# New Inhibitors of the Tat–TAR RNA Interaction Found with a "Fuzzy" Pharmacophore Model

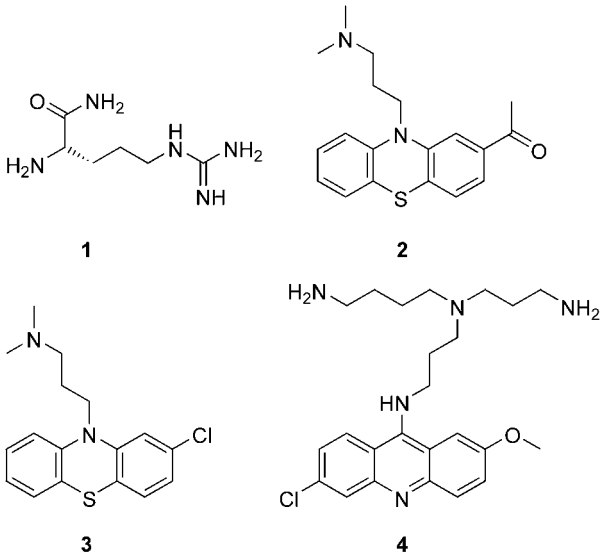
Steffen Renner, Verena Ludwig, Oliver Boden, Ute Scheffer, Michael Göbel, and Gisbert Schneider*[a]

*TAR RNA is a potential target for AIDS therapy. Ligand-based virtual screening was performed to retrieve novel scaffolds for RNA-binding molecules capable of inhibiting the Tat–TAR interaction, which is essential for HIV replication. We used a "fuzzy" pharmacophore approach (SQUID) and an alignment-free pharmacophore method (CATS3D) to carry out virtual screening of a vendor database of small molecules and to perform "scaffold-hopping". A small subset of 19 candidate molecules were experimentally tested for TAR RNA binding in a fluorescence resonance energy transfer (FRET) assay. Both methods retrieved molecules that exhibited activities comparable to those of the reference molecules acetylpromazine and chlorpromazine, with the best molecule showing ten times better binding behavior ($IC_{50}$ = 46 $\mu$M). The hits had molecular scaffolds different from those of the reference molecules.*

## Introduction

In recent years it has become clear that large numbers of RNAs fold into well defined three-dimensional structures, providing interfaces for specific intermolecular protein–RNA and small molecule–RNA interactions, which are essential parts of regulatory networks or antiinfective responses.[1] These findings have led to constantly increasing interest in RNA as a potential drug target with a plethora of potential applications,[2] and several natural and synthetic small molecules have been reported to interact specifically with RNA.[3] One of the best characterized RNA-based regulatory systems is the transactivation response element (TAR) of HIV mRNA.[4] Specific binding of the Tat protein to TAR is essential for virus replication. The TAR RNA therefore represents a potential target for HIV therapy as well as a model system to deepen understanding of RNA–small molecule interactions and the development of drugs for RNA targets in general. The structure of TAR consists of two rigid double-strand stems connected by a flexible bulge of three bases, which provides a specific binding pocket for the Tat protein.[4] A variety of molecules that inhibit the Tat–TAR interaction, and consequently virus replication, have been found;[5] such molecules include peptidic derivatives of the Tat binding motif, such as argininamide (**1**; Scheme 1), antibiotics such as neomycin, and a set of small molecules with unnatural scaffolds. Interestingly, most classes of bulge-binding ligands for which structures have been determined bind in distinct regions and stabilize very different conformations of the bulge.[6]

In searching for new drugs or leads, one is usually interested in finding druglike molecules with favorable pharmacokinetic profiles and oral bioavailability.[7] These conditions are not met by the majority of the ligands that inhibit the Tat–TAR interaction found so far;[8] most bear multiple charges and some are very large in comparison to commercial drugs. One interesting class of ligands found by virtual screening are promazines such as acetylpromazine (**2**), for which an NMR structure has been published,[6] and chlorpromazine (**3**).[6,9] RNA binding of this molecule is not dominated by charged interactions, and so its binding pocket might facilitate binding of classes of ligands with drug-like properties. We therefore decided to focus on finding new molecular scaffolds with binding modes similar to that of acetylpromazine.




**Scheme 1.** TAR–Tat interaction inhibitors. Argininamide (**1**), acetylpromazine (**2**), chlorpromazine (**3**) and CGP40336A (**4**).

[a] *Dipl.-Biol. S. Renner, Dipl.-Chem. V. Ludwig, Dipl.-Chem. O. Boden, Dr. U. Scheffer, Prof. Dr. M. Göbel, Prof. Dr. G. Schneider*
*Beilstein Endowed Chair for Cheminformatics*
*Johann-Wolfgang-Goethe-Universität*
*Institut für Organische Chemie und Chemische Biologie*
*Marie-Curie-Straße 11, 60439 Frankfurt am Main (Germany)*
*Fax: (+49) 69-798-29826*
*E-mail: gisbert.schneider@modlab.de*

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

 **DOI: 10.1002/cbic.200400376**  

Virtual screening has been shown to be useful for finding small, enriched sets of candidate molecules for further experimental testing.[10] For TAR, only structure-based virtual screening has so far been reported, an automated docking approach including a scoring function optimized for RNA resulted in the identification of acetylpromazine.[9,11] Other studies indicated that the inherent flexibility of RNA structures might limit the applicability of entirely structure-based approaches.[2c,12]

Ligand-based approaches are an alternative to structure-based virtual screening.[10] In particular, methods including the active-analogue model of pharmacophores have been shown to be suited for scaffold-hopping.[13] In this work we applied two of our recently published approaches—SQUID fuzzy pharmacophores and CATS3D two-point pharmacophore similarity searching—to find new Tat–TAR inhibitors.[14,15] Both approaches are based on correlation vectors (CVs) and are consequently invariant to translation and rotation of molecules, allowing rapid screening of large databases. The screening approaches were complemented by the use of a neural network drug-likeness filter.[16] CATS3D encodes the three-dimensional conformations of molecules in the form of scaled histograms ("fingerprints") which represent the frequencies of pairs of pharmacophoric features within defined distance ranges. Similar fingerprints code for molecules with pharmacophoric features in common and which might be isofunctional. SQUID fuzzy pharmacophore models approximate three-dimensional alignments of several ligands by spatial Gaussian probability densities for the presence of pharmacophoric features. Features that are more conserved in the alignment are assigned higher weights than less conserved features. The "fuzziness" parameter can be used to represent the underlying alignment by different resolutions. For virtual database screening these probability densities are transformed into two-point correlation vectors.

## Computational Methods

### Preparation of the virtual screening database

We screened the SPECS commercial library (June 2003 version), consisting of 229658 compounds.[17] To obtain higher quality results and to restrict the calculation of 3D conformations we selected the 20000 most druglike molecules as predicted by an artificial neural network.[16] For each of these molecules up to 20 3D conformations were calculated in MOE[18] with default settings and the MMFF94 forcefield.[19] All molecules were neutralized before descriptors were calculated.

### Alignment of inhibitors

For the alignment of the known reference Tat–TAR interaction inhibitors we used the flexible alignment tool of MOE[18] with default settings and the MMFF94 forcefield.[19]

### Calculation of the CATS3D descriptor

CATS3D encodes the conformation of a molecule in the form of a histogram or correlation vector that contains the normalized frequencies of all possible pairs of atom types in a molecule.[15] All atom pairs were subdivided into groups characterized by atom–atom distance ranges and six different pharmacophore types (Figure 1). We used 20 equal distance bins, from 0–20 Å. One of the pharmacophore types—cation, anion, hydrogen-bond acceptor, hydrogen-bond donor, polar (hydrogen-bond acceptor AND hydrogen-bond donor), or hydrophobic—was assigned to each atom with the ph4_aType function of MOE.[18] Atoms with no matching pharmacophore type were not considered further. The use of 20 distance bins for each of the 21 possible combinations of pairs of pharmacophore points resulted in a descriptor of 420 dimensions. The value stored in each bin was scaled by the added incidences of the two respective features. Each dimension ("bin") of the CATS3D CV was calculated according to Equation 1,

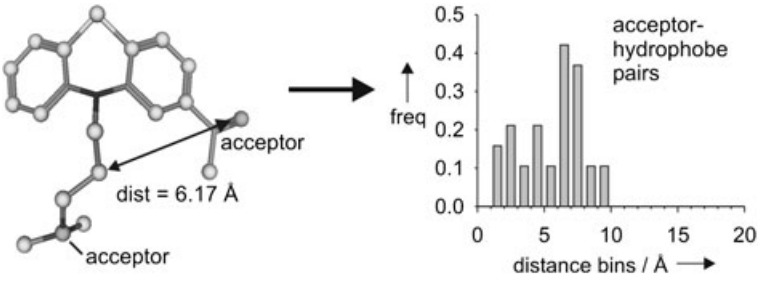



**Figure 1.** Calculation of the CATS3D correlation vector. Spheres denote potential pharmacophore points (two hydrogen bond acceptors, 17 hydrophobic PPPs). Distances are measured between all pairs of atoms, and frequencies (freq) of pairs are determined for all pairs of pharmacophoric types and for defined distance ranges ("bins"). As an example, a section of the resulting CV representing hydrogen-bond acceptor–hydrophobe pairs is shown.

$$CV_d^T = \frac{1}{N_1 + N_2} \sum_i \sum_j \frac{1}{2} \delta_{ij,d}^T \qquad (1)$$

where $i$ and $j$ are atom indices, $d$ is a distance range, $T$ is the pair of pharmacophoric types of atoms $i$ and $j$, $N_1$ and $N_2$ are the total number of atoms of types of $i$ and $j$ present in a molecule, and $\delta_d^T$ (Kronecker delta) evaluates to 1 for all pairs of atoms of type $T$ within the distance range $d$. The factor of 0.5 in the sum avoids double counting of pairs. Pairs of atoms with themselves were not considered. When CATS3D was used to encode molecules for SQUID database screening, the final descriptor vector was scaled to a maximum of 1.

### Calculation of the fuzzy pharmacophore model

A SQUID fuzzy pharmacophore model approximates the spatial distribution of pharmacophoric points in an alignment of molecules by a set of generalized potential pharmacophore points (PPPs) of Gaussian probability densities.[14] Atoms in the alignment comprising the same pharmacophoric features are clus-

tered into PPPs for a more general and "fuzzy" representation of the major characteristics of the alignment. The resolution of the model is defined by the cluster radius, which is the parameter that affects how strictly features are clustered into PPPs. The ideal resolution of the pharmacophore model has to be determined separately for each set of aligned ligands.

Each PPP in the pharmacophore model is represented by four attributes. The first of these is the pharmacophore type of the atoms represented by the PPP, the second is the PPP position in 3D space, while the third is the standard deviation ($\sigma$) characterizing the width of the distribution of the atoms represented by a PPP (in graphical illustrations of SQUID pharmacophore models $\sigma$ is represented by the radii of the PPPs). The fourth attribute (the conservation weight $w$) weights each PPP by its conservation among the molecules of the alignment (in graphical illustrations of SQUID pharmacophore models $w$ is represented by the intensity of the color of a PPP). This is done under the assumption that more conserved features of a set of molecules binding to the same receptor with comparable affinity are more important for the binding than less conserved features.

A schematic overview of the calculation of a SQUID pharmacophore is given in Figure 2. The starting point is an alignment of known active reference compounds. Assignment of pharmacophoric types (cation, anion, hydrogen-bond acceptor, hydrogen-bond donor, polar, or hydrophobic) transforms the alignment into a field of pharmacophoric features. Maxima in the local feature densities (LFDs) were used as cluster seeds to cluster the features into PPPs for a more general representation of the underlying alignment. For each atom $k$ of type $t$ in the alignment the LFD was calculated by Equation 2,

$$\mathrm{LFD}(\mathrm{atom}_k^t) = \sum_i \max\left\{0,\ 1-\frac{D_2(\mathrm{atom}_k^t,\mathrm{atom}_i^t)}{r_c}\right\} \qquad (2)$$

where $i$ are all atoms of type $t$ in the molecular ensemble, $D_2$ is the Euclidean distance between two atoms, and $r_c$ is the cluster radius. Positions of atoms of type $t$ for which no other atom of type $t$ within $r_c$ yielding a higher LFD was found were taken as cluster seeds for PPPs of type $t$. All atoms were subsequently clustered to the nearest cluster seed of their respective type. The geometric centre of the atoms in a cluster was taken as the position of the resulting PPP, whilst the median distance from all atoms contributing to a PPP to the centre of the PPP was taken as the value of the standard deviation ($\sigma$) of the PPP. For this value a minimum of 0.5 Å was used. The conservation weights of the PPPs were calculated by Equation 3,

$$w(\mathrm{PPP}_k) = \sum_{i=1}^{m} \min\left\{\frac{1}{m}, \frac{\text{no. atoms from molecule}_i \text{ of PPP}_k}{\text{no. atoms of PPP}_k}\right\} \qquad (3)$$

where $m$ is the number of molecules in the model. This function returns a maximum value of 1 for PPPs representing the same number of atoms from all molecules of the ensemble, and a minimum of $m^{-1}$ for PPPs consisting only of atoms of one molecule.

For virtual screening the three-dimensional distribution of the PPPs is transformed into a two-point PPP CV (Figure 2), arranged exactly like the CATS3D CV. The PPP CV represents the

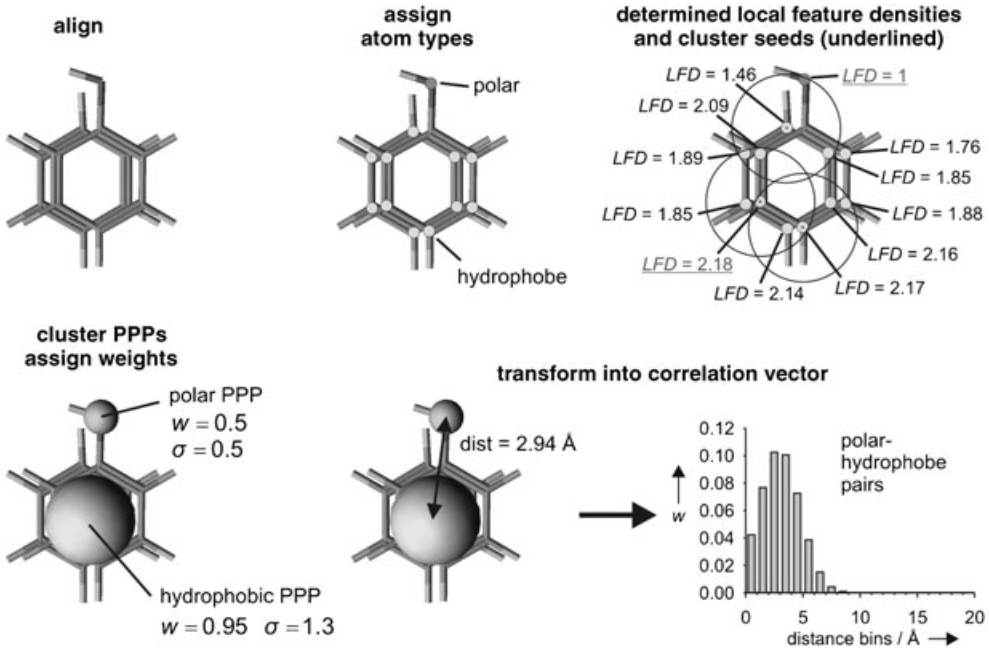



**Figure 2.** Calculation of the SQUID fuzzy pharmacophore correlation vector. Pharmacophore atom types are assigned to all atoms of a set of aligned molecules. Maxima in the LFDs are determined for use as cluster seeds. In this example a cluster radius ($r_c$) of 1.5 Å was used. Standard deviations ($\sigma$) and conservation weights ($w$) are calculated for each PPP resulting from the clustering procedure. Finally, distances between all pairs of PPPs are measured and the three-dimensional representation is transformed into a correlation vector by application of Equation 4. This results in conservation weights ($w$) for pairs of PPPs in the CV. A section of the resulting CV representing polar–hydrophobe pairs is shown as an example.

three-dimensional distribution of Gaussian densities in the form of the distribution of pairs of PPPs over the distance bins and over the feature types. The transformation is calculated according to Equation 4,

$$CV_d^T = \frac{1}{\text{no. pairs}(T)} \sum_p \sum_q \frac{1}{2}\delta_{pq}^T \left( \frac{w_p w_q}{\sqrt{2\pi}(\sigma_p \sigma_q)} \exp\left( -\frac{1}{2} \frac{(D_2(p,q) - centre_d)^2}{(\sigma_p + \sigma_q)^2} \right) \right) \quad (4)$$

where $p$ and $q$ are PPPs, $d$ is a distance range ("bin"), $T$ is the pair of pharmacophoric types of $p$ and $q$ (e.g., Figure 2: $p$ = hydrophobic, $q$ = polar), $w$ are the PPP conservation weights, $\sigma$ is the standard deviation of a PPP, $centre_d$ is the centre of the distance range $d$, and $\delta^T$ (Kronecker delta) evaluates to 1 for all pairs of PPPs of types $T$. $D_2$ is the Euclidean distance metric. The factor of 0.5 in the sum avoids double counting of pairs. Pairs of PPPs with themselves were not considered. The values of each dimension were scaled by the total number of possible pairs of PPPs of the two features considered. Finally the CV was scaled to a maximum value of 1.

Like the CATS3D descriptor, the SQUID CV consisted of 420 dimensions, representing the same distance bins and pairs of atom types. The SQUID CV was used to rank molecules encoded with the CATS3D descriptor.

To obtain optimal virtual screening results, additional weights ("feature-type weights") were used to weight the importance of each of the pharmacophoric feature types (e.g., hydrophobic or hydrogen bond donor) in the CV. The sums of the single feature-type weights were used to weight the importance of each pair of feature types in the CV, and the sum of the probabilities in the CV for each pair of features over all distance bins was scaled to the value of the feature-type weights. Finally the whole CV was scaled to a maximum of 1. It was found that a simple optimization by permutation of all combinations of the weight values {0.1, 0.2, 0.3, 0.4, 0.5} for each of the single features and subsequent testing of these weights in virtual screening was sufficient to retrieve good virtual screening results.[14]

#### Virtual screening

For virtual screening all entries in the screening database were scored according to their similarity to the query. For each molecule all but the best scoring conformation were removed. For screening with CATS3D we used the Manhattan distance [Eq. (5)] to rank database compounds according to their distance to the query structures. Lower distance values resulted in higher (better) ranking positions.

$$D_{A,B} = \sum_{i=1}^{i=N} |x_{iA} - x_{iB}| \quad (5)$$

In Equation (5), $A$ and $B$ indicate the CATS3D vectors of two molecules, $x_i$ is the value of vector element $i$, and $N$ is the total number of vector elements ($N$ = 420).

For ranking of the molecules of the screening database by SQUID pharmacophore models we used the SQUID similarity score [Eq. (6)],

$$S_{A,B} = \frac{\sum_{i=1}^{i=N} (x_{iA} x_{iB})}{1 + \sum_{i=1}^{i=N} ((1 - x_{iA}) x_{iB})} \quad (6)$$

where $A$ is the CV of the SQUID pharmacophore and $B$ is the CATS3D CV of a molecule, $x_i$ is the $i$th element of a vector and $N$ is the total number of vector elements. The value $x_{iA}$ could be regarded as the "idealized probability" of the presence of atom features in $x_{iB}$. This results in high scores for molecules with many features in regions of the query descriptor which have a high probability. The denominator, in contrast, penalizes the presence of such pairs in regions with a low probability in active molecules.

## Results and Discussion

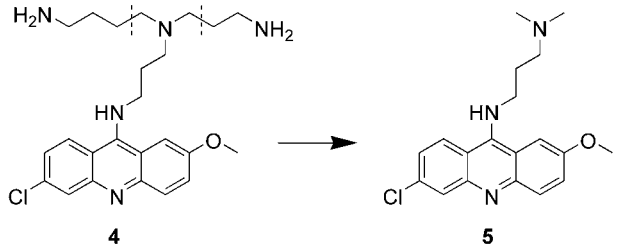
The SPECS compound set, containing 229658 screening compounds, was virtually screened for potential inhibitors of the Tat–TAR interaction. Virtual screening consisted of three steps: i) calculation of a "drug-likeness" score by an artificial neural network as a prescreening step, ii) CATS3D pharmacophore similarity searching, and iii) SQUID pharmacophore similarity searching based on the flexible alignment of known active reference molecules. Steps ii) and iii) were performed independently for the 20000 most "druglike" compounds.

#### Calculation of an alignment of reference compounds

Acetylpromazine (**2**)[9] and CGP40336A (**4**)[21] (Scheme 1) were chosen as reference ligands from the literature with reported nanomolar $IC_{50}$ values. Binding to the bulge had been experimentally verified for both molecules, although detailed structural data were only available for **2**. As **4** contains a ring system—which might be involved in stacking interactions as in **2**—and a charged flexible part—which might interact similarly to a potential charge–$\pi$ interaction of **2** with C24,[6] we assumed that **4** might have a binding mode comparable to that of **2**. For calculation of a SQUID model the two ligands had to be aligned to each other. One possibility would be to dock the reference ligands into the TAR binding pocket, whilst the other would be to perform a flexible ligand-based alignment. Since we were not able to reproduce the experimentally determined TAR-bound conformation of acetylpromazine (**2**) within the binding pocket either by MOE[18] docking or by the AUTODOCK approach[22] (results not shown), we decided to align CGP40336A (**4**) to the NMR conformation of **2** with the aid of the flexible alignment tool in MOE. Interestingly, fruitless attempts to reproduce the NMR structure of **2** complexed to TAR RNA were also reported by Detering and Varani, who successfully reproduced many other RNA–ligand complexes with AUTODOCK, but failed to reproduce the acetylpromazine binding

mode with an RMSD value below 2 Å.[23] Their study supports our decision to follow the ligand-based alignment approach. For the alignment calculation we used the first NMR model of the Protein Database entry (PDB code: 1 LVJ).[6] Since it was not possible to predict reasonable conformations of the aliphatic amino groups of **4** from flexible alignment alone, we decided to cut off these groups and to use molecule **5** instead




**Scheme 2.** Modification of ligand **4** for the alignment.

(Scheme 2) for the alignment and virtual screening. The top scoring solutions of the flexible alignment were visually inspected, and we selected the conformation in which the ligand appeared to fit best into the receptor (Figure 1). Stacking and polar interactions of **4** occupy the same parts of the binding pocket as acetylpromazine (**2**), so we think that a reasonable starting solution was found (Figure 3).

### Calculation of pharmacophores and virtual screening

For virtual screening with CATS3D we calculated the CATS3D descriptor from those conformations of the reference molecules that resulted from the flexible alignment. For screening with the SQUID pharmacophore model the best resolution of

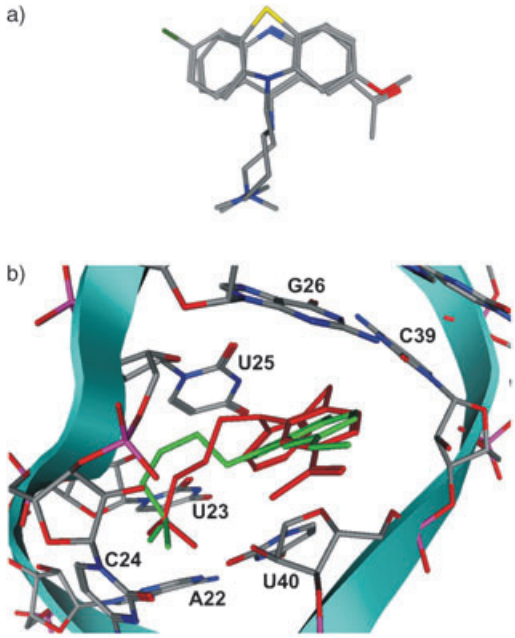



**Figure 3.** a) Alignment of **5** to the NMR conformation of **2** (PDB-code: 1 LVJ). b) The alignment shown in the binding pocket of TAR, with **2** in red and **5** in green.

the model—that is, the optimal PPP cluster radius—and the best weights for the different features had to be determined. The performance of the different parameter sets was determined by their ability to rank the molecules from the pharmacophore model to top positions in comparison to molecules from the COBRA reference dataset (version 3.12) of bioactive molecules,[24] as described.[14] The best performing pharmacophore model is shown in Figure 4.

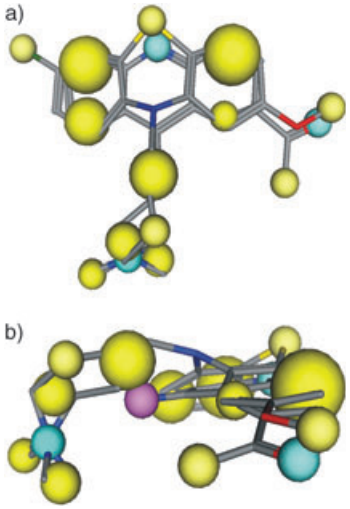



**Figure 4.** SQUID fuzzy pharmacophore model derived from **2** and **5** in a) top view and b) side view. The spheres represent the Gaussian feature densities describing the pharmacophore. The radii of the PPPs are the standard deviations of the underlying feature distribution. The intensities of the colors denote the conservation of the underlying features in the alignment. Yellow PPPs represent hydrophobic interactions, magenta PPPs represent hydrogen bond donor interactions, and cyan PPPs represent hydrogen bond acceptor interactions.

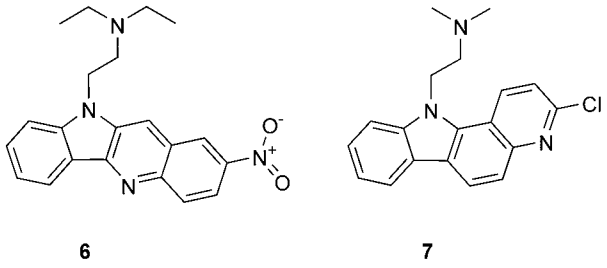
Three virtual screening experiments were performed with different queries: i) +, ii) the two CATS3D CVs calculated from molecules **2** and **5**, and iii) the CV from the optimized SQUID pharmacophore model. The top scoring database molecules from the results were visually inspected, and a set of 19 molecules (ten molecules from SQUID and ten molecules from CATS3D, one molecule overlap) was selected for experimental testing.

### FRET determination of the inhibition constants

All 19 molecules were tested for their potency in a Tat–TAR inhibition assay. As reference we determined the $IC_{50}$ values of argininamide,[25] acetylpromazine,[9] and chlorpromazine[9]—three inhibitors from the literature with reported values of $K_i \approx 1$ mM for argininamide and $IC_{50} < 1$ μM for acetylpromazine and chlorpromazine. The $IC_{50}$ values found in our assay were 1.4 mM for argininamide and 500 μM for acetylpromazine and chlorpromazine. The large discrepancies in the $IC_{50}$ values for acetylpromazine and chlorpromazine in relation to the reported values are consistent with a recently published article that reported a discrepancy of the same order of magnitude for acetylpromazine ($K_D = 270$ μM compared to $IC_{50} \sim 1$ μM, as pre-

viously stated).[8] As a first prescreening of the compounds we performed single-point measurements of the inhibition potency with three fixed concentrations (10, 100, and 1000 μM) of the candidate molecule. Molecules **6**[26]and **7**[27] (Scheme 3) showed stronger inhibition than argininamide in the single-point measurements (cf. Supporting Information). Multipoint measurements yielded $IC_{50}$ values of 500 μM and 46 μM for **6** and **7**, respectively. Compound **6** was found by CATS3D with reference molecule **5**, whilst compound **7** was found by SQUID.



**6** **7**

**Scheme 3.** Best molecules found by CATS3D (**6**) and SQUID (**7**).

To some degree, the two ligand-based pharmacophore methods were able to perform "scaffold-hopping", retrieving isofunctional but slightly different molecular scaffolds from the SPECS catalogue. Each new ligand contains a central structure consisting of three rings with an aliphatic amide side-chain, as in the reference compounds. An additional aromatic ring is present at different positions in each of the two molecules, extending the original ring systems to four concatenated rings. Flexible alignments of **6** and **7** (Figure 5) revealed that **7** fits better to the reference alignment than **6**. Also, the aliphatic amide side-chain of **7** was more closely aligned to the corresponding side-chains of the references. The nitrogen of the additional pyridine ring of **7** was positioned directly above the potential hydrogen bond acceptors of the reference molecules. In both **6** and **7** the additional ring might be used for more favorable stacking interactions with the receptor. In **6** this potentially favorable effect might have been compensated by steric stress due to an unfavorable orientation of the ring or the amide side-chain. Still the $IC_{50}$ value is comparable to those of acetylpromazine and chlorpromazine.

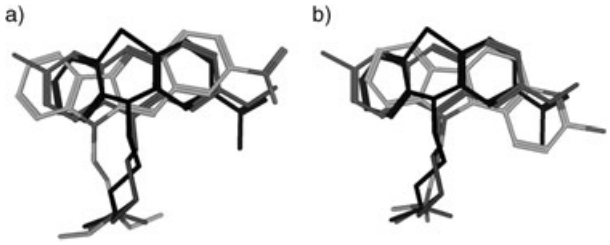



**Figure 5.** Flexible alignment of a) **6** and b) **7** to the aligned reference molecules **2** (black) and **5** (dark grey).

## Conclusion

In this study we present the application of two ligand-based virtual screening approaches for the compilation of a small focused library containing potential TAR RNA ligands. Among the 19 molecules tested we found two molecules capable of inhibiting the Tat–TAR interaction in a FRET assay. The SQUID fuzzy pharmacophore approach yielded the more potent molecule, with an improved activity of one order of magnitude relative to acetylpromazine (**2**) or chlorpromazine (**3**). This could be an effect of the incorporation of information from multiple active reference molecules into the pharmacophore-based search for new TAR ligands.[14]

Ligand-based approaches provide a complementary concept to structure-based design, which might be hampered by the high inherent flexibility of RNA targets. Though it has been shown that specific parameterization of scoring functions is not essential for ligand docking to RNA, it is still significantly slower than a ligand-based approach.[23] We have demonstrated that ligand-based pharmacophore approaches are capable of finding new RNA ligands. Although the best molecule gave a moderate $IC_{50}$ of only 46 μM in our assay this molecule might provide a starting point for further improvement. Certainly, other assay types will be needed to confirm these findings and to scrutinize them further. We also wish to stress that the new inhibitors might not represent ideal candidates for starting a lead optimization project. Additional experiments would have to be performed, addressing the question of the role the additional ring system actually plays for RNA recognition and binding affinity. Furthermore, structures **6** and **7** might be intercalating agents and exhibit unspecific binding to both RNA and DNA targets, due to their planar ring systems and relatively high lipophilicity. Such issues could also be addressed in a different setting of the virtual screening approach; to obtain selectivity towards RNA, for example, known DNA binders and intercalators could be used as negative examples for similarity searching. This tactic is currently being pursued in our laboratory.

Irrespective of the outcome of such analyses, both ligand-based methods have proved useful for finding new molecules within the activity range of known reference compounds. Notably, both approaches were originally developed for protein ligands, but they also seem to be applicable to virtual screening for RNA ligands. To the best of our knowledge this study presents the first inhibitors of an RNA–protein complex found by ligand-based virtual screening.

## Experimental Section

**Materials**: Argininamide was purchased from Sigma Chemical Corp. (St. Louis, USA). The molecules resulting from virtual screening were purchased from SPECS (Delft, The Netherlands) as stock solutions in DMSO (10 mM), and were diluted for binding assays with DEPC-treated water to 1 mM or 100 μ. Fluorescence-based binding assays[20] were performed in 96-well microplates at 37 °C. Reader: FluoStar Galaxy (BMG Labtechnologies, Offenburg, Germany), excitation wavelength 540 nm, emission wavelength 590 nm. Microplates: Corning 6860, black, nonbinding surface. The dye-labeled $Tat_{49-57}$-sequence fluoresceine-AAARKKRRQRRRAAAC-rhodamine (1 μM stock solution) was purchased from the Thermo Electron Corporation (Ulm, Germany). Oligonucleotides were obtained from Biospring (Frankfurt, Germany).

**In vitro transcription**: Equimolar amounts of T7-primer (5′-TAATACGACTCACTATAG-3′) and TAR template (5′-GGCCAGAGAGCTCCCAGGCTCAGATCTGGCCCTATAGTGAGTCGTATTA-3′) were mixed in TE buffer (10 mM Tris-HCl, 1 mM EDTA; pH 7.4) to give a final concentration of 50 pmol $\mu L^{-1}$ in a volume of 100 μL. After having been heated to 90 °C for 5 min, the reaction mixture was allowed to cool down slowly to RT. All in vitro transcriptions were performed with the RiboMax Large Scale RNA Production Systems Kit (P1300; Promega, Mannheim, Germany) as described by the manufacturer. Subsequent to transcription the DNA template was removed as follows: after the transcription mixture had been heated at 95 °C for 5 min it was chilled immediately on ice. RQ1 DNase buffer (10 μL, Promega) and RQ1 RNase-free DNase (20 μL, 20 U, Promega) were added and the mix was incubated for 30 min at 37 °C. After phenol/chloroform extraction, RNA was precipitated with three volumes of ethanol in the presence of sodium acetate (0.3 M, pH 5.2). The RNA was desalted on a NAP™ column (Amersham Biosciences, Freiburg, Germany). After lyophilisation, the RNA pellet was redissolved in DEPC-treated water to a final concentration of 100 μM (stock solution) or 1 μM (final dilution).

**FRET assay**: The following stock solutions were used in the assay: labeled Tat-peptide (1 μM), TAR-RNA (1 μM), TK buffer (500 mM Tris-HCl, 200 mM KCl, 0.1% Triton-X 100, pH 7.4). The final volume per well was 100 μL. The fluorescence of pure Tat peptide was determined first: stock solutions of Tat (10 μL) and TK buffer (10 μL) were filled with DEPC-treated water to a final volume of 100 μL. Tat solution (10 μL, 1 μM), TAR solution (10 μL, 1 μM), TK buffer (10 μL) and DEPC-treated water (70 μL) were then mixed in a second well to measure the emission of the Tat–TAR complex. After establishing the numbers for free and bound peptide, single-point measurements of potential inhibitors were carried out at 1000, 100, and 10 μM by using 10 μL of the stock solutions (10 mM, 1 mM, and 100 μM). RNA and peptide concentrations were kept constant at 100 nM in each well (10 μL Tat, 10 μL TAR, 10 μL TK buffer, 10 μL inhibitor, and 60 μL DEPC-treated water). Addition of DMSO strongly increases the fluorescence of rhodamine independently of peptide–RNA binding. To eliminate this effect, samples of Tat and of Tat–TAR (each 100 nM) were also measured in the presence of 10%, 1%, or 0.1% DMSO. Division of these figures by the value obtained in pure water generated the correction factors. For compounds that showed an inhibitory effect in the preliminary test, complete titration curves were determined from eleven data points. The molecular concentration at which the fitted titration curve intersected with the mean value of the fluorescence counts of the Tat–TAR complex and uncomplexed Tat was used as the $IC_{50}$ value of a molecule.

## Acknowledgements

*Dr. K. Langer kindly provided measuring time on the FluoStar. This research was supported by the Sonderforschungsbereich (SFB) 579 of the Deutsche Forschungsgemeinschaft and the Beilstein-Institut zur Förderung der Chemischen Wissenschaften, Frankfurt. S.R. is grateful for a fellowship from Merz Pharmaceuticals.*

**Keywords:** chemoinformatics · drug design · FRET · RNA · virtual screening

[1] Special journal issues on RNA as drug target: a) *Biopolymers* **2003**, *70*, 1–119; b) *ChemBioChem* **2003**, *4*, 913–1106.
[2] a) G. J. R. Zaman, P. J. A. Michiels, C. A. A. van Boeckel, *Drug Discovery Today* **2003**, *8*, 297–306; b) M. J. Drysdale, G. Lentzen, N. Matassova, A. I. H. Murchie, F. Aboul-ela, M. Afshar, *Prog. Med. Chem.* **2002**, *39*, 73–119; c) J. Gallego, G. Varani, *Acc. Chem. Res.* **2001**, *34*, 836–843; d) S. J. Sucheck, C. H. Wong, *Curr. Opin. Chem. Biol.* **2000**, *4*, 678–686.
[3] T. Hermann, *Biopolymers* **2003**, *70*, 4–18.
[4] J. Karn, *J. Mol. Biol.* **1999**, *293*, 135–154.
[5] a) A. Krebs, V. Ludwig, O. Boden, M.-W. Göbel, *ChembBioChem* **2003**, *4*, 972–978; b) M. Froeyen, P. Herdewijn, *Curr. Top. Med. Chem.* **2002**, *2*, 1123–1145.
[6] Z. Du, K. E. Lind, T. L. James, *Chem. Biol.* **2002**, *9*, 707–712.
[7] E. Byvatov, U. Fechner, J. Sadowsky, G. Schneider, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1882–1889.
[8] M. Mayer, T. L. James, *J. Am. Chem. Soc.* **2004**, *126*, 4453–4460.
[9] K. E. Lind, Z. Du, K. Fujinaga, B. M. Peterlin, T. L. James, *Chem. Biol.* **2002**, *9*, 185–193.
[10] a) G. Schneider, H.-J. Böhm, *Drug Discovery Today* **2002**, *7*, 64–70; b) *Virtual Screening for Bioactive Molecules* (Eds.: H.-J. Böhm, G. Schneider) Wiley-VCH, Weinheim, **2000**.
[11] A. V. Filikov, V. Mohan, T. A. Vickers, R. H. Griffey, P. D. Cook, R. A. Abagyan, T. L. James, *J. Comput.-Aided Mol. Des.* **2000**, *14*, 593–610.
[12] a) N. Leulliot, G. Varani, *Biochemistry* **2001**, *40*, 7947–7956; b) J. R. Williamson, *Nat. Struct. Biol.* **2000**, *7*, 834–837.
[13] G. Schneider, W. Neidhard, T. Giller, G. Schmid, *Angew. Chem.* **1999**, *111*, 3068–3070; *Angew. Chem. Int. Ed.* **1999**, *38*, 2894–2896.
[14] S. Renner, G. Schneider, *J. Med. Chem.* **2004**, *47*, 4653–4664.
[15] U. Fechner, L. Franke, S. Renner, P. Schneider, G. Schneider, *J. Comput.-Aided Mol. Des.* **2003**, *17*, 687–698.
[16] G. Schneider, P. Schneider in *Chemogenomics in Drug Discovery* (Eds.: H. Kubinyi, G. Müller), Wiley-VCH, Weinheim, **2004**, pp. 341–376.
[17] Specs, Delftechpark 30, 2628 XH Delft, The Netherlands; http://www.specs.net.
[18] MOE, Molecular Operating Environment, version 2004.03, Chemical Computing Group Inc., 1010 Sherbrooke St. West, 910, Montreal, Canada.
[19] T. A. Halgren, *J. Comput. Chem.* **1996**, *5 & 6*, 490–512.
[20] C. Matsumoto, K. Hamasaki, H. Mihara, A. Ueno, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1857–1861.
[21] F. Hamy, V. Brondani, A. Florsheimer, W. Stark, M. J. Blommers, T. Klimkait, *Biochemistry* **1998**, *37*, 5086–5095.
[22] 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.
[23] C. Detering, G. Varani, *J. Med. Chem.* **2004**, *47*, 4188–4201. The authors used a threshold of 2.5 Å in their docking study comprising many RNA–ligand complexes. For the TAR–acetylpromazine complex some of the best docking solutions showed an RMSD between 2.0 and 2.5 Å. See Supplementary Information to the reference.
[24] P. Schneider, G. Schneider, *QSAR Comb. Sci.* **2003**, *22*, 713–718.
[25] J. Tao, A. D. Frankel, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 2723–2726.
[26] N. Z. Tugusheva, S. Y. Ryabova, N. P. Solov'eva, V. G. Granik, *Chem. Heterocycl. Comp.* **1998**, *34*, 216–221.
[27] A. K. Shanazarov, V. G. Granik, N. I. Andreeva, S. M. Golovina, M. D. Mashkovskii, *Khim.-Farm. Zh.* **1989**, *23*, 1197–1200.

Received: October 21, 2004