RESEARCH ONLINE

University of Wollongong
Research Online

Faculty of Science - Papers (Archive)

Faculty of Science, Medicine and Health

2003

Novel Pharmacophore Based Methods Reveal Gossypol as a Reverse Transcriptase Inhibitor

P. A. Keller
University of Wollongong, keller@uow.edu.au

C. Birch
Victorian Infectious Diseases Reference Laboratory

S. P. Leach
University of Wollongong

D. Tyssen
Victorian Infectious Diseases Laboratory

R. Griffith
University of Newcastle

Publication Details

This article was originally published as: Keller, PA, Birch, C, Leach, SP, et al, Novel Pharmacophore Based Methods Reveal Gossypol as a Reverse Transcriptase Inhibitor, Journal of Molecular Graphics and Modelling, 2003, 21(5), 365-373. The journal homepage can be found here through Elsevier.

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

RESEARCH ONLINE

Novel Pharmacophore Based Methods Reveal Gossypol as a Reverse Transcriptase Inhibitor

Abstract

In a program to identify new structural entities for the inhibition of the HIV-1 reverse transcriptase (RT) enzyme via database searching, a series of RT pharmacophores were developed. By utilising a novel filtering technique, the National Cancer Institute database of compounds was scanned producing 15 compounds to be screened for activity. A notable inclusion was a series of gossypol derivatives. The testing of a series of compounds revealed the parent compound gossypol to be an HIV-1 reverse transcriptase inhibitor. These results suggest that at least part of its anti-HIV activity is due to gossypol targeting the non-nucleoside inhibitor binding pocket of RT.

Keywords

Gossypol, pharmacophores, database searching, HIV, reverse transcriptase, Medicinal Chemistry

Disciplines

Life Sciences | Physical Sciences and Mathematics | Social and Behavioral Sciences

Publication Details

This article was originally published as: Keller, PA, Birch, C, Leach, SP, et al, Novel Pharmacophore Based Methods Reveal Gossypol as a Reverse Transcriptase Inhibitor, Journal of Molecular Graphics and Modelling, 2003, 21(5), 365-373. The journal homepage can be found here through Elsevier.

This journal article is available at Research Online: http://ro.uow.edu.au/scipapers/43

Novel Pharmacophore Based Methods Reveal Gossypol as a Reverse Transcriptase Inhibitor

Paul A. Keller,*[a] Chris Birch,[c] Scott P. Leach,[a] David Tyssen,[c] Renate Griffith*[a,b]

a Department of Chemistry, University of Wollongong, Wollongong,

NSW, Australia, 2522

+61 2 4221 4692 (ph) +61 2 4221 4287 (fax)

paul_keller@uow.edu.au

b School of Biological and Chemical Sciences, University of Newcastle, Callaghan,

NSW 2308, Australia

+61 2 4921 6990 (ph) +61 2 4921 6923 (fax)

renateg@mail.newcastle.edu.au

c Victorian Infectious Diseases Reference Laboratory (VIDRL), Private Bag 815,

Carlton South, Victoria, Australia, 3053

Abstract

In a program to identify new structural entities for the inhibition of the HIV-1 reverse transcriptase (RT) enzyme *via* database searching, a series of RT pharmacophores were developed. By utilising a novel filtering technique, the National Cancer Institute database of compounds was scanned producing 15 compounds to be screened for activity. A notable inclusion was a series of gossypol derivatives. The testing of a series of compounds revealed the parent compound gossypol to be an HIV-1 reverse

transcriptase inhibitor. These results suggest that at least part of its anti-HIV activity is due to gossypol targeting the non-nucleoside inhibitor binding pocket of RT.

Keywords

Gossypol, pharmacophores, database searching, reverse transcriptase

Introduction

The use of ligand-based computer-aided molecular modelling techniques to investigate the non-nucleoside inhibitor binding pocket (NNIBP) of the reverse transcriptase (RT) enzyme has been documented. The breadth of coverage encompasses QSAR methodologies,[1,2,3] molecular field-based similarity studies,[4,5] neural networks[6] and pharmacophore generation.[6] Each study maintains its identity whether it be by limiting its training set to a single class of inhibitors,[6] or by comparing two different ligand-based approaches[3] to name just two.

The RT enzyme presents an interesting target for computer-aided modelling studies as the NNIBP is known to be flexible, moving both to accommodate inhibitors, and to take on different shapes depending on the inhibitor bound.[7] This must be considered in any pharmacophore-based approach as a rigid pocket geometry is usually assumed for this computational aspect. Thus, pharmacophore models generated for the NNIBP are usually limited to a single class of compound. While these models could be used for database searching and subsequent docking experiments to reveal new structural entities as possible inhibitors, they do limit the potential versatility of the technique by simply ignoring one of the vital and significant features of the site, *i.e.* the pocket flexibility. We have developed our pharmacophore models to incorporate the possibility of pocket flexibility by the implementation of a novel filtering system during database searching. Therefore, the aim was not to produce the traditional 'common' pharmacophores derived from a range of different structural classes of inhibitors, but rather to produce a range of pharmacophores corresponding to each inhibitor class.

Essentially, we are not using 'pharmacophores' in the traditional sense, i.e. trying to produce a quantitative pharmacophore that precisely defines a space occupied by a congeneric set of ligands. We feel that this has been done already for HIV-RT. The purpose of our study was to generate pharmacophores, not to be used for predictions, but rather to try and search a different molecular space

in database searching, and therefore give rise to a different set of hits for testing. Our approach here was to not eliminate possibilities (e.g. by using a "shape" option within the pharmacophore software), but rather, to use separately generated pharmacophores, defined for each inhibitor class, as filters. Therefore, the end result would be a series of molecules (hits) which should fit all potential pocket shapes, as currently reported. This procedure would take into account not only ligand flexibility/adaptability, but enzyme flexibility because it does not try and define one (quantitative) pharmacophore for a series of different pocket shapes – rather, it would try to find molecules that will fit all pocket shapes. Furthermore, this methodology would unearth compounds that can be 'induced' to fit several pharmacophore models through conformational changes, or by different fittings within the model. This is an additional advantage of this method, as examples have already been reported whereby different derivatives of the same lead compound bind in a different orientation within the NNIBP. Therefore, by allowing these alternative positions, we are starting to take into account a degree of enzyme flexibility as it is currently defined in the literature. Furthermore, the pharmacophore software incorporates the different mappings (binding modes) for each molecule and will therefore also take into account possible enzyme flexibility.

This technique was adopted in an attempt to find new structural entities as potential NNIBP inhibitors by searching different spaces within compound databases.

Methodology

The automated pharmacophore development program Catalyst[8,9] was used to generate separate pharmacophore models, using default parameters, for a series of four distinct NNI classes: nevirapine, hydroxyethoxymethylphenylthiothymine (HEPT), tetrahydroimidazobenzodiazepinone (TIBO), and α-anilinophenylacetamide (APA) as well as an additional 'combined model'. These pharmacophores were used as a filter by sequentially running them through the NCI database. This approach did not aim to refine all the data into generating one hypothesis, but to keep the hypotheses separated into compound classes. These multiple hypotheses were then used to search what is probably a more diverse structural space within the database. This could possibly, at least indirectly, take into account that the NNIBP is flexible and changes shape with different structural entities. The compounds provided by the NCI were tested for their HIV-1 reverse transcriptase inhibition.

Results and Discussion

Pharmacophore Generation

Catalyst generated the pharmacophore models for the four distinct NNI classes: nevirapine, HEPT, TIBO, and APA and the additional 'combined model', essentially by determining common features (such as H-bond acceptors, hydrophobic regions, *etc*) and correlating their location in space to the activity of a series (training set) of inhibitor compounds. The actual hypotheses scores were calculated according to the number of bits required to completely describe a hypothesis. Simpler hypotheses require fewer bits for a complete description, and the assumption is made that simpler hypotheses are better.

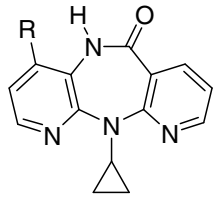
The entire program was developed for use on rigid pocket systems, a constraint that the RT enzyme does not satisfy. It does, however, model ligand flexibility by considering the conformations of each inhibitor used in the training set within a set energy range. Further, ligand based drug design programs assume that the inhibitors used within the training set interact with the active site in the same manner. The NNIs to be analysed may orientate themselves slightly differently upon binding and interact with slightly different amino acid sets within the active site. This is suggested by the structural diversity shown by the inhibitors and the fact that different NNRTIs do not necessarily show cross resistance to one another.[10] For this reason, Catalyst was allowed to generate hypotheses for each of the reverse transcriptase inhibitor (RTI) classes in turn, allowing for the assumption that within each class, all inhibitors will bind in the same fashion.

Catalyst limits the choice of different features to a maximum of five (a feature can be used more than once in a pharmacophore). We chose the following features: hydrophobic aliphatic, aromatic π-stacking, hydrogen bond donor and hydrogen bond acceptor.

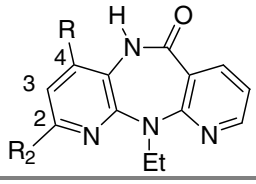
Separate pharmacophores were generated for a series of nevirapine, TIBO, HEPT, and APA analogues, and a further 'combined hypothesis' was also generated. The compounds used in the training sets and their activities[11,12,13] are listed in Tables 1-4.

Table 1: Data used for the generation of the nevirapine hypotheses.[11]

4-Substituted Derivatives

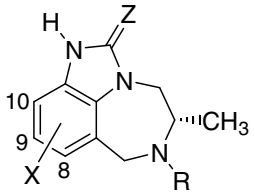


2,4-Disubstituted Derivatives



Entry	R	R_2	IC_{50} (nM)
1	CH_3	H	80
2	CH_2OH	H	3000
3	CH_2OEt	H	1300
4	CH_2OPh	H	120
5	CH_2OCH_2Ph	H	270
6	$CH_2O(CH_2)_2Ph$	H	420
7	CH_2NHPh	H	60
8	CH_2SPh	H	1030
9	CH_2O(Ph-*o*-Me)	H	290
10	CH_2O(Ph-*m*-Me)	H	2430
11	CH_2O(Ph-*p*-Me)	H	580
12	CH_2O(Ph-*o*-OH)	H	460
13	CH_2O(Ph-*o*-Cl)	H	3050
14	CH_2O(Ph-*p*-NH_2)	H	100
15	CH_2O(Ph-*p*-NHEt)	H	80
16	CH_2O(Ph-*p*-OMe)	H	1040
17	CH_2O(Ph-*p*-CN)	H	410
18	CH_2O(Ph-*p*-NO_2)	H	1130
19	CH_2Ph	H	140
20	CH_2OCH_2Ph	OMe	920
21	CHO	OMe	240
22	CH_2OH	OMe	50
23	CH_2OPh	OMe	4000
24	CH_2Ph	OH	340
25	CH_2OCH_2Ph	OH	110
26	CH_2OH	OH	420

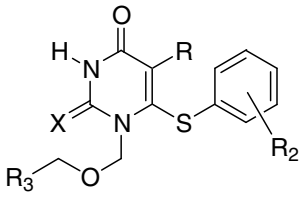
Table 2: Data used for the generation of the TIBO hypotheses.[13]



Entry	R	X	Z	IC_{50} (nM)
1	DMA[a]	9-Cl	S	34
2	DMA	8-Cl	S	4.3
3	DMA	8-F	S	5.8
4	DMA	8-SMe	S	5
5	DMA	8-OMe	S	33.9
6	DMA	8-OEt	S	95.9
7	DMA	8-CN	S	56.3
8	DMA	8-CHO	S	188
9	DMA	8-$CONH_2$	O	6360
10	DMA	8-Br	O	47.3
11	DMA	8-I	O	88
12	DMA	8-I	S	47.4
13	DMA	8-C≡CH	O	437.6
14	DMA	8-C≡CH	S	29.6
15	DMA	8-Et	O	>5940
16	DMA	8-Me	O	989
17	DMA	8-Me	S	13.6
18	DMA	10-OMe	O	6630
19	DMA	10-OMe	S	4725
20	DMA	9,10-diCl	S	25.5
21	DMA	10-Br	S	1075
22	DMA	H	S	44
23	DMA	8-Br	S	3
24	propyl		O	>500000

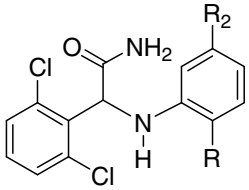
[a] DMA = dimethylallyl

Table 3: Data used for the generation of the HEPT hypotheses.[12]



Entry	X	R	R_2	R_3	IC_{50} (nM)
1	O	Me	H	CH_2OH	6500
2	O	Me	H	Me	330
3	O	Me	H	Ph	93
4	O	C_2H_5	H	CH_2OH	120
5	S	Et	H	CH_2OH	110
6	O	*iso*-propyl	H	CH_2OH	63
7	S	*iso*-propyl	H	CH_2OH	59
8	O	Et	H	Me	22
9	O	Et	H	Ph	4.9
10	O	Et	3,5-diMe	CH_2OH	16
11	S	Et	3,5-diMe	CH_2OH	5
12	O	*iso*-propyl	3,5-diMe	CH_2OH	2.7
13	S	*iso*-propyl	3,5-diMe	CH_2OH	5
14	O	Et	3,5-diMe	Me	6.2
15	S	Et	3,5-diMe	Me	5.6
16	O	Et	3,5-diMe	Ph	2.4
17	S	Et	3,5-diMe	Ph	7.6

Table 4: Data used for the generation of the α-APA hypotheses.[12]



Entry	R	R_2	optical rotation	IC_{50} (nM)
1	NO_2	H	(+)	1731
2	NO_2	H	(–)	33
3	COMe	H	(+)	6506
4	COMe	H	(–)	19
5	COMe	Me	(+)	2109
6	COMe	Me	(–)	5

Nevirapine Pharmacophores

Initial attempts at producing pharmacophores based upon the nevirapine series of inhibitors produced a correlation coefficient 'r' (actual *vs* predicted activity) of 0.40. It has been reported that the lack of correlation between actual and predicted RT inhibitory activities might have been due to the 4-substituted nevirapine analogues interacting in a different manner with the pocket than the 2,4-disubstituted derivatives.[14] Consequently, separate pharmacophores were generated for the 4-substituted and the 2,4-disubstituted nevirapine analogues. This resulted in a strong increase in the correlation between the actual and predicted activities, with the hypothesis for the 2,4-disubstituted compounds improving from 0.40 to 0.82. The 4-substituted analogues also show a slight improvement with the correlation increasing from 0.40 to 0.55. Both these hypotheses are illustrated in Figure 1 (**A** and **B**).

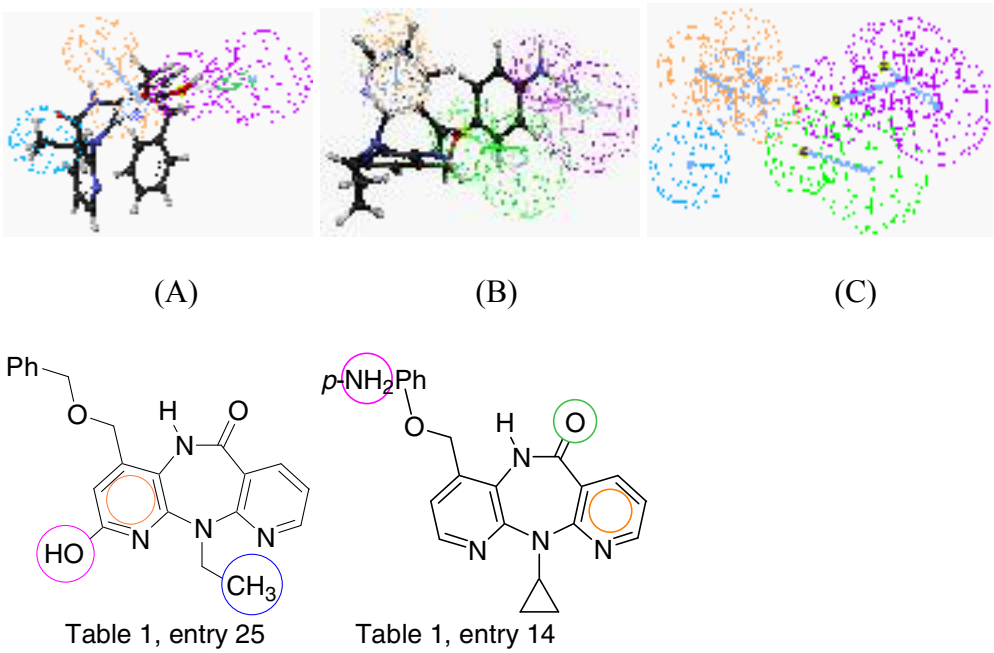


***Figure 1**: The highest scoring hypotheses generated for the 2,4-disubstituted nevirapine training set (**A**) and the 4-substituted nevirapine training set (**B**). Sample compounds overlaid on each pharmacophore are entry **25** and entry **14** respectively (see table 1). **C** shows the highest scoring catalyst hypotheses for the 4-substituted and 2,4-disubstituted nevirapine analogues overlaid. The mesh spheres illustrate the location constraints of each feature and are colour coded: Hydrogen Bond Donor = magenta, with the proposed positioning of the H-bond shown by an arrow or cone pointing to the proposed position of the interacting group on the protein (larger sphere); Hydrogen Bond Acceptor = green, likewise with the proposed position of the H-Bond illustrated; Hydrophobic Aliphatic regions = blue; Hydrophobic Aromatic region = orange with proposed π-stacking interaction shown by an arrow and the plane of the aromatic ring of the ligand constrained by a blue grid. Note in these hypotheses, the hydrophilic region (right) with common hydrogen bond donor features and the hydrophobic region (left) with common π-stacking regions.*

This result confirms that the compounds possessing a 2-substituent bind in a different orientation to those compounds that lack the 2-substituent. It is possible that the dipyridodiazepinone ring system rotates to access the enzymatic residues near the position of the 2-substituent. This rotation would almost certainly lead to a change in the binding environment of the 4-position.

In an attempt to further characterise the different binding orientations of the 4-substituted nevirapine compounds compared with the 2,4-disubstituted analogues, their respective hypotheses were overlaid (Figure 1**C**). Both hypotheses share a common aromatic region (orange) and a common hydrogen bond donor region (magenta). Differences in binding include a hydrogen bond acceptor region (green) that is present for the 4-substituted derivatives, whereas the 2,4-disubstituted derivatives seem to favour another hydrophobic interaction (blue). A further notable feature of these hypotheses is the separation of hydrophobic and hydrophilic regions.

TIBO, HEPT and APA Pharmacophores

Further pharmacophores generated were based upon TIBO, HEPT and APA inhibitors and their derivatives. Figure 2 illustrates the highest scoring hypotheses for each, together with an example inhibitor overlaid. Tables 2, 3 and 4 list the compounds used in the training sets.[12,13]

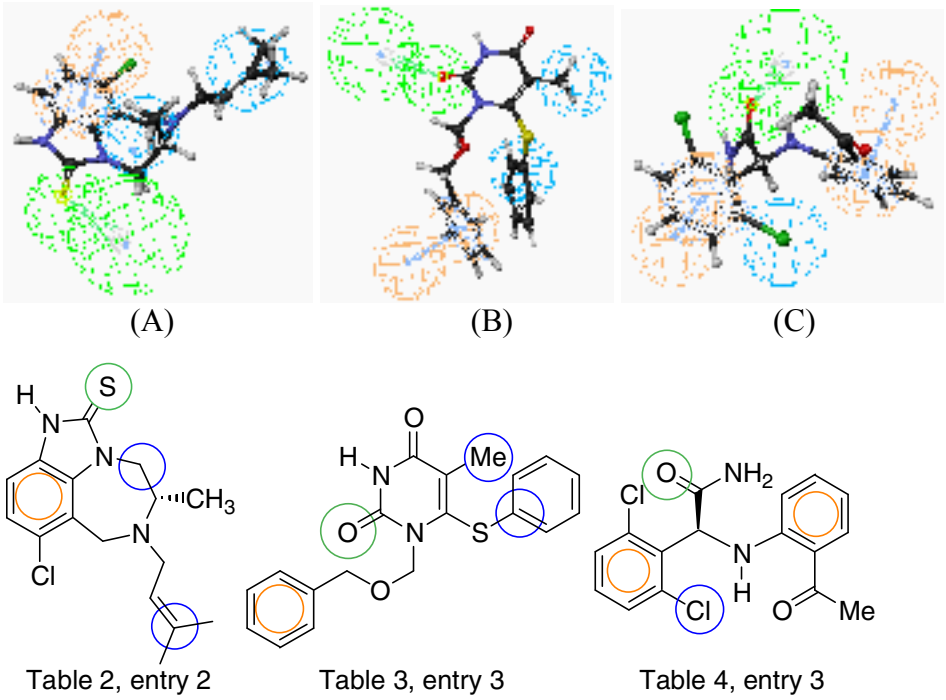


Figure 2: *The highest scoring hypotheses for TIBO (**A**), HEPT (**B**) and APA (**C**) and derivatives. Overlaid on each are example inhibitors (entries **2**, **3** and **3** from Tables 2, 3, and 4 respectively). Colour coding of pharmacophore features is as described for Figure 1.*

The training set of TIBO derivatives has inhibitory values spanning 5 orders of magnitude, ranging from 4 nM to 500,000 nM. It is highly recommended for training sets in Catalyst pharmacophores to span 5 orders of magnitude. The pharmacophore itself (Fig 2**A**) consists of one aromatic π-stacking interaction (orange), which maps to the single phenyl ring present in all TIBO derivatives, two hydrophobic groups (blue) which correspond approximately to the 8-membered ring and to the dimethylallyl (DMA) substituent, and one hydrogen bond acceptor region (green) which overlays precisely with the sulfur substituent attached to the 5-membered ring system. The hypothesis has a correlation of 0.84, with most of the significant deviation observed for predicted IC_{50} values greater than 500 nM.

The hypothesis generated for the HEPT derivatives (Fig 2**B**) consists of one aromatic region (orange), two hydrophobic regions (blue) and one hydrogen bond acceptor (green). Analysis of the hypothesis mapped reveals that significant interaction occurs around the central uracil ring with the hydrogen bond acceptor region located on the O atom on one side, and on the opposite side, a hydrophobic region which maps to the alkyl substituent. Of the two phenyl groups at the extremities of the molecule, one has a π-stacking interaction with the receptor and the other a hydrophobic (aliphatic) interaction. The hypothesis estimates the activities of the training set well, giving a correlation of 0.95.

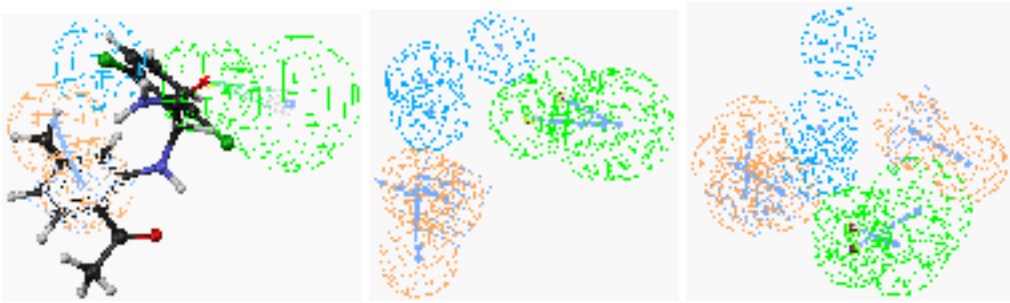
APA is the last of the non-nucleoside inhibitors where the binding location has been determined *via* a crystal structure. Only six compounds were included in the training set. This was the result of the limited amount of published literature available on APA derivatives, as well as difficulties in relating the absolute configuration of the stereogenic carbons present to reported activities. For this reason the hypothesis must be regarded with a degree of caution if it was to be used to predict activity. Nevertheless, it predicts (see Figure 2**C**) that the most important interactions for APA-RT inhibitors will be two aromatic π-stacking interactions mapping onto the two phenyl rings of the molecules, one hydrophobic centre, which maps to the *ortho* substituent on the di-substituted ring, and one electron donating (hydrogen bond acceptor) interaction which overlays with the amide oxygen atom. The correlation coefficient for the APA series is 0.999, and this is likely a result of the structurally congeneric training set used.

The Combined Non-Nucleoside Inhibitor Pharmacophore Model

The four non-nucleoside inhibitor classes analysed so far have all been shown by X-ray crystallography to share a common binding location, the NNIBP.[15] Structure based drug design has found that all of these inhibitors seem to adopt a similar general conformation when bound to the pocket of RT.[16] It has also been speculated that the interactions between the inhibitors and the enzyme are similar for all four of the NNIs analysed, but not necessarily identical. If this is so, then analysis of the similarities of the pharmacophore models produced should show that at least some of the interactions identified in each hypothesis model are common to all, and in a similar spatial location. It is not envisaged that all interactions will be in precisely the same location, due to the flexibility of the NNIBP. For example, the inhibitors are all shown, or proposed, to be interacting with Tyr 188, but it is extremely unlikely that Tyr 188, and indeed any of the other pocket residues, will occupy exactly the same spatial location in every case.

To investigate this idea, a further pharmacophore model was generated by combining all of the inhibitors used previously into a single training set. This model was used as a basis for determining the common features and general locations of important interactions within the pocket.

The combined hypothesis (Fig 3**A**) identifies two hydrophobic interactions, one of which is speculated to be involved in π-stacking interactions with the receptor, the other to be a hydrophobic aliphatic interaction. The other type of interaction believed important is a hydrogen bond acceptor.



(A) (B) (C)

Figure 3: ***A****: The combined hypothesis for the NNIs. The molecule used to illustrate the hypothesis is Entry* ***3*** *in Table 4 (see Fig. 2(C) for structure).* ***B****: The HEPT hypothesis overlaid with the combined model hypothesis showing a good likeness.* ***C****: APA and TIBO hypotheses overlaid showing a good likeness, typical of any two of the models overlaid. Common to all of the hypotheses produced so far, seems to be the location of hydrophobic (features coloured blue and orange) and hydrophilic interactions (features coloured green and magenta), at opposite ends of the pharmacophores. Colour coding of features is as described in Figure 1.*

As was the case with the nevirapine derivatives, it is likely that the low correlation observed occurs as a result of some of the inhibitors interacting with the NNIBP in a subtly different way.

In order to test the validity of the combined model, it was overlaid with each of the original pharmacophores to determine the extent of the similarities between the individual structural models and the combined hypothesis. In a typical example, Figure 3**B** shows the HEPT model overlaid with the combined model and illustrates the general similarities. Further, when the individual pharmacophore models were overlaid with each other, it further exemplified the similarities that existed with the pharmacophores. As one example, Figure 3**C** shows the hypotheses generated for the APA and TIBO series overlaid, a result typical for all cases when two of these NNI pharmacophores were superimposed. The aromatic regions overlay in approximately the same location, suggesting that the π-stacking between ligand and receptor occurs in the same fashion irrespective of the ligand. Slight deviations in the positions of the aromatic rings could also be accounted for by the three different π-stacking interactions that can occur; parallel displaced, perpendicular and parallel tilted, as well as by the flexibility of the pocket.

Figure **3C** also shows the hydrogen bond acceptor and hydrophobic regions located in similar spatial positions. Again this seems to be a fairly general result, as any two of the hypotheses overlaid always shared at least one hydrogen bond acceptor (green) or hydrophobic region (blue), along with the aromatic π-stacking (orange).

As stated before, it is not expected that different classes of inhibitors will interact with the NNIBP in precisely the same manner. This would be unreasonable, due to the structural diversity present in the inhibitors and the known flexibility of the pocket. Thus, small differences in the locations of important features are tolerable, as we are primarily concerned with determining the approximate location of these 'overlapping' interactions. The fact that there is such variation between the inhibitors suggests that provided that the locations of their binding features are positioned in approximately the correct location, the NNIBP will accommodate structurally different molecules leading to individual induced fit mechanisms in the RT enzyme based on the specific nature of the ligand.

Database Filtering

The combined NNI hypothesis was used to search the NCI database.[17] This initial examination yielded 47,028 hits. In an attempt to refine the number of NCI search hits down to something more manageable, it was decided to filter the NCI database hits by using all of the pharmacophores generated

successively, resulting in an end product that would only contain compounds that fitted all hypotheses. This, in theory, should produce compounds that possess all of the essential features necessary for potent RT inhibitory activity.

Table 5 shows how the number of database search hits was progressively reduced as each pharmacophore model was passed through the reduced database – providing the same hypotheses are used, the order with which the database is filtered is irrelevant as the same outcome will result. After utilising all of the pharmacophores for NNIs of known binding site as a series of filters, 540 compounds remained which satisfied all of the constraints on the spatial location of features imposed by the models. In an attempt to further refine the database results into a manageable number, manual filtering was employed, using the following criteria:

- Compounds lacking hydrophilic character were excluded from further analysis, since it has been speculated that certain hydrophilic interactions play a significant role in 'navigating' the substrate into the pocket.
- Compounds possessing bulky, or unusually long side chains were omitted from the analysis. This was a result of Catalyst mapping compounds to the selected pharmacophores, which were in reality far too large to fit into the NNIBP. Excluded volume restraints were not utilised in this study.
- Molecules with no conformational potential for adopting anything of a 'butterfly' conformation were also removed from further selection.
- Structural novelty

Table 5: *Database filtering using the different pharmacophore hypotheses. Note, the order with which the (same) hypotheses are used is irrelevant to the final outcome.*

Pharmacophore Model	*NCI hits (cumulative)*
Combined	47028
HEPT	6773
TIBO	3261
APA	1161
Nevirapine 2,4-disubstituted	652
Nevirapine 4-substituted	**540**

Analysis of Selected Structural Elements from Database Searching and the RT Inhibitory Action of Gossypol

In support of the quality of the filters used, two hydrophobic regions were always present in these compounds, presenting an array of conformations, emulating known non-nucleoside inhibitors of HIV-1 reverse transcriptase, and their well established hydrophobic interactions. Further support for the usefulness of the pharmacophores were the presence of numerous ADAM (alkenyldiarylmethane) type structures amongst the search hits (Fig 4). This class of compounds are potent RT inhibitors.[18]

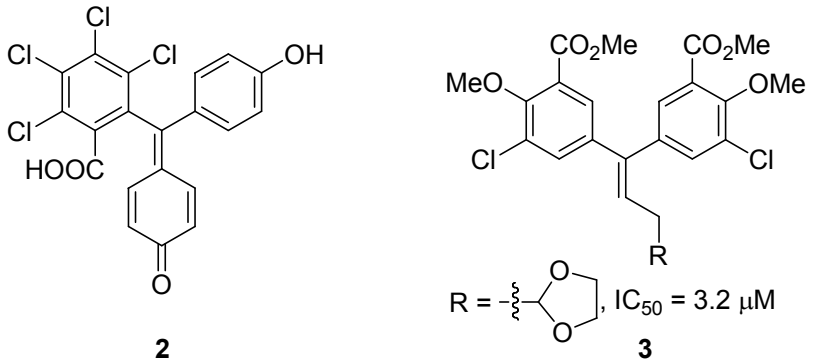


Figure 4: *An example of an ADAM (arenediarylmethane) structure (**2**), which emerged as a hit from our database searching. Also illustrated is an example of a known ADAM (**3**) with its activity as a reverse transcriptase inhibitor.*

A series of gossypol analogues comprising of long alkyl chain derivatives also emerged as hits from our database searches. Since our pharmacophore models contained no volume constraints, and these long chain hydrophobic imine functionalities did not overlay on any of the features contained within the hypotheses generated, it was not considered necessary to synthesize these molecules, instead the parent compound (**1**) was used for testing. Figure 5 also shows gossypol overlaid on the 'combined' pharmacophores hypothesis. An important aspect to note is the observation that sections from both halves of the dimeric gossypol map onto the pharmacophores, suggesting that the basic monomeric unit is not sufficient for adequate activity.

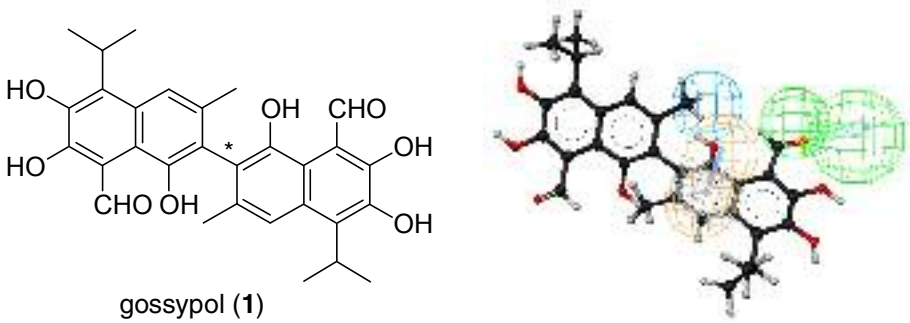


Figure 5: *Structure of the parent compound gossypol (**1**) and gossypol overlaid on the combined hypothesis.*

The presence of gossypol derivatives in the database search results posed an interesting question. The parent gossypol compound was reported in 1989 as an anti-HIV agent,[19] however, it was stated that it was unlikely to be primarily a reverse transcriptase inhibitor, although no data was presented for the effects of gossypol in a reverse transcriptase assay. In a later article (1995),[20] the effects of gossylic iminolactone (GIL) were compared to those of AZT to ascertain whether they have different targets. From the outset, this publication never considered the possibility that gossypol might act at a different site on the reverse transcriptase enzyme (*e.g.* the non-nucleoside inhibitor binding pocket) when compared to AZT (a known nucleoside inhibitor). This paper concludes that GIL does not inhibit RT, but also suggests that gossypol derivatives may exert their antiviral effects in a different manner to the parent compound (**1**) due to repeated observations that the two sets of compounds exhibit markedly different chemical and biological properties.

Considering that the base structure of this compound arose as a hit from our database searching, and that there has been no data published showing the inhibitory effects (or otherwise) of gossypol on the reverse transcriptase enzyme, we included the structure (**1**) in our testing for anti-RT activity.

Biological Testing

A series of sixteen compounds arising from our pharmacophore studies, including 15 donated by the NCI (see Fig 6) and racemic gossypol, were tested for their anti-RT activity using nevirapine and foscarnet as positive controls and unphosphorylated AZT as a negative control. The only structure to emerge as possessing activity was gossypol (Fig. 7). At the highest concentration tested (100mg/mL), it inhibited reverse transcriptase by greater than 99% of its normal activity. This inhibition is dose dependant: at the lower concentrations of 20 and 4 μg/mL, it restricted RT activity to 73.2% and 76.3% of normal activity respectively, but at a concentration of 0.8 μg/mL, had lost its inhibitory effects. These results suggest that gossypol is, in part, a reverse transcriptase inhibitor. Certainly, the anti-HIV activity of gossypol cannot be attributed solely to its RT inhibition – indeed notable compounds have been shown to have multiple modes of action, including the michellamine class of compounds that are also reverse transcriptase inhibitors, but are known to act at two additional sites to account for their anti-HIV activity.[21] The results from our study conclusively demonstrate that gossypol is a weak reverse transcriptase inhibitor, and strongly suggest that the site of this action is the non-nucleoside inhibitor binding pocket.

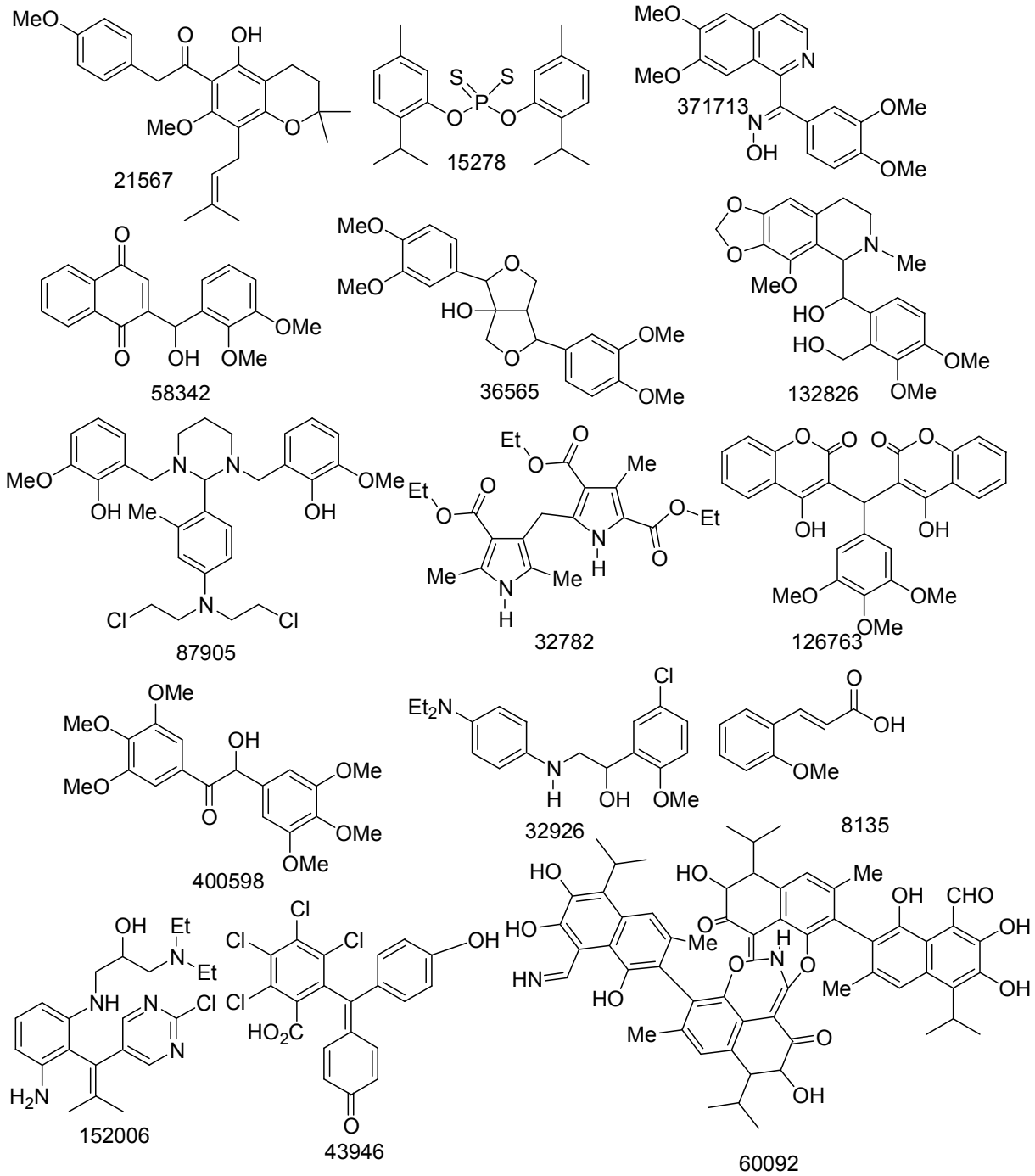


Figure 6*: Structures of compounds donated by the NCI and tested.*

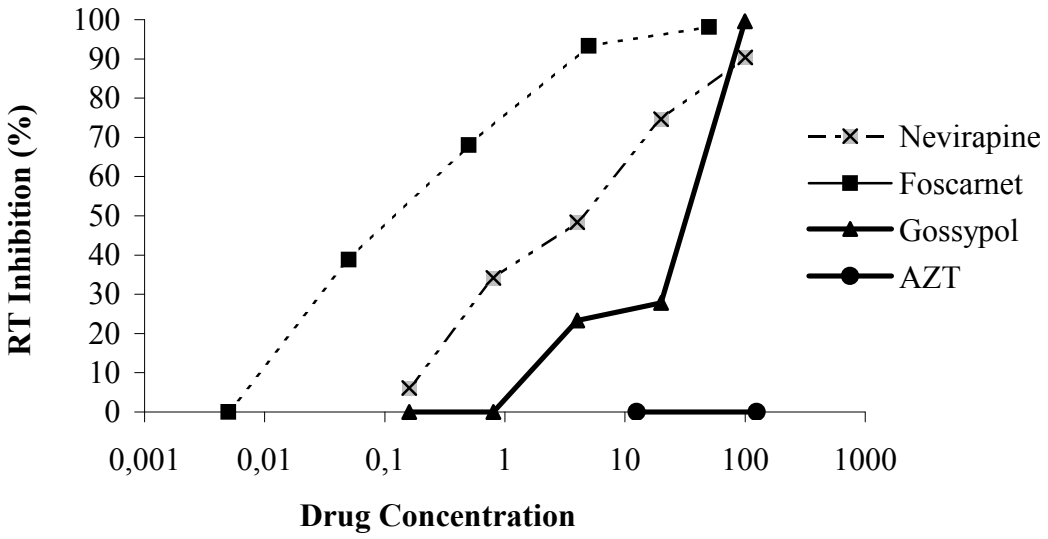


Figure 7: *In vitro inhibition of HIV-1 reverse transcriptase by gossypol. Nevirapine and foscarent are used as the positive controls while unphosphorylated AZT is used as the negative control.*

Conclusions

A series of pharmacophores were developed for the NNIBP of the RT enzyme of HIV-1. These pharmacophore hypotheses were designed for database searching in such a fashion as to incorporate some consideration of the known flexibility of this pocket. A search of the NCI database resulted in a number of hits, a selection of which were tested for their RT inhibition activity. Emerging from these results was gossypol, already reported as an anti-HIV agent, however earlier reported as unlikely to be acting on the RT enzyme. Our testing revealed gossypol to be an RT inhibitor of significance, and that there is a strong suggestion that the site of action is the non-nucleoside inhibitor binding pocket.

Experimental

Compounds Tested

Gossypol was purchased from Sigma-Aldrich and was tested as the racemic mixture of atropisomers without further purification. All other compounds were generously supplied by the National Cancer Institute and tested as supplied.

Biological Assay

The mechanism of action of gossypol at the level of HIV-RT was determined in an *in vitro* assay using solubilised HIV as the source of the enzyme. Nevirapine and foscarnet were included as positive (active against RT) controls. Unphosphorylated AZT was included as a negative control. The RT

inhibition reaction was performed in a 96 well U-bottomed microtitre plate (Nunc). Equal volumes of solubilised virus and the test or control compound dilutions were mixed and incubated at room temperature for 15 minutes before being added in duplicate to wells containing reaction mix (reaction mix contains 50mM Tris HCl pH 7.8, 10mM $MgCl_2$, 5mM dithiothreitol, 10 units poly(rA).poly(dT)10, 2 mg/mL dATP, 5mCi ^{3}HTTP).

After 3 hours incubation at 37°C, the DNA product was precipitated by addition of trichloroacetic acid. Precipitated products were collected and washed using a cell-harvester (Filtermate 196). After addition of liquid scintillant (40 microscint), counts per minute (cpm) were determined on a Top Count Microplate Scintillant Counter. The average cpm for each drug dilution was obtained and compared to the virus only (no drug) control. A dose-response curve (% RT inhibition *vs* drug concentration) was then plotted for each drug.

References

1. Gussio, R.; Pattabiraman, N.; Kellog, G. E.; Zaharerevitz, D. W. Use of 3D QSAR Methodology for Data Mining the National Cancer Institute Repository of Small Molecules: Application to HIV-1 Reverse Transcriptase Inhibition. *Methods 1998*, **14**, 255-263.
2. Hopfinger, A. J.; Wang S.; Tokarski, J. S.; Jin, B.; Albuquerque, M.; Madhav, P. J.; Duraiswami, C. Construction of 3D-QSAR Models Using the 4D-QSAR Analysis Formalism *J. Am. Chem: Soc. 1997*, **119**, 10509-10524.
3. Gussio, R.; Pattabiraman, N.; Zaharerevitz, D. W.; Kellog, G. E.; Igor, A.; Rice, W. G.; Schaeffer, C.; Erickson, J. W.; Burt, S. K. All-atom Models for the Non-nucleoside Binding Site of HIV-1 Reverse Transcriptase Complexed with Inhibitors: a 3D QSAR Approach. *J. Med. Chem. 1996*, **39**, 1645-1650.
4. Mestres, J.; Rohrer, D. C.; Maggiora, G. M. A Molecular Field-based Similarity Study of Non-nucleoside HIV-1 Reverse Transcriptase Inhibitors. *J. Comput.-Aided Mol. Des. 1999*, **13**, 79-93.
5. Mestres, J.; Rohrer, D. C.; Maggiora, G. M. MIMIC: A Molecular-field Matching Program. Exploiting Applicability of Molecular Similarity Approaches. *J. Comput. Chem. 1997*, **18**, 934-954.
6. Tronchet, J. M. J.; Grigorov, M.; Dolatshahi, N.; Moriaud, F.; Weber, J. A QSAR Study Confirming the Heterogeneity of the HEPT Derivative Series Regarding Their Interaction with HIV Reverse Transcriptase. *Eur. J. Med. Chem. 1997*, **32**, 279-299.

7. Jager, J.; Smerdon, S. J.; Wang, J.; Boisvert, D. C.; Steitz, T. A. Comparison of Three Different Crystal Forms Shows HIV-1 Reverse Transcriptase Displays an Internal Swivel Motion. *Structure 1994*, **2**, 869-876.
8. Catalyst, version 4.0, 1998, Molecular Simulations, Inc. (now Accelrys), San Diego, CA, USA; and accompanying documentation.
9. Sprague, P.W., Hoffmann, R. CATALYST pharmacophore models and their utility as queries for searching 3D databases. in *Computer-Assisted Lead Finding and Optimization: Current Tools for Medicinal Chemistry*, van der Waterbeemd, H., Testa, B., and Folkers, G., Eds., VHCA Wiley-VCH, Weinheim, Germany, 1997, pp 225-240.
10. Ding, J.; Das, K.; Tantillo, C.; Zhang, W.; Clark, A. D.; Jessen, S.; Lu, X.; Hsiou, Y.; Jacobomolina, A.; Andries, K.; Pauwels, R.; Moereels, H.; Koymans, L.; Janssen, P. A. J.; Smith, R. H.; Koepke, M. K.; Michejda, C. J.; Hughes, S. H.; Arnold, E. Structure of HIV-1 Reverse Transcriptase in a Complex with the Non-Nucleoside Inhibitor α-APA R 95845 at 2.8Å Resolution. *Structure 1995*, **3**, 365-379.
11. Kelly, T. A.; Proudfoot, J. R.; McNeil, D. W.; Patel, U. R.; David, E.; Hargrave, K. D.; Grob, P. M.; Cardozo, M.; Agarwal, A.; Adams, J. Novel Non-Nucleoside Inhibitors of Human Immunodeficiency Virus Type 1 Reverse Transcriptase. 5. 4-Substituted and 2,4-Disubstituted Analogs of Nevirapine. *J. Med. Chem. 1995*, **38**, 4839-4847.
12. De Clercq, HIV-1-Specific RT Inhibitors: Highly Selective Inhibitors of Human Immunodeficiency Virus Type 1 that are Specifically Targeted at the Viral Reverse Transcriptase. E. *Med. Res. Rev. 1993*, **13**, 229-258.
13. Ho, W.; Kukla, M. J.; Breslin, H. J.; Ludovici, D. W.; Grous, P. P.; Diamond, C. J.; Miranda, M.; Rodgers, J. D.; Ho, C. Y.; De Clercq, E.; Pauwels, R.; Andries, K., Janssen, M. A. C; Janssen, P. A. J. Synthesis and Anti-HIV-1 Activity of 4,5,6,7-Tetrahydro-5-methylimidazo[4,5,1-jk][1,4]benzodiazepin-2(1H)-one (TIBO) Derivatives. 4. *J. Med. Chem. 1995*, **38**, 794-802.
14. The 2- and 2,4-substituted nevirapine derivatives have been noted to bind in marginally different orientations, see reference 11; however, docking studies have shown that, in principle, they interact with the NNIBP in the same manner, see reference 16.
15. Keller, P. A.; Leach, S. P.; Luu, T. T. T.; Titmuss, S. J.; Griffith, R. Development of Computational and Graphical Tools for Analysis of Movement and Flexibility in Large Molecules. *J. Mol. Graphics Mod. 2000*, **18**, 235-241. and references cited therein.

16. Titmuss, S. J.; Keller, P. A.; Griffith, R. Docking Experiments in the Flexible Non-Nucleoside Inhibitor Binding Pocket of HIV-1 Reverse Transcriptase. *Bioorg. Med. Chem. 1999*, **7**, 1163-1170.
17. The NCI database as implemented in Catalyst, see reference 8.
18. Casimiro-Garcia, A.; Micklatcher, M.; Turpin, J. A.; Stup, T. L.; Watson, K.; Buckheit, R. W.; Cushman, M. Novel Modifications in the Alkenyldiarylmethane (ADAM) Series of Non-Nucleoside Reverse Transcriptase Inhibitors. *J. Med. Chem. 1999*, **42**, 4861-4874 and references sited therein.
19. Lin, T. S.; Schinazi, R.; Griffith, B. P.; August, E. M.; Eriksson, B. F. H.; Zheng, D. K.; Huang, L.; Prusoff, W. H. Selective Inhibition of Human Immunodeficiency Virus Type 1 Replication by the (-) but not the (+) Enantiomer of Gossypol. *Antimicrob. Agents Chemother. 1989*, **33**, 2149-2151.
20. Royer, R. E.; Mills, R. G.; Younf, S. A.; Vander Jagt, D. L. Comparison of the Antiviral Activities of 3'-Azido-3'-deoxythymidine (AZT) and Gossylic Iminolactone (GIL) Against Clinical Isolates of HIV-1. *Pharmacol. Res. 1995*, **31**, 49-52.
21. McMahon, J. B.; Currens, M. J.; Gulakowski, R. J.; Buckheit, R. W.; Lackmansmith, C.; Hallock, Y. F.; Boyd, M. R.; Michellamine B, a Novel Plant Alkaloid, Inhibits Human Immunodeficiency Virus-Induced Cell Killing by at Least Two Distinct Mechanisms, *Antimicrob. Agents Chemother. 1995*, **39**, 484-488.