Structure-Based Pharmacophore Identification of New Chemical Scaffolds as Non-Nucleoside Reverse Transcriptase Inhibitors

Maria Letizia Barreca,*,† Laura De Luca,*,† Nunzio Iraci,† Angela Rao,† Stefania Ferro,† Giovanni Maga,‡ and Alba Chimirri†

Dipartimento Farmaco-Chimico, Università di Messina, Viale Annunziata, 98168 Messina, Italy, and Istituto di Genetica Molecolare IGM-CNR, via Abbiategrasso 207, 27100 Pavia, Italy

Received July 28, 2006

A structure-based molecular modeling approach was performed to identify novel structural characteristics and scaffolds that might represent new classes of HIV-1 non-nucleoside reverse transcriptase inhibitors (NNRTIs). The software LigandScout was used for identification and visualization of protein−ligand interaction sites and pharmacophore model generation. In the next step virtual screening of 3D multiconformational databases together with docking experiments allowed the identification of promising candidates for biological testing. The positive biological results obtained confirm the validity of our work strategy.

INTRODUCTION

The reverse transcriptase (RT) of the human immunodeficiency virus type 1 (HIV-1) is one of the major targets of the antiretroviral drug therapies that are used in the treatment of AIDS.[1] Non-nucleoside RT inhibitors (NNRTIs) bind in a noncompetitive manner to a unique site on the enzyme (the non-nucleoside inhibitor-binding pocket, or NNIBP), altering its ability to function and achieving a highly selective suppression of HIV-1 replication with little cytotoxicity.[2−4] Presently, three NNRTIs, namely nevirapine, delavirdine, and efavirenz, are available in clinical practice.

A combination of these drugs with nucleoside RT inhibitors and protease inhibitors leads to a considerable decrease of the viral load in most HIV-infected patients. However, in view of the increasing incidence of resistance to current drug regimens and the frequency of adverse events, the development of additional novel, selective, and potent NNRTIs remains a high priority for effective antiretroviral therapy. In attempts to gather useful information for the rational design of new leads as NNRTIs, we have recently generated a three-dimensional ligand-based pharmacophore model using the X-ray structure of RT/non-nucleoside inhibitor complexes.[5] Starting from this pharmacophore hypothesis, which consisted of 5 features (a hydrogen bond acceptor, a hydrogen bond donor, and three hydrophobic groups), a new potent class of NNRTIs was designed and synthesized, containing the 1,3-dihydro-2H-benzimidazol-2-one system.[5]

We have now developed a new pharmacophore model using a structure-based approach with a different aim. In fact, this method is able to present the interactions of a ligand to the target protein in a very specific way, and hence it is capable of providing selective and specific pharmacophore models.

In particular, considering that diarylpyrimidine (DAPY) compounds currently represent the next-generation NNRTIs showing the most promising results in clinical trials,[6,7] we decided to develop a structure-based pharmacophore model for this class of NNRTIs. The model obtained, together with computational virtual screening techniques, allowed new chemical scaffolds as potential NNRTIs to be identified. Some hits were selected and subjected to biological testing. The evaluation of their RT inhibitory activity confirmed the strength of our rational approach.

METHODS

Structure-Based Pharmacophore Modeling. The software LigandScout[8] was used for the detection and interpretation of crucial interaction patterns between the non-nucleoside inhibitor R185545 and RT protein (PDB code 1SUQ). LigandScout is a tool that allows the automatic construction and visualization of 3D pharmacophores from structural data of macromolecule/ligand complexes. For the LigandScout algorithm, chemical features include hydrogen-bond donors and acceptors as directed vectors and positive and negative ionizable regions as well as lipophilic areas represented by spheres. Moreover, in order to increase the selectivity, the LigandScout model includes spatial information regarding areas inaccessible to any potential ligand thus reflecting possible steric restrictions. In particular, excluded volume spheres placed in positions that are sterically forbidden are automatically added to the generated pharmacophore model.

In our study, the protein−ligand interactions (i.e., the pharmacophore model) identified by LigandScout for R185545/RT complex were exported and translated by a script into a Catalyst pharmacophore file.[9] The model obtained was further refined by the modification of the constraint tolerance of the spheres in accordance with the default values of Catalyst software.

In Silico Screening. A 3D query of two molecule databases using the generated structure-based pharmacophore

* Corresponding author phone: (++39) 090-6766464; fax: (++39) 090-355613; e-mail: barreca@pharma.unime.it (M.L.B.), ldeluca@pharma.unime.it (L.D.L.).

† Università di Messina.

‡ Istituto di Genetica Molecolare IGM-CNR.

10.1021/ci600320q CCC: $37.00
Published on Web 02/03/2007

model was accomplished by means of Catalyst with the aim of validating the generated hypothesis and identifying structural templates for biological testing. The Fast Flexible search routine was thus selected to screen the Derwent World Drug Index 2005 (Derwent-WDI2005)[10] and Chemicals Available for Purchase (CAP)[11] Complete 2004 databases. A molecule had to fit all the features of our Catalyst query to be retrieved as a hit.

The resulting hit molecules were ranked using the Fast Fit option. Compounds with fit values greater than or equal to 3.00 were exported to a SD format file and then processed by Advanced Chemistry Development (ACD)/Labs software,[12] in order to calculate the chemical physical properties of each compound and remove those molecules that did not satisfy the well-known Lipinski rules, describing properties of druglike compounds.[13] The remaining hits were submitted to docking calculations.

Automated Molecular Docking Experiments. The crystal structure of RT complexed with the inhibitor R185545 was retrieved from the RCSB Protein Data Bank (entry code 1SUQ) and used as target for our modeling studies. First-Discovery Protein Preparation[14] procedure was used to obtain a satisfactory starting structure for docking studies. This facility is designed to ensure chemical correctness and to optimize the protein structure for further analysis. The process adds hydrogens, neutralizes appropriate amino acid chains, and relieves steric clashes. In particular, it performs a series of restrained, partial minimizations on the cocrystallized structure, each of which employs a limited number of minimization steps. It is not intended to minimize the system completely. In our study, the minimization (OPLS 2001 force field[15]) was stopped when rmsd of the nonhydrogen atoms reached 0.30 Å, the specified limit by default.

Docking studies were performed using the Glide program,[14,16] since it has been successful in reproducing several experimental binding modes from cocrystallized PDB complexes.[14,16,17] The prepared structure was used to generate the receptor grid, and no scaling was done for van der Waals radii of nonpolar receptor atoms.

An enclosing box was used as docking space, centered on the allosteric site using the R185545 crystallographic position as reference; the ligand diameter midpoint box was set to default value (10 Å).

The structures of the molecules selected via the in silico screening approach were processed using the Schrodinger LigPrep utility. This generates a number of low-energy 3D-structures from each input molecule, with various ionization states, tautomers, stereochemistries, and ring conformations. For our study a pH range of 6−8 was set, specified chiralities were retained, desalt was applied, tautomeric forms were obtained, and only one low-energy ring conformation per ligand was generated. Coordinates for R185545 were taken directly from the X-ray structure of its complex with the HIV-1 RT.

Docking experiments were performed using 0.80 to scale the VdW radii of the nonpolar ligand atoms with a charge cutoff of 0.15. Poses were discarded as duplicates if both rms deviation in the ligand-all atom was less than 0.5 Å and maximum atomic displacement was less than 1.3 Å. At most, 2 poses per ligand were retained. GlideScore SP was used as the scoring method: this scoring function is a modified and extended version of the empirically based ChemScore function devised by Eldridge et al.[18] This function accounts for several contributions, such as lipophilic contacts, hydrogen bonding, metal binding, etc.

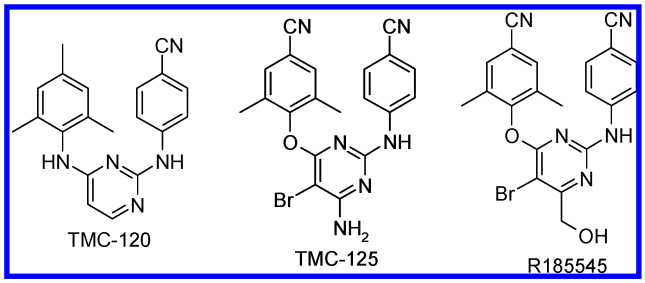

Figure 1. Chemical structures of DAPY TMC-120, TMC-125, and R185545 analogues.

HIV-1 RT DNA Polymerase Activity Assay. RNA-dependent DNA polymerase activity was assayed as described[19] in the presence of 0.5 μg of poly(rA)/oligo(dT)$_{10:1}$ (0.3 μM 3′-OH ends), 10 μM [^{3}H]-dTTP (1 Ci/mmol), and 2−4 nM RT in the presence of 4% final concentration of DMSO.

Steady-State Kinetic Measurements. Reactions were performed under the conditions described for the HIV-1 RT RNA- and DNA-dependent DNA polymerase activity assay.[19] Time-dependent incorporation of radioactive nucleotides into the different template-primers at different nucleotide substrate concentrations was monitored by removing 25 μL aliquots at 2-min time intervals. Initial velocities of the reaction were then plotted against the corresponding substrate concentrations. For ID_{50} determination, an interval of inhibitor concentrations between 0.2 ID_{50} and 5 ID_{50} was used in the inhibition assays. Data were then plotted in accordance with Lineweaver-Burke and Dixon.

RESULTS AND DISCUSSION

Diarylpyrimidines (DAPYs) are novel potent NNRTIs that show great promise for the treatment of AIDS. Overall, the in vitro potency profiles of some DAPY derivatives are better than those of all currently approved NNRTI drugs.[6] Das et al. have recently described the X-ray crystal structures of HIV-1 RT in complex with the three different DAPY compounds TMC-120, TMC-125, and R185545 (Figure 1), providing important information of binding requirements for this class of NNRTIs.[6]

The aim of this study was thus to use these structural data to generate a structure-based RT pharmacophore for DAPY analogues that could enhance the identification process of new chemical scaffolds as potential NNRTIs. In particular, we selected the complex HIV-1 RT/R185545 for our molecular modeling studies (pdb entry 1SUQ), since R185545 is the most active DAPY analogue cocrystallized with the wild-type RT enzyme.

This structure was thus used as the starting point for an extensive investigation of the chemical features important for DAPY-enzyme interaction and thus for pharmacophore generation by means of the software LigandScout.[8] The obtained hypothesis contained five features: one hydrogen bond donor and four hydrophobic groups describing the interactions between the protein RT and the ligand R185545; moreover, the software automatically included 13 excluded volumes in the model (Figure 2).

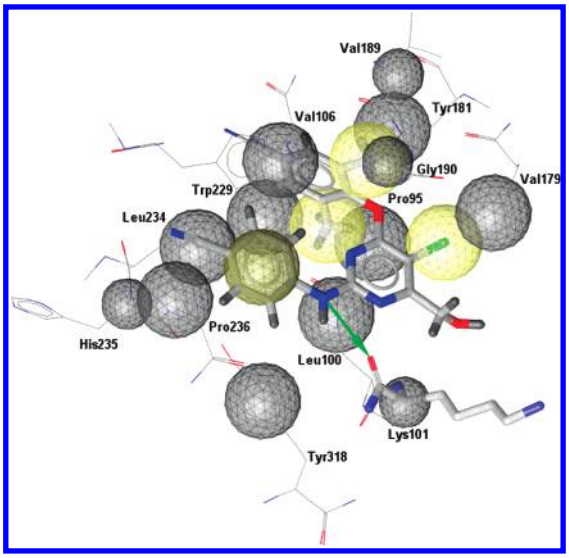

Figure 2. Pharmacophore model generated within LigandScout from R185545-RT complex. Four hydrophobic groups (light yellow spheres), one hydrogen bond donor (green projected point), and 13 excluded volumes (gray spheres) are shown.

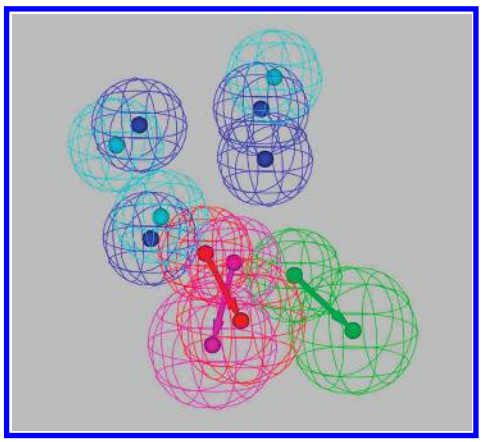

Figure 3. Comparison between our previously developed ligand-based pharmacophore model generated with Catalyst 4.9. (green spheres, hydrogen bond acceptor; magenta, hydrogen bond donor; cyan, hydrophobic sites) and the new structure-based pharmacophore model generated with LigandScout (red spheres, hydrogen-bond donor; blue, hydrophobic sites).

In particular, the hydrogen bond donor with the projected point characterized the NH group (linking the pyrimidine and 4-cyanophenyl ring) of R185545, and this interacted with Lys101 carbonyl oxygen. The four hydrophobic spheres were instead occupied by the aromatic ring of the 4-cyanophenyl moiety, the two methyl groups of the 2,6-dimethyl-4-cyanophenyl substituent, and the bromine atom, respectively.

The 13 excluded volumes reflect potential steric restriction and correspond to the positions that are sterically claimed by the macromolecular environment surrounding the ligand (i.e., residues Pro95, Leu100, Lys101, Val106, Val179, Tyr181, Val189, Gly190, Trp229, Leu234, His235, Pro236, and Tyr318). Excluded volume spheres included in the model enhance steric selectivity, thus eliminating database compounds that are not allowed to penetrate these features.

The comparison between this automated structure-based pharmacophore model and our previously developed ligand-based pharmacophore model[5] highlighted some differences in the chemical features (Figure 3). In particular, the new hypothesis shows an additional hydrophobic region and lacks the hydrogen bond acceptor group. This mismatch can be explained considering that the LigandScout pharmacophore model describes the predominant interactions of the DAPY ligands to the protein. In contrast the previous hypothesis related to the interactions that were common among several highly active compounds belonging to different chemical classes of inhibitors.

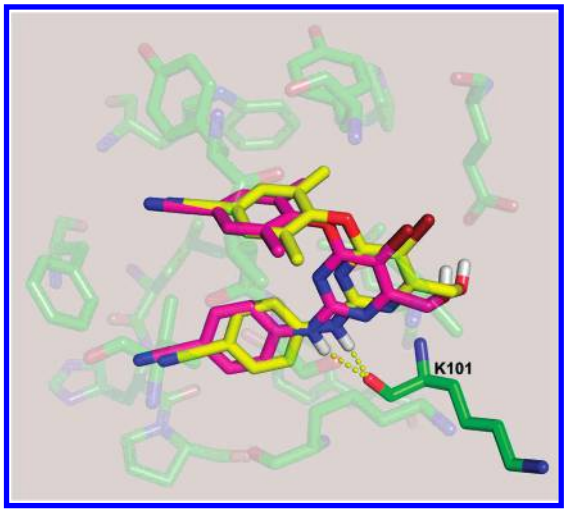

Figure 4. Comparison between predicted (yellow) and experimental (magenta) binding mode of R185545.

The new pharmacophore model is obviously more selective in database queries, searching for compounds with spatial and functional properties similar to known NNRTIs DAPY derivatives.

The 3D coordinates of the R185545-RT interaction points were exported into Catalyst pharmacophore format. They were then used as a search query against two Catalyst multiconformational three-dimensional structure databases (Derwent-WDI2005 and CAP Complete 2004) in order to validate the hypothesis and obtain new chemical scaffolds with putative activity against the RT enzyme.

Derwent-WDI2005 is an authoritative index for marketed and development drugs;[10] it contains 67 050 compounds including approximately 160 known NNRTIs. A set of 9232 molecules, 84 of which belong to the NNRTI class, were obtained as hits from our initial screening of this compound database. In order to further increase the probability that a hit is a lead,[5,20] a fitness score ≥ 3.00 was used as a second filter, leading to 521 hit molecules; among them, TMC-120, TMC-125, and other DAPY derivatives such as TMC-278 were retrieved. The same screening experiment was also carried out on the CAP Complete 2004 database,[11] which is a collection of 1 619 612 compounds from chemical suppliers, resulting in 11 273 hit molecules with a fitness score ≥ 3.00.

The chemical physical properties of these compounds were calculated by means of ACD/labs software,[12] and then the Lipinski filter[13] was applied in order to likely exclude products with poor ADME properties. A final hit list of 9435 CAP molecules was thus obtained from this screening.

Even if the hit molecules returned from Derwent-WDI2005 and CAP Complete 2004 databases presented the chemical features and the shape suggested by the structure-based pharmacophore model, their possible binding mode and interactions within the NNIBP were further analyzed via automated docking calculations using the Glide program.[14]

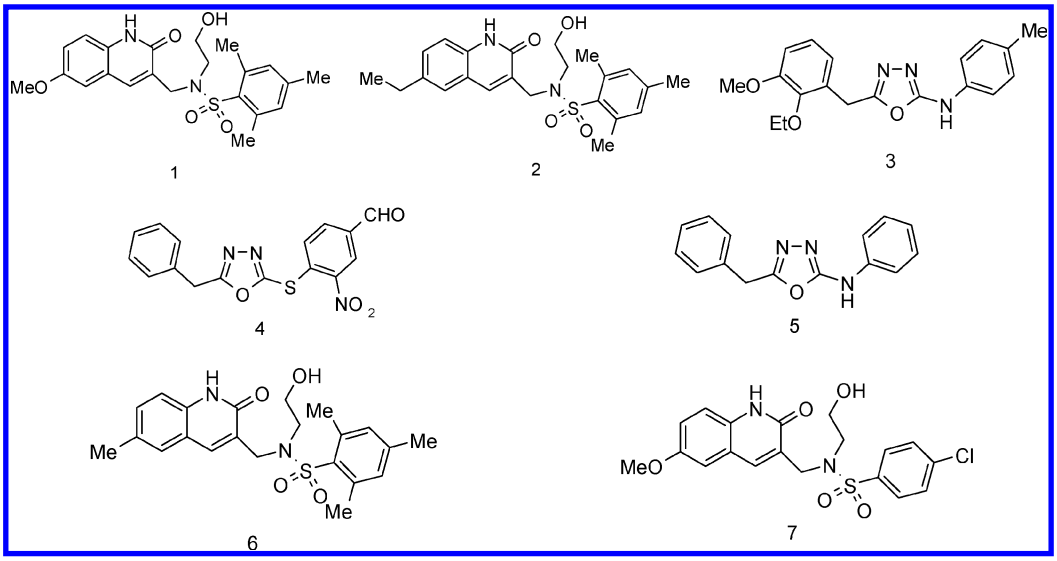

Figure 5. Compounds selected for biological experiments as potential NNRTIs.

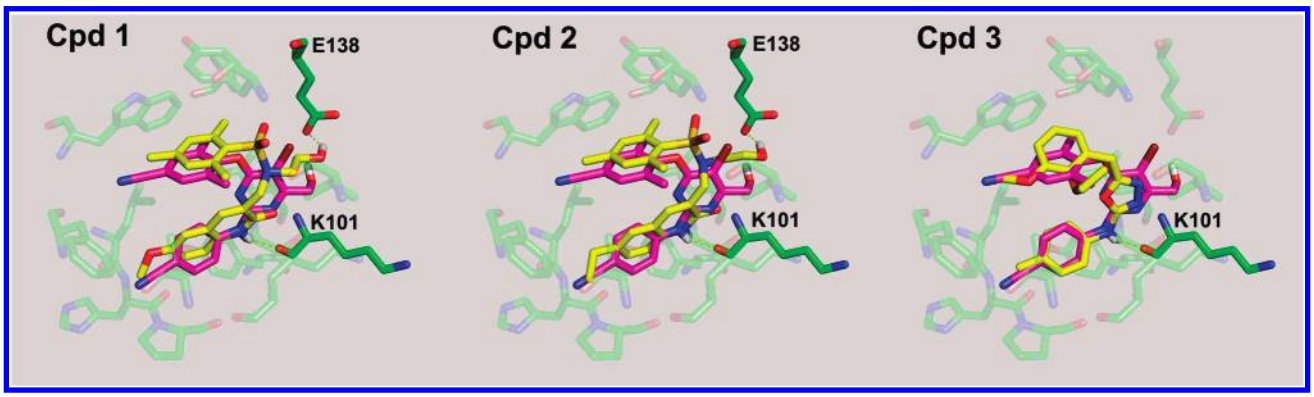

Figure 6. Glide-predicted binding mode of compounds **1**−**3** (yellow) compared to the experimental position of R185545 (magenta). Important residues of RT binding site are showed in green. Hydrogen bonds (shown as dashed yellow lines) are formed between the ligands and RT. This figure was prepared using the program PyMOL.[22]

First of all, since the ligands are flexible in the docking calculations, while the protein coordinates are fixed, test docking calculations using R185545 were carried out to validate the docking protocol. The ligand was thus extracted from the corresponding RT complex[6] and then docked back into the allosteric pocket of the enzyme crystal structure, with the aim of comparing experimental and predicted binding modes and controlling the program performance.

The best docking pose of R185545 (selected on the basis of the highest GlideScore) agreed well with the experimental binding mode of this NNRTI, with a root-mean square deviation (rmsd) value of 0.883 (Figure 4). Later, docking of 9956 compounds (selected from Derwent-WDI2005 and CAP Complete 2004 databases) into the RT allosteric site generated a number of possible binding conformations. Docked conformations with a rmsd less than 0.5 Å were clustered together and ranked by their GlideScore; then the first 1000 docking poses with the highest fitness scores were visually inspected. Not surprisingly, most compounds within this set were structurally similar to known DAPY derivatives. However, the docking results also revealed some interesting structures: **1** and **2** (from CAP Complete 2004) and **3** (from Derwent-WDI2005) (Figure 5) which interacted in a fashion similar to R185545 with RT residues of the NNIBP (Figure 6) and showed high GlideScore values.

In particular, analogously to most NNRTIs, these compounds were engaged in hydrogen bonding interactions with the backbone of the amino acid Lys101. It is worth noting that compounds **1** and **2** were not structurally related to any known class of NNRTIs, thus giving hope to the idea that they could represent a new class of NNRTIs. Moreover, compound **3** had a core structure similar to that of DAPY derivatives, providing a good chance that this molecule could bind to the same biological system, i.e., RT enzyme, and block its action by a noncompetitive inhibition.

In an attempt to assess the docking results with regard to our developed 3D structure-based pharmacophore model, compounds **1**−**3** in their "predicted" conformation (i.e., bound to RT) were mapped onto the LigandScout hypothesis using the "Fast fit" option within Catalyst. The selected poses were able to fit the model (Figure 7), confirming the plausibility of the suggested binding modes.

Finally, we used SciFinder,[21] a software available from Chemical Abstracts Service (CAS), to search for the commercial availability of these compounds. We first checked compound **3** returned by Derwent-WDI2005. This molecule has previously been studied as an anticonvulsant agent and was not available for purchase. However, the SciFinder Substructure Module (SSM) allowed the substructure search on the desired product to be carried out within most current chemical substance databases, and two commercially available molecules (**4** and **5**) structurally related to **3** were found (Figure 5). Even if compounds (**4** and **5**) did not show all the features highlighted by the model, we decided to purchase

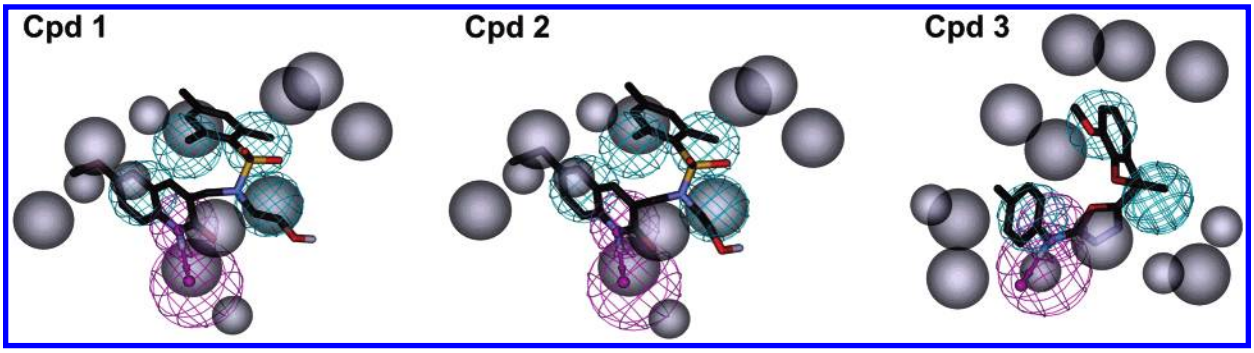

Figure 7. The best docking poses of compounds **1**−**3** mapped onto the 3D structure-based pharmacophore model.

Table 1. RT Inhibitory Activity for Compounds **1**, **2**, and **4**−**7**

compd	IC_{50}[a] (μM)
1	0.2 ± 0.03
2	1.5 ± 0.3
4	>20
5	4.2 ± 0.2
6	4 ± 0.7
7	2.8 ± 0.3
nevirapine	0.18 ± 0.02

[a] Concentration required to inhibit by 50% the in vitro RNA-dependent DNA polymerase activity of recombinant RT.

and test them in order to check the potential RT inhibitory activity of the oxadiazole skeleton and of the other frames.

Compounds **1** and **2** were commercially available and were purchased together with their analogues **6** and **7** (Figure 5) in order to study how the nature and number of substituents could influence biological activity. Compounds **1**, **2**, and **4**−**7** were evaluated in enzymatic tests for their ability to inhibit RT activity (Table 1). Interestingly, all quinolin-2(1H)-one derivatives **1**, **2**, **6**, and **7** proved to be potent RT inhibitors. In particular compound **1** showed an IC_{50} value comparable to that of nevirapine, one of the three NNRTIs currently available in antiretroviral therapy.

From a structure−activity relationship viewpoint, RT inhibition was decreased by the replacement of the methoxy group at C6 of quinolin-2(1H)-one system with alkyl substituents, such as ethyl (**2**) or methyl (**6**). Moreover, the presence of the 2,4,6-trimethylphenylsulfonyl substituent on the nitrogen atom of the aliphatic chain (**1**) positively influenced the inhibitory activity against the RT enzyme with respect to the corresponding 4-chlorophenylsulfonyl moiety (**7**). In contrast, within the class of oxadiazole derivatives (**4** and **5**), compound **4** showed no activity in RT inhibition, probably due the lack of the NH group as hydrogen bond donor for Lys101 interaction, while the **5** analogue proved to be an active compound.

In summary, we have developed a 3D structure-based pharmacophore model for DAPY NNRTIs that was used to screen large chemical databases. The hits obtained were further filtered by taking into account the fitness score value and by using Lipinski's rule of five in addition to molecular docking studies. Six compounds were finally selected for determination of their inhibitory effects against HIV-1 RT activity. In particular compound **1**, which belongs to the quinolin-2(1H)-one family and is not related to any known class of NNRTIs, exhibited RT inhibitory activity at nanomolar concentration with an IC_{50} value comparable to that of nevirapine. Following these promising results further studies are in progress to develop a lead optimization strategy.

ACKNOWLEDGMENT

Financial support for this research by Fondo Ateneo di Ricerca (2004, Messina, Italy), European TRIoH Consortium (LSHB-CT-2003-503480) is gratefully acknowledged.

REFERENCES AND NOTES

(1) Jonckheere, H.; Anne, J.; De Clercq, E. The HIV-1 reverse transcription (RT) process as target for RT inhibitors. *Med. Res. Rev.* **2000**, *20*, 129−154.
(2) De Clercq, E. Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs): Past, Present, and Future. *Chem. Biodiversity* **2004**, *1*, 44−64.
(3) Silvestri, R.; Maga, G. Current state-of-the-art in preclinical and clinical development of novel non-nucleoside HIV-1 reverse transcriptase inhibitors. *Expert Opin. Ther. Pat.* **2006**, *16*, 939−962.
(4) Balzarini, J. Current status of the non-nucleoside reverse transcriptase inhibitors of human immunodeficiency virus type 1. *Curr. Top. Med. Chem.* **2004**, *4*, 921−944.
(5) Barreca, M. L.; Rao, A.; De Luca, L.; Zappala, M.; Monforte, A. M. et al. Computational strategies in discovering novel non-nucleoside inhibitors of HIV-1 RT. *J. Med. Chem.* **2005**, *48*, 3433−3437.
(6) Das, K.; Clark, A. D., Jr.; Lewi, P. J.; Heeres, J.; De Jonge, M. R. et al. Roles of conformational and positional adaptability in structure-based design of TMC125-R16533.5 (etravirine) and related non-nucleoside reverse transcriptase inhibitors that are highly potent and effective against wild-type and drug-resistant HIV-1 variants. *J. Med. Chem.* **2004**, *47*, 2550−2560.
(7) Boone, L. R. Next-generation HIV-1 non-nucleoside reverse transcriptase inhibitors. *Curr. Opin. Invest. Drugs* **2006**, *7*, 128−135.
(8) Wolber, G.; Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model* **2005**, *45*, 160−169.
(9) Catalyst, version 4.9; Accelrys Inc.: San Diego, CA, 2006.
(10) Accelrys: Catalyst Databases. http://www.accelrys.com/products/catalyst/catalystproducts/cat_dbs.html#wdi (accessed Oct 2, 2006).
(11) Accelrys: Catalyst Databases. http://www.accelrys.com/products/catalyst/catalystproducts/cat_dbs.html#cap (accessed Oct 2, 2006).
(12) ACD/Structure Design Suite, version 9.00; Advanced Chemistry Development, Inc.: Toronto, ON, Canada, 2005.
(13) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. Experimental and Computational Approaches to Estimate Solubility and Permeability in a Drug Discovery and Development Setting. *Adv. Drug Delivery Rev.* **1997**, *23*, 3−25.
(14) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J. et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739−1749.
(15) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(16) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L. et al. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750−1759.
(17) Kontoyianni, M.; McClellan, L. M.; Sokol, G. S. Evaluation of docking performance: comparative data on docking algorithms. *J. Med. Chem.* **2004**, *47*, 558−565.

(18) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425−445.

(19) Maga, G.; Amacker, M.; Ruel, N.; Hubscher, U.; Spadari, S. Resistance to nevirapine of HIV-1 reverse transcriptase mutants: loss of stabilizing interactions and thermodynamic or steric barriers are induced by different single amino acid substitutions. *J. Mol. Biol.* **1997**, *274*, 738−747.

(20) Gopalakrishnan, B.; Aparna, V.; Jeevan, J.; Ravi, M.; Desiraju, G. R. A virtual screening approach for thymidine monophosphate kinase inhibitors as antitubercular agents based on docking and pharmacophore models. *J. Chem. Inf. Model.* **2005**, *45*, 1101−1108.

(21) SciFinder Scholar, version 2006; The American Chemical Society: Washington, DC, 2005.

(22) The PyMOL Molecular Graphics System, version 0.99; DeLano Scientific: San Carlos, CA, U.S.A., 2002.

CI600320Q