

Small Molecule Inhibitors Targeting HIV-1 Reverse Transcriptase Dimerization

Dina Grohmann,^[a] Valentina Corradi,^[b] Mira Elbasyouny,^[a] Annika Baude,^[a] Florian Horenkamp,^[a] Sandra D. Laufer,^[a] Fabrizio Manetti,^[b] Maurizio Botta,^{*,[b]} and Tobias Restle^{*,[a]}

The enzymatic activities of human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT) are strictly correlated with the dimeric forms of this vital retroviral enzyme. Accordingly, the development of inhibitors targeting the dimerization of RT represents a promising alternative antiviral strategy. Based on mutational studies, we applied a structure-based ligand design approach generating pharmacophoric models of the large subunit connection subdomain to possibly identify small molecules from the ASINEX database, which might interfere with the RT subunit interaction. Docking studies of the selected compounds identified several candidates, which were initially tested in an in vitro subunit association assay. One of these molecules (MASO) strongly re-

duced the association of the two RT subunits p51 and p66. Most notably, the compound simultaneously inhibited both the polymerase as well as the RNase H activity of the retroviral enzyme, following preincubation with $t_{1/2}$ of about 2 h, indicative of a slow isomerization step. This step most probably represents a shift of the RT dimer equilibrium from an active to an inactive conformation. Taken together, to the best of our knowledge, this study represents the first successful rational screen for a small molecule HIV RT dimerization inhibitor, which may serve as attractive hit compound for the development of novel therapeutic agents.

Introduction

Human immunodeficiency virus 1 (HIV-1) reverse transcriptase (RT) plays a pivotal role in the retrovirus life cycle and hence many of the clinically used drugs are targeted at this enzyme.^[1] However, particularly with regard to the emergence of highly resistant viral strains, the development of new and more potent inhibitors remains indispensable. The biologically relevant form of HIV-1 RT is a heterodimer consisting of a p51 and p66 subunit.^[2] The small subunit p51 is a N-terminal proteolytic cleavage product of the large subunit p66. The corresponding polypeptide chains can be divided into a N-terminal polymerase domain, a C-terminal RNase H domain, which is missing in p51, and a so called connection domain.^[3] Whereas p66 harbors all enzymatic activities, the RNA- and DNA-dependent DNA polymerase activity and RNase H activity, p51 is crucial for the structural integrity of the enzyme. Several years ago, we showed that the enzymatic activities of this retroviral polymerase are strictly correlated with the dimeric forms.^[2,4,5] Thus, the development of inhibitors targeting the dimerization of the RT represents a highly promising alternative antiviral strategy.^[6–8]

Several molecules have been described to modulate the dimer stability of RT. Among them, certain non-nucleoside RT inhibitors (NNRTIs) have been shown to stabilize the dimer,^[9] whereas others impair the stability of RT.^[10–12] In addition, peptides have been described to prevent dimerization of RT in vitro and one of them (Pep-7) could be shown to abolish the production of viral particles in infected cells.^[13–15] More recently, it could be shown that the latter peptide interacts preferentially with the p51 subunit within the heterodimer and thereby


destabilizes the dimer conformation which eventually triggers dissociation.^[16]

Mechanistic studies revealed that dimer formation occurs in a two-step process,^[17,18] which involves the rapid association of the two subunits into an inactive dimer, followed by a slow conformational change yielding the mature enzymatically active form. The subunit association is predominantly mediated by hydrophobic interactions between the two connection subdomains. Here a tryptophan repeat motif, an extraordinary cluster of six tryptophans, is important for RT dimerization as shown by mutational studies.^[19,20] Interestingly, this tryptophan cluster is highly conserved amongst primate lentiviral RTs.

This study describes a structure-based ligand design approach aimed to identify small molecules, which interfere with HIV-1 RT heterodimer stability and thereby eventually inhibit

[a] Dr. D. Grohmann, M. Elbasyouny, A. Baude, F. Horenkamp, S. D. Laufer, Prof. T. Restle
Institut für Molekulare Medizin
Universitätsklinikum Schleswig–Holstein, Campus Lübeck
Ratzeburger Allee 160, 23538 Lübeck (Germany)
Fax: (+49) 451-5002729
E-mail: restle@imm.uni-luebeck.de

[b] Dr. V. Corradi, Dr. F. Manetti, Prof. M. Botta
Dipartimento Farmaco Chimico Tecnologico
Università degli Studi di Siena
Via Alcide de Gasperi, 2, 53100 Siena (Italy)
Fax: (+39) 0577-234333
E-mail: botta@unisi.it

 Supporting information for this article is available on the WWW under <http://www.chembiochem.org> or from the author.

the retroviral enzyme. We provide a “proof of principle” that a large dimeric macromolecule such as HIV-1 RT with an interface of $>4500 \text{ \AA}^2$ can be successfully targeted by a small molecule dimerization inhibitor.

Results

Effect of mutations in the p51 subunit on HIV-1 RT dimerization and polymerase activity

To start with, we performed computational and mutational studies to unravel the contribution of individual amino acids for the HIV-1 RT heterodimer stability, which then served as the starting point for the development of small molecules that block interactions in the connection subdomain (see section below). Analysis of crystallographic structures of HIV-1 RT revealed amino acids K331, N363, and D364 in the p51 subunit with putative contributions to subunit interaction. According to our studies, K331 forms a salt bridge to D364 and N363 undergoes a π - π interaction with W410 (p66), which in addition interacts via an edge to face interaction with W401 (p51) (Figure 1). Accordingly, the residues K331, N363, and D364 in p51 were individually mutated to an alanine and the proteins were expressed and purified separately. Heterodimer formation was induced by mixing equimolar amounts of both subunits and the dimerization process was followed by HPLC gel filtration analysis (Figure S1a) and polymerase activity assays (Figure S1b). The mutant K331A showed a drastically reduced ability to associate with the p66 subunit. The N363A mutation showed a similar effect but to a lesser extent, whereas the D364A mutation had no effect on the dimerization nor the activity of the enzyme.

Structure-based ligand design approach

Based on the mutational studies described above, we followed a structure-based ligand design approach for the identification of novel compounds disturbing the RT subunit interaction by interrupting the interactions described above which target the p66 tryptophan cluster. Starting from the crystallographic structure of the p66 subunit (1RTH),^[21] three different three-dimensional pharmacophoric models of the p66 connection subdomain were generated taking into account the receptor flexibility. In particular, p66 was first submitted to molecular dynamics (MD) calculations from which six snapshots were chosen on the basis of the maximum root-mean-square deviations affecting dihedral angles of both W402 and W410 residues. Next, for each snapshot, a grid centered on W402, also including W410 and additional residues, was coded by means of a GRID Molecular Interaction Field^[22,23] approach into profitable interaction regions with three probes (namely DRY, N1, and O), corresponding to hydrophobic contacts, hydrogen bond donor, and hydrogen bond acceptor groups, respectively. Points corresponding to the best interactions between amino acid groups and probes were transformed into parts of the pharmacophores, termed pharmacophoric features (Figure S3). The pharmacophoric models were then used as three-dimen-

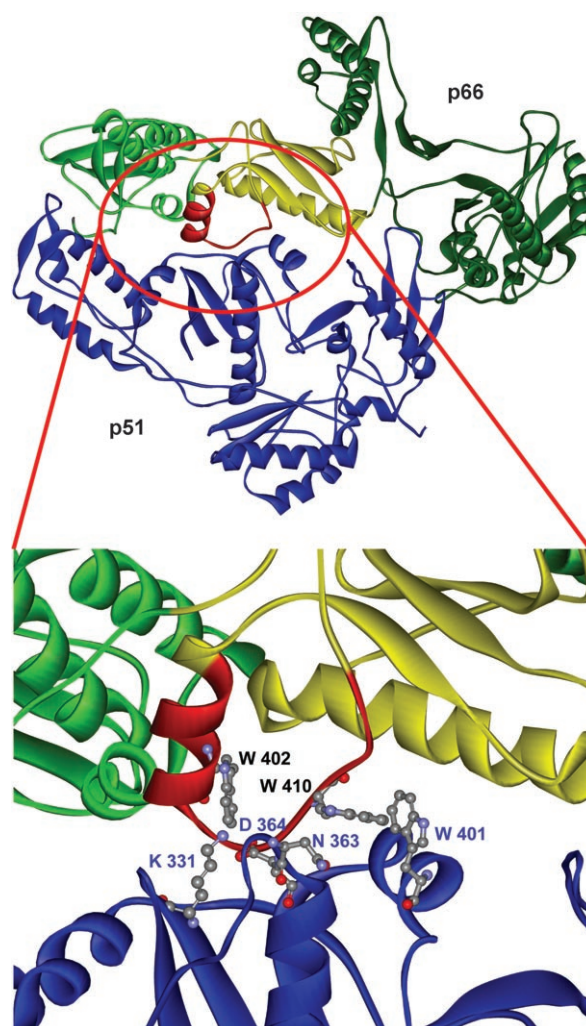


Figure 1. X-ray structure of HIV-1 RT (1RTH). The p66 subunit is color coded as follows: polymerase domain (dark green), connection domain (yellow), and RNase H domain (light green). The p51 subunit is shown in blue and the p66 tryptophan cluster (aa 398–414) in red. Key amino acids for dimer stability W402^{p66}, W410^{p66}, K331^{p51}, N363^{p51}, D364^{p51}, and W401^{p51} are depicted as balls and sticks.

sional queries to perform a virtual screening of databases of commercially available compounds. With this aim, entries of the ASINEX Gold Collection able to fit at least one of the pharmacophoric models and to fully satisfy Lipinski's Rule-of-Five,^[24,25] were chosen as hit compounds and docked into the p66 connection subdomain (Figure 2A and Figure S4).

Effect of selected compounds on HIV-1 RT subunit association and enzymatic activities

The structure-based ligand design approach described above allowed us to identify ten candidates, which were initially tested in an in vitro subunit association assay. For this purpose the heterodimeric form of HIV-1 RT was reversibly dissociated by the addition of acetonitrile (final concentration 12–14%). Upon reduction of the acetonitrile concentration to 0.8% by a simple dilution of the samples, reassociation of RT was initiated

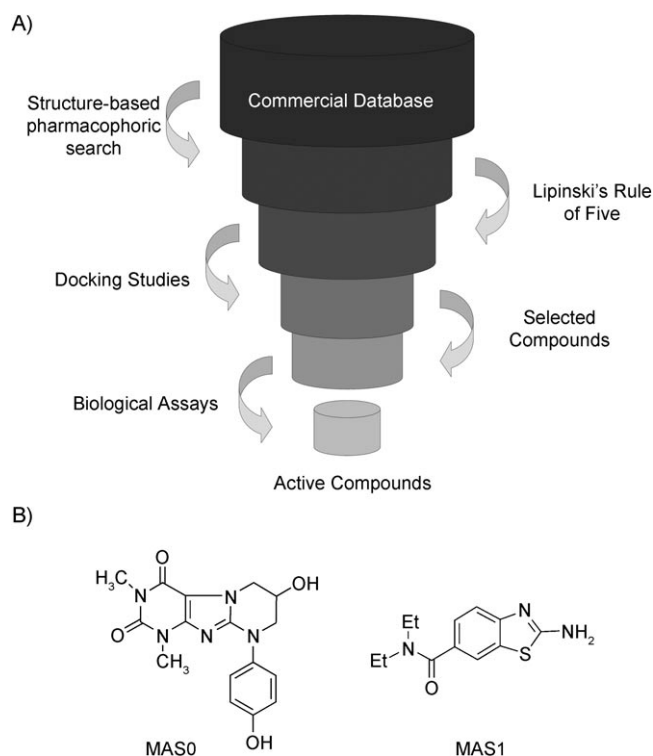


Figure 2. A) Flowchart of the rational drug design procedure. B) Structure of the selected compounds.

and could be easily followed by size-exclusion chromatography or the increase of the enzymatic activities.^[4] Applying this procedure, we identified two compounds, namely MAS0 and MAS1 (Figure 2B), which showed a dose-dependent inhibition of the dimerization process (Figure S2). MAS0 proved to be about five times more active than MAS1 in this assay. Incubation of the compounds with heterodimeric RT for up to a week at 4 °C or 25 °C on the other hand did not induce monomerization of the protein. Next, we tested whether these molecules do interfere with the enzymatic activities of RT. Figure 3A shows a dose-dependent, simultaneous inhibition of both the polymerase and the RNase H activity by MAS0 (MAS1 did not show any effect) yielding IC_{50} values of 155 and 111 μM , respectively. Most notably, inhibition of RT could only be observed following a preincubation of enzyme and MAS0 with $t_{1/2}$ of about 2 h (Figure 3B).

Effect of MAS0 on HIV-1 RT heterodimer stability

To examine if MAS0 had an effect on dimer stability, for example the equilibrium between enzymatically inactive and active dimers, which could explain the results described above, we performed transient kinetic studies following changes of the intrinsic RT fluorescence upon acetonitrile-induced heterodimer dissociation in the presence or absence of the compound using a stopped flow device. As reported earlier, the intrinsic fluorescence emission of RT increases up to 25 % upon dissociation of the heterodimer.^[17,18,26] As shown in Figure 4, there was doubling of the observed RT heterodimer dissociation rate

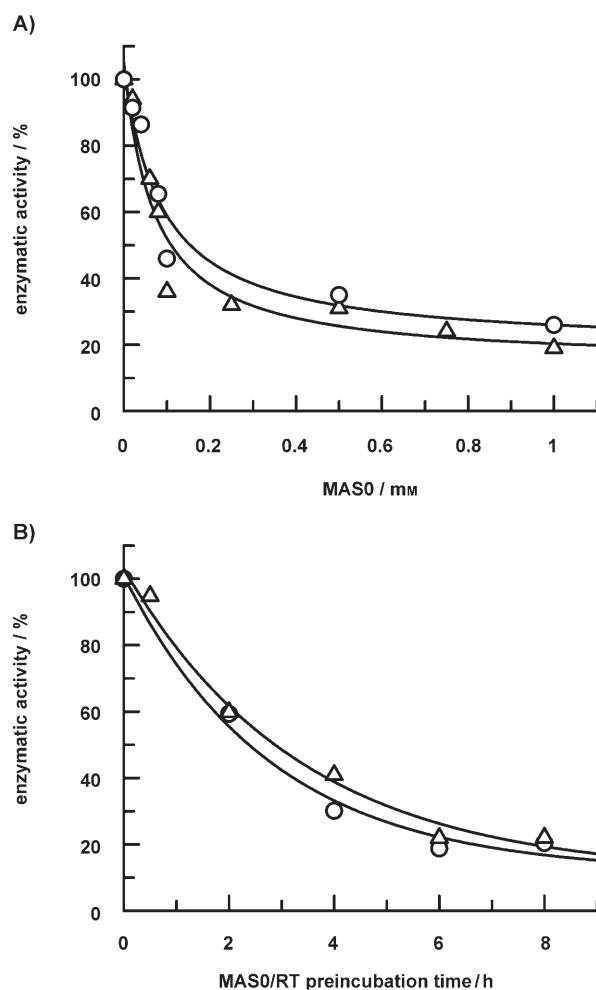


Figure 3. Simultaneous inhibition of HIV-1 RT polymerase and RNase H activity by MAS0. A) HIV-1 RT (100 nM) was incubated for 6 h at 20 °C in the presence of increasing amounts of MAS0 before polymerase (○) or RNase H (△) activity was determined. Data were analyzed using a hyperbolic equation yielding an IC_{50} value of 155 μM for the polymerase and 111 μM for the RNase H activity. The fits reach for 19% (± 9) and 14% (± 7) residual polymerase and RNase H activity, respectively. At MAS0 concentrations far beyond 1 mM, solubility problems were encountered causing varying and uncontrollable inhibitor concentrations. B) Effect of RT/MAS0-preincubation time on both enzymatic activities. Data were analyzed using an exponential equation yielding rates of 0.35 (± 0.08) and 0.3 (± 0.07) h^{-1} for the polymerase (○) and the RNase H (○) activity, respectively. During preincubation no significant reduction of enzymatic activities was observed in the absence of the compound (data not shown).

constants from $0.077(\pm 0.0002) \text{ s}^{-1}$ to $0.188(\pm 0.0004) \text{ s}^{-1}$ caused by MAS0.

Though, as already described in the previous section, this phenomenon could only be observed after preincubation of RT with the compound. By plotting the apparent dissociation rate constants on the RT/MAS0 preincubation time we found a time dependent increase of these rates with a rate of $0.47(\pm 0.06) \text{ h}^{-1}$. This number translates into $t_{1/2}$ of about 1.5 h (Figure 4B) and is in remarkably good agreement with the rate derived for the correlation of preincubation time and degree of inhibition of enzymatic activities.

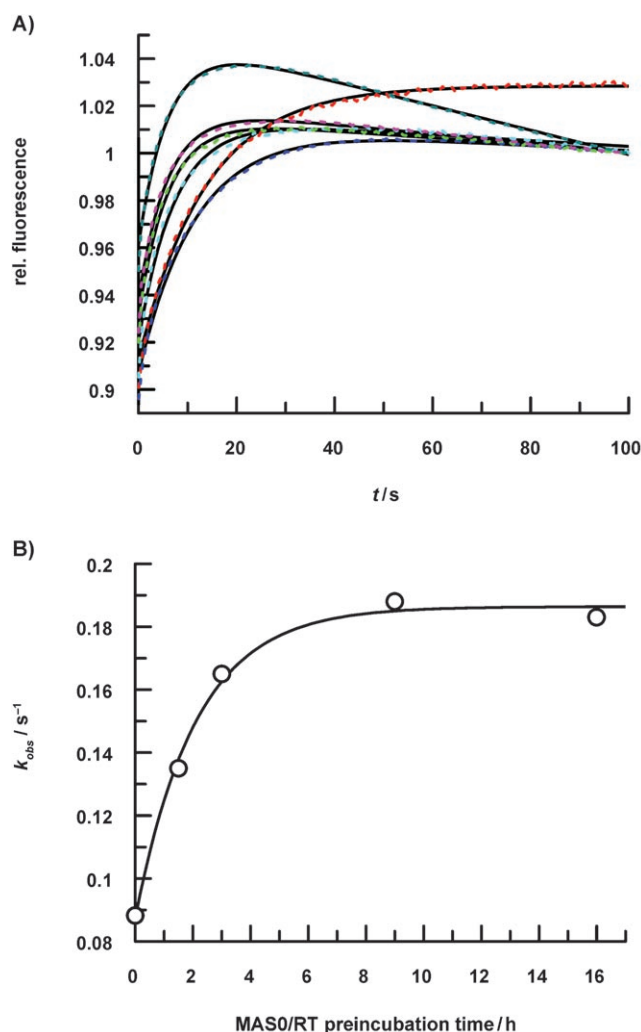


Figure 4. Transient kinetics of acetonitrile-induced HIV-1 RT dissociation in the presence or absence of MASO. A) RT (2.5 μ M) and MASO (1 mM) were preincubated at 20 °C for the following time intervals (t): red: control without MASO, $t=0$ h; blue: $t=0$ h; cyan: $t=1.5$ h; green: $t=3.0$ h; magenta: $t=9.0$ h; and dark cyan: $t=16$ h, following a rapid mix with 28 % acetonitrile (final concentration of 14 % after mixing) in a stopped flow device. The increase in intrinsic tryptophan fluorescence of RT upon dissociation of the heterodimer was followed. Excitation was at 290 nm, and emission was detected via a cut-off filter (320 nm). The experimental data were fitted to a double exponential equation plus offset (single exponential in case of the control) yielding rates of $0.077(\pm 0.0002) s^{-1}$ (control without MASO, $t=0$ h), $0.088(\pm 0.0002) s^{-1}$ ($t=0$ h), $0.135(\pm 0.0005) s^{-1}$ ($t=1.5$ h), $0.165(\pm 0.0005) s^{-1}$ ($t=3.0$ h), $0.188(\pm 0.0004) s^{-1}$ ($t=9.0$ h), and $0.183(\pm 0.0007) s^{-1}$ ($t=16$ h) for the first phase. The second phase observed in the presence of MASO is most likely due to quenching effects of the compound interacting with the fully exposed binding pocket. To exclude any effects on dissociation independent of MASO, control experiments were performed. No significant change of the rate after up to 16 h of preincubation of RT in the absence of the compound could be observed (data not shown). B) Dependence of the observed dissociation rate constants on the RT/MASO preincubation time. Data were analyzed using an exponential equation yielding a rate of $0.47(\pm 0.06) h^{-1}$.

Specificity of inhibition by MASO

Finally, we analyzed the specificity of MASO for the HIV-1 enzyme. For this matter, polymerase assays were performed with HIV-2 RT, Avian Myeloblastosis Virus (AMV) RT, and *E. coli*

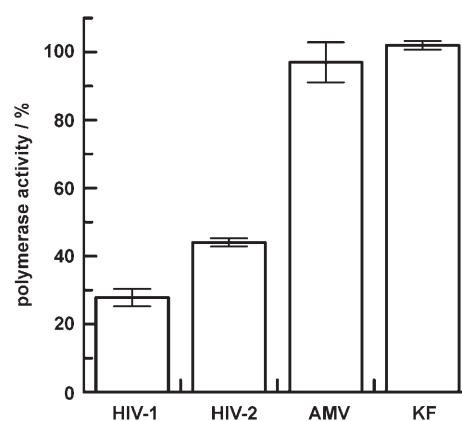


Figure 5. Determination of polymerase activity of different enzymes in the presence of MASO. Polymerases (60–150 nM) were preincubated with 1 mM MASO for 16 h at 20 °C. Polymerase activities were analyzed as described in the Experimental Section.

DNA Pol I (Klenow fragment, KF). As shown in Figure 5, MASO decreased the polymerase activity of HIV-2 by about 60%, whereas the activity of the AMV RT and KF remained unaffected.

Discussion and Conclusion

With an interface of $>4500 \text{ \AA}^2$ ^[27] and a binding affinity in the range of 10^{-10} M ^[17] the HIV-1 RT heterodimer represents a particularly challenging target especially for small molecule drugs. This might be one of the reasons why, although initially proposed some 17 years ago,^[4] nobody has so far succeeded in successfully developing drugs interfering with RT dimerization. As briefly outlined in the Introduction, there are several examples of molecules which do in fact modulate dimer stability to some extent. However, none of them is capable of really blocking the enzyme. Certain peptides, on the other hand, could be identified as powerful inhibitors. However, peptides possess inherent properties that make them highly unattractive for downstream drug development. Thus, the intention of the present study was the identification of small molecules which affect HIV-1 RT dimer stability by applying a structure-based ligand design approach. As opposed to earlier studies we focused strictly on compounds obeying Lipinski's Rule-of-Five as a prerequisite for potential further development towards a clinical application.

The structure-based ligand design approach revealed ten compounds, which were initially screened applying a well established redimerization assay.^[4] Herein, we could identify two molecules (MASO and MAS1), which affected the p51/p66 sub-unit association reaction while being incapable of promoting any dissociation of the dimer even after extended incubation. A careful examination of the effect of these two compounds on the enzymatic activities of the retroviral enzyme showed an identical simultaneous dose-dependent inhibition pattern of both the polymerase and the RNase H activity by MASO whereas the second compound was inactive.

Remarkably, the degree of inhibition of both enzymatic activities was strictly correlated with the preincubation time of enzyme and inhibitor with a $t_{1/2}$ of about 2 h. This observation strongly supports the notion that inhibition is mediated by a slow equilibrium, conformational change of the enzyme rather than by a simultaneous direct interaction of the compound with the two active sites of the enzyme, which are approximately 65 Å apart. Moreover, in the latter case the compound would have to bind both active sites with equal affinities, which would be rather implausible.

A highly likely scenario to rationalize the underlying mechanism of MASO-mediated inhibition is a shift in dimer equilibrium from an active dimer to an inactive dimer (see Figure 6 for

tryptophan repeat motif, Figure S4), the data provided in Figure 5 are in strong support of this notion. Apart from HIV-1 RT, we also observed inhibition of the closely related HIV-2 RT, albeit to a slightly reduced level. Both enzymes possess a tryptophan cluster which is conserved amongst primate lentiviral RTs. The different levels of inhibition could be explained by some minor sequence variation in this region (Figure S5). The likewise heterodimeric AMV RT,^[28] on the other hand, is not affected by MASO. Even though all three RTs (HIV-1, HIV-2,^[29] and AMV) do presumably share a similar overall folding (the X-ray structure of AMV RT is not known), only enzymes containing a tryptophan cluster (AMV RT has no such motif) are affected.

Why MAS1 failed to inhibit polymerase and RNase H activity remains unclear. At least two scenarios are likely: 1) the two compounds do have different binding modes and only MASO is able to trap the enzyme in its inactive dimer form or 2) it is simply a matter of the binding constant. MAS1 appears to bind about five times weaker than MASO (see Figure S2). Likewise, it might be a combination of both.

In conclusion, to the best of our knowledge, this study represents the first successful rational screen for a small molecule HIV RT dimerization inhibitor, which may serve as an attractive hit compound for the development of novel therapeutic agents. Obviously, concentrations in the low three-digit μM range required to cause an inhibitory effect are still fairly high. Then again, there are numerous examples in the literature where the initial hit compounds showed

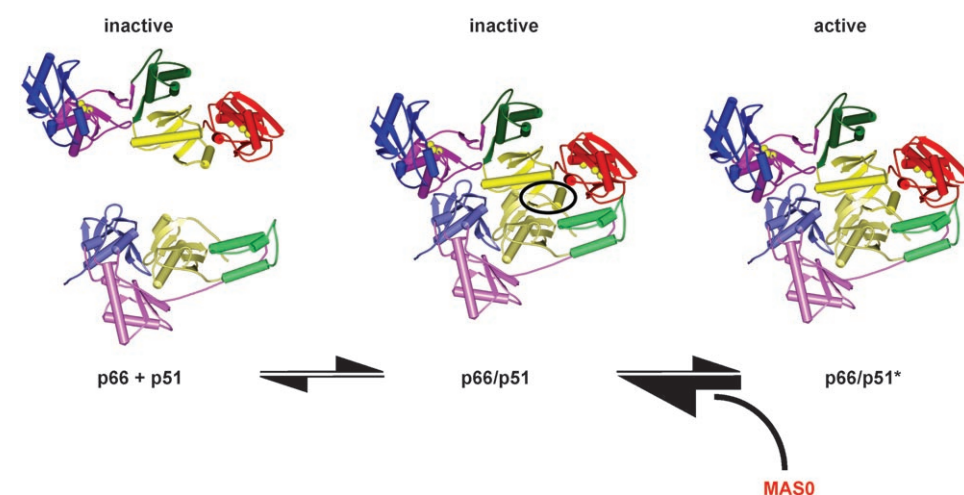


Figure 6. Proposed mode of MASO action. MASO is supposed to interfere with the dimer equilibrium^[17,18] by shifting the equilibrium from an active to an inactive dimer by trapping the enzyme in the inactive conformation. As the structures of the intermediate inactive dimer and the monomeric species are not known, the structures shown here for illustration purposes are all derived from the structure of an active heterodimer. The protein is color coded as follows: fingers (blue), palm (magenta), thumb (green), connection (yellow), and RNase H (red). Active site residues (polymerase as well as RNase H) are depicted as yellow spheres. The proposed binding site of MASO is encircled. The star distinguishes structural states of the system that otherwise have the same composition.

illustration). This would explain the simultaneous inhibition of both enzymatic activities as both activities are confined exclusively to the mature heterodimer.^[2,4,5] If indeed the compound interferes with the dimer equilibrium this should affect dimer stability and such an effect should in principle be measurable. To address this question we performed transient dissociation studies. For this purpose, we analyzed acetonitrile-induced heterodimer dissociation in the presence or absence of the compound. As anticipated, MASO increased the apparent rate of subunit dissociation. Again, the observed effect, increase in dissociation rate constant, was strictly correlated with the preincubation time of enzyme and inhibitor with a $t_{1/2}$ of about 1.5 h. This is in remarkably good agreement with the value obtained for the inhibition of the enzymatic activities and strongly supports the proposed concept that MASO traps the enzyme in an inactive state.

Whereas we currently do not have direct proof that MASO does in fact exactly bind to the pocket it was selected for (p66

inhibition in the micromolar range (for example, HIV-1 RT NNRTIs^[30]), which eventually turned into powerful inhibitors with IC_{50} values in the low nanomolar range by rather few modifications, considerably increasing affinity.

Experimental Section

Mutagenesis of the p51 subunit of HIV-1 RT: The desired mutations were introduced as described previously^[31] using the plasmid p6HRT51^[32] followed by transformation of *E. coli* M15/pDML1 cells.^[33]

Protein purification: Recombinant heterodimeric wild-type HIV-1_{BH10} as well as p66 subunits and HIV-2_{D194} RTs were expressed in *E. coli* and purified as described.^[34,35] Purification of His-tagged HIV-1 p51 was carried out according to a protocol described previously.^[36] Enzyme concentrations were determined using an extinction coefficient at 280 nm of 260 450 $\text{M}^{-1}\text{cm}^{-1}$ (HIV-1 RT), 238 150 $\text{M}^{-1}\text{cm}^{-1}$ (HIV-2 RT), 136 270 $\text{M}^{-1}\text{cm}^{-1}$ (p66), and

$124180\text{ M}^{-1}\text{ cm}^{-1}$ (p51). Avian Myeloblastosis Virus (AMV) RT was purchased from Invitrogen (Karlsruhe, Germany). *E. coli* DNA Pol I (Klenow fragment, KF) was purified by a Ni-NTA column and enzyme concentration was routinely determined using an extinction coefficient at 280 nm of $55330\text{ M}^{-1}\text{ cm}^{-1}$.

Polymerase assay: RNA-dependent DNA polymerase activity was measured at 37°C for 10 min with poly(rA)/oligo(dT)_{12–18} as substrate and 10–20 nM of enzyme.^[4] All experiments were routinely performed in reaction buffer containing Tris/HCl (50 mM pH 8.0), KCl (80 mM), MgCl₂ (8 mM), DTT (0.1 mM), BSA (10 $\mu\text{g mL}^{-1}$), and [³H]TTP (20 μM). In the case of KF, DNA-dependent DNA polymerase activity was measured at 37°C for 10 min with DNase I activated salmon sperm DNA (Invitrogen, Karlsruhe Germany) as substrate in a reaction buffer containing Tris/HCl (50 mM pH 8.0), DTT (1 mM), and MgCl₂ (10 mM). $2 \times 10\text{ }\mu\text{L}$ of the reaction mixtures were loaded onto approximately 2 cm^2 DEAE filters (Whatman) and allowed to dry. The filters were washed twice with SSC buffer (0.15 M NaCl, 15 mM sodium citrate pH 7.0) to remove free dNTPs. After a final wash with ethanol (100%), filters were dried again, and subsequently filter retained radioactivity was quantified in a scintillation counter.

RNase H assay: RNase H activity was measured at 37°C for 15 min in a buffer containing Tris/HCl (50 mM pH 8.0), KCl (80 mM), MgCl₂ (8 mM), DTT (0.1 mM), and BSA (10 $\mu\text{g mL}^{-1}$) with 5'-³²P-labeled 35/52 mer RNA/DNA substrate (130 nM) and enzyme (15–20 nM). Products were analyzed by denaturing PAGE (15% polyacrylamide/8 M urea) and quantified by scanning the dried gel using a phosphorimager (TyphoonTM 8600, GE Healthcare).

HPLC size exclusion chromatography: Association of separately expressed and purified HIV-1 RT subunits was followed after mixing of equimolar amounts of both (1.175 μM) in a final volume of 100 μL at 4°C . Subsequent association of the two subunits was monitored by HPLC size exclusion chromatography. Chromatography was performed using a Superdex 200 HR 10/30 column (GE Healthcare). The column was eluted with Bis-Tris Propane/HCl (10 mM pH 7.0) and ammonium sulfate (100 mM) at 0.5 mL min^{-1} . Prior to mixing of the p66 subunits with p51, preformed p66 homodimers were dissociated by treatment with acetonitrile (15%) on ice followed by dilution of the samples in acetonitrile free buffer.^[4] Reassociation of acetonitrile treated HIV-1 RT heterodimers was followed after complete dissociation of the heterodimers in a buffer containing MES (50 mM pH 6.0) by adding acetonitrile to a final concentration of 12–14% and incubation at 20°C for 10 min. Reassociation of the subunits was initiated in the absence or presence of MAS compounds by a 15-fold dilution into acetonitrile free polymerase reaction buffer (see above) with a final enzyme concentration of 1 μM and followed by HPLC size exclusion chromatography.

Stopped flow measurements: Experiments on the kinetics of acetonitrile-induced HIV-1 RT heterodimer dissociation were performed using a stopped flow apparatus (SX 20, Applied Photophysics Ltd, Leatherhead, England). 2.5 μM of RT (1.25 μM final concentration) in Tris/HCl (50 mM pH 8.0), MgCl₂ (10 mM), KCl (50 mM), and DTT (1 mM) were rapidly mixed 1:1 with a solution of acetonitrile (28% in H₂O). Excitation of the samples was at 290 nm using a Xe high-pressure arc lamp and detection was through a filter with a cutoff at 320 nm. As reported earlier, the intrinsic fluorescence emission of RT increases up to 25% upon dissociation of the heterodimer.^[17] Data were collected using the software package provided by Applied Photophysics and analyzed using the program "GraFit" (Erithacus software).

Computational studies: To build structure-based pharmacophoric models of the p66 connection subdomain, a molecular dynamics (MD) simulation was performed on the p66 subunit (taken from the crystallographic structure of the HIV-1 RT, entry 1RTH^[21] of the Brookhaven Protein Data Bank), using the software packages NAMD^[37] (version 2.5) and CHARMM27^[38] force field. Hydrogen atoms were added by means of the *psfgen* package. The p66 subunit was embedded in a sphere of water molecules (60 Å radius) applying spherical boundary conditions. The starting structure was optimized with 1000 steps of conjugate gradient energy minimization to remove unfavorable contacts. MD simulation was carried out at 310 K for 1 ns, collecting snapshot structures every 1 ps. The Langevin Dynamics procedure, with a dumping factor of 5 ps^{-1} , was used to control the temperature. From the MD trajectory, six snapshots, characterized by different conformations of relevant residues of the connection subdomain (in particular, W402 and W410), were chosen. For each snapshot, a $39 \times 33 \times 14\text{ }\text{\AA}$ grid (NPLA parameter set to 0.25 Å), centered on W402, was defined. Molecular interaction fields (MIFs, by means of the software GRID, version 21,^[22,23] Molecular Discovery Ltd.: Pinner, Middlesex, UK) between residues within the grid and three probes (hydrophobic, DRY; hydrogen bond donor, N1; hydrogen bond acceptor, O) were computed to describe hydrophobic interactions and hydrogen bond contacts. The best interaction points (that is, minimum energy points of the MIFs calculated for each probe) between probes and both W402 and W410 were identified for each snapshot (using MINIM and FILMAP programs included in the GRID package) and converted by means of the DS Viewer Pro 6 software (Accelrys, Inc., San Diego, CA, USA) into portions (called features) of pharmacophoric models. In particular, the center of each pharmacophoric feature (a sphere with a 1.5 Å radius) was placed at the same coordinates of each best interaction point. Minima identified by computing MIFs with the DRY probe corresponded to hydrophobic features (defined by aliphatic groups and aromatic moieties), whereas minima for N1 and O probes were replaced by the center of hydrogen bond donor and acceptor features, respectively. Excluded-volume spheres (corresponding to residue W398, W402, W410, and W414) were added to each of the six models to better define size and shape of the binding site and to avoid the identification of compounds during the next virtual screening procedure that may overlap portions of the protein, thus causing steric protein–ligand clashes.

Finally, the six pharmacophoric models derived from the selected MD snapshots were merged into three pharmacophoric hypotheses using the Merge Hypothesis/Features option of the View Hypothesis Workbench module of the software Catalyst 4.8 (Catalyst 4.8; Accelrys, Inc., San Diego, CA, USA). The merging was carried out on the basis of the distance tolerance value. As a result, Hypo1 was obtained by merging pharmacophoric models derived from three snapshots, Hypo2 merging models from two snapshots, whereas Hypo3 corresponded to one pharmacophoric model as obtained from the corresponding MD snapshot (Figure S3).

Catalyst was also used to apply the virtual screening procedure. All the compounds of the Asinex Gold Collection (www.ASINEX.com/prod/gold.html) were converted in a Catalyst database by means of *catDB* (maxconfs option set to 100). The screening was carried out using the Fast Flexible Search module. Compounds satisfying all the pharmacophoric features of at least one model were retrieved. Selected compounds were then filtered on the basis of Lipinski's Rule-of-Five,^[24,25] to retain only drug-like entries. Finally, only compounds showing the highest fit value to the pharmacophoric models were kept.

Selected compounds were docked into the p66 connection subdomain using AutoDock 3.0.5 (The Scripps Research Institute, La Jolla, CA, USA) as a docking simulation tool. To prepare the input structures of the selected compounds for docking calculations, a geometry optimization was performed using the Gaussian03 program (semiempirical hamiltonian AM1,^[39] Gaussian, Inc., Wallingford, CT, USA). Charges were computed by a Hartree–Fock calculation with a 6–31G(d)^[40] basis-set, according to the Merz–Singh–Kollman procedure.^[41,42] Finally, the structures of the compounds, together with charge values, were imported into AutoDockTools to automatically define rigid root and rotatable bonds.

To prepare the protein structure for docking calculations, three snapshots (one for each pharmacophoric model) were chosen. Their structures were imported into AutoDockTools and manipulated by removing nonpolar hydrogens, while Kollman united-atom partial charges and solvent parameters were added.

The Lamarkian Genetic Algorithm (LGA)^[43] was used to perform docking simulations. For each compound, the following protocol was applied: 200 independent LGA runs, a population size of 400 individuals, and a maximum number of 1 000 000 energy evaluations. Results differing by less than 1 Å in positional root mean square deviations were clustered together. Results of the docking simulations were analyzed on the basis of the cluster analyses and the values of the binding/docking energy. Residues involved in the binding of MAS0 are shown in Figure S4.

Acknowledgements

We thank Eliane Schweizer, François Diederich, and Roger S. Goody for their help and advice during initial stages of the project. The authors acknowledge funding by EC grant LSHG-CT-2003-503480–TRIoH. T.R. additionally acknowledges funding by a DFG grant.

Keywords: antiviral agents • dimerization • drug design • HIV • reverse transcriptase

- [1] V. Vivet-Boudou, J. Didierjean, C. Isel, R. Marquet, *Cell. Mol. Life Sci.* **2006**, *63*, 163–186.
- [2] T. Restle, M. Pawlita, G. Sczakiel, B. Müller, R. S. Goody, *J. Biol. Chem.* **1992**, *267*, 14654–14661.
- [3] L. A. Kohlstaedt, J. Wang, J. M. Friedman, P. A. Rice, T. A. Steitz, *Science* **1992**, *256*, 1783–1790.
- [4] T. Restle, B. Müller, R. S. Goody, *J. Biol. Chem.* **1990**, *265*, 8986–8988.
- [5] T. Restle, B. Müller, R. S. Goody, *FEBS Lett.* **1992**, *300*, 97–100.
- [6] T. Berg, *Angew. Chem.* **2003**, *115*, 2566–2586; *Angew. Chem. Int. Ed.* **2003**, *42*, 2462–2481.
- [7] M. J. Camarasa, S. Velazquez, A. San Felix, M. J. Perez-Perez, F. Gago, *Antiviral Res.* **2006**, *71*, 260–267.
- [8] S. Srivastava, N. Sluis-Cremer, G. Tachedjian, *Curr. Pharm. Des.* **2006**, *12*, 1879–1894.
- [9] G. Tachedjian, M. Orlova, S. G. Sarafianos, E. Arnold, S. P. Goff, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 7188–7193.
- [10] N. Sluis-Cremer, G. I. Dmitrienko, J. Balzarini, M. J. Camarasa, M. A. Parniak, *Biochemistry* **2000**, *39*, 1427–1433.
- [11] N. Sluis-Cremer, D. Arion, M. A. Parniak, *Mol. Pharmacol.* **2002**, *62*, 398–405.
- [12] M. J. Camarasa, S. Velazquez, A. San Felix, M. J. Perez-Perez, M. C. Bonache, S. De Castro, *Curr. Pharm. Des.* **2006**, *12*, 1895–1907.
- [13] G. Divita, T. Restle, R. S. Goody, J. C. Chermann, J. G. Baillon, *J. Biol. Chem.* **1994**, *269*, 13080–13083.
- [14] M. C. Morris, C. Berducou, J. Mery, F. Heitz, G. Divita, *Biochemistry* **1999**, *38*, 15097–15103.
- [15] M. C. Morris, V. Robert-Hebmann, L. Chaloin, J. Mery, F. Heitz, C. Devaux, R. S. Goody, G. Divita, *J. Biol. Chem.* **1999**, *274*, 24941–24946.
- [16] J. Depollier, M. L. Hourdou, G. Aldrian-Herrada, P. Rothwell, T. Restle, G. Divita, *Biochemistry* **2005**, *44*, 1909–1918.
- [17] G. Divita, T. Restle, R. S. Goody, *FEBS Lett.* **1993**, *324*, 153–158.
- [18] G. Divita, K. Rittinger, C. Geourjon, G. Deleage, R. S. Goody, *J. Mol. Biol.* **1995**, *245*, 508–521.
- [19] G. Tachedjian, H. E. Aronson, M. de Los Santos, J. Seehra, J. M. McCoy, S. P. Goff, *J. Mol. Biol.* **2003**, *326*, 381–396.
- [20] A. Mulky, S. G. Sarafianos, Y. Jia, E. Arnold, J. C. Kappes, *J. Mol. Biol.* **2005**, *349*, 673–684.
- [21] J. Ren, R. Esnouf, E. Garman, D. Somers, C. Ross, I. Kirby, J. Keeling, G. Darby, Y. Jones, D. Stuart, *Nat. Struct. Biol.* **1995**, *2*, 293–302.
- [22] P. J. Goodford, *J. Med. Chem.* **1985**, *28*, 849–857.
- [23] D. N. Boobbyer, P. J. Goodford, P. M. McWhinnie, R. C. Wade, *J. Med. Chem.* **1989**, *32*, 1083–1094.
- [24] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Delivery Rev.* **1997**, *23*, 3–25.
- [25] C. A. Lipinski, *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235–249.
- [26] G. Divita, K. Rittinger, T. Restle, U. Immendorfer, R. S. Goody, *Biochemistry* **1995**, *34*, 16337–16346.
- [27] J. Wang, S. J. Smerdon, J. Jager, L. A. Kohlstaedt, P. A. Rice, J. M. Friedman, T. A. Steitz, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 7242–7246.
- [28] S. Werner, B. M. Wöhr, *J. Virol.* **2000**, *74*, 3245–3252.
- [29] J. Ren, L. E. Bird, P. P. Chamberlain, G. B. Stewart-Jones, D. I. Stuart, D. K. Stammers, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14410–14415.
- [30] R. Pauwels, K. Andries, J. Desmyter, D. Schols, M. J. Kukla, H. J. Breslin, A. Raeymaeckers, J. Van Gelder, R. Woestenborghs, J. Heykants, K. Schellekens, M. Janssen, E. De Clercq, P. A. J. Janssen, *Nature* **1990**, *343*, 470–474.
- [31] O. Kensch, T. Restle, B. M. Wöhr, R. S. Goody, H. J. Steinhoff, *J. Mol. Biol.* **2000**, *301*, 1029–1039.
- [32] S. F. Le Grice, T. Naas, B. Wohlgensinger, O. Schatz, *EMBO J.* **1991**, *10*, 3905–3911.
- [33] U. Certa, W. Bannwarth, D. Stuber, R. Gentz, M. Lanzer, S. F. Le Grice, F. Guillot, I. Wendler, G. Hunsmann, H. Bujard, *EMBO J.* **1986**, *5*, 3051–3056.
- [34] B. Müller, T. Restle, S. Weiss, M. Gautel, G. Sczakiel, R. S. Goody, *J. Biol. Chem.* **1989**, *264*, 13975–13978.
- [35] B. Müller, T. Restle, H. Kühnel, R. S. Goody, *J. Biol. Chem.* **1991**, *266*, 14709–14713.
- [36] B. M. Wöhr, R. Krebs, S. H. Thrall, S. F. Le Grice, A. J. Scheidig, R. S. Goody, *J. Biol. Chem.* **1997**, *272*, 17581–17587.
- [37] L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, K. Schulten, *J. Comput. Phys.* **1999**, *151*, 283–312.
- [38] A. D. Mackerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kucera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- [39] M. J. S. Dewar, E. G. Zebisch, E. F. Healy, J. P. Stewart, *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- [40] G. A. Petersson, A. Bennett, T. G. Tensfeldt, M. A. Al Laham, W. A. Shirley, J. Mantzaris, *J. Chem. Phys.* **1988**, *89*, 2193–2218.
- [41] B. H. Besler, K. M. Merz, Jr., P. A. Kollman, *J. Comput. Chem.* **1990**, *11*, 431–439.
- [42] U. C. Singh, P. A. Kollman, *J. Comput. Chem.* **1984**, *5*, 129–145.
- [43] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639–1662.

Received: November 7, 2007

Published online on March 3, 2008