8. Corbett, J. W.; Ko, S. S.; Rodgers, J. D.; Jeffrey, S.; Bacheler, L. T.; Klabe; R, M.; Diamond, S.; Lai, C.-M.; Rabel, S. R.; Saye, J. A.; Adams, S. P.; Trainor, G. L.; Anderson, P. S.; Erickson-Viitanen, S. K. *Antimicrob. Agents Chemother.* **1999**, *43*, 2893.
9. Cocuzza, A. J.; Chidester, D. R.; Cordova, B. C.; Jeffrey, S.; Parsons, R. L.; Bacheler, L. T.; Erickson-Viitanen, S.; Trainor, G. L.; Ko, S. S. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1177.
10. Cocuzza, A. J.; Chidester, D. R.; Cordova, B. C.; Klabe, R. M.; Jeffrey, S. J.; Diamond, S.; Weigelt, C. A.; Ko, S. S.; Bacheler, L. T.; Erickson-Viitanen, S. K.; Rogers, J. D. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1389.
11. Corbett, J. W.; Senliang, P.; Markwalder, J. A.; Cordova, B. C.; Klabe, R. M.; Garber, S.; Rodgers, J. D.; Erickson-Viitanen, S. K. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 211.
12. Jorgensen, W. L.; Jenson, C. *J. Comput. Chem.* **1998**, *19*, 1179.
13. (a) Ren, J.; Milton, J.; Weaver, K. L.; Short, S. A.; Stuart, D. I.; Stammers, D. K. *Structure* **2000**, *8*, 1089. (b) Lindberg, J.; Sigurdsson, S.; Löwgren, S.; Andersson, H. O.; Sahlberg, C.; Noréen, R.; Fridborg, K.; Hong, Z.; Unge, T. *Eur. J. Biochem.* **2002**, *269*, 1670.
14. Jorgensen, W. L. *GenMol, version 1.7*; Yale University: New Haven, 2002.
15. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
16. Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Aided. Mol. Des.* **1995**, *9*, 87.
17. Jorgensen, W. L. *MCPRO, version 1.68*; Yale University: New Haven, CT, 2002.
18. Udier-Blagovic, M.; Tirado-Rives, J.; Jorgensen, W. L. *J. Med. Chem.* Submitted for publication.
19. Balzarini, J.; Perez-Perez, M. J.; Velazquez, S.; San-Felix, A.; Camarasa, M. J.; De Clercq, E.; Karlsson, A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5470.