

Discovery of a Tat HIV-1 inhibitor through computer-aided drug design

Didier Esquieu^a, Jean-Marie Péloponèse^a, Sandrine Opi^a, Catherine Gregoire^a, Jean de Mareuil^a, Grant Campbell^d, Jean-Pierre Dunot^a, James Sturgis^a, Myriam Witvrouw^b, Christophe Pannecouque^b, Erik de Clercq^b, Mickaël Montembault^c, Vo-Thanh Giang^c, Monique Villiéras^c, Valérie Fargeas^c, Jacques Lebreton^c and Erwann P. Loret^{a,*}

^a CNRS UPR 9027, Institut de Biologie Structurale et Microbiologie, 31 Chemin Joseph Aiguier, 13402 Marseille, France

^b Rega Institute for Medical Research, K.U. Leuven Minderbroedersstraat 10, B-3000 Leuven, Belgium

^c CNRS UMR 6512, Faculté des Sciences et des Techniques, 2 rue de la Houssinière, 44222 Nantes, France

^d CNRS UMR 6032, Faculté de Pharmacie, 27 Bd Jean Moulin, 13385 Marseille, France

Abstract. Tat is a regulatory HIV-1 protein, which has the particularity to be secreted very early by HIV-infected cells. The extra cellular roles of Tat are suspected to be the main cause of the maintenance of reservoirs of HIV-infected cells and the failure of actual AIDS therapies to eradicate HIV. This study describes the rationale used to design molecules that bind to a target area containing an hydrophobic pocket identified in the 2D-NMR structure of Tat. Molecules were synthesized and the derivative named TDS2 was shown to be a Tat inhibitor. Fluorescence revealed that TDS2 binds in the target area, which is conserved across five different Tat variants representative of the main HIV-1 subtypes. TDS2 inhibited *in vitro* HIV-1 replication in human T-cells. Further chemical modifications remain necessary to enhance affinity to Tat and reduce cytotoxicity.

1. Introduction

AIDS is due to the loss of CD4 cells resulting from HIV cytotoxicity but is also related to the extra-cellular effect of Tat, which is a toxic protein secreted early by HIV-infected cells [1,2]. Tat was first identified as a regulatory protein essential for the HIV viral cycle due to its ability to dramatically increase the HIV gene expression [3]. However, this function in the viral cycle, called transactivation, is now considered of diminished significance since the discovery of the extracellular functions of the Tat protein [4]. Tat acts on different cell lines such as macrophages and cytotoxic T-lymphocytes (CTL), which are essential for the cellular response of the immune system to eliminate virus-infected cells [2]. The extracellular roles of Tat are probably the major reason for the maintenance of HIV-infected cells (or reservoir cells), and the failure of current antiviral therapies to eradicate HIV [1,2,5].

Tat is a short protein and its size can vary from 86 to 105 residues, although the majority of Tat variants have a size of 101 residues [2]. Tat is a flexible protein and this particularity provides a structural heterogeneity for this protein in hydrophilic versus hydrophobic environments and explains its capacity to cross membranes [6]. The main functional regions in Tat are a cysteine-rich region and a basic region [2]. The trans-activation activity of Tat requires the binding of Tat to a nascent leader RNA hairpin, called TAR, located at the 5' end of all HIV-1 mRNAs, thus enabling the synthesis of full-length viral transcripts

*Corresponding author. E-mail: loret@ibsm.cnrs-mrs.fr.

by RNA polymerase II [2]. Tat binds to TAR with a nanomolar affinity and its basic region needs to fit into the TAR major groove for transactivation [7]. The two regions adjacent to the basic regions are highly folded and could play a hinge-like role to allow the basic region to fit into TAR [8]. It was therefore postulated that Tat needs to change its structure in order to bind to TAR [8]. This hypothesis, suggesting that structural changes are required for Tat function, led to the proposition that a ligand could block Tat function by disturbing its structural flexibility [8].

Preliminary work was required before proceeding to the design of our Tat inhibitors. The first step was to obtain large amounts of various Tat proteins with biological activity to do structural studies. We developed a procedure, using solid phase synthesis, which made it possible to obtain six different Tat variants with biological activities [6]. The second step was to determine the three-dimensional structure of active Tat proteins. A 2D-NMR study was carried out on an European Tat variant [9] and on an African Tat variant [10]. A similar folding was observed for these two variants. However, local variations became apparent such as a short α -helix of seven residues in the African variant while there is no such α -helix in the European variant [10].

2. Materials and methods

2.1. Protein synthesis, purification and characterisation

Tat variants were synthesised with an ABI 433A peptide synthesizer (Perkin Elmer, Applied Biosystem Inc.) with FASTMoc chemistry according to the method of Barany and Merrifield [11] on 4-hydroxymethyl-phenoxy-methyl-copolystyrene-1% divinylbenzene preloaded resin (HMP; 0.50–0.65 mmol; Perkin Elmer, Applied Biosystem Inc., Foster City, CA), as previously described [6]. Purity and integrity of proteins were confirmed by amino acid composition (6300 Beckman analyser), partial sequence analyses (473A Protein Sequencer, Applied Biosystem) and by MALDI-TOF mass spectrometry (Perspective Biosystems, Voyager DE-RP).

2.2. Molecular modeling of the Tat–TDS2 complex

The model of the Tat–TDS2 complex was built with the Insight II software from MSI Technologies, Inc. (San Diego, CA), running on a Silicon Graphics O2 workstation. Tat structures (PDB entries 1FKU and 1JFK) were obtained from NMR studies [9,10]. Docking of the two molecules was optimized with the Consistent Valence Force Field (CVFF) in terms of the internal energies, using the van der Waals energy to monitor each step of the model. Minimisation was performed with steepest descent and conjugate gradient algorithms. Dynamic was performed at 200 K for 1.1 ps using 1000 steps.

2.3. TDS binding monitored by fluorescence

Fluorescence spectra of compounds in the presence and absence of the Tat protein were obtained with a Safas (Monaco) spectrofluorometer. An excitation wavelength of 280 nm was used, and emission spectra were recorded from 350 to 500 nm. A slit width of 5 nm was used for the fluorescence excitation and emission recordings. All fluorescence experiments were carried out in 100 mM phosphate buffer and 100 mM NaCl, pH 6, at room temperature. The fluorescence emission was recorded after the spectroscopic signal had been left to stabilise for 30 min. Baseline spectra of Tat were subtracted from the resulting spectra of Tat with the different compounds.

2.4. Inhibition of HIV-1 infection

Human T-cell lines H9 ($3 \cdot 10^5$ /ml) or PBMCs ($2 \cdot 10^5$ /200 μ l) were plated in the presence of different concentrations of compounds and were infected respectively with HIV-1_{NL4-3} or HIV-1_{IIIB} at 1000 CCID₅₀/ml. At days 4 post infection, the supernatant of the infected cultures was removed and replaced with fresh medium containing the compounds at the appropriate concentrations. At days 7 post infection, p24 antigen was detected in the culture supernatant by an enzyme-linked immunosorbent assay (NEN, Dreieich, Germany). The 50% cytotoxic concentration (CC₅₀) was determined by cell counting and trypan-blue exclusion assay and by tetrazolium-based colorimetric assay. The 50% virus-inhibitory concentration (IC₅₀) required to inhibit HIV replication in CEM and PBMCs was determined by p24 measurements in the cell culture supernatant. Experiments were done in triplicate.

3. Results and discussion

In this study, we report the synthesis of four new Tat variants with biological activity. It was necessary to synthesise these new Tat variants because the Tat variants previously synthesised were representative of only B and D HIV-1 subtypes [6]. At the beginning of the nineties the AIDS epidemic was mainly situated in Central Africa, North America and Western Europe. D subtypes tended to be prevalent in Africa, while B subtype were (and still are) prevalent in North America and Western Europe. Changes appeared in the AIDS epidemic during the last five years with a burst of contamination in different parts of the world that are associated to other HIV-1 subtypes, such as subtype A (Central Africa), subtype C (South America, South-Africa and India) or subtype E (South-East Asia) [12]. Figure 1 shows the sequences of the five Tat variants selected to represent the actual predominant HIV-1 subtypes all over the world. These are long-size variants from 99 to 101 residues since the short-size variants from 86 to 87 residues tend to have disappeared [2]. These five synthetic Tat variants have a biological activity similar to wild Tat, which was revealed by their capacity to cross membranes and the transactivation of the HIV LTR in HeLa P4 cells (Opi et al., in press).

The choice of the target area on Tat was based on the identification of a cluster of well-conserved residues in the Tat 3D structure. Figure 1 shows that Tat can tolerate up to 38% mutations and still retains its biological activity. However there is a sequence from Leu 43 to Arg 53 that remains almost unchanged. This sequence adopts a similar folding in the two Tat 3D structures [9,10], which is characterised by a type II β -turn that includes the sequence ⁴⁴GISY. Figure 2 shows that in the 3D structure, this sequence

Tat Exon 1	I	II	III	IV	V
HXB2	MEPVDPRLPEPWKHPGSQPKTA	CTNCYCKKCCFHCQVCF	ITKALGISYGR	KKRRQRRAH	QNSQTHQASLSKQ
92BRN.....N.....R.SY..L..	Q..G.....S.PPS.ED..	NPIP..
ELI	.D.....N.....N.....R.P.	NK.H.....Y..P..	LN.G.....GPP.	GG.A..VPIP..
CM240	..L.....N.....N.....T..	SK.....W...L..	LK.G.....KH..GTP.	S.KD..NPIP..
UG11RP	.D.....NI...N.....T.P.	NK.....V..Y..L..	QS.G.....K...GPT.	SNKQ..NPIP..

Tat Exon 2	VI	
HXB2	PTSQPRGDPTGPKE-KKKVERETETDPPD	100
92BR	.LP.T...Q..SE.S....SK.....	101
ELI	.S.....Q.....S.A....	99
CM240	.LP.II.RN..D...S..E.ASKA...QC.	101
UG11RP	.IPRTQ.IS...E.S.....DK...RR.	101

Fig. 1. Tat variant sequences representative of different HIV-1 subtypes. These five Tat variants were obtained by solid phase synthesis with a procedure previously described [6] and possess Tat biological activity. Tat is divided in six regions that are shown inside boxes and between boxes.

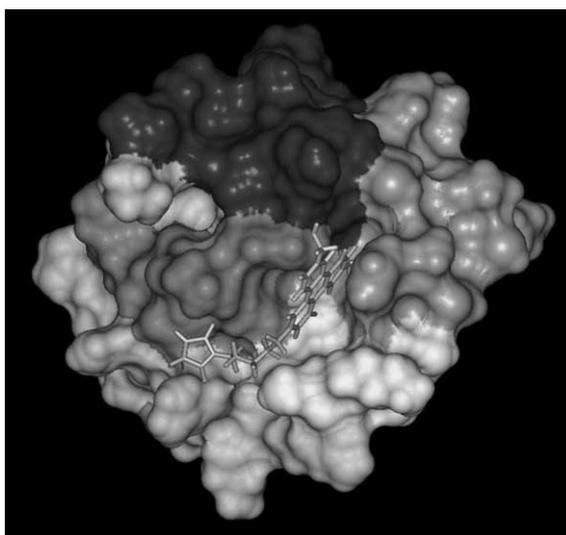


Fig. 2. Tat–TDS2 complex. The skeleton of TDS2 is represented in the pharmacological binding site of Tat, which is shown with a Connolly surface representation. Tat was divided in six regions regarding chemical properties of amino acid residues (see Fig. 1). Region 1 is colored in red, region 2 in orange, region 3 in yellow, region 4 in green, region 5 in light blue and region 6 in blue. The main functional regions of Tat are the cysteine-rich (orange) and the basic (green) regions.

is located near the N-terminal region containing strictly conserved residues such as Trp 11 (Fig. 1). This cluster of well-conserved residues constitute an hydrophobic pocket, both in the European and the African Tat variants [9,10], which is located between the two functional regions of Tat (coloured in orange and green in Fig. 2). It was interesting to use this hydrophobic environment to position a ligand interacting with the two functional regions of Tat, as previously proposed [8].

The triphenylene ring was selected from a set of core chemical compounds (or scaffolds) constituting parts of drugs that passed successfully pre-clinical toxicity tests. The triphenylene ring had the appropriate size to fit into the hydrophobic pocket of Tat (Fig. 2). Triphenylene is not soluble in physiological conditions and functional groups (or probes) are added to triphenylene to afford aqueous solubility and specificity to Tat. It was expected that triphenylene could stabilize the binding due to a hydrophobic effect resulting from an entropy increase [13]. Probes were selected to be proton donor or acceptor and to induce hydrogen bonds with the basic region and the cysteine-rich regions. The first probe to be selected was a succinimide group and it was connected to the triphenylene ring in R1 by an aliphatic chain of five carbons (Fig. 3). The two oxygen of the succinimide had the appropriate orientation for two hydrogen bonds within the basic region and the N-terminal region (Fig. 2). The second probe was a hydroxyl group, which was connected by only one aliphatic carbon to the triphenylene ring in R2. The hydroxyl group was in the appropriate location to create an hydrogen bond with the cysteine-rich region.

This molecule, called TDS2, was synthesised with other derivatives possessing the triphenylene ring (Montembault et al., manuscript submitted). To verify the binding on Tat, we chose to monitor the interaction by fluorescence associated with the triphenylene ring. Figure 4 shows that triphenylene has a high fluorescence signal that can be slightly modified regarding the compounds described in Fig. 3. The fluorescence signal of Tat was much lower (Fig. 4) and it was difficult to use the fluorescence quenching of Trp 11 to monitor the TDS2 binding to Tat. Moreover, the quenching of the fluorescence of a Trp by a ligand can be observed to a distance up to 10 nm [14] and could not be due necessarily to a binding with the triphenylene in the hydrophobic pocket. A high number of fluorescence molecules have a dramatic

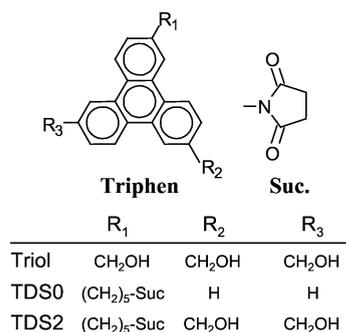


Fig. 3. Structure of the TDS compounds. The scaffold of the TDS molecules is based on the triphenylene with the variation of the probes R₁, R₂ and R₃ shown in the table. Suc: succinimide. Triol: 2,6,10-tri(hydroxymethyl)triphenylene, TDS0: 2-(5-Nsuccinimide-pentyl)triphenylene and TDS2: 2-(5Nsuccinimide-pentyl)-6,10-di(hydroxymethyl)triphenylene.

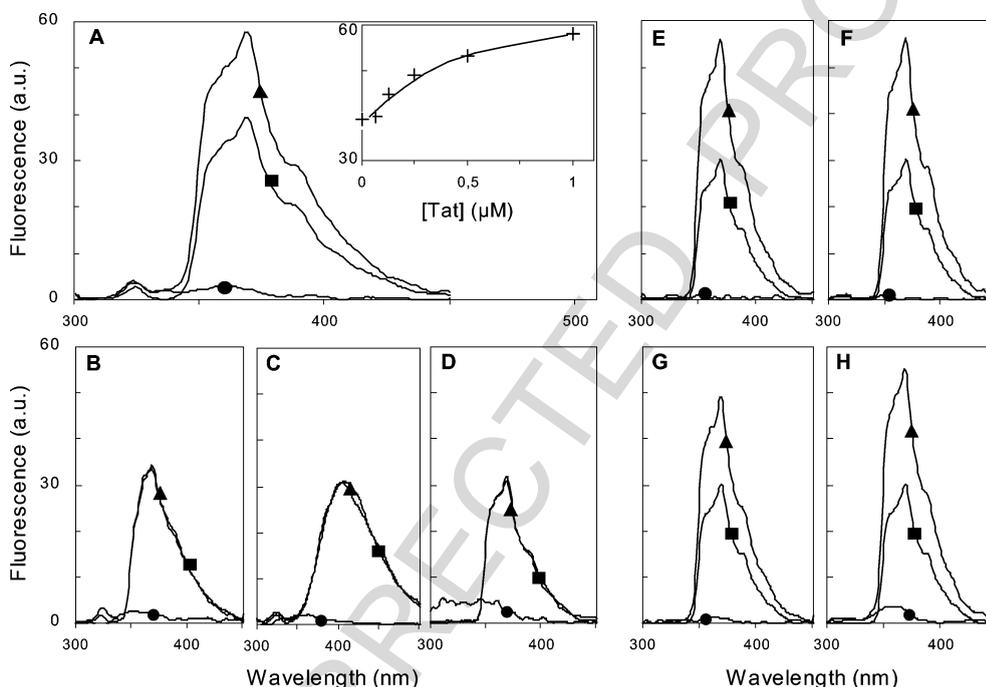


Fig. 4. Fluorescence emission spectra of TDS2 (panels A, D, E, F, G and H), TDS0 (panel B) and Triol (panel C). Squares represent compounds alone at 1 μM , triangles represent the compounds with different Tat variants at 1 μM (Tat HXB2: A, B and C; Tat Ug11RP: E; Tat 92Br: F; Tat CM240: G; Tat Eli: H) and circles represent Tat variants alone. Panel D represents TDS2 emission in presence of human serum albumin at 1 μM . The fluorescence emission at 369 nm of TDS2 (1 μM) with varying concentration of Tat HXB2 from 0.05 μM to 1 μM is shown in the corner of panel A.

enhancement of fluorescence when they pass from a polar to a non-polar environment [14]. This enhancement was observed with TDS2 in the presence of Tat (Fig. 4A) and only the binding of TDS2 with its triphenylene ring in the hydrophobic pocket (Fig. 2) can explain this enhancement of fluorescence. TDS2 binds to Tat with an apparent K_a of 0.33 μM^{-1} (Fig. 4A). Interestingly, no fluorescence enhancement was observed with TDS0 and Triol (Fig. 4B,C). Moreover, only TDS2 showed activity against HIV laboratory strains as determined on lymphocyte T-cell lines (H9) and peripheral blood mononuclear cells (PBMCs)

Table 1
Anti-HIV effect of the TDS compounds

	H9		PBMC	
	IC ₅₀ μ M	CC ₅₀ μ M	IC ₅₀ μ M	CC ₅₀ μ M
Triol	ND	ND	ND	ND
TDS2	0.5	10	0.8	1.2
TDS0	>100	>100	ND	ND

ND: Not determined.

IC₅₀, concentration of TDS compounds which inhibited 50% of infection, and CC₅₀, 50% cytotoxic concentration, were determined in human H9 T-cell lines and peripheral blood monocyte cells (PBMC).

(Table 1). No fluorescence enhancement was observed with TDS2 in the presence of human serum albumin (Fig. 4D). Fluorescence experiments indicated that TDS2 can bind similarly to five Tat variants (Fig. 4A,E–H), suggesting that the hydrophobic pocket is well-conserved among Tat variants (Fig. 2).

The first compound to be synthesized (TDS1) had the succinimide probe in R1 and a methyl in R2 and R3, and turned out to be a hit with a K_a of $33 \mu\text{M}^{-1}$ on Tat (Montembault et al., manuscript submitted). It is interesting to note that although a methyl cannot be considered as a functional group, if there is a hydrogen in R2 and R3, there is no binding to Tat (Fig. 4). We tried to improve the TDS binding with the synthesis of derivatives with variable aliphatic chains connecting the succinimide probe with no success (data not shown). Different scaffolds were tested with the same probes but no binding to Tat was observed except for a derivative with a benzene ring, which had a very low affinity for Tat compared to TDS1 (data not shown). It is probable that the triphenylene ring provides an affinity for the target and does not play just the role of a scaffold. The second step was to test a hydroxyl probe in R2, and we obtained our best hit up to now, which is TDS2 (Table 1). Our third step will be to test a third probe in R3 able to hydrogen bonding with the guanidinium group of Arg 78. This residue located in the RGD motif of Tat is not strictly conserved since a mutation of Arg 78 \rightarrow Gln is observed in E-subtypes (Fig. 1). However, molecular modeling shows that the amine of the lateral chain of Gln 78 could be in a similar location as one of the amines of the guanidinium group for Arg 78 Tat variants (data not shown).

Our goal was to create a low-molecular-weight compound that binds specifically to Tat. TDS2 (MW 455) has no amide group, no charged group and no stereogenic centre. Moreover, TDS2 binds to different Tat variants and its binding site consists of well-conserved residues that should prevent development of HIV resistance. We have shown that TDS2 is capable of blocking HIV infection *in vitro*, however, toxicity was observed, particularly in PBMCs (Table 1). This toxicity was due to the hydroxyl groups in R2 and R3 since no toxicity, up to $100 \mu\text{M}$, was observed for TDS1 (Montembault et al., manuscript submitted) or TDS0 (Table 1). We envisage that, by replacing the hydroxyl group in R3 with another group, we will be enabling new hydrogen bonding, thereby increasing Tat affinity and lowering cytotoxicity.

Acknowledgements

We thank the American Scientific Committee represented by Drs David Ho, Michael Saag and John Bartlett to have selected the TDS project for the GlaxoSmithKline (GSK) Drug Discovery and Development Award. We thank GSK, Région Provence Alpes Côte D'Azur and the Agence Nationale de la Recherche sur le SIDA (ANRS) for their financial support. JL and ELP thank l'Agence Nationale de la Valorisation de la Recherche (ANVAR) for its financial support. The authors are deeply indebted to the Centre National de la Recherche Scientifique (CNRS) for their continuous support of the TDS project.

References

- [1] R.C. Gallo, *Proc. Natl. Acad. Sci. USA* **96** (1999), 8224–8226.
- [2] K.T. Jeang, H. Xiao and E.A. Rich, *J. Biol. Chem.* **274** (1999), 28827–28840.
- [3] A. Fisher, M. Feinberg, S. Josephs, M. Harper, L. Marselle, G. Reyes, M. Gonda, A. Aldovini, C. Debouk, R.C. Gallo et al., *Nature* **220** (1986), 267–271.
- [4] B. Ensoli, G. Barillari, S.Z. Salahuddin, R.C. Gallo and F. Wong-Staal, *Nature* **345** (1990), 84–86.
- [5] B. Ramratnam, J.E. Mittler, L. Zhang, D. Boden, A. Hurley, F. Fang, C.A. Macken, A.S. Perelson, M. Markowitz and D.D. Ho, *Nat. Med.* **1** (2000), 82–85.
- [6] J.-M. Péloponèse, Y. Collette, C. Gregoire, C. Bailly, D. Campèse, E. Meurs, D. Olive and E.P. Loret, *J. Biol. Chem.* **274** (1999), 11472–11479.
- [7] E.P. Loret, P. Georgel, W.C. Johnson and P.S. Ho, *Proc. Natl. Acad. Sci. USA* **89** (1992), 9724–9728.
- [8] C. Grégoire and E.P. Loret, *J. Biol. Chem.* **271** (1996), 22641–22646.
- [9] J.M. Péloponèse, C. Gregoire, S. Opi, D. Esquieu, J. Sturgis, E. Lebrun, E. Meurs, Y. Collette, D. Olive, A.M. Aubertin et al., *C. R. Acad. Sci. III* **323** (2000), 883–894.
- [10] C. Grégoire, J.M. Péloponèse, D. Esquieu, M. Solomiac, S. Opi, E. Lebrun, J. Lebreton and E.P. Loret, *Biopolymers* **62** (2001), 324–335.
- [11] G. Barany and R.B. Merrifield, in: *The Peptide: Analysis, Synthesis, Biology*, E. Gross and J. Meinhofer, eds, Vol. 2, Academic Press, New York, 1980, pp. 1–284.
- [12] S. Osmanov, C. Pattou, N. Walker, B. Schwardlander and J. Esparza, *J. Acquir. Immune Defic. Syndr.* **29** (2002), 184–190.
- [13] C.K. Mathews and K.E. van Holde, in: *Biochemistry*, S. Weisberg, ed., Benjamin/Cummings Inc., Redwood City, CA, 1990, pp. 192–193.
- [14] C.R. Cantor and E.R. Schimmel, in: *Biophysical Chemistry. Part II: Techniques for the Study of Biological Structure and Function*, W.H. Freeman Company, San Francisco, 1980, pp. 444–448.