Comparative Quantitative Structure–Activity Relationship Studies on Anti-HIV Drugs. (Chem. Rev. 1999 , 99 , 3525–3601. Published on the Web November 17, 1999)

ARTICLE *in* CHEMICAL REVIEWS · SEPTEMBER 2002

Impact Factor: 46.57 · DOI: 10.1021/cr020468f

Comparative Quantitative Structure−Activity Relationship Studies on Anti-HIV Drugs

Rajni Garg,[†] Satya P. Gupta,[*,‡] Hua Gao,[§] Mekapati Suresh Babu,[||] Asim Kumar Debnath,[⊥] and Corwin Hansch[*,†]

Department of Chemistry, Pomona College, Claremont, California 91711, Departments of Chemistry and Pharmacy, Birla Institute of Technology and Science, Pilani 333031, India, Pharmacia & Upjohn, 301 Henrietta Street, Kalamazoo, Michigan 49007, and Biochemical Virology Laboratory, Lindsley F. Kimball Research Institute of The New York Blood Center, 310 E. 67th Street, New York, New York 10021

Received May 21, 1999

Contents

I. Introduction

In the present era, acquired immunodeficiency syndrome (AIDS) is the most fatal disorder for which no completely successful chemotherapy has been developed so far. The pandemic spread of this disease has prompted an unprecedented scientific and clinical effort to understand and combat it. The causative agent of AIDS has been identified as a retrovirus of the *Lentiviridae* family.[1,2] Originally referred to as HTLV−III or LAV, this enveloped single-stranded RNA virus is now called human immunodeficiency virus (HIV)[3,4] and two genetically distinct subtypes, HIV-1 and HIV-2, have been characterized,[5−7] of which the former has been found to be prevalent in causing the disease.

In the present review, the QSAR studies available or derivable on anti-HIV chemicals are discussed. We have compared the optimum Clog *P* values (log P_0) observed in correlation equations and then compared them with the Clog *P* values (calculated log *P*) of those anti-HIV chemicals which are in the market. We found surprising uniformity in these values. Equally surprising is the finding that most of the nucleoside reverse transcriptase inhibitors, e.g., zidovudine, zalcitabine, stavudine, etc., in use are hydrophilic. Could it be due to the toxicity of more hydrophobic compounds? This seems to conform to our principle of minimal hydrophobicity in drug design[179] that one wants to make drugs as hydrophilic as possible commensurate with efficacy. These values are so low that one wonders if somewhat more hydrophobic drugs might be more effective in reaching into hydrophobically protected regions. Our results seem to agree well with the fact that FDA in the United States recently approved[178] a few nonnucleoside reverse transcriptase HIV inhibitors, e.g., nevirapine, delavirdine, and efavirenz, which are more hydrophobic and their Clog *P* values are in good agreement with our results. Our studies on HIV-protease also show that commercial protease inhibitors such as saquinavir, ritanovir, indinavir, and nelfinavir have high log P_0 values in line with our findings.

The HIV-1 infection, which targets monocytes expressing surface CD4 receptors, eventually produces profound defects in cell-mediated immunity.[8] Over time, infection leads to serve depletion of $CD4^+$ T-lymphocytes (T-cells), resulting in opportunistic infections, neurologic and neoplastic diseases, and ultimately death. Besides T-cells, other cells expressing CD4 on their surface may also harbor HIV-1 and thereby act as a reservoir for the virus, thus extending the latency period associated with the infection. These include macrophages, monocytes, and lymphoid cells.[9] The AIDS chemotherapy now depends on the identification of the molecular events critical to virus replication, and this depends on the detailed

* To whom correspondence should be addressed.
[†] Pomona College.
[‡] Department of Chemistry, Birla Institute of Technology and Science.
[§] Pharmacia & Upjohn.
[||] Department of Pharmacy, Birla Institute of Technology and Science.
[⊥] Lindsley F. Kimball Research Institute of The New York Blood Center.

10.1021/cr9703358 CCC: $35.00
Published on Web 11/17/1999

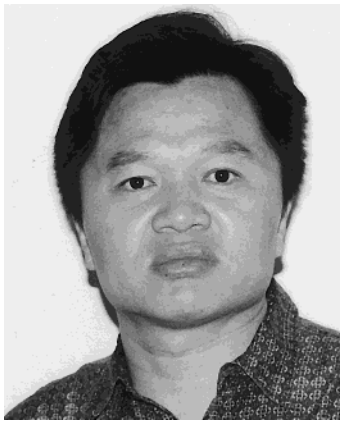
Rajni Garg received her M.Sc. in Chemistry (1984) from Meerut University and M.Phil. (1988) degree from Delhi University, India. Her M.Phil. dissertation work was on peptide synthesis. She was a faculty member in the Chemistry Department of Birla Institute of Technology and Science, Pilani, India, from 1991 to 1996, where she taught organic and physical chemistry. She received her Ph.D. degree in 1996 under the supervision of Professor S. P. Gupta. Her doctoral work was on QSAR studies on anti-HIV agents. In February 1997, she joined Professor Corwin Hansch as a postdoctoral researcher, and she is currently involved in building a C-QSAR databank. Her research interests include QSAR and computer-assisted drug design.

Satya P. Gupta is a professor of Chemistry at Birla Institute of Technology and Science (BITS), Pilani. He obtained his M.Sc. degree in Physical Chemistry with a gold medal in 1967 from the University of Allahabad, Allahabad, and conducted his D.Phil. degree under the supervision of Professor Balkrishna at the same University in 1971. In his doctoral work he developed a new molecular orbital model known as IOC-ω-technique (inclusion of overlap charges in ω-technique), which was then found very useful in dealing with the problems of conjugated systems. After his D.Phil. degree, Gupta moved to Tata Institute of Fundamental Research (TIFR), Bombay, where he worked with Professor G. Govil on molecular biology. He joined BITS in 1973, and since then he has been there only working on theoretical aspects of drug design. For the excellence of his work in this area, he was made a fellow of the National Academy of Science of the country and was awarded the Ranbaxy Research Foundation award, a prestigious national award for outstanding work in medicinal and pharmaceutical sciences. Gupta has been associated with the American Chemical Society in various ways: as a member, as an author of series of reviews, as a reviewer, and once as a member of the task force to evaluate *Chemical Reviews*. Besides his original contributions to science, Gupta has also authored two books, namely, *Quantum Biology* and *Elements of Biophysics*.

knowledge of the structure and the life cycle of the virus.

II. Structural Components and Life Cycle of HIV-1

High-resolution electron microscopy has illustrated that HIV-1 is an enveloped virus of about 100 nm



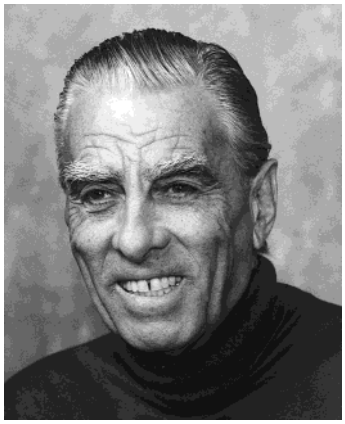
Hua Gao received his Ph.D. in Pharmaceutical Sciences at the University of Southern California. He joined Professor Corwin Hansch in 1995 as a postdoctoral research associate and worked at BioByte Corporation as a scientist. After working in MDS Panlabs as a scientist, he joined Pharmacia & Upjohn as a research scientist. His research interests include QSAR, computer-assisted drug design (CADD), cheminformatics, and combichem informatics.

Born in 1969, Mekapati S. Babu received his B.Pharm. degree in 1990 from Annamalai University, Chidambaram, and M.Pharm. degree in 1992 from Birla Institute of Technology and Science (BITS), Pilani. In the same year, he joined the faculty of BITS but was deputed to Hyderabad to monitor the Practice School Programme of the Institute. After a couple of years, he returned to BITS and joined the research group of Professor S. P. Gupta for his doctoral degree. He is now involved in the studies related to anti-HIV drugs.

diameter.[10] It contains an outer lipid bilayer, derived from the host cell during maturation, and consists of two major viral glycoproteins, the external gp120 and the transmembrane gp41 (gp stands for glycoprotein and the number refers to the mass of protein in thousands of Dalton). Immediately beneath the outer envelope is a membrane-associated protein p18, which provides a matrix for the viral structure and is vital for the integrity of the virion. The matrix surrounds a characteristic dense, cylindrical nucleoid containing capsid protein p24. Inside this nucleoid are two identical RNA strands with which the viral RNA-dependent DNA polymerase (pol) p66/p55, called reverse transcriptase, is in association with nucleoprotein p9, integrase protein p12, and protease p15 components.

The HIV life cycle begins with high-affinity binding of gp120 envelope protein to its receptor CD4 on the host cell surface (Figure 1).[11] The CD4 receptor is a protein molecule found predominantly on a subset of T-lymphocytes responsible for helper or inducer func-

Asim Kumar Debnath was born in 1954 in India. He obtained his B.S. and M.S. degrees in Pharmacy and his Ph.D. degree in Pharmaceutical Chemistry from Jadavpur University, Calcutta, India. He joined Professor Hansch's laboratory at Pomona College as a postdoctoral fellow in 1987. In 1993 he joined the Lindsley F. Kimball Research Institute of the New York Blood Center in New York. Currently, he is an Assistant Member in the Biochemical Virology Laboratory and his focus of research is drug design against HIV-1 targeted to the envelope glycoproteins (gp41 and gp120). His research interests include 2D and 3D QSAR, molecular modeling, structure-based drug design, database mining, and protein modeling.



Corwin Hansch received his undergraduate education at the University of Illinois and his Ph.D. degree in Organic Chemistry from New York University in 1944. After working with the DuPont Company, first on the Manhattan Project and then in Wilmington, DE, he joined the Pomona College faculty in 1946. He has remained at Pomona except for two sabbaticals: one at the Federal Institute of Technology in Zurich with Professor Prelog and the other at the University of Munich with Professor Huisgen. The Pomona group published the first paper on the QSAR approach relating chemical structure with biological activity in 1962. Since then, QSAR has received widespread attention. In *Current Contents* (1981/41) he was named as one of the 300 most cited scientists out of over one million publishing in all fields of science for the period 1965–1978. In 1986, Hansch was cited in *Current Contents* as being one of the 250 most cited primary authors for 1984. He is an honorary fellow of the Royal Society of Chemistry and recently received the ACS Award for Computers in Chemical and Pharmaceutical Research for 1999.

tion in the immature response. Following binding, the fusion of virus with host cell membrane occurs via the gp41 molecules and the HIV genomic RNA is uncoated and internalized. The enzyme reverses transcription of genomic RNA into double-stranded DNA. The DNA migrates to the nucleus to be integrated into the host cell chromosome through the action of virally encoded enzyme, integrase. The incorporation of this "provirus" into the cell genome is permanent. The provirus may remain transcriptionally inactive (latent) or manifest a high level of gene expression with active production of virus.

The activation of provirus (the gene expression) from the latent state by selective and constitutive host transcription factors, notably the NF−Kβ family of DNA enhancer binding proteins, leads to the sequential production of various viral m-RNAs. These m-RNAs are translated into regulatory proteins—Tat, Rev, and Nef. The viral core is formed by the assembly of these proteins, enzymes, and genomic RNA at the plasma membrane of the cells. Budding of the progeny virion occurs through the host cell membrane, where the core acquires its external envelope. During the final budding process, the cleavage of gag−pol polyprotein precursor by HIV protease occurs, leading to morphological maturation of virions.

Thus, the replicative cycle of HIV-1 presents several viable targets that could be exploited for the development of anti-HIV chemotherapy. Ideally, an anti-HIV agent should arrest the virulence and further infection of healthy cells without displaying toxicity toward normal cellular physiology. To achieve this goal, attention has been focused on several intervention strategies and various kinds of inhibitors as described in the following sections.

III. Intervention Strategies and Inhibitors

Theoretically, an anti-HIV agent may exert its activity by inhibiting a variety of steps in the life cycle of the virus. However, medicinal chemists have focused their attention predominantly on the following stages: (A) Viral binding to target cells, (B) Virus cell fusion, (C) Virus uncoating, (D) Reverse transcription of genomic RNA, (E) Viral integration, (F) Gene expression, (G) Cleavage event, (H) Virion maturation, By hitting any of these stages, the viral replication can be terminated. Brief descriptions of the inhibitors of these stages are given below. For details, readers may see a recent article by De Clercq.[12]

A. Viral Binding Inhibitors

It was demonstrated that a truncated CD4 (sCD4) molecule was capable of inhibiting the binding of gp120 to CD4 receptor and thus the viral replication in cell cultures.[13] However, further clinical studies of sCD4 with viral isolates were disappointing. The reasons were attributed to the insensitivity of the latter for the former and the difficulty in attaining sufficient therapeutic plasma levels due to the short half-life of sCD4.

Some polyanionic compounds, whatever anion they are based upon, have also been found to inhibit the virus adsorption. Suramin, a hexasulfonatenaphthylurea derivative, was the first compound to enter clinical trials as a possible chemotherapeutic agent against AIDS in the United States.[14] However, due to insufficient immunological benefit, it was dropped in favor of the presently used drugs.

Polyanionic substances suffer from a number of pharmacokinetic and toxicological drawbacks, which seem to mar their clinical utility. They are poorly

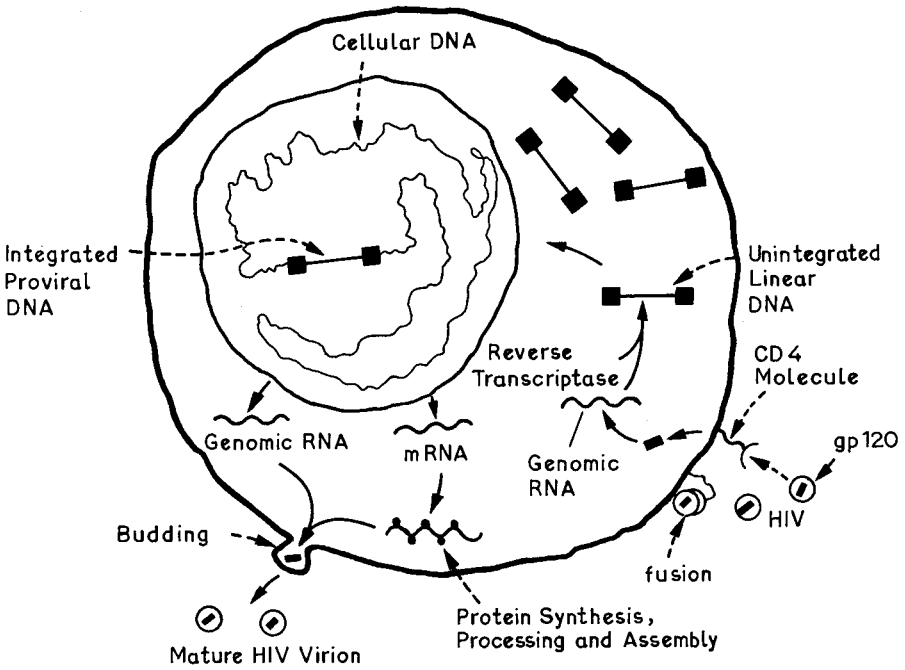


Figure 1. Life cycle of HIV-1. (Reprinted with permission from ref 11. Copyright 1988 Science.)

absorbed following oral administration[15] (typical of molecules with very low log *P* values) and even if directly delivered in the blood stream, they do not easily cross lipophilic barriers. Because of these drawbacks, medicinal chemists could not sustain their interest in polyanionic compounds.

B. Virus Cell Fusion Inhibitors

A number of compounds such as mannose-specific plant lecitins,[16,17] the polypeptidepolyphemusin,[18] negatively charged albumins,[19,20] and triterpene derivatives[21a] have been postulated to interfere specifically with virus−cell fusion. Recently, Debnath et al. reported inhibitors against the gp41 core structure of HIV-1.[21b] The virus−cell fusion depends on the interaction of the envelope glycoproteins gp120 and gp141 with the cell membrane, but it is as yet not clear with which region(s) of gp120 or gp41 the fusion inhibitors actually interact. Further, it has been difficult, so far, to assess the clinical usefulness of the virus−cell fusion inhibitors, as the toxicological and pharmacokinetic profiles for most of these compounds remain to be established.

C. Virus Uncoating Inhibitors

The virus uncoating has been regarded as an appropriate target for antiviral agents. It has been speculated that HIV p24 capsid protein can interact with the virus uncoating inhibitors.[22] At present, however, there is only one group of compounds that have been found to inhibit the virus uncoating. These are the bicyclams, of which the prototypes are JM 2763 and JM 3100.

D. Reverse Transcriptase Inhibitors

The process of reverse transcription of genomic RNA into double-stranded DNA by the enzyme reverse transcriptase (RT) is central to the replication of HIV. Therefore, the inhibition of this key biochemical event in the viral life cycle provides the most attractive target for anti-HIV drug development. Most of the compounds approved so far by the FDA in the United States for the treatment of HIV infections are RT inhibitors.[178] Among them zidovudine (AZT) (**1**), zalcitabine (DDC) (**2**), didanosine (DDI) (**3**), stavudine (D4T) (**4**), lamivudine (3TC) (**5**), and abacavir succinate (**6**) belong to the class of 2′,3′-dideoxynucleoside (ddN) analogues while nevirapine (**7**), delavirdine (**8**), and efavirenz (**9**) belong to the non-nucleoside (NN) class. Besides, several other non-nucleoside reverse transcriptase inhibitors (NNRTIs) have proceeded onto clinical development such as tivirapine (**11**) and the HEPT derivative MKC-442 (**12**) to name a few.[182] NNRTIs have recently gained an increasingly important role in the therapy of HIV infections, Chart 1.

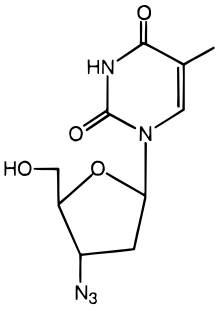
The ddN analogues are substrate analogues and hence interact at the substrate binding sites of the enzyme characterized by its catalytic triad (D110, D185, and D186), where the deoxynucleoside triphosphates normally bind. To be able to interact with the enzyme, the ddNs need to undergo in vivo phosphorylation to generate the 5′-triphosphate derivatives (ddNTP) so that they can compete with natural substrates (dTTP, dCTP, dATP, and dGTP).

While ddN analogues, after being converted to the corresponding triphosphates, compete with natural substrates to interact with the enzyme, non-nucleoside (NN) analogues have been found to interact noncompetitively with an allosteric site, leading to inactivation of the enzyme.[23,24] A detailed discussion of both ddN and NN classes of RT inhibitors will be presented in the next section.

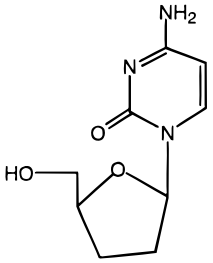
E. Integrase Inhibitors

Incorporation of viral DNA into the host cell genome could be translated as the basis of life-long

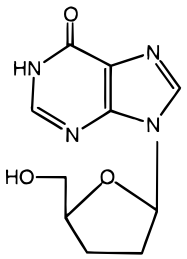
Chart 1



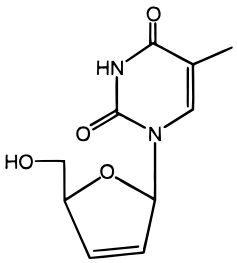
1
Zidovudine



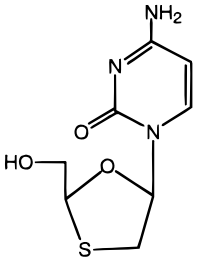
2
Zalcitabine



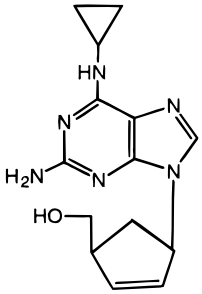
3
Didanosine



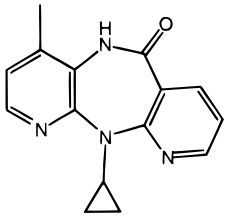
4
Stavudine



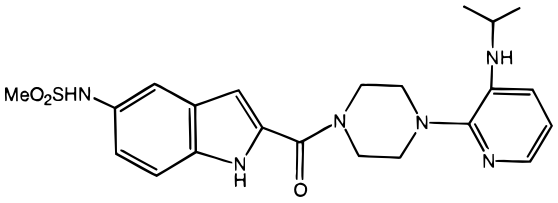
5
Lamivudine



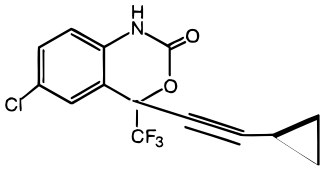
6
Abacavir



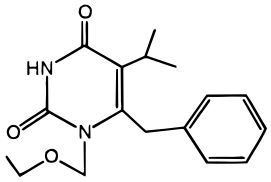
7
Nevirapine



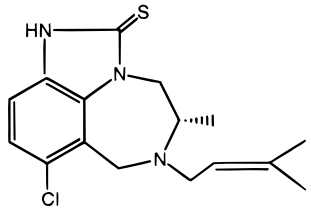
8
Delavirdine



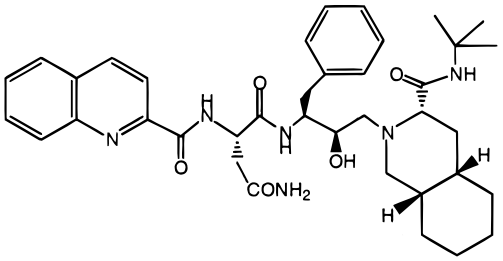
9
Efavirenz



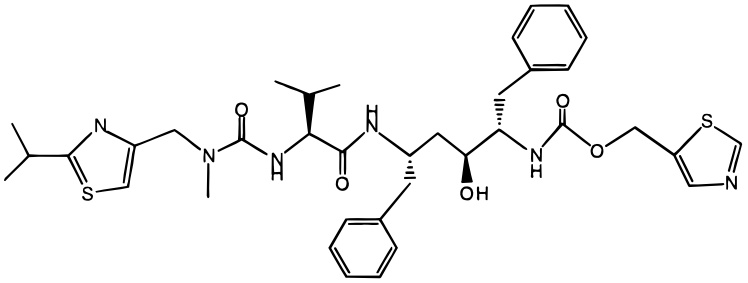
10
MKC-442
HEPT derivative



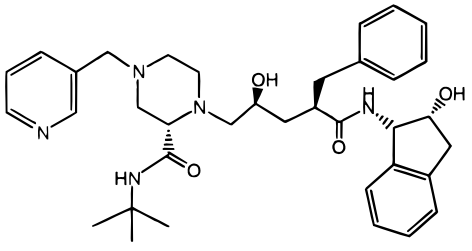
11
Tivirapine
8-Cl-TIBO



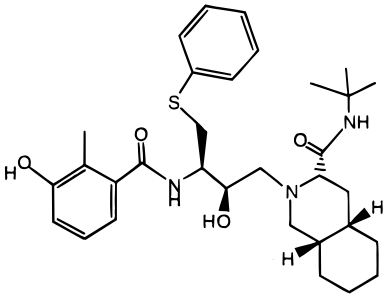
12
Saquinavir



13
Ritonavir



14
Indinavir



15
Nelfinavir

infection. Therefore, this biochemical event, catalyzed by the enzyme integrase, is a pivotal step in viral life cycle and thus worthy of being exploited to develop for anti-HIV chemotherapy. The enzyme integrase is produced by protease-mediated cleavage of the gag-pol precursor during virion maturation. A wide array of compounds has been speculated to act as integrase inhibitors. Several DNA binding agents were found to inhibit HIV-1 integrase, probably due to a nonspecific interaction with the DNA binding domain of the enzyme.[25] Catechol derivatives have also been found to act as integrase inhibitors, but they have been postulated to elicit their effects by interfering with the coordination of the metal ions that are required for the phosphoryl transfer.[26] However, the catechol derivatives do not exhibit much antiviral specificity in the cell culture and hence are no longer considered to be worth pursuing.[27]

F. Gene Expression Inhibitors

The viral integration into host cell genome becomes a sure cause of chronic infection, as the replicative machinery of the host cell will continue producing viral gene products (mRNA). If the translation of mRNA (gene expression) is inhibited, it may lead to the prevention of the spread of infection. Antisense oligonucleotides are generally thought to be the plausible inhibitors of this process due to their capacity to form stable duplexes with complimentary sequences of the viral mRNA. However, some pertinent problems related to the cost of synthesis, bioavailability, site-specific delivery, and hybridization at a desired location will have to be addressed in order to fully realize the therapeutic utility of antisense oligonucleotides.[28]

G. Protease Inhibitors

The cleavage of large polypeptide precursors into smaller, functional protein fragments required for packaging and infectivity of budding virions needs HIV protease. HIV protease is a viral encoded homodimeric aspartyl protease with C_2 symmetry. A catalytic triad of Asp-Thr-Gly contributed by each monomer comprises the active site of this enzyme. The inhibition of this enzyme in vitro results in the production of progeny virions that are immature and noninfectious.[29,30] Since lots of structural information is available on this enzyme, it has become an attractive target for computer-aided drug design strategies[31,32] and consequently a prime focus for the development of anti-HIV chemotherapy.[33] A plethora of peptidal (substrate-based) and nonpeptidal HIV-1 protease inhibitors have been described.[31,34] The FDA has approved several HIV-protease inhibitors for the treatment of HIV, e.g., saquinavir (**12**), ritanovir (**13**), indinavir (**14**), and nelfinavir (**15**). The development of anti-HIV chemotherapy based on protease inhibition will always be an ongoing need because the virus has the ability to rapidly generate resultant mutants.[35]

H. Glucosidase Inhibitors

The final step in the viral replication, leading to virion maturation, involves the processing of surface glycoproteins by the enzyme HIV glucosidase. This enzyme cleaves off glucose units from the oligosaccharide chain and thus helps the maturation of infectious virion. The inhibition of this enzyme, therefore, will lead to the inhibition of virion maturation. Polyhydroxylated compounds such as castenospermine and *N*-butyldeoxynojirimycin have demonstrated inhibitory potential in preclinical evaluation.[36] However, the selectivity of these compounds and their ability to distinguish between cellular and viral glycosylation has to be confirmed before widespread use.

Thus, efforts have been made to exploit all the above-mentioned intervention stages in the viral life cycle to develop anti-HIV chemotherapy. However, because of the availability of structural information on reverse transcriptase (RT) and virus protease (HIV-1-PR) enzymes, the structure−activity relationship (SAR) studies were mainly focused on only the inhibitors of these two enzymes. The quantitative SAR study simultaneously picked up and greatly facilitated the design and synthesis of prospective, therapeutically useful anti-HIV agents. We discuss here QSAR studies available on them and also whatever was available or derivable for other intervention strategies.

The purpose of this review is limited to those studies that have examined a significant number of "congeners" to allow quantitative analysis. Many studies focusing on a few molecules or molecules that are obviously not congeners have not been considered. For the sake of completeness, we have included QSAR that are not as complete as we would want, for a stand alone article, although they are statistically valid in terms of F-statistics. Most researchers in QSAR like to see a minimum of five well-spaced data points per term. Many of those covered do not meet this standard; nevertheless, we believe that these studies are useful starting points for the next round of synthesis and testing.

IV. Methods

All the anti-HIV data has been collected from the literature (see individual data sets for detailed references). The anti-HIV activity of the compounds has been expressed either by the compound's ability to inhibit the enzyme or by the compound's ability to protect MT-4 or CEM cells against the cytopathic effect of the virus. In either case, the concentration of the compound leading to 50% effect has been measured and expressed as IC_{50} for the former and EC_{50} in mol/L or mol/g for the latter. For the selectivity of the compounds, their cytotoxic effect has also been measured in terms of CC_{50}, the concentration of the compound required to reduce by 50% the number of mock-infected MT-4 or CEM cells. The logarithms of the inverse of these parameters have been used as biological end points (log 1/*C*) in the QSAR studies.

All the physicochemical parameters are automatically loaded from our C-QSAR database, and the QSAR regression analyses were executed with the C-QSAR program. The utility of the QSAR program in comparative correlation analysis has been discussed.[37−40,45,66,136]

Included in the program are all the commonly used substituent parameters.[39]

The parameters used in this report have been discussed in detail along with their applications.[40] Here we provide a brief definition. Es is the classic Taft parameter derived from the rate of hydrolysis of aliphatic esters. It is normally most useful for intramolecular steric effects but with relatively small substituents sometimes accounts for intermolecular interactions. CMR is the calculated molar refractivity for the whole molecule. MR is calculated as follows: $((n^2 - 1)/(n^2 + 2))\cdot(MW/d)$, where n is the refractive index, MW is the molecular weight, and d is the density of a substance. Since there is very little variation in n, MR is largely a measure of volume with a small correction for polarizability. We have scaled our MR values by 0.1. MR can be used for a substituent or for the whole molecule. MgVol is the molar volume calculated by the methods of McGowan. B1, B5, and L are the Verloop's sterimol parameters for substituents.[41] B1 is a measure of the width of the first atom of a substituent, B5 is an attempt to define the overall volume, and L is the substituent length. Clog P is the calculated[40] octanol/water partition coefficient of the molecule, and π is that of a substituent. Log P applies to the *neutral form* of partially ionized compounds. The electronic Hammett parameters σ, σ^-, and σ^+ apply to substituent effects on aromatic systems, and Taft's σ^* applies to aliphatic systems. The number in parentheses in the QSAR equations are for 95% confidence intervals.

V. QSAR Results and Discussion

A. Reverse Transcriptase (RT) Inhibitors

The RT inhibitors can be broadly put into two categories: (1) 2′,3′-dideoxynucleoside analogues (ddNs) or substrate analogues and (2) non-nucleoside RT inhibitors (NNRTIs) or nonsubstrate analogues. The ddN inhibitors are first metabolized to the corresponding 5′-phosphates and then act as competitive inhibitors/alternative substrates of the enzyme. Binding at the substrate binding site they act as chain terminators, as following their incorporation into the growing DNA chain they do not permit further chain elongation.[42]

However, ddN analogues have also been found to elicit some toxic side effects that may be attributed to the interference of their metabolites (5′-mono-, di-, and triphosphates) with 2′-deoxynucleoside metabolism and, in particular, to the interference of their triphosphate metabolites with the cellular DNA polymerization process. Therefore, nonsubstrate analogues that do not interact with the substrate binding site of DNA polymerases, whether DNA dependent or RNA dependent, may be expected not to cause any side effects that compromise the clinical utility of the ddN analogues.[23,43]

The non-nucleoside RT inhibitors interact noncompetitively with an allosteric site of the enzyme and thus do not directly impair the function of the substrate binding site.[44] In fact, NNRTIs have a comparatively higher binding affinity for the enzyme−substrate complex than for the free enzyme itself.

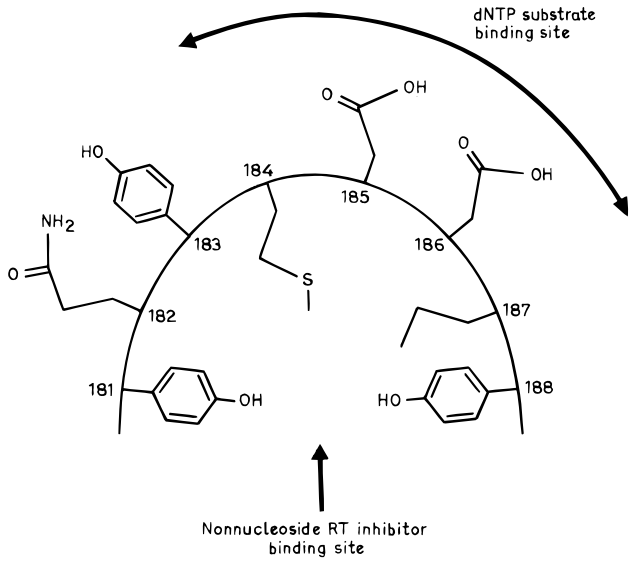


Figure 2. Scheme for non-nucleoside RT inhibitor binding. (Reprinted with permission from ref 23. Copyright 1993 John Wiley & Sons, Inc.)

Their interaction with the enzyme leads to a conformational change in the enzyme, resulting in a decrease in the affinity of the active site for the substrate. However, NNRTIs are active against the RT of only HIV-1 and not of HIV-2 or any other retrovirus. This specificity of NNRTIs for the HIV-1-RT is due to the presence in HIV-1-RT, and not in other RTs or DNA polymerases, of a flexible highly hydrophobic pocket in which a nonsubstrate analogue can fit snugly.[46,47,164]

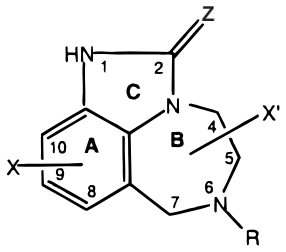
The hydrophobic pocket in HIV-1-RT is formed by the hydrophobic residues of the Y181−Y188 region (Figure 2).[23] The hydrophilic residues, in particular the D184-D186 dipeptide, which are essential for nucleotide substrate binding are located at the outside of the pocket.

Several classes of compounds could be considered as NNRTIs, which are specifically targeted at HIV-1 RT. They can be put into the following categories: (1) Tetrahydroimidazobenzodiazepinone (TIBO) derivatives,[48,49] (2) Hydroxyethoxymethylphenylthiothymine (HEPT) derivatives,[50,51] (3) Dihydropyridodiazepinone such as nevirapine derivatives,[52] (4) Pyridinone derivatives,[53] (5) Bis(heteroaryl)piperazine (BHAP) derivatives,[54] (6) Tertiarybutyldimethylsilylspiroaminooxathioledioxide (TSAO) pyrimidine nucleosides,[55] and (7) α-Anilinophenylacetamide (α-APA) derivatives.[56]

These NNRTIs, as demonstrated with some representatives,[57] are able to completely suppress virus replication in cell cultures for at least 3 months (and probably longer),[58] while under the same conditions, ddN analogues cannot prevent the virus from breaking through even after a few days in the continued presence of the compound.[58,59] Because of this, SAR studies are more directed NNRTIs than toward ddN analogues. However, we present here QSAR results, whenever available or derivable, on both categories of inhibitors. Since from a QSAR point of view NNRTIs are more widely studied than ddN analogues, we present here the first QSARs of NNRTIs.

1. Non-nucleoside RT Inhibitors

a. TIBO Derivatives. *(i) IC_{50} Inhibition Data of TIBO Derivatives (**16**) in MT-4 Cells (Table 1).*[60−63]



16

Kukla et al.[60−63] reported anti-HIV activity (IC_{50}, minimum inhibitory concentration of the compound required to achieve 50% protection of MT-4 cells against the cytopathic effect of virus) for three different series of TIBO (4,5,6,7-tetrahydro-5-methylimidazo[4,5,1-*j*,*k*][1,4] benzodiazepin-2(1*H*)-one) derivatives (**16**). We combined all three series and derived eq 1.

$$\log 1/C = 0.39(\pm 0.20)\text{Clog } P + 1.16(\pm 0.43)B1_8 + 1.49(\pm 0.28)I_z + 0.88(\pm 0.36)I_R + 1.80(\pm 0.78)$$

$$n = 82,\ r^2 = 0.861,\ s = 0.55,\ q^2 = 0.840 \qquad (1)$$

Outliers 7

Outliers not included in the derivation of a QSAR are indicated in the data tables. In eq 1, n is the number of data points, r is the correlation coefficient, s is the standard deviation, q^2 is calculated as described by Cramer et al.,[63b] and the data within the parentheses are for the 95% confidence intervals. Z at ring C was either S or O. The indicator variable I_Z is used to account for this variation with a value of unity for the former and zero for the latter. Its high positive coefficient indicates that the sulfur leads to better activity than oxygen. It may be that both sulfur and the oxygen are involved in some charge-transfer phenomenon with the receptor in which the former may act as a better electron donor than the latter. The variation in R is accounted for by the variable I_R with a value of unity for 3,3-dimethylallyl (DMA) and zero for others. Its positive coefficient indicates that DMA is most suitable at the 6-position. Gupta et al.[64] proposed that the substituents at this position may be involved in some hydrophobic interaction with the receptor in which the DMA may be expected to possess the optimum hydrophobicity. The X-substituents at ring A have wide variation and are at different positions. All of them are shown to effect the activity by their hydrophobic property, but those at the 8-position are shown to also have some additional positive steric effect. This additional effect is described by the Verloop's sterimol parameter $B1_8$. The Clog P values have been obtained using BioByte software.[37] The coefficient with Clog P is within the normally expected range of 0.2−1.2.[66]

Gupta and Garg[64] also published an extensive QSAR study on TIBO derivatives and suggested that a hydrophobic X-substituent will be beneficial to the activity and that it would be more advantageous if it is at the 8-position. Similarly, they also suggested that at position 2 Z = S was more favorable than Z = O and that at position 6 a 3,3-dimethylallyl group would be preferred.

*(ii) EC_{50} Activity of TIBO Derivatives (**16**) in MT-4 Cells (Table 2).*[65] Pauwels et al.[65] reported EC_{50} (effective concentration of the compound required to achieve 50% protection of MT-4 cells against the cytopathic effect of virus) data for another small series of **16** where X′ was either 4-Me or 5-Me. We derived eq 2 for that data and found that the activity is significantly correlated to hydrophobicity of the molecule and an indicator variable I_Z, which has the same meaning as in eq 1.

$$\log 1/C = 0.93(\pm 0.34)\text{ Clog } P + 1.79(\pm 0.56)I_Z + 3.50(\pm 1.46)$$

$$n = 22,\ r^2 = 0.820,\ s = 0.57,\ q^2 = 0.767 \qquad (2)$$

Outlier 1

*(iii) CC_{50} Activity of TIBO Derivatives (**16**) in MT-4 Cells (Table 3).*[65] The cytotoxic data CC_{50} (cytotoxic concentration to reduce the viability of 50% mock infected MT-4 cells) of the same series[65] for which eq 2 was derived was found to be correlated with hydrophobicity in a parabolic fashion (eq 3).

$$\log 1/C = 2.40(\pm 1.49)\text{ Clog } P - 0.22(\pm 0.18)(\text{Clog } P)^2 - 1.88(\pm 3.02)$$

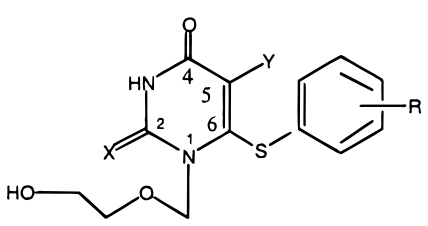
$$n = 19,\ r^2 = 0.813,\ s = 0.27,\ q^2 = 0.753 \qquad (3)$$

$\log P_0 = 5.55$ (4.84−12.73), Outliers 3

This parabolic model, however, gives an optimum value of log P equal to 5.55. The coefficient with the log P terms in eqs 1 and 2 are in the normally expected range.[66a,b] The hydrophobicity of the molecule in TIBO derivatives seems to play an important role in the inhibition of HIV-1 reverse transcriptase.

b. HEPT Derivatives. Excellent correlations were obtained for anti-HIV activity of several series of HEPT (1-[(2-hydroxyethoxy)methyl]-6-(phenylthio)-thymine) derivatives.

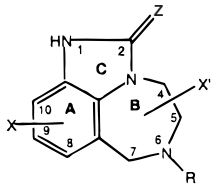
*(i) EC_{50} Activity of HEPT Derivatives (**17**) in MT-4 Cells (Table 4).*[67a] Tanaka et al.[67a] synthesized and



17

measured EC_{50} activity data for a series of HEPT

Table 1. IC_{50} Activity of TIBO Derivatives (16)[60−63]



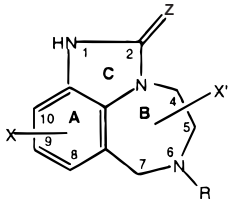
	substituents				log 1/*C*						
no.	X	Z	R	X′	obsd	calcd (eq 1)	Δ	Clog *P*	$B1_8$	I_Z	I_R
1	H	S	DMA[a]	5-Me(S)	7.36	6.72	0.65	3.53	1.00	1.00	1.00
2	9-Cl	S	DMA	5-Me(S)	7.47	6.99	0.48	4.24	1.00	1.00	1.00
3	8-Cl	S	DMA	5-Me(S)	8.37	7.92	0.45	4.24	1.80	1.00	1.00
4	8-F	S	DMA	5-Me(S)	8.24	7.18	1.06	3.67	1.35	1.00	1.00
5	8-SMe	S	DMA	5-Me(S)	8.30	7.75	0.55	4.09	1.70	1.00	1.00
6	8-OMe	S	DMA	5-Me(S)	7.47	7.09	0.38	3.45	1.35	1.00	1.00
7	8-OC_2H_5	S	DMA	5-Me(S)	7.02	7.30	−0.28	3.98	1.35	1.00	1.00
8	8-CN	O	DMA	5-Me(S)	5.94	6.01	−0.07	3.76	1.60	0.00	1.00
9	8-CN	S	DMA	5-Me(S)	7.25	7.19	0.06	2.96	1.60	1.00	1.00
10	8-CHO	S	DMA	5-Me(S)	6.73	7.16	−0.43	2.89	1.60	1.00	1.00
11	8-$CONH_2$	O	DMA	5-Me(S)	5.20	5.44	−0.24	2.59	1.50	0.00	1.00
12	8-Br	O	DMA	5-Me(S)	7.33	6.87	0.46	4.91	1.95	0.00	1.00
13	8-Br	S	DMA	5-Me(S)	8.52	8.16	0.36	4.39	1.95	1.00	1.00
14	8-I	O	DMA	5-Me(S)	7.06	7.21	−0.15	5.17	2.15	0.00	1.00
15	8-I	S	DMA	5-Me(S)	7.32	8.49	−1.17	4.66	2.15	1.00	1.00
16	8-C≡−CH	O	DMA	5-Me(S)	6.36	6.15	0.21	4.11	1.60	0.00	1.00
17	8-C≡−CH	S	DMA	5-Me(S)	7.53	7.52	0.01	3.80	1.60	1.00	1.00
18	8-Me	O	DMA	5-Me(S)	6.00	6.15	−0.15	4.34	1.52	0.00	1.00
19	8-Me	S	DMA	5-Me(S)	7.87	7.52	0.36	4.03	1.52	1.00	1.00
20	9-NO_2	O	CPM[b]	5-Me(S)	4.48	4.25	0.23	3.30	1.00	0.00	0.00
21	8-NH_2	O	CPM	5-Me(S)	3.07	4.11	−1.04	1.88	1.35	0.00	0.00
22	8-NMe_2	O	CPM	5-Me(S)	5.18	4.65	0.53	3.27	1.35	0.00	0.00
23	9-NH_2	O	CPM	5-Me(S)	4.22	3.70	0.52	1.88	1.00	0.00	0.00
24	9-NMe_2	O	CPM	5-Me(S)	5.18	4.24	0.94	3.27	1.00	0.00	0.00
25	9-NHCOMe	O	CPM	5-Me(S)	3.80	3.52	0.28	1.42	1.00	0.00	0.00
26	9-NO_2	S	CPM	5-Me(S)	5.61	5.45	0.16	2.55	1.00	1.00	0.00
27	9-F	S	DMA	5-Me(S)	7.60	6.77	0.83	3.67	1.00	1.00	1.00
28	9-CF_3	O	DMA	5-Me(S)	5.23	5.84	−0.61	5.09	1.00	0.00	1.00
29	9-CF_3	S	DMA	5-Me(S)	6.31	7.06	−0.75	4.41	1.00	1.00	1.00
30	9-Me	O	DEA[c]	5-Me(S)[e]	6.50	5.01	1.49	5.22	1.00	0.00	0.00
31	10-OMe	O	DMA	5-Me(S)	5.18	5.30	−0.12	3.71	1.00	0.00	1.00
32	10-OMe	S	DMA	5-Me(S)[e]	5.33	6.62	−1.29	3.28	1.00	1.00	1.00
33	9,10-di-Cl	S	DMA	5-Me(S)	7.60	7.23	0.37	4.84	1.00	1.00	1.00
34	10-Br	S	DMA	5-Me(S)	5.97	7.05	−1.08	4.39	1.00	1.00	1.00
35	H	O	$CH_2CH{=}CH_2$	5-Me(S)	4.15	4.10	0.05	2.91	1.00	0.00	0.00
36	H	O	2-MA[d]	5-Me(S)	4.33	4.26	0.07	3.31	1.00	0.00	0.00
37	H	O	CH_2CO_2Me	5-Me(S)	3.07	3.78	−0.71	2.09	1.00	0.00	0.00
38	H	O	$CH_2C{=}{-}CH$	5-Me(S)	3.24	3.84	−0.60	2.24	1.00	0.00	0.00
39	H	O	CH_2-2-furanyl	5-Me(S)	3.97	4.03	−0.06	2.72	1.00	0.00	0.00
40	H	O	$CH_2CH{=}CH_2$[S(+)]	5-Me(S)	4.18	4.10	0.08	2.91	1.00	0.00	0.00
41	H	O	$CH_2CH_2CH{=}CH_2$	5-Me(S)	4.30	4.23	0.07	3.24	1.00	0.00	0.00
42	H	O	$CH_2CH_2CH_3$	5-ME(S)	4.05	4.15	−0.10	3.02	1.00	0.00	0.00
43	H	O	2-MA[S(+)]	5-ME(S)	4.72	4.26	0.46	3.31	1.00	0.00	0.00
44	H	O	CPM	5-Me(S)	4.36	4.18	0.18	3.11	1.00	0.00	0.00
45	H	O	$CH_2CH{=}CHMe(E)$	5-Me(S)	4.24	4.24	0.00	3.27	1.00	0.00	0.00
46	H	O	$CH_2CH{=}CHMe(Z)$	5-Me(S)	4.46	4.24	0.22	3.27	1.00	0.00	0.00
47	H	O	$CH_2CH_2CH_2Me$	5-Me(S)	4.00	4.35	−0.35	3.55	1.00	0.00	0.00
48	H	O	DMA	5-Me(S)	4.90	5.35	−0.45	3.84	1.00	0.00	1.00
49	H	O	$CH_2C(Br){=}CH_2$	5-Me(S)	4.21	4.37	−0.16	3.60	1.00	0.00	0.00
50	H	O	$CH_2C(Me){=}CHMe(E)$	5-Me(S)	4.54	4.40	0.14	3.67	1.00	0.00	0.00
51	H	O	DMA[*R*(+)]	5-Me(S)	4.66	5.35	−0.69	3.84	1.00	0.00	1.00
52	H	O	DMA[*S*(+)]	5-Me(S)	5.40	5.35	0.05	3.84	1.00	0.00	1.00
53	H	O	$CH_2C(C_2H_5){=}CH_2$	5-Me(S)	4.43	4.40	0.03	3.67	1.00	0.00	0.00
54	H	O	$CH_2CH{=}CHC_6H_5(Z)$	5-Me(S)	3.91	4.73	−0.82	4.51	1.00	0.00	0.00
55	H	O	$CH_2C(CH{=}CH_2){=}CH_2$	5-Me(S)	4.15	4.30	−0.15	3.41	1.00	0.00	0.00
56	8-Cl	S	DMA	H	7.34	7.66	−0.32	3.56	1.80	1.00	1.00
57	9-Cl	S	DMA	H	6.80	6.73	0.08	3.56	1.00	1.00	1.00
58	H	O	2-MA	5,5-di-Me	4.64	4.40	0.25	3.66	1.00	0.00	0.00
59	H	O	2-MA	4-Me	4.50	4.14	0.36	3.00	1.00	0.00	0.00
60	9-Cl	S	2-MA	4-Me(S)	6.17	5.91	0.26	3.72	1.00	1.00	0.00
61	9-Cl	S	CPM	4-Me(R)	5.66	5.83	−0.17	3.52	1.00	1.00	0.00
62	H	O	C_3H_7	4-$CHMe_2$	4.13	4.51	−0.38	3.95	1.00	0.00	0.00
63	H	O	2-MA	4-$CHMe_2$	4.90	4.62	0.28	4.24	1.00	0.00	0.00
64	H	O	2-MA	4-C_3H_7	4.32	4.67	−0.35	4.37	1.00	0.00	0.00

Table 1. (Continued)

	substituents				log 1/*C*						
no.	X	Z	R	X′	obsd	calcd (eq 1)	Δ	Clog *P*	$B1_8$	I_Z	I_R
65	H	O	DMA	7-Me	4.92	5.35	−0.43	3.84	1.00	0.00	1.00
66	8-Cl	O	DMA	7-Me	6.84	6.64	0.20	4.76	1.80	0.00	1.00
67	9-Cl	O	DMA	7-Me	6.80	5.71	1.09	4.76	1.00	0.00	1.00
68	H	S	C_3H_7	7-Me	5.61	5.52	0.10	2.72	1.00	1.00	0.00
69	H	S	DMA	7-Me	7.11	6.72	0.40	3.53	1.00	1.00	1.00
70	8-Cl	S	DMA	7-Me	7.92	7.93	−0.01	4.24	1.80	1.00	1.00
71	9-Cl	S	DMA	7-Me	7.64	6.99	0.65	4.24	1.00	1.00	1.00
72	H	O	DMA	4,5-di-Me(*cis*)	4.25	5.55	−1.30	4.36	1.00	0.00	1.00
73	H	S	DMA	4,5-di-Me(*cis*)	5.65	6.92	−1.27	4.05	1.00	1.00	1.00
74	H	S	CPM	4,5-di-Me(*trans*)	4.87	5.75	−0.88	3.32	1.00	1.00	0.00
75	H	S	DMA	4,5-di-Me(*trans*)[e]	4.84	6.92	−2.08	4.05	1.00	1.00	1.00
76	H	S	DMA	5,7-di-Me(*trans*)	7.38	6.92	0.46	4.05	1.00	1.00	1.00
77	H	S	DMA	5,7-di-Me(*cis*)	5.94	6.92	−0.098	4.05	1.00	1.00	1.00
78	9-Cl	O	DMA	5,7-di-Me(R,R-*trans*)	6.64	5.91	0.73	5.28	1.00	0.00	1.00
79	9-Cl	S	DMA	5,7-di-Me(R,R-*trans*)	6.32	7.20	−0.88	4.77	1.00	1.00	1.00
80	H	S	DMA	4,7-di-Me(*trans*)[e]	4.59	6.92	−2.33	4.05	1.00	1.00	1.00
81	9-Cl	O	DMA	5-Me(S)	6.74	5.71	1.03	4.76	1.00	0.00	1.00
82	9-Cl	S	CPM	5-Me(S)[e]	7.47	5.83	1.64	3.52	1.00	1.00	0.00
83	H	S	CPM	5-ME(S)[e]	7.22	5.55	1.67	2.80	1.00	1.00	0.00
84	H	O	C_3H_7	5-Me	4.22	4.15	0.07	3.02	1.00	0.00	0.00
85	H	S	C_3H_7	5-Me	5.78	5.52	0.27	2.72	1.00	1.00	0.00
86	H	O	2-MA	5-Me	4.46	4.26	0.20	3.31	1.00	0.00	0.00
87	H	S	DMA	5-Me	7.01	6.72	0.30	3.53	1.00	1.00	1.00
88	H	O	DMA	5-Me(S)	5.48	5.35	0.13	3.84	1.00	0.00	1.00
89	H	S	2-MA	5-Me(S)[e]	7.59	5.63	1.96	3.00	1.00	1.00	0.00

[a] 3,3-Dimethylallyl. [b] Cyclopropylmethyl. [c] 3,3-Diethylallyl. [d] 2-Methylallyl. [e] Data points not included in deriving equation.

Table 2. EC_{50} Activity of TIBO Derivatives (16)[65]

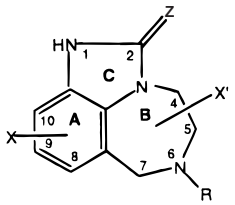


	substituents				log 1/*C*				
no.	X	Z	R	X′	obsd	calcd (eq 2)	Δ	Clog *P*	I_Z
1	H	O	$CH_2CH{=}CH_2$	5-Me	4.23	4.41	−0.18	2.91	1.00
2	H	O	$CH_2CH{=}CH_2$	5-Me	4.17	4.41	−0.24	2.91	1.00
3	H	O	$CH_2C(Me){=}CH_2$	5-Me	4.85	4.79	0.07	3.31	1.00
4	H	O	$CH_2CH{=}CMe_2$	5-Me	5.38	5.28	0.10	3.84	1.00
5	9-Cl	O	$CH_2C(Me){=}CH_2$	5-Me	5.33	5.65	−0.32	4.23	1.00
6	9-Me	O	$CH_2CH{=}(C_2H_5)_2$	5-Me	6.51	6.57	−0.06	5.22	1.00
7	9-Cl	O	$CH_2CH{=}CMe_2$	5-Me	7.60	6.97	0.64	5.65	1.00
8	H	S	$CH_2CH{=}CMe_2$	5-Me	7.36	6.79	0.57	3.53	0.00
9	H	S	$CH_2CH{=}CMe_2$	5-Me	7.41	7.27	0.14	4.05	0.00
10	7-Me	S	$CH_2CH{=}CMe_2$	5-Me	6.10	7.27	−1.17	4.05	0.00
11	H	S	C_3H_7	5-Me	6.62	6.03	0.59	2.72	0.00
12	9-Cl	S	$CH_2CH{=}CMe_2$	5-Me	7.48	7.46	0.03	4.24	0.00
13	9-Cl	S	$CH_2CH_2C_3H_5$	5-Me	6.48	7.27	−0.79	4.04	0.00
14	9-Cl	S	$CH_2CH{=}CMe_2$	5-Me	7.04	7.46	−0.42	4.24	0.00
15	9-Cl	S	$CH_2C_4H_7$	5-Me	7.37	7.30	0.07	4.08	0.00
16	9-Cl	S	$CH_2CH{=}C(C_2H_5)_2$	5-Me	7.82	8.28	−0.46	5.13	0.00
17	9-Cl	S	$CH_2CH(Me){=}CH_2$	4-Me	6.35	6.96	−0.62	3.72	0.00
18	9-Cl	S	$CH_2CH_2C_3H_5$	4-Me[a]	5.66	7.27	−1.61	4.04	0.00
19	9,10-di-Cl	S	$CH_2CH{=}CMe_2$	5-Me	7.60	8.01	−0.41	4.84	0.00
20	8-Cl	S	$CH_2CH{=}CMe_2$	5-Me	8.34	7.46	0.88	4.24	0.00
21	8-Cl	S	$CH_2CH{=}C(C_2H_5)_2$	5-Me	8.29	8.28	0.01	5.13	0.00
22	8-Br	S	$CH_2C{=}CMe_2$	5-Me	8.52	7.59	0.93	4.39	0.00
23	8-Me	S	$CH_2CH{=}CMe_2$	5-Me	7.85	7.21	0.65	3.98	0.00

[a] Data point not used in deriving equation.

derivatives (**17**). Hansch and Zhang[67b] published the first HIV QSAR on the data reported by Tanaka et al. We modified our previously published QSAR to include Clog *P* and obtained eq 4.

Table 3. CC_{50} Activity of TIBO Derivatives (16)[65]



	substituents				log 1/*C*			
no.	X	Z	R	X′	obsd	calcd (eq 3)	Δ	Clog *P*
1	H	O	$CH_2CH{=}CH_2$	5-Me	3.21	3.27	−0.06	2.91
2	H	O	$CH_2CH{=}CH_2$	5-Me	3.17	3.27	−0.10	2.91
3	H	O	$CH_2C(Me){=}CH_2$	5-Me	3.96	3.64	0.32	3.31
4	H	O	$CH_2CH{=}CMe_2$	5-Me[a]	3.33	−1.75	5.08	3.84
5	9-Cl	O	$CH_2C(Me){=}CH_2$	5-ME	4.77	4.41	0.36	4.23
6	9-Me	O	$CH_2CH{=}(C_2H_5)_2$	5-Me	4.70	4.77	−0.07	5.22
7	9-Cl	O	$CH_2CH{=}CMe_2$	5-Me	4.66	4.68	−0.02	5.65
8	H	S	$CH_2CH{=}CMe_2$	5-Me	3.26	3.84	−0.58	3.53
9	H	S	$CH_2CH{=}CMe_2$	5-Me	4.09	4.28	−0.19	4.05
10	7-Me	S	$CH_2CH{=}CMe_2$	5-Me	4.13	4.28	−0.15	4.05
11	H	S	C_3H_7	5-Me	3.25	3.09	0.17	2.72
12	9-Cl	S	$CH_2CH{=}CMe_2$	5-Me	4.47	4.42	0.05	4.24
13	9-Cl	S	$CH_2CH_2C_3H_5$	5-Me	4.44	4.27	0.17	4.04
14	9-Cl	S	$CH_2CH{=}CMe_2$	5-Me	4.72	4.42	0.30	4.24
15	9-Cl	S	$CH_2C_4H_7$	5-Me	4.55	4.30	0.26	4.08
16	9-Cl	S	$CH_2CH{=}C(C_2H_5)_2$	5-Me	4.92	4.77	0.15	5.13
17	9-Cl	S	$CH_2CH(Me){=}CH_2$	4-Me[a]	4.62	−1.68	6.30	3.72
18	9,10-di-Cl	S	$CH_2CH{=}CMe_2$	5-Me[a]	4.35	4.72	−0.38	4.84
19	8-Cl	S	$CH_2CH{=}CMe_2$	5-Me	3.85	−1.99	5.85	4.24
20	8-Cl	S	$CH_2CH{=}C(C_2H_5)_2$	5-Me	4.92	4.77	0.15	5.13
21	8-Br	S	$CH_2C{=}CMe_2$	5-Me	4.28	4.52	−0.25	4.39
22	8-Me	S	$CH_2CH{=}CMe_2$	5-Me	4.10	4.22	−0.13	3.98

[a] Data points not used in deriving equation.

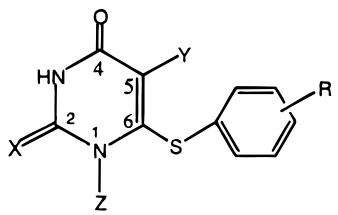
$$\log 1/C = 14.58(\pm3.97)L_Y - 1.97(\pm0.53)(L_Y)^2 - 1.25(\pm0.77)B1_{R,3} + 0.94(\pm0.36)\text{Clog } P + 1.51(\pm0.73)Es_{R,2} - 20.03(\pm7.32)$$

$$n = 32,\ r^2 = 0.911,\ s = 0.45,\ q^2 = 0.863 \qquad (4)$$

$$(L_Y)_0 = 3.70\ (3.58-3.83),\ \text{Outliers } 2$$

Equation 4 suggests that the hydrophobicity of the molecule enhances the activity; however, the ortho and meta position R-substituents have detrimental steric effects. Also, the length of Y-substituent is favorable to the activity, but it has an optimum value of 3.70.

*(ii) EC_{50} Activity of HEPT Derivatives (**18**) in MT-4 Cells (Table 5).*[68] For another equally large series of



18

deoxy analogues of HEPT (**18**) studied by Tanaka et al.,[68] where the variation in the N_1-substituent was reported, eq 5 was derived by us for the anti-HIV data.

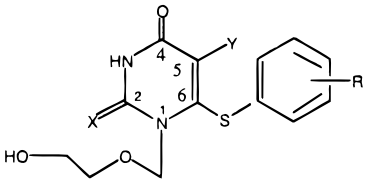
$$\log 1/C = -1.22(\pm0.88)\text{MgVol} + 2.48(\pm0.66)MR_Y + 1.49(\pm0.50)I_{3,5} + 1.44(\pm0.52)I_Z + 6.99(\pm1.91)$$

$$n = 33,\ r^2 = 0.842,\ s = 0.47,\ q^2 = 0.783 \qquad (5)$$

$$\text{Outliers } 2$$

In eq 5, the negative MgVol (calculated molecular volume by McGowan method) term indicates that the big molecules are not suitable for activity. I_Z is an indicator variable used with a value of unity for a Z-substituent that contains a phenyl moiety. Thus, the Z-substituents produce steric effects but a phenyl-containing group seems to have a positive effect, probably due to a good stacking of the planar phenyl ring with the receptor. The indicator variable $I_{3,5}$ is used with a value of unity for 3,5-Me_2 or 3,5-Cl_2 in the phenyl ring, which shows a positive effect of these substituents. The positive MR_Y term shows the steric effect of Y-substituents, similar to eq 4. There is no hydrophobic term in this equation. We believe that may be due to insufficient variation in too rigid a phenyl ring (R-substituents), where a hydrophobic binding site has been proposed by Garg et al.[70] Alkyl R groups might show a different effect. The term $I_{3,5}$

Table 4. EC_{50} Activity of HEPT Derivatives (17)[67a]

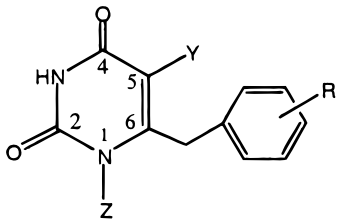


	substituents			log 1/*C*						
no.	R	Y	X	obsd	calcd (eq 4)	Δ	L_Y	$B1_{R,3}$	Clog *P*	$Es_{R,2}$
1	2-Me	Me	O	4.15	4.26	−0.11	2.87	1.00	1.87	−1.24
2	2-NO_2	Me	O	3.85	3.90	−0.05	2.87	1.00	1.11	−1.01
3	2-OMe	Me	O	4.72	4.38	0.34	2.87	1.00	0.89	−0.55
4	3-Me	Me	O	5.59	5.48	0.11	2.87	1.52	1.87	0.00
5	3-C_2H_5	Me	O	5.57	5.97	−0.40	2.87	1.52	2.40	0.00
6	3-CMe_3	Me	O	4.92	5.37	−0.45	2.87	2.60	3.20	0.00
7	3-CF_3	Me	O	4.35	4.78	−0.43	2.87	1.99	1.75	0.00
8	3-F	Me	O	5.48	5.36	0.12	2.87	1.35	1.51	0.00
9	3-Cl	Me	O	4.89	5.33	−0.44	2.87	1.80	2.08	0.00
10	3-Br	Me	O	5.24	5.28	−0.04	2.87	1.95	2.23	0.00
11	3-I	Me	O	5.00	5.28	−0.28	2.87	2.15	2.49	0.00
12	3-NO_2	Me	O	4.47	4.55	−0.08	2.87	1.70	1.13	0.00
13	3-OH	Me	O	4.09	4.60	−0.51	2.87	1.35	0.70	0.00
14	3-OMe	Me	O	4.66	5.15	−0.49	2.87	1.35	1.29	0.00
15	4-Me	Me	O[a]	3.66	6.13	−2.47	2.87	1.00	1.87	0.00
16	3,5-di-Me	Me	O	6.59	5.95	0.64	2.87	1.52	2.37	0.00
17	3,5-di-Cl	Me	O	5.89	6.00	−0.11	2.87	1.80	2.80	0.00
18	3,5-di-Me	Me	S	6.66	6.12	0.54	2.87	1.52	2.56	0.00
19	3-CO_2Me	Me	O	5.10	4.83	0.27	2.87	1.64	1.34	0.00
20	3-COMe	Me	O	5.14	4.39	0.75	2.87	1.60	0.81	0.00
21	3-CN	Me	O	5.00	4.38	0.62	2.87	1.60	0.80	0.00
22	H	$CH_2CH{=}CH_2$	O[a]	5.60	3.72	1.88	5.11	1.00	1.94	0.00
23	H	C_2H_5	S	6.96	7.40	−0.44	4.11	1.00	2.09	0.00
24	H	C_3H_7	S	5.00	5.32	−0.32	4.92	1.00	2.62	0.00
25	H	$CHMe_2$	S	7.23	7.77	−0.54	4.11	1.00	2.49	0.00
26	3,5-di-Me	C_2H_5	S	8.11	7.68	0.43	4.11	1.52	3.09	0.00
27	3,5-di-Me	$CHMe_2$	S	8.30	8.06	0.24	4.11	1.52	3.49	0.00
28	3,5-di-Cl	C_2H_5	S	7.37	7.73	−0.36	4.11	1.80	3.52	0.00
29	H	C_2H_5	O	6.92	7.22	−0.30	4.11	1.00	1.90	0.00
30	H	C_3H_7	O	5.47	5.15	0.32	4.92	1.00	2.43	0.00
31	H	$CHMe_2$	O	7.20	7.59	−0.39	4.11	1.00	2.30	0.00
32	3,5-di-Me	C_2H_5	O	7.89	7.51	0.39	4.11	1.52	2.90	0.00
33	3,5-di-Me	$CHMe_2$	O	8.57	7.88	0.69	4.11	1.52	3.30	0.00
34	3,5-di-Cl	C_2H_5	O	7.85	7.56	0.29	4.11	1.80	3.33	0.00

[a] Data points not used in deriving equation.

may be pointing toward some hydrophobic interaction.

*(iii) EC_{50} Activity of HEPT Derivatives (**19**) in MT-4 Cells (Table 6).*[69] Almost the same conclusions were



19

drawn when Tanaka et al.'s data[69] on a small series of **19,** where the 6-phenylthio group was replaced by a benzyl group, were analyzed by us to obtain eq 6, which agrees well with eq 4.

$$\log 1/C = 0.46(\pm 0.43)\text{Clog } P + 1.19(\pm 0.73)I_{3,5} + 1.71(\pm 1.20)\text{MR}_Y + 3.75(\pm 1.31)$$

$$n = 13,\ r^2 = 0.887,\ s = 0.48,\ q^2 = 0.739 \qquad (6)$$

Garg and Gupta[70] also performed extensive QSAR studies on HEPT derivatives[67−69] and obtained similar results.

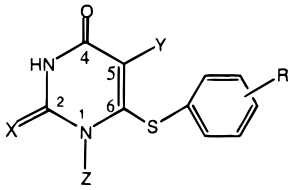
*(iv) CC_{50} Activity of HEPT Derivatives (**18**) in MT-4 Cells (Table 7).*[68] For the series of HEPT compounds (**18**) for which eq 5 was obtained, we also found the cytotoxic data reported for some compounds to be well correlated with the hydrophobicity of the molecules as in eq 7.

$$\log 1/C = 0.35(\pm 0.06)\text{Clog } P + 2.82(\pm 0.22)$$

$$n = 23,\ r^2 = 0.870,\ s = 0.14,\ q^2 = 0.841 \qquad (7)$$

Outliers 2

Table 5. EC_{50} Activity of HEPT Derivatives (18)[68]



	substituents				log 1/*C*						
no.	Z	X	Y	R	obsd	calcd (eq 5)	Δ	MgVol	MR_Y	$I_{3,5}$	I_Z
1	$CH_2OCH_2CH_2OMe$	O	Me	H	5.06	5.54	−0.48	2.35	0.57	0.00	0.00
2	CH_2OMe	O	Me	H	5.68	5.95	−0.27	2.00	0.57	0.00	0.00
3	$CH_2OC_2H_5$	O	Me	H	6.48	5.78	0.70	2.14	0.56	0.00	0.00
4	$CH_2OC_3H_7$	O	ME	H	5.44	5.61	−0.17	2.29	0.57	0.00	0.00
5	$CH_2OC_4H_9$	O	Me	H	5.33	5.44	−0.11	2.43	0.57	0.00	0.00
6	$CH_2OCH_2C_6H_5$	O	Me	H	7.06	6.65	0.40	2.61	0.57	0.00	1.00
7	$CH_2OC_2H_5$	S	C_2H_5	H[a]	7.59	6.64	0.95	2.39	1.03	0.00	0.00
8	$CH_2OC_2H_5$	S	C_2H_5	3,5-Me_2	8.36	7.79	0.57	2.67	1.03	1.00	0.00
9	$CH_2OC_2H_5$	S	C_2H_5	3,5-Cl_2	7.89	7.83	0.05	2.64	1.03	1.00	0.00
10	CH_2CHMe_2	S	C_2H_5	H	6.66	6.54	0.12	2.47	1.03	0.00	0.00
11	$CH_2OC_6H_{11}$	S	C_2H_5	H	5.80	6.08	−0.29	2.85	1.03	0.00	0.00
12	$CH_2OCH_2C_6H_{11}$	S	C_2H_5	H	6.46	5.91	0.54	2.99	1.03	0.00	0.00
13	$CH_2OCH_2C_6H_5$	S	C_2H_5	H	8.11	7.51	0.60	2.86	1.03	0.00	1.00
14	$CH_2OCH_2C_6H_5$	S	C_2H_5	3,5-Me_2	8.16	8.66	−0.50	3.14	1.03	1.00	1.00
15	$CH_2OCH_2C_6H_4$-4-Me	S	C_2H_5	H	7.11	7.34	−0.23	3.00	1.03	0.00	1.00
16	$CH_2OCH_2C_6H_4$-4-Cl	S	C_2H_5	H	7.92	7.36	0.56	2.98	1.03	0.00	1.00
17	$CH_2OCH_2CH_2C_6H_5$	S	C_2H_5	H	7.04	7.34	−0.30	3.00	1.03	0.00	1.00
18	$CH_2OC_2H_5$	S	$CHMe_2$	H	7.85	7.62	0.23	2.53	1.50	0.00	0.00
19	$CH_2OCH_2C_6H_5$	S	$CHMe_2$	H	8.17	8.50	−0.33	3.00	1.50	0.00	1.00
20	$CH_2OC_2H_5$	S	Cy-C_3H_5	H	7.02	7.40	−0.38	2.42	1.35	0.00	0.00
21	$CH_2OC_2H_5$	O	C_2H_5	H	7.72	6.76	0.96	2.29	1.03	0.00	0.00
22	$CH_2OC_2H_5$	O	C_2H_5	3,5-Me_2[a]	8.27	6.93	1.34	2.43	0.57	1.00	0.00
23	$CH_2OC_2H_5$	O	C_2H_5	3,5-Cl_2	8.13	7.96	0.17	2.53	1.03	1.00	0.00
24	CH_2OCHMe_2	O	C_2H_5	H	6.47	6.59	−0.12	2.43	1.03	0.00	0.00
25	$CH_2OC_6H_{11}$	O	C_2H_5	H	5.40	6.21	−0.08	2.74	1.03	0.00	0.00
26	$CH_2OCH_2C_6H_{11}$	O	C_2H_5	H	6.35	6.04	0.31	2.88	1.03	0.00	0.00
27	$CH_2OCH_2C_6H_5$	O	C_2H_5	H	8.23	7.64	0.59	2.75	1.03	0.00	1.00
28	$CH_2OCH_2C_6H_5$	O	C_2H_5	3,5-Me_2	8.50	8.79	−0.29	3.04	1.03	1.00	1.00
29	$CH_2OCH_2CH_2C_6H_5$	O	C_2H_5	H	7.02	7.47	−0.45	2.89	1.03	0.00	1.00
30	$CH_2OC_2H_5$	O	$CHMe_2$	H	7.92	7.75	0.17	2.43	1.50	0.00	0.00
31	$CH_2OCH_2C_6H_5$	O	$CHMe_2$	H	8.57	8.62	−0.05	2.89	1.50	0.00	1.00
32	$CH_2OC_2H_5$	O	Cy-C_3H_5	H	7.00	7.53	−0.53	2.32	1.35	0.00	0.00
33	C_2H_5	O	Me	H	5.66	6.02	−0.37	1.95	0.57	0.00	0.00
34	C_4H_9	O	Me	H	5.92	5.68	0.24	2.23	0.57	0.00	0.00
35	$CH_2OCH_2CH_2OH$	O	Me	H	5.16	5.71	−0.55	2.20	0.57	0.00	0.00

[a] Data points not used in deriving equation.

*(v) CC_{50} Activity of HEPT Derivatives (**18**) in MT-4 Cells (Table 8).*[71] Baba et al.[71] reported cytotoxic activity on another series of HEPT derivatives (**18**) for which we derived eq 8, where again a hydrophobic interaction seems to be involved between receptor−ligand.

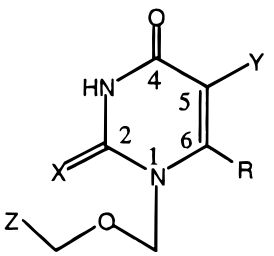
$$\log 1/C = 0.47(\pm 0.09)\text{Clog } P + 2.49(\pm 0.26)$$
$$n = 14,\ r^2 = 0.919,\ s = 0.09,\ q^2 = 0.899 \qquad (8)$$

*(vi) EC_{50} Activity of HEPT Derivatives (**18**) in MT-4 Cells (Table 9).*[71] Baba et al.[71] also reported EC_{50} data on another series of **18**, for which we derived eq 9.

$$\log 1/C = 0.55(\pm 0.22)\text{Clog } P + 1.15(\pm 0.40)B5_Z + 2.35(\pm 0.99)$$
$$n = 19,\ r^2 = 0.894,\ s = 0.37,\ q^2 = 0.860 \qquad (9)$$
Outlier 1

Equation 9 again shows that hydrophobicity of the molecule favors activity. The coefficient with log *P* is in the normally expected range. The positive $B5_Z$ term also shows that Z-substituents have favorable steric effects. A low intercept in eqs 7−9 indicate nonspecific toxicity.

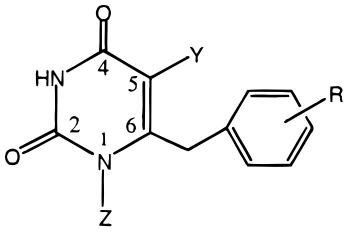
*(vii) EC_{50} Activity of HEPT Derivatives (**20**) in CEM Cells (Table 10).*[72] For a set of compounds where the



20

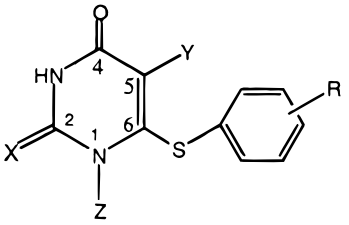
change in substituents was studied by Balzarini et al.[72] at a number of positions and anti-HIV activities (EC_{50}) were evaluated against HIV-1 and a panel of

Table 6. EC_{50} Activity of HEPT Derivatives (19)[69]



	substituents			log 1/*C*					
no.	Y	R	Z	obsd	calcd (eq 6)	Δ	Clog *P*	$I_{3,5}$	MR_Y
1	C_2H_5	H	$CH_2OCH_2CH_2OH$	6.46	6.22	0.24	1.58	0.00	1.03
2	C_2H_5	3,5-Me_2	$CH_2OCH_2CH_2OH$	7.89	7.85	0.03	2.58	1.00	1.03
3	C_2H_5	H	$CH_2OC_2H_5$	7.39	6.77	0.62	2.80	0.00	1.03
4	C_2H_5	3,5-Me_2	$CH_2OC_2H_5$	8.80	8.41	0.39	3.80	1.00	1.03
5	$CHMe_2$	H	$CH_2OCH_2CH_2OH$	7.20	7.19	0.01	1.98	0.00	1.50
6	$CHMe_2$	3,5-Me_2	$CH_2OCH_2CH_2OH$	8.57	8.83	−0.26	2.98	1.00	1.50
7	$CHMe_2$	H	$CH_2OC_2H_5$	8.38	7.75	0.63	3.20	0.00	1.50
8	$CHMe_2$	3,5-Me_2	$CH_2OC_2H_5$	9.22	9.38	−0.16	4.20	1.00	1.50
9	C_2H_5	H	C_4H_9	6.68	7.05	−0.38	3.43	0.00	1.03
10	$CHMe_2$	H	C_4H_9	7.38	8.03	−0.65	3.83	0.00	1.50
11	C_2H_5	H	CH_2H_2OMe	6.60	6.42	0.19	2.02	0.00	1.03
12	$CHMe_2$	H	CH_2CH_2OMe	7.28	7.39	−0.11	2.42	0.00	1.50
13	Me	H	$CH_2OCH_2CH_2OH$	4.63	5.18	−0.54	1.05	0.00	0.56

Table 7. CC_{50} Activity of HEPT Derivatives (18)[68]



	substituent				log 1/*C*			
no.	R	Y	X	Z	obsd	calcd (eq 7)	Δ	Clog *P*
1	H	Me	O	$CH_2OCH_2CH_2OMe$	3.52	3.52	0.00	1.99
2	H	Me	O	$CH_2OCH_2CH_2OC_5H_{11}$	4.26	4.27	−0.01	4.11
3	H	Me	O	$CH_2OCH_2CH_2OCH_2C_6H_5$	4.35	4.22	0.13	3.96
4	H	Me	O	CH_2OMe	3.61	3.55	0.06	2.07
5	H	Me	O	$CH_2OC_2H_5$	3.64	3.74	−0.10	2.60
6	H	Me	O	$CH_2OC_3H_7$	3.83	3.92	−0.09	3.13
7	H	Me	O	$CH_2OC_4H_9$	4.08	4.11	−0.03	3.66
8	H	Me	O	$CH_2OCH_2CH_2SiMe_3$	4.50	4.31	0.19	4.22
9	H	Me	O	$CH_2OCH_2C_6H_5$	4.02	4.06	−0.03	3.51
10	H	C_2H_5	S	$CH_2OC_2H_5$	4.09	3.99	0.10	3.32
11	3,5-di-Cl	C_2H_5	S	$CH_2OC_2H_5$	4.35	4.49	−0.14	4.74
12	H	Cy-C_3H_5	S	$CH_2OC_2H_5$[a]	4.34	3.96	0.38	3.23
13	H	C_2H_5	O	$CH_2OC_2H_5$	3.79	3.92	−0.13	3.13
14	3,5-di-Cl	C_2H_5	O	$CH_2OC_2H_5$	4.35	4.42	−0.08	4.55
15	H	C_2H_5	O	CH_2OCHMe_2	3.85	4.03	−0.19	3.44
16	H	C_2H_5	O	$CH_2OCH_2C_6H_{11}$	4.77	4.67	0.10	5.25
17	H	C_2H_5	O	$CH_2OCH_2C_6H_5$	4.47	4.24	0.23	4.04
18	H	C_2H_5	O	$CH_2OCH_2CH_2C_6H_5$	4.42	4.47	−0.05	4.69
19	H	$CHMe_2$	O	$CH_2OC_2H_5$	3.98	4.06	−0.09	3.53
20	H	Cy-C_3H_5	O	$CH_2OC_2H_5$	3.65	3.89	−0.24	3.04
21	H	Me	O	H	3.60	3.40	0.20	1.64
22	H	Me	O	Me	3.82	3.59	0.24	2.17
23	H	Me	O	C_2H_5[a]	4.03	3.59	0.44	2.17
24	H	Me	O	C_4H_9	4.05	3.96	0.09	3.22
25	H	Me	O	$CH_2OCH_2CH_2OH$	3.13	3.31	−0.17	1.37

[a] Data points not used in deriving equation.

their mutant strains containing single mutations in their RTs as indicated in the footnote of Table 10. We found that the activity depends not only on the hydrophobic character of the molecule, but also on

their various structural characteristics that could be accounted for by different indicator parameters (eqs 10−16).

Activity against HIV-1 wild type:

$$\log 1/C = 4.25(\pm1.67)\text{Clog } P - 0.50(\pm0.22)(\text{Clog } P)^2 - 0.41(\pm0.28)I_5 + 0.43(\pm0.28)I_6 - 0.32(\pm3.12)$$

$$n = 18,\ r^2 = 0.839,\ s = 0.26,\ q^2 = 0.650 \quad (10)$$

$$\log P_0 = 4.23\ (4.04-4.61),\ \text{Outliers } 5$$

Activity against HIV-1 mutant strain 100-Leu→Ile:

$$\log 1/C = 2.80(\pm1.52)\text{Clog } P - 0.30(\pm0.20)(\text{Clog } P)^2 + 0.47(\pm0.28)I_6 + 1.47(\pm2.76)$$

$$n = 21,\ r^2 = 0.813,\ s = 0.30,\ q^2 = 0.691 \quad (11)$$

$$\log P_0 = 4.60\ (4.18-6.42),\ \text{Outliers } 2$$

Activity against HIV-1 mutant strain 103-Lys→Asn:

$$\log 1/C = 0.29(\pm0.16)\text{Clog } P + 0.86(\pm0.23)I_6 + 4.81(\pm0.61)$$

$$n = 18,\ r^2 = 0.864,\ s = 0.22,\ q^2 = 0.821 \quad (12)$$

Outliers 2

Activity against HIV-1 mutant strain 106-Val→Ala:

$$\log 1/C = 2.45(\pm1.38)\text{Clog } P - 0.26(\pm0.18)(\text{Clog } P)^2 - 0.41(\pm0.26)I_5 + 1.31(\pm2.50)$$

$$n = 20,\ r^2 = 0.847,\ s = 0.26,\ q^2 = 0.721 \quad (13)$$

$$\log P_0 = 4.82\ (4.31-7.62),\ \text{Outliers } 3$$

Activity against HIV-1 mutant strain 138-Glu→Lys:

$$\log 1/C = 0.29(\pm0.14)\text{Clog } P + 0.60(\pm0.21)I_6 + 0.43(\pm0.27)I_6' + 5.88(\pm0.52)$$

$$n = 18,\ r^2 = 0.857,\ s = 0.19,\ q^2 = 0.798 \quad (14)$$

Outliers 5

Activity against HIV-1 mutant strain 181-Tyr→Cys:

$$\log 1/C = 0.55(\pm0.13)\text{Clog } P + 0.44(\pm0.19)I_6 + 0.68(\pm0.26)I_6' + 4.02(\pm0.49)$$

$$n = 20,\ r^2 = 0.894,\ s = 0.19,\ q^2 = 0.817 \quad (15)$$

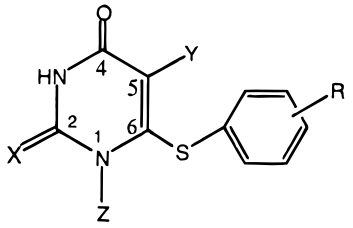
Outliers 3

Activity against HIV-1 mutant strain 181-Tyr→Ile:

$$\log 1/C = 0.47(\pm0.15)\text{Clog } P - 1.06(\pm0.25)I_5 + 0.23(\pm0.23)I_6 + 4.20(\pm0.57)$$

$$n = 18,\ r^2 = 0.910,\ s = 0.22,\ q^2 = 0.868 \quad (16)$$

Table 8. CC_{50} Activity of HEPT Derivatives (18)[71]



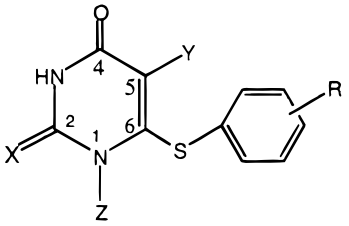
	substituents				log 1/*C*			
no.	R	Y	X	Z	obsd	calcd (eq 8)	Δ	Clog *P*
1	3-Me	Me	O	CH_2OH	3.38	3.37	0.01	1.87
2	3,5-Me_2	Me	O	CH_2OH	3.62	3.60	0.02	2.37
3	H	Me	O	Me	3.64	3.71	−0.07	2.60
4	H	Me	O	C_6H_5	4.20	4.14	0.06	3.51
5	H	Me	O	C_2H_5	3.85	3.96	−0.10	3.13
6	H	C_2H_5	O	CH_2OH	3.40	3.38	0.02	1.90
7	H	C_2H_5	O	Me	3.84	3.96	−0.12	3.13
8	H	C_2H_5	O	C_6H_5	4.52	4.39	0.14	4.04
9	3,5-Me_2	C_2H_5	O	CH_2OH	3.81	3.85	−0.04	2.90
10	H	$CHMe_2$	O	CH_2OH	3.65	3.57	0.08	2.30
11	H	$CHMe_2$	O	Me	4.00	4.15	−0.14	3.53
12	3,5-Me_2	$CHMe_2$	O	CH_2OH	3.99	4.04	−0.05	3.30
13	H	C_3H_7	O	CH_2OH	3.70	3.63	0.07	2.43
14	H	C_2H_5	S	Me	4.18	4.05	0.13	3.32

Equations 10−16 clearly indicate a role for hydrophobicity. Among all these equations, eqs 10 and 14 are poor equations with 5 outliers and a low q^2 but they do bring out the importance of hydrophobicity in these interactions. Equations 10, 11, and 13 represent parabolic correlation with Clog *P* showing that in some cases the activity can be optimized with an optimum value of Clog $P = 4.50$. In deriving eqs 10−16, indicator parameter I_5, I_6, and I_6' have been used. Where, $I_5 = 1.0$ for 5-$CHMe_2$, $I_6 = 1.0$ for the presence of 3,5-di-Me in 6-benzyl or 6-thiophenyl ring, and $I_6' = 1.0$ for 6-benzyl. The parameter I_6 occurs in all the equations, except in eq 13, and has a positive coefficient, suggesting that the presence of 3,5-di-Me is favorable to the activity. In eqs 14 and 15, I_6' suggests that a 6-benzyl derivative is preferred over 6-thiophenyl for better activity. However, in eqs 10, 13, and 16, the presence of I_5 with a negative coefficient indicates that a 5-$CHMe_2$ group would produce an adverse effect.

In the derivation of eqs 10−16, certain compounds, as indicated in Table 10, were not included as they were found to be misfit in the correlations. Since different compounds were misfits in different equations, it was difficult to explain uniformly such aberrations, but for all such compounds the corresponding equations highly overestimated their activities. It is of interest that the role for log *P* in the equations varies considerably from one mutant strain to another.

*(viii) EC_{50} Activity of HEPT Derivatives (**21**) in MT-4 Cells (Table 11).*[73] In the above analysis, the 6-position substituents of the type $CH_2C_6H_5$, $CH_2C_6H_3$-3,5-di-Me or 6-SC_6H_3-3,5-di-Me have been shown to be conducive to activity. Tanaka et al.[73] reported data for a small series of 5- or 6-substituted HEPT

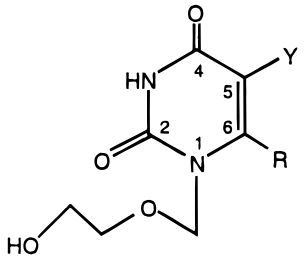
Table 9. EC_{50} Activity of HEPT Derivatives (18)[71]



	substituents				log 1/*C*				
no.	R	Y	X	Z	obsd	calcd (eq 9)	Δ	Clog *P*	$B5_Z$
1	H	Me	O	CH_2OH	5.19	5.46	−0.27	1.37	2.04
2	3-Me	Me	O	CH_2OH	5.59	5.73	−0.15	1.87	2.04
3	3,5-di-Me	Me	O	CH_2OH	6.59	6.01	0.58	2.37	2.04
4	H	Me	O	Me	6.48	6.13	0.35	2.60	2.04
5	H	Me	O	C_6H_5	7.03	6.64	0.40	3.51	2.04
6	H	Me	O	C_2H_5	5.52	6.43	−0.90	3.13	2.04
7	H	C_2H_5	O	CH_2OH	6.92	7.05	−0.13	1.90	3.17
8	H	C_2H_5	O	Me	7.66	7.73	−0.07	3.13	3.17
9	H	C_2H_5	O	C_6H_5	8.31	8.23	0.08	4.04	3.17
10	3,5-di-Me	C_2H_5	O	CH_2OH	7.80	7.60	0.19	2.90	3.17
11	3,5-di-Me	C_2H_5	O	Me	8.21	8.28	−0.07	4.12	3.17
12	3,5-di-Me	C_2H_5	O	C_6H_5	8.62	8.78	−0.16	5.03	3.17
13	H	$CHMe_2$	O	CH_2OH	7.14	7.27	−0.13	2.30	3.17
14	H	$CHMe_2$	O	Me	8.09	7.95	0.14	3.53	3.17
15	H	$CHMe_2$	O	C_6H_5	8.47	8.45	0.02	4.44	3.17
16	3,5-di-Me	$CHMe_2$	O	CH_2OH	8.48	7.82	0.66	3.30	3.17
17	H	C_3H_7	O	CH_2OH[a]	5.44	7.71	−2.27	2.43	3.49
18	H	C_2H_5	S	Me	7.59	7.83	−0.25	3.32	3.17
19	3,5-di-Me	C_2H_5	S	Me	8.25	8.39	−0.13	4.32	3.17
20	H	$CHMe_2$	S	Me	7.92	8.06	−0.13	3.72	3.17

[a] Data point not used in deriving equation.

derivatives (**21**) that were found to be correlated with the STERIMOL width parameter B5 of 6-substitu-



21

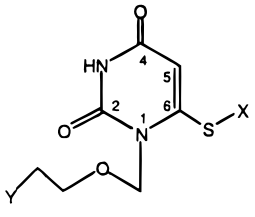
ents (R), as shown by eq 17. This shows that R-substituents have a favorable steric effect.

$$\log 1/C = 2.07(\pm 0.85)B5_R - 7.88(\pm 5.27)$$

$$n = 8,\ r^2 = 0.856,\ s = 0.29,\ q^2 = 0.734 \qquad (17)$$

Outlier 1

*(ix) EC_{50} Activity of HEPT Derivatives (**22**) in CEM-SS Infected Cells (Table 12).*[74] Pontkis et al.[74] reported EC_{50} activity on the inhibition of HIV-1 multiplication in CEM-SS infected cells for a small series of N^1 side chain modified analogues of HEPT (**22**). We developed eq 18 for the set in which activity was found to be correlated with hydrophobicity parabolically. Equation 18 shows that the hydrophobicity of the molecule is important but it has an optimum Clog *P* of 3.81.



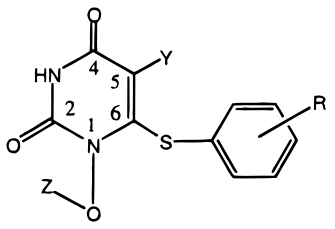
22

$$\log 1/C = 3.66(\pm 1.49)\text{Clog } P - 0.48(\pm 0.20)(\text{Clog } P)^2 + 0.18(\pm 2.29)$$

$$n = 8,\ r^2 = 0.889,\ s = 0.42,\ q^2 = 0.725 \qquad (18)$$

$\log P_0 = 3.81$ (3.51−4.18), Outliers 3

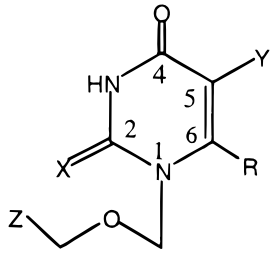
*(x) EC_{50} Activity of HEPT Derivatives (**23**) in CEM-SS Infected Cells (Table 13).*[75] For a comparatively



23

larger series of HEPT analogues where N_1-substituents were widely varied, Kim et al.[75] reported the

Table 10. Physicochemical Parameters and EC_{50} Activity (log 1/*C*) of HEPT Derivatives (20)[72]



(a) Physicochemical Parameters

no.	substituents R	Y	X	Z	Clog *P*	I_5	I_6	I_6'
1	SC_6H_5	C_2H_5	O	Me	3.13	0.00	0.00	0.00
2	SC_6H_3-3,5-di-Me	C_2H_5	O	Me	4.12	0.00	1.00	0.00
3	SC_6H_5	C_2H_5	S	Me	3.32	0.00	0.00	0.00
4	SC_6H_3-3,5-di-Me	C_2H_5	S	Me	4.32	0.00	1.00	0.00
5	$CH_2C_6H_5$	C_2H_5	S	Me	2.81	0.00	0.00	1.00
6	$CH_2C_6H_3$-3,5-di-Me	C_2H_5	O	Me	3.80	0.00	1.00	1.00
7	SC_6H_5	C_2H_5	O	C_6H_5	4.04	0.00	0.00	0.00
8	SC_6H_5-3,5-di-Me	C_2H_5	O	C_6H_5	5.03	0.00	1.00	0.00
9	SC_6H_5	C_2H_5	S	C_6H_5	4.23	0.00	0.00	0.00
10	SC_6H_3-3,5-di-Me	C_2H_5	S	C_6H_5	5.23	0.00	1.00	0.00
11	SC_6H_3-3,5-di-Me	C_2H_5	O	CH_2OH	2.90	0.00	1.00	0.00
12	SC_6H_3-3,5-di-Me	C_2H_5	S	CH_2OH	3.09	0.00	1.00	0.00
13	$CH_2C_6H_3$-3,5-di-Me	C_2H_5	O	CH_2OH	2.58	0.00	1.00	1.00
14	SC_6H_5	$CHMe_2$	O	Me	3.53	1.00	0.00	0.00
15	SC_6H_5	$CHMe_2$	S	Me	3.72	1.00	0.00	0.00
16	$CH_2C_6H_5$	$CHMe_2$	O	Me	3.20	1.00	0.00	1.00
17	SC_6H_5	$CHMe_2$	O	C_6H_5	4.44	1.00	0.00	0.00
18	SC_6H_5	$CHMe_2$	S	C_6H_5	4.63	1.00	0.00	0.00
19	SC_6H_5	$CHMe_2$	O	CH_2OH	2.30	1.00	0.00	0.00
20	SC_6H_3-3,5-di-Me	$CHMe_2$	O	CH_2OH	3.30	1.00	1.00	0.00
21	SC_6H_5	$CHMe_2$	S	CH_2OH	2.49	1.00	0.00	0.00
22	SC_6H_3-3,5-di-Me	$CHMe_2$	S	CH_2OH	3.49	1.00	1.00	0.00
23	$CH_2C_6H_3$-3,5-di-Me	$CHMe_2$	O	CH_2OH	2.98	1.00	1.00	1.00

(b) EC_{50} Activity (log 1/*C*)

no.	obsd[a] (eq 10)	calcd (eq 10)	Δ (eq 10)	obsd[b] (eq 11)	calcd (eq 11)	Δ (eq 11)	obsd[c] (eq 12)	calcd (eq 12)	Δ (eq 12)	obsd[d] (eq 13)	calcd (eq 13)	Δ (eq 13)
1	8.22	8.05	0.18	7.48	7.24	0.24	5.81	5.71	0.11	6.19	6.49	−0.30
2	8.40[e]	9.08	−0.68	8.22	8.29	−0.07	7.16	6.85	0.31	7.05	7.09	−0.04
3	7.96	8.24	−0.28	7.22	7.40	−0.18	5.96	5.76	0.20	6.51	6.64	−0.13
4	8.52[e]	9.08	−0.56	8.00	8.34	−0.34	6.96	6.90	0.06	7.16	7.15	0.01
5	7.80	7.64	0.16	7.10	6.92	0.18	5.54	5.62	−0.08	6.28	6.18	0.10
6	9.00	8.99	0.01	8.10	8.17	−0.07	6.82	6.76	0.07	7.22	6.95	0.27
7	8.70	8.64	0.06	7.70	7.80	−0.10	6.09	5.97	0.12	6.92	7.06	−0.14
8	9.00	8.76	0.24	8.70	8.30	0.40	7.22	7.11	0.11	7.00	7.10	−0.20
9	8.52	8.66	−0.14	7.52	7.86	−0.33	6.19	6.02	0.17	6.47[e]	7.12	−0.65
10	8.22	8.59	−0.37	8.10	8.24	−0.15	6.96	7.16	−0.20	7.40	7.17	0.23
11	8.52	8.19	0.33	7.70	7.48	0.22	6.46	6.50	−0.04	6.60	6.27	0.33
12	8.40	8.43	−0.03	7.52	7.67	−0.15	6.48	6.55	−0.07	6.42	6.45	−0.03
13	7.55	7.71	−0.16	6.77	7.12	−0.35	5.70[e]	6.41	−0.71	5.82	5.93	−0.11
14	7.60	8.00	−0.40	6.96[e]	7.55	−0.59	5.30	5.82	−0.52	5.70[e]	6.38	−0.68
15	8.05	8.12	−0.07	7.52	7.66	−0.14	6.00	5.88	0.12	6.80	6.49	0.30
16	8.70[e]	7.72	0.98	7.70	7.31	0.39	5.89	5.73	0.16	6.64[e]	6.14	0.50
17	8.52	8.23	0.29	8.00	7.89	0.11	6.13	6.08	0.04	6.60	6.76	−0.16
18	8.40	8.17	0.23	8.00	7.90	0.10	5.82	6.14	−0.31	6.70	6.79	−0.09
19	7.55[e]	6.38	1.18	6.55	6.29	0.26				5.57	5.19	0.38
20	8.40	8.24	0.16	8.30	7.85	0.45	6.39	6.61	−0.23	6.48	6.21	0.27
21	6.70	6.73	−0.03	6.00	6.54	−0.54				5.00	5.42	−0.42
22	8.22	8.40	−0.18	8.05	7.99	0.06	6.22[e]	6.67	−0.44	6.47	6.35	0.12
23	6.77[e]	7.89	−1.12	6.10[e]	7.56	−1.46				5.55	5.94	−0.39

(c) EC_{50} Activity (log 1/*C*)

no.	obsd[f] (eq 14)	calcd (eq 14)	Δ (eq 14)	obsd[g] (eq 15)	calcd (eq 15)	Δ (eq 15)	obsd[h] (eq 16)	calcd (eq 16)	Δ (eq 16)
1	6.89	6.77	0.11	6.19	5.75	0.44	5.96	5.66	0.30
2	7.52	7.65	−0.13	6.52	6.75	−0.22	6.22	6.36	−0.14
3	6.89	6.83	0.06	5.82	5.86	−0.03	5.51	5.75	−0.24
4	7.52	7.71	−0.19	7.05	6.85	0.20	6.24	6.45	−0.20
5	7.05	7.11	−0.06	5.92	6.22	−0.32	5.16	5.51	−0.35
6	8.05	7.99	0.06	7.40	7.25	0.15	6.46	6.21	0.25

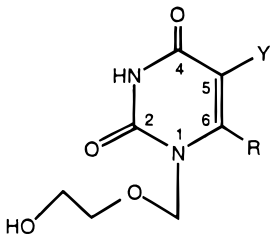
Table 10. (Continued)

(c) EC_{50} Activity (log 1/*C*) (Continued)

no.	obsd[f] (eq 14)	calcd (eq 14)	Δ (eq 14)	obsd[g] (eq 15)	calcd (eq 15)	Δ (eq 15)	obsd[h] (eq 16)	calcd (eq 16)	Δ (eq 16)
7	7.22	7.03	0.19	6.39	6.25	0.14	6.46	6.09	0.37
8	8.05	7.91	0.13	7.10	7.25	−0.15	6.96	6.78	0.17
9	7.05	7.09	−0.04	6.41	6.36	0.05	6.11	6.18	−0.07
10	8.00	7.97	0.03	6.50[e]	7.35	−0.86	6.62	6.87	−0.25
11	7.52	7.30	0.22	6.06	6.07	−0.01	5.85	5.78	0.07
12	7.40	7.36	0.04	6.26	6.17	0.09	6.00	5.87	0.13
13	6.70[e]	7.64	−0.94	5.89[e]	6.57	−0.68	5.60	5.63	−0.03
14	6.35	6.89	−0.54	5.89	5.97	−0.08			
15	7.05	6.94	0.11	6.05	6.08	−0.03			
16	7.22	7.22	0.00	6.66	6.47	0.18	4.57	4.64	−0.08
17	7.70[e]	7.15	0.55	6.52	6.47	0.05	5.30	5.22	0.08
18	7.30	7.20	0.10	6.40	6.58	−0.18	5.30	5.31	−0.01
19	6.60	6.54	0.07	5.26	5.29	−0.03			
20	8.00[e]	7.42	0.58	6.16	6.29	−0.13	4.92	4.92	0.01
21	5.82[e]	6.59	−0.77	5.22	5.40	−0.17			
22	7.30	7.47	−0.17	6.47	6.39	0.08	5.00	5.00	0.00
23	6.28[e]	7.75	−1.48	5.47[e]	6.79	−1.32			

[a] Against HIV-1 wild type. [b−d] Against HIV-1 mutant strains: [b] 100-Leu→Ile, [c] 103-Lys→Asn, [d] 106-Val→Ala. [e] Data points not included in deriving respective equations. [f−h] Against HIV-1 mutant strains: [f] 138-Glu→Lys, [g] 181-Tyr→Cys, [h] 181-Tyr→Ile.

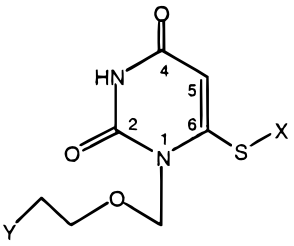
Table 11. EC_{50} Activity of HEPT Derivatives (21)[73]



	substituents		log 1/*C*			
no.	Y	R	obsd	calcd (eq 17)	Δ	$B5_R$
1	Me	SC_6H_5	5.16	5.42	−0.26	6.42
2	Me	SC_4H_9	3.89	3.74	0.15	5.61
3	Me	SC_6H_{11}	5.09	5.15	−0.06	6.29
4	Me	OC_6H_5	4.07	4.32	−0.25	5.89
5	Me	$CH_2C_6H_5$	4.64	4.59	0.05	6.02
6	I	SC_6H_5	5.44	5.42	0.03	6.42
7	$CH{=}C(C_6H_5)_2$	SC_6H_5[a]	6.08	5.42	0.66	6.42
8	$CH{=}CHC_6H_5$	SC_6H_5	5.22	5.42	−0.20	6.42
9	$CH{=}CH_2$	SC_6H_5	5.96	5.42	0.54	6.42

[a] Data point not used in deriving equation.

Table 12. EC_{50} Activity of HEPT Derivatives (22)[74]



	substituents		log 1/*C*			
no.	X	Y	obsd	calcd (eq 18)	Δ	Clog *P*
1	Ph[a]	Br	6.12	6.66	−0.54	2.80
2	Ph	I[c]	5.66	6.97	−1.31	3.19
3	Ph	$N(CH_2CN)_2$	4.56	4.45	0.11	1.44
4	Ph	NHC_6H_5	7.26	7.13	0.13	3.62
5	Ph	$N(C_6H_5)_2$	5.06	5.07	−0.01	5.89
6	Ph	NH_2	4.51	4.46	0.05	1.44
7	Ph	$NHCO(CH_2)_2Cl$	4.94	5.45	−0.50	1.93
8	Ph	S-Py[c]	6.08	7.06	−0.98	3.38
9	Py[b]	NHC_6H_5	6.80	6.35	0.45	2.52
10	Py	SC_6H_5[c]	6.00	7.06	−1.06	3.38
11	Py	S-Py	6.36	6.03	0.33	2.28

[a] Ph = phenyl. [b] Py = pyridyl. [c] Data points not used in deriving equation.

EC_{50} data. We formulated eq 19 for the set.

$$\log 1/C = 1.23(\pm1.02)L_Z - 0.01(\pm0.08)(L_Z)^2 - 0.58(\pm0.15)B5_Z + 1.28(\pm0.32)\pi_{R-3,5} + 3.46(\pm3.32)$$

$$n = 27,\ r^2 = 0.841,\ s = 0.29,\ q^2 = 0.761 \quad (19)$$

$$(L_Z)_0 = 6.44,\ \text{Outliers } 7$$

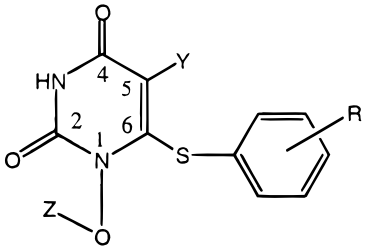
In eq 19 the STERIMOL length (L) and width parameters (B5) of Z-substituents were found to dominate the potency. Equation 19 also exhibits a hydrophobic effect of meta-substituents at the phenyl ring of 6-phenylthio group. In the derivation of eq 19, however, seven compounds were not included. The equation predicts very high activity for them as compared to their corresponding observed activity.

*(xi) EC_{50} Activity of HEPT Derivatives (**18**) in MT-4 Cells (Table 14).*[76,77] For a very large series of HEPT analogues represented by the general structure **18**, in which the sets of compounds for which eqs 4 and 5 were obtained were also included along with the compounds from other sources,[76] Luco and Ferretti[77] correlated the anti-HIV data (EC_{50} is 50% protection of MT-4 cells against the cytopathic effect of HIV-1 (HTLV−IIIB strain)) as in eq 20.

$$\log 1/C = 0.48(\pm0.12)\Sigma\pi(\mathbf{R}+\mathbf{Y}) + 162.5(\pm17.6)^1\chi^N(\mathbf{Y}) - 239.8(\pm25.1)^1\chi^N(\mathbf{Y})^2 - 0.85(\pm0.19)B1(\mathbf{R,3}) + 1.52(\pm0.28)Es(\mathbf{R,2}) + 52.06(\pm7.54)^4\chi^N{}_P + 0.78(\pm0.34)MgVol - 2.22(\pm0.28)I(\mathbf{Z}) - 0.56(\pm0.21)^0\Delta\chi(\mathbf{Z}) - 35.99(\pm2.41)$$

$$n = 79,\ r^2 = 0.949,\ s = 0.44,\ q^2 = 0.745 \quad (20)$$

Table 13. EC_{50} Activity of HEPT Derivatives (23)[75]



	substituents			log 1/*C*					
no.	R	Y	Z	obsd	calcd (eq 19)	Δ	L_Z	$B5_Z$	$\pi_{R-3,5}$
1	H	$CHMe_2$	C_3H_7	5.17	5.17	0.00	4.92	3.49	0.00
2	2-Me	$CHMe_2$	C_3H_7	4.86	5.17	−0.31	4.92	3.49	0.00
3	3-Me	$CHMe_2$	C_3H_7	6.01	5.89	0.12	4.92	3.49	0.56
4	3,5-di-Me	$CHMe_2$	C_3H_7	7.19	6.60	0.59	4.92	3.49	1.12
5	3-F	$CHMe_2$	C_3H_7	5.72	5.35	0.37	4.92	3.49	0.14
6	3,5-di-F	$CHMe_2$	C_3H_7	5.62	5.53	0.09	4.92	3.49	0.28
7	H	$CHMe_2$	$(CH_2)_3OH$	4.72	4.77	−0.05	6.17	4.54	0.00
8	3-Me	$CHMe_2$	$(CH_2)_3OH$	5.33	5.49	−0.16	6.17	4.54	0.56
9	3,5-di-Me	$CHMe_2$	$(CH_2)_3OH$	6.72	6.21	0.52	6.17	4.54	1.12
10	3,5-di-F	$CHMe_2$	$(CH_2)_3OH$	5.12	5.13	−0.01	6.17	4.54	0.28
11	3,5-di-Me	$CHMe_2$	C_4H_9	6.38	6.21	0.17	6.17	4.54	1.12
12	3,5-di-Me	$CHMe_2$	$CH_2C_6H_5$	5.07	5.04	0.04	4.62	6.02	1.12
13	3,5-di-Me	$CHMe_2$	$CH_2CH_2C_6H_5$	6.31	6.43	−0.12	8.33	3.58	1.12
14	3,5-di-Me	$CHMe_2$	$CH_2CH_2C_6H_4$-3-Me	6.08	5.90	0.18	8.33	4.49	1.12
15	3,5-di-Me	$CHMe_2$	CH_2CH_2OMe	6.37	6.17	0.20	5.55	4.49	1.12
16	3,5-di-Me	$CHMe_2$	$CH_2CH_2OC_6H_5$	6.34	6.14	0.19	9.00	3.58	1.12
17	3,5-di-Me	C_2H_5	C_3H_7	6.46	6.60	−0.15	4.92	3.49	1.12
18	3,5-di-Me	C_2H_5	$(CH_2)_3OH$	6.27	6.21	0.06	6.17	4.54	1.12
19	3,5-di-Me	C_2H_5	C_4H_9	6.01	6.21	−0.19	6.17	4.54	1.12
20	3,5-di-Me	C_2H_5	$CH_2C_6H_5$	5.24	5.04	0.20	4.62	6.02	1.12
21	3,5-di-Me	C_2H_5	$CH_2CH_2C_6H_5$	6.34	6.43	−0.09	8.33	3.58	1.12
22	3,5-di-Me	C_2H_5	CH_2CH_2OMe	5.71	6.17	−0.46	5.55	4.49	1.12
23	3,5-di-Me	Cy-C_3H_5	C_3H_7[a]	5.21	6.60	−1.39	4.92	3.49	1.12
24	3,5-di-Me	Cy-C_3H_5	$(CH_2)_3OH$[a]	4.82	6.21	−1.39	6.17	4.54	1.12
25	3,5-di-Me	Cy-C_3H_5	$CH_2C_6H_5$	4.80	5.04	−0.24	4.62	6.02	1.12
26	3,5-di-Me	Cy-C_3H_5	$CH_2CH_2C_6H_5$[a]	5.22	6.43	−1.21	8.33	3.58	1.12
27	3,5-di-Me	Cy-C_3H_5	C_3H_7[a]	5.59	6.60	−1.01	4.92	3.49	1.12
28	3,5-di-Me	C_3H_7	$(CH_2)_3OH$[a]	5.09	6.21	−1.12	6.17	4.54	1.12
29	3,5-di-Me	C_3H_7	$CH_2C_6H_5$	4.99	5.04	−0.05	4.62	6.02	1.12
30	3,5-di-Me	C_3H_7	$CH_2CH_2C_6H_5$	6.16	6.43	−0.27	8.33	3.58	1.12
31	3,5-di-Me	Me	C_3H_7	6.03	6.60	−0.57	4.92	3.49	1.12
32	3,5-di-Me	Me	$(CH_2)_3OH$[a]	5.26	6.21	−0.95	6.17	4.54	1.12
33	3,5-di-Me	Me	$CH_2C_6H_5$	5.00	5.04	−0.04	4.62	6.02	1.12
34	3,5-di-Me	Me	$CH_2CH_2C_6H_5$[a]	5.51	6.43	−0.92	8.33	3.58	1.12

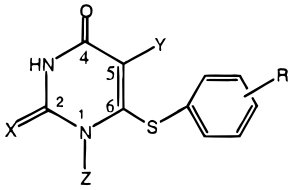
[a] Data points not used in deriving equation.

In this equation, in addition to the hydrophobic constant π and Taft's steric constant Es, there are a number of structural or topological parameters. The STERIMOL parameter B1 refers to the minimum width of a substituent. $^m\chi^N$ is Kier's molecular connectivity index of *m*th order ($^m\chi$)[78] which has been divided by the number of atoms (N) involved in its calculation in order to minimize the role of the substituent size. Thus, it is a bond additive description in which the influence of substituent size has been minimized, and the presence of $^m\chi^N$ in the equation shows that the activity is strongly dependent on the structural variation of Y-substituents.

$^4\chi_P$ is a connectivity index that describes a weighted count of all fragments or subgraphs consisting of four bonds joined as a path.[78] MgVol is the McGowan volume which can be easily calculated for any solute simply from a knowledge of its molecular structure.[79] The positive dependence of the activity on this parameter and on $^4\chi_P$ indicates an important role of molecular size and the different aspects of the molecular shape in the action of the anti-HIV drugs. However, the negative dependence of the activity on differential molecular connectivity of zeroth order $^0\Delta\chi_Z$ for Z-substituents, selected as an electronic parameter based on graph theory,[80] and on an indicator variable I_Z which has been used with a value of 1 or 0 for the presence or absence of a six-membered saturated ring in Z suggests that the Z-substituents, because of their electronic character, will be detrimental to the activity and would be of further disadvantage if they possess a six-membered saturated ring.

Luco and Ferretti[77] also performed a partial least-squares (PLS) analysis on the same series which consisted of the same variables as those used in eq 20. The PLS model resulted in a significant three-component model with the following statistics: $r^2 = 0.889$, $r_{cv} = 0.927$, $s = 0.44$, and $F = 202.27$, where r_{cv} is the cross-validated *r*, which describes the predictive power of the model.[81]

Table 14. EC_{50} Activity of HEPT Derivatives (18) Studied by Luco et al.[76,77]



	substituents				log 1/*C*							
no.	R	Y	X	Z	obsd	calcd (eq 21)	Δ	$B1_{R,2}$	$B5_{R,5}$	$B1_Y$	Clog *P*	MR_Y
1	2-Me	Me	O	OCH_2CH_2OH	4.15	4.35	−0.20	1.52	1.00	1.52	1.87	0.57
2	2-NO_2	Me	O	OCH_2CH_2OH	3.85	3.87	−0.02	1.70	1.00	1.52	1.11	0.57
3	2-OMe	Me	O	OCH_2CH_2OH	4.72	4.39	0.33	1.35	1.00	1.52	0.89	0.57
4	3-Me	Me	O	OCH_2CH_2OH	5.59	5.20	0.39	1.00	1.00	1.52	1.87	0.57
5	3-C_2H_5	Me	O	OCH_2CH_2OH	5.57	5.33	0.24	1.00	1.00	1.52	2.40	0.57
6	3-CMe_3	Me	O	OCH_2CH_2OH	4.92	5.52	−0.60	1.00	1.00	1.52	3.20	0.57
7	3-CF_3	Me	O	OCH_2CH_2OH	4.35	5.29	−0.94	1.00	1.00	1.52	2.25	0.57
8	3-F	Me	O	OCH_2CH_2OH	5.48	5.11	0.37	1.00	1.00	1.52	1.51	0.57
9	3-Cl	Me	O	OCH_2CH_2OH	4.89	5.25	−0.36	1.00	1.00	1.52	2.08	0.57
10	3-Br	Me	O	OCH_2CH_2OH	5.24	5.29	−0.05	1.00	1.00	1.52	2.23	0.57
11	3-I	Me	O	OCH_2CH_2OH	5.00	5.35	−0.35	1.00	1.00	1.52	2.49	0.57
12	3-NO_2	Me	O	OCH_2CH_2OH	4.47	5.01	−0.54	1.00	1.00	1.52	1.11	0.57
13	3-OH	Me	O	OCH_2CH_2OH	4.09	4.91	−0.82	1.00	1.00	1.52	0.70	0.57
14	3-OMe	Me	O	OCH_2CH_2OH	4.66	5.06	−0.40	1.00	1.00	1.52	1.29	0.57
15	3,5-di-Me	Me	O	OCH_2CH_2OH	6.59	6.42	0.17	1.00	2.04	1.52	2.37	0.57
16	3,5-di-Cl	Me	O	OCH_2CH_2OH	5.89	6.27	−0.38	1.00	1.80	1.52	2.80	0.57
17	3,5-di-Me	Me	S	OCH_2CH_2OH	6.66	6.47	0.19	1.00	2.04	1.52	2.56	0.57
18	3-CO_2Me	Me	O	OCH_2CH_2OH	5.10	5.07	0.03	1.00	1.00	1.52	1.34	0.57
19	3-COMe	Me	O	OCH_2CH_2OH	5.14	4.94	0.20	1.00	1.00	1.52	0.81	0.57
20	3-CN	Me	O	OCH_2CH_2OH	5.00	4.94	0.06	1.00	1.00	1.52	0.80	0.57
21	H	$CH_2CH{=}CH_2$	O	OCH_2CH_2OH[a]	5.60	7.69	−2.09	1.00	1.00	1.52	1.94	1.45
22	H	C_2H_5	S	OCH_2CH_2OH	6.96	6.55	0.41	1.00	1.00	1.52	2.09	1.03
23	H	C_3H_7	S	OCH_2CH_2OH[a]	5.00	7.99	−2.99	1.00	1.00	1.52	2.62	1.50
24	H	$CHMe_2$	S	OCH_2CH_2OH	7.23	7.23	0.01	1.00	1.00	1.90	2.49	1.50
25	3,5-di-Me	C_2H_5	S	OCH_2CH_2OH	8.11	7.90	0.21	1.00	2.04	1.52	3.09	1.03
26	3,5-di-Me	$CHMe_2$	S	OCH_2CH_2OH	8.30	8.57	−0.27	1.00	2.04	1.90	3.49	1.50
27	3,5-di-Cl	C_2H_5	S	OCH_2CH_2OH	7.37	7.75	−0.38	1.00	1.80	1.52	3.52	1.03
28	H	C_2H_5	O	OCH_2CH_2OH	6.92	6.51	0.41	1.00	1.00	1.52	1.90	1.03
29	H	C_3H_7	O	OCH_2CH_2OH[a]	5.47	7.94	−2.47	1.00	1.00	1.52	2.43	1.50
30	H	$CHMe_2$	O	OCH_2CH_2OH	7.20	7.18	0.02	1.00	1.00	1.90	2.30	1.50
31	3,5-di-Me	C_2H_5	O	OCH_2CH_2OH	7.89	7.85	0.04	1.00	2.04	1.52	2.90	1.03
32	3,5-di-Me	$CHMe_2$	O	OCH_2CH_2OH	8.57	8.52	0.05	1.00	2.04	1.90	3.30	1.50
33	3,5-di-Cl	C_2H_5	O	OCH_2CH_2OH	7.85	7.70	0.149	1.00	1.80	1.52	3.33	1.030
34	4-Me	Me	O	OCH_2CH_2OH	3.66	5.20	−1.54	1.00	1.00	1.52	1.87	0.57
35	H	Me	O	OCH_2CH_2OH	5.15	5.08	0.07	1.00	1.00	1.52	1.37	0.57
36	H	Me	S	OCH_2CH_2OH	6.01	5.12	0.89	1.00	1.00	1.52	1.56	0.57
37	H	I	O	OCH_2CH_2OH	5.44	6.42	−0.98	1.00	1.00	2.15	2.33	1.39
38	H	$CH{=}CH_2$	O	OCH_2CH_2OH	5.69	6.47	−0.78	1.00	1.00	1.60	1.60	1.10
39	H	$CH{=}CHC_6H_5$	O	OCH_2CH_2OH[a]	5.22	13.34	−8.12	1.00	1.00	1.60	3.16	3.42
40	H	$CH_2C_6H_5$	O	OCH_2CH_2OH[a]	4.37	12.27	−7.90	1.00	1.00	1.52	2.94	3.00
41	H	Me	O	OCH_2CH_2OMe	5.06	5.23	−0.17	1.00	1.00	1.52	1.99	0.57
42	H	Me	O	OCH_2CH_2OCOMe	5.17	5.30	−0.12	1.00	1.00	1.52	2.27	0.57
43	H	Me	O	$OCH_2CH_2OCOC_6H_5$	5.12	5.71	−0.59	1.00	1.00	1.52	3.96	0.57
44	H	Me	O	OCH_2Me	6.48	5.38	1.11	1.00	1.00	1.52	2.60	0.57
45	H	Me	O	OCH_2CH_2Cl	5.82	5.39	0.43	1.00	1.00	1.52	2.66	0.57
46	H	Me	O	$OCH_2CH_2N_3$	5.24	5.54	−0.30	1.00	1.00	1.52	3.27	0.57
47	H	Me	O	OCH_2CH_2F	5.96	5.31	0.65	1.00	1.00	1.52	2.32	0.57
48	H	Me	O	OC_3H_7	5.48	5.50	−0.02	1.00	1.00	1.52	3.13	0.57
49	H	Me	O	$OCH_2C_6H_5$	7.06	5.60	1.46	1.00	1.00	1.52	3.51	0.57
50	H	C_2H_5	O	OC_2H_5	7.72	6.80	0.92	1.00	1.00	1.52	3.13	1.03
51	H	C_2H_5	S	OC_2H_5	7.58	6.85	0.73	1.00	1.00	1.52	3.32	1.03
52	3,5-di-Me	C_2H_5	O	OC_2H_5	8.24	8.15	0.09	1.00	2.04	1.52	4.12	1.03
53	3,5-di-Me	C_2H_5	S	OC_2H_5	8.30	8.20	0.10	1.00	2.04	1.52	4.32	1.03
54	H	C_2H_5	O	$OCH_2C_6H_5$	8.23	7.03	1.20	1.00	1.00	1.52	4.04	1.03
55	3,5-di-Me	C_2H_5	O	$OCH_2C_6H_5$	8.55	8.37	0.18	1.00	2.04	1.52	5.03	1.03
56	H	C_2H_5	S	$OCH_2C_6H_5$	8.09	7.07	1.02	1.00	1.00	1.52	4.23	1.03
57	3,5-di-Me	C_2H_5	S	$OCH_2C_6H_5$	8.14	8.42	−0.28	1.00	2.04	1.52	5.23	1.03
58	H	$CHMe_2$	O	OC_2H_5	7.99	7.48	0.51	1.00	1.00	1.90	3.53	1.50
59	H	$CHMe_2$	O	$OCH_2C_6H_5$	8.51	7.70	0.81	1.00	1.00	1.90	4.44	1.50
60	H	$CHMe_2$	S	OC_2H_5	7.89	7.52	0.37	1.00	1.00	1.90	3.72	1.50
61	H	$CHMe_2$	S	$OCH_2C_6H_5$	8.14	7.74	0.40	1.00	1.00	1.90	4.63	1.50
62	H	Me	O	OMe	5.68	5.25	0.43	1.00	1.00	1.52	2.07	0.57
63	H	Me	O	OC_4H_9	5.33	5.63	−0.30	1.00	1.00	1.52	3.66	0.57
64	H	Me	O	OMe	5.66	5.27	0.39	1.00	1.00	1.52	2.17	0.57

Table 14. (Continued)

	substituents				log 1/C							
no.	R	Y	X	Z	obsd	calcd (eq 21)	Δ	$B1_{R,2}$	$B5_{R,5}$	$B1_Y$	Clog *P*	MR_Y
65	H	Me	O	OC_3H_7	5.92	5.53	0.39	1.00	1.00	1.52	3.22	0.57
66	3,5-di-Cl	C_2H_5	S	OC_2H_5	7.89	8.05	−0.16	1.00	1.80	1.52	4.74	1.03
67	H	C_2H_5	S	$OCHMe_2$	6.66	6.93	−0.27	1.00	1.00	1.52	3.63	1.03
68	H	C_2H_5	S	OC_6H_{11}	5.79	7.22	−1.43	1.00	1.00	1.52	4.82	1.03
69	H	C_2H_5	S	$OCH_2C_6H_{11}$	6.45	7.37	−0.92	1.00	1.00	1.52	5.44	1.03
70	H	C_2H_5	S	$OCH_2C_6H_4$-4-Me	7.11	7.20	−0.08	1.00	1.00	1.52	4.73	1.03
71	H	C_2H_5	S	$OCH_2C_6H_4$-4-Cl	7.92	7.25	0.68	1.00	1.00	1.52	4.94	1.03
72	H	C_2H_5	S	$OCH_2CH_2C_6H_5$	7.04	7.23	−0.19	1.00	1.00	1.52	4.89	1.03
73	3,5-di-Cl	C_2H_5	O	OC_2H_5	8.13	8.00	0.13	1.00	1.80	1.52	4.55	1.03
74	H	C_2H_5	O	$OCHMe_2$	6.47	6.88	−0.41	1.00	1.00	1.52	3.44	1.03
75	H	C_2H_5	O	OC_6H_{11}[a]	5.40	7.17	−1.77	1.00	1.00	1.52	4.63	1.03
76	H	C_2H_5	O	$OCH_2C_6H_{11}$	6.35	7.32	−0.97	1.00	1.00	1.52	5.25	1.03
77	H	C_2H_5	O	$OCH_2CH_2C_6H_5$	7.02	7.19	−0.17	1.00	1.00	1.52	4.70	1.03
78	H	Cy-C_3H_5	S	OC_2H_5	7.02	7.68	−0.66	1.00	1.00	1.55	3.23	1.35
79	H	Cy-C_3H_5	O	OC_2H_5	7.00	7.63	−0.63	1.00	1.00	1.55	3.04	1.35

[a] Data points not used in deriving equation.

However, although eq 20 or the PLS model can be of predictive value, they do not throw much light on the modes of drug−receptor interaction. As far as structural parameters are concerned, we were able to show that fewer parameters are sufficient to give a modest approximation of the activity (eq 21). In the derivation of eq 21, however, a few outliers were excluded.

$$\log 1/C = -1.64(\pm1.35)B1_{R,2} + 1.06(\pm0.38)B5_{R,5} - 1.92(\pm1.62)B1_Y + 0.24(\pm0.15)\text{Clog } P + 2.80(\pm0.78)MR_Y + 6.65(\pm2.79)$$

$$n = 73,\ r^2 = 0.815,\ s = 0.60,\ q^2 = 0.783 \quad (21)$$

Outliers 6

Although the quality of fit is not as good as eq 20, the meaning of the parameters is more readily apparent. Equation 21 again emphasizes that in HEPT series of HIV-1 reverse transcriptase inhibitors, the hydrophobicity of the molecule is significant for activity. Also, the R-substituents in the 6-phenylthio or 6-benzyl ring show steric effects as shown by eqs 4−6.

*(xii) EC_{50} Activity of HEPT Derivatives (**17**) in MT-4 Cells.*[82,83] Kireev et al.[83] derived the correlation for EC_{50} activity for another large series of HEPT compounds, including all those compounds for which eqs 5, 6, 16, and 17 were obtained and a few compounds studied by Hopkins et al.[82]

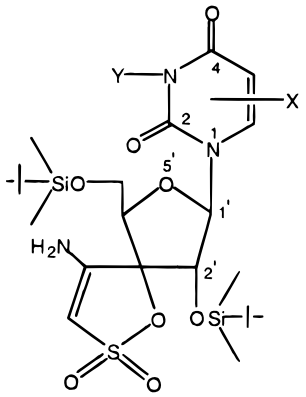
$$\log 1/C = 1.64 - 1.92(\pm0.28)\sum(q^-)_6 + 1.43(\pm0.11)({}^{1/2}W_5) - 0.47(\pm0.07)W_{cis} + 0.41(\pm0.05)(RB5)_{cis}$$

$$n = 87,\ r^2 = 0.884,\ s = 0.46 \quad (22)$$

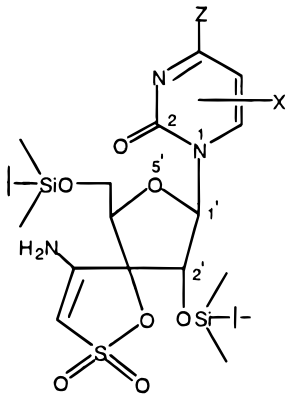
where $(\sum q^-)_6$ refers to the total negative charge calculated by using AM1 method over the atoms of the C-6 substituent (refer to **17**), W_5 refers to the width of plane C-5 substituent, W_{cis} is the width parameter of the whole molecule when the rotable bond S−C1′ is cis to the N1−C6 bond, and $(RB5)_{cis}$ is the rotation barrier to the rotable bond (S−C1′ numbered as 5) when cis to the N1−C6 bond. Thus, besides the electronic and molecular size effect, the effect of the conformational flexibility of the molecule is also exhibited. Kireev et al. suggested that the stability of the S−C1′ cis conformation would be one of the most important properties to check when planning a synthesis of new congeners. There is no hydrophobic term in the equation derived by Kireev et al., though it contains more or less the same molecules studied in deriving eq 21. It seems they did not consider hydrophobic interactions. Collinearity probably obscures the picture.

From the results of eqs 4−22 on HEPT derivatives, we can see that HIV-1 reverse transcriptase has a hydrophobic binding domain, though it seems that the site has an optimum size as indicated by the presence of parabolic correlations in eqs 10, 11, 13, and 18 with an optimum Clog *P* ranging from 3.81 to 4.82. Also, there are certain limited steric effects of certain specific substituents.

c. TSAO Pyrimidine- and Pyrimidine-Modified Nucleosides. *(i) EC_{50} Activity of TSAO Derivatives (**24, 25**) in MT-4 Cells (Table 15).*[85] For the



24

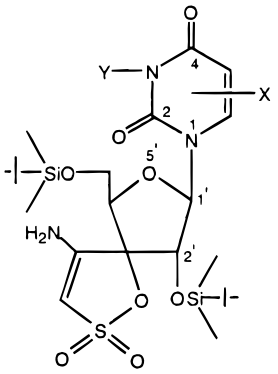


25

antiviral activity, some authors focused their attention on *tert*-butyldimethylsilylspiroaminooxathioledioxide (TSAO) pyrimidine and pyrimidine-modified nucleosides. Compounds **24**, having a thymine ring,

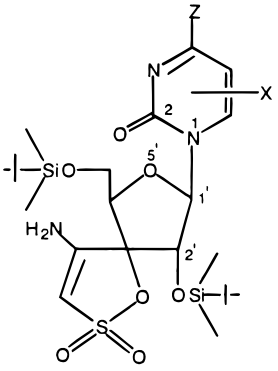
Table 15. EC_{50} Activity of TSAO−T (24) and TSAO−C Derivatives (25)[85]

(a) EC_{50} Activity of TSAO−T Derivatives



	substituents		log 1/*C*					
no.	X	Y	obsd	calcd (eq 23b)	Δ	π_Y	σ_X	L_Y
1	5-CH_3	H	7.24	7.04	0.21	0.56	−0.07	2.06
2	H	H	6.70	6.56	0.14	0.00	0.00	2.06
3	5-C_2H_5	H	7.18	7.34	−0.16	1.02	−0.07	2.06
4	5-CH_3	CH_3	7.23	6.94	0.29	0.56	−0.07	2.87
5	5-CH_3	C_2H_5	6.91	6.79	0.12	0.56	−0.07	4.11
6	5-CH_3	$CH_2CH{=}CH_2$	6.63	6.67	−0.04	0.56	−0.07	5.11
7	5-CH_3	$CH_2CH{=}C(CH_3)_2$	6.42	6.52	−0.10	0.56	−0.07	6.39
8	H	$CH_2CH{=}CH_2$	6.21	6.20	0.01	0.00	0.00	5.11
9	5-C_2H_5	H	7.22	7.34	−0.12	1.02	−0.07	2.06
10	5-F	H	6.30	6.18	0.12	0.14	0.34	2.06
11	5-Br	H	6.50	6.59	−0.09	0.86	0.39	2.06
12	5-I	H	7.05	6.82	0.23	1.12	0.35	2.06
13	6-CF_3	H	6.09	6.55	−0.46	0.88	0.43	2.06
14	6-CN	H	5.44	5.25	0.19	−0.57	0.66	2.06
15	6-C(=NH)OCH_3	H	5.57	5.91	−0.34	−0.57	0.00	2.06

(b) EC_{50} Activity of TSAO−C Derivatives



	substituents		log 1/*C*				
no.	X	Y	obsd	calcd (eq 23c)	Δ	*I*	σ_Z
1	4-NHCOMe	H	6.80	6.76	0.04	0.00	0.00
2	4-NH_2	5-Me	6.90	6.93	−0.03	1.00	−0.66
3	4-NH_2	H	6.00	6.19	−0.19	0.00	−0.66
4	4-NHMe	H	6.17	6.16	0.01	0.00	−0.70
5	4-NMe_2	H	6.18	6.04	0.14	0.00	−0.83
6	4-NHMe	5-Me	6.92	6.90	0.02	1.00	−0.70
7	-4-NMe_2	5-Me	6.80	6.79	0.01	1.00	−0.83

are referred to as TSAO−T nucleosides and **25**, having a cytosine ring, as TSAO−C nucleosides. Garg and Gupta[84] did QSAR studies on some data of Camarasa et al.[85] on these two types of TSAO derivatives, and for a series of TSAO−T analogues (Table 15a), they correlated the anti-HIV data with hydrophobic and electronic parameters. In this correlation, a set of seven TSAO−C derivatives (**25**) (Table 15b), where X = H or 5-CH_3 and where in some compounds 4-NH_2 was replaced by NHMe, NMe_2, or NHCOMe, could be easily incorporated, using a indicator variable *I* with a value of unity for all TSAO−T derivatives and zero for all TSAO−C derivatives (eq 23a).

$$\log 1/C = 0.70(\pm 0.23)\pi_X - 1.40(\pm 0.64)\sigma_X - 0.25(\pm 0.21)\pi_Y + 0.21(\pm 0.30)I + 6.33$$

$$n = 22,\ r^2 = 0.856,\ s = 0.25 \qquad (23a)$$

Since I was not found to be significant at the 95% confidence level, little difference could be assigned to the activity contributions of the thymine and cytosine rings. Instead, at both the rings, a hydrophobic and electron-donating X-substituent was found to be very conducive to activity along with a marginal effect of π_Y.

From a mechanistic point of view it is sometimes interesting to divide a big set into smaller sets for study. When we studied only the TSAO−T derivatives (**24**) (Table 15a), the following correlation was observed.

$$\log 1/C = 0.67(\pm0.24)\pi_X - 1.41(\pm0.65)\sigma_X - 0.12(\pm0.11)L_Y + 6.81(\pm0.40)$$

$$n = 15,\ r^2 = 0.864,\ s = 0.24,\ q^2 = 0.30 \quad (23b)$$

TSAO−C derivatives (Table 15b) gave following correlation

$$\log 1/C = 0.74(\pm0.27)I + 0.86(\pm0.51)\sigma_Z + 6.76(\pm0.32)$$

$$n = 7,\ r^2 = 0.940,\ s = 0.12,\ q^2 = 0.30 \quad (23c)$$

in which indicator variable I was used with a value of unity to account for the presence of 5-Me. It is of interest to note that eqs 23b and 23c bring out some of the same properties as those shown in eq 23a, except that instead of a negative hydrophobic term a steric parameter, Verloop's L, seems to fit better at position 3. One needs to pay attention to a negative hydrophobic term as it might be indicating a steric rather hydrophilic effect.

*(ii) CC_{50} Activity of TSAO Derivatives (**24, 25**) in MT-4 Cells (Table 16).*[85] The cytotoxic data were also reported[85] for most of the compounds used in eq 23a (Table 15). Adding to the series a few more TSAO−T derivatives, where in some cases both *O-tert*-butyldimethylsilyl [Si−] groups were replaced by H or OH, and few compounds (1−4 in Table 16b) with a sugar ring in the β-D-xylo configuration, Garg and Gupta[84] obtained eq 24a.

$$\log 1/C = 0.73(\pm0.45)I_2 + 1.14(\pm0.37)I_5 - 1.12(\pm0.42)I_{X,Y} + 0.34(\pm0.32)I + 2.81$$

$$n = 28,\ r^2 = 0.810,\ s = 0.31 \quad (24a)$$

In this correlation, I_2 and I_5 take the value of unity each for [Si−] group at the 2′- and 5′-positions, respectively. Similarly, $I_{X,Y}$ has been used with a value of unity for a thymine ring that has both X- and Y-substituents. For a thymine ring that has only one substituent either X or Y or no substituent at all, this parameter is zero. The parameter I has the same meaning as in eq 23. Equation 24a thus showing that [Si−] groups would increase the toxicity of the compounds but a disubstituted thymine ring would drastically reduce it, though the thymine ring itself would be slightly more toxic than the cytosine ring.

We also studied the TSAO derivatives used in deriving eq 24a separately and observed the following correlations.

CC_{50} activity of TSAO−T derivatives (**24**) (Table 16a)

$$\log 1/C = -2.25(\pm0.65)B1_Y + 7.32(\pm0.74)$$

$$n = 13,\ r^2 = 0.840,\ s = 0.23,\ q^2 = 0.76 \quad (24b)$$

Outlier 1

It is of interest to note that this equation compares well with eq 24a. The negative coefficient of $B1_Y$ in eq 24b is similar to $I_{X,Y}$ in eq 24a, L_Y in eq 23b, and π_Y in eq 23a. The variation in the activity seem to be totally due to variation in Y-substituents.

CC_{50} activity of TSAO−T *derivatives* (**24**) (Table 16b)

$$\log 1/C = 0.49(\pm0.23)\text{Clog } P - 1.55(\pm1.15)I + 4.72(\pm0.70)$$

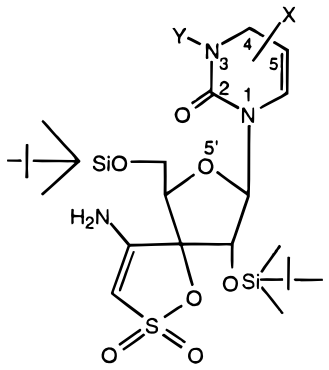
$$n = 8,\ r^2 = 0.856,\ s = 0.42,\ q^2 = 0.597 \quad (24c)$$

Outlier 1

In eq 24c, indicator variable I was used with a value of 1.0 for the presence of a 2′-[Si] group. Its negative coefficient shows that the presence of only a 2′-[Si] group is not conducive to cytotoxic activity. Both 2′,5′-positions of the sugar ring substituted by [Si] group are more effective; the same observation was made in eq 24a. However, eq 24c also shows some hydrophobic interactions, not seen in eq 24a. It is not clear why there is no hydrophobic term in eq 24a.

Thus, for the selectivity of the compounds, a disubstituted thymine ring seems to play a crucial role but replacement of [Si−] groups by H or OH is not advisable, as for the anti-HIV activity the [Si−] groups at both the positions have been found to be essential.[85]

*(iii) EC_{50} Activity of TSAO Derivatives (**26**) in Different Cells (Table 17).*[86] For a small set of

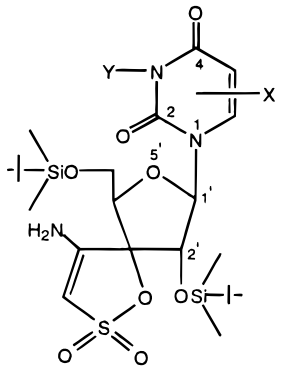


26

TSAO−T and TSAO−C derivatives (**26**) reported by Balzarini et al.,[86] we found, however, that hydrophobic substituents on the pyrimidine ring are essential

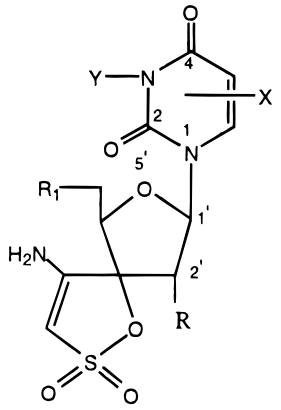
Table 16. CC_{50} Activity of TSAO−T Derivatives (24)[85]

(a)



	substituents		log 1/*C*			
no.	X	Y	obsd	calcd (eq 24b)	Δ	$B1_Y$
1	5-CH_3	H	4.88	5.07	−0.19	1.00
2	H	H	4.84	5.07	−0.23	1.00
3	5-C_2H_5	H	5.26	5.07	0.19	1.00
4	5-CH_3	CH_3	3.62	3.90	−0.28	1.52
5	5-CH_3	C_2H_5	3.91	3.90	0.01	1.52
6	5-CH_3	$CH_2CH{=}C(CH_3)_2$	4.16	3.90	0.26	1.52
7	H	$CH_2CH{=}CH_2$[a]	5.05	3.90	1.15	1.52
8	5-C_2H_5	H	5.28	5.07	0.21	1.00
9	5-F	H	5.30	5.07	0.23	1.00
10	5-Br	H	5.44	5.07	0.37	1.00
11	5-I	H	4.92	5.07	−0.15	1.00
12	6-CF_3	H	4.85	5.07	−0.22	1.00
13	6-CN	H	4.96	5.07	−0.11	1.00
14	6-C(=NH)OCH_3	H	4.96	5.07	−0.11	1.00

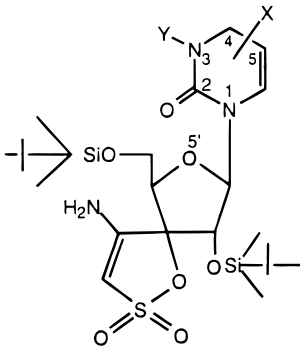
(b)



	substituents			log 1/*C*				
no.	X	R	R_1	obsd	calcd (eq 24c)	Δ	*I*	Clog *P*
1[b]	5-Me	2-[Si][c]	5-[Si]	4.52	5.07	−0.55	1.00	3.92
2[b]	H	2-[Si]	5-[Si]	5.27	4.83	0.44	1.00	3.42
3[b]	H	2-[OH]	5-OH]	3.04	3.34	−0.30	0.00	−2.84
4[b]	H	2-[Si]	5-OH	3.32	3.18	0.14	1.00	0.04
5	5-Me	2-[OH]	5-OH	3.52	3.58	−0.06	0.00	−2.34
6	5-Me	2-[OH]	5-[Si][a]	4.03	5.22	−1.19	0.00	1.04
7	5-Me	H	5-[Si]	5.49	5.13	0.36	0.00	0.85
8	5-Me	2-[Si]	5-OH	3.64	3.43	0.22	1.00	0.54
9	5-Me	2-[Si]	5-Obz[a]	4.44	4.69	−0.25	1.00	3.14

[a] Data points not included in deriving the equation. [b] Sugar ring in β-D-xylo configuration. [c] [Si] = *O*-*tert*-butyldimethylsilyl. [d] $OCH_2C_6H_5$.

Table 17. EC_{50} Activity of TSAO Derivatives (26)[86]



	substituents		log 1/*C*						
no.	Y	X	obsd[a] (eq 25)	calcd (eq 25)	Δ (eq 25)	obsd[b] (eq 26)	calcd (eq 26)	Δ (eq 26)	Clog *P*
1	H	4-OH,5-Me	7.77	7.44	0.33	7.57	7.38	0.19	3.95
2	H	4-OH	6.76	6.98	−0.22	6.81	6.84	−0.03	3.39
3	H	4-NH_2	6.58	6.51	0.07	5.87	5.83	0.04	2.83
4	H	4-NH_2,5-Me	6.92	6.98	−0.05	6.73	6.84	−0.11	3.39
5	3-$CH_2CH{=}CH_2$	4-OH,5-Me				7.10	7.08	0.03	5.05
6	3-Me	4-OH,5-Me	7.77	7.90	−0.13	7.34	7.45	−0.11	4.51

	substituents		log 1/*C*						
no.	Y	X	obsd[c] (eq 27)	calcd (eq 27)	Δ (eq 27)	obsd[d] (eq 28)	calcd (eq 28)	Δ (eq 28)	Clog *P*
1	H	4-OH,5-Me	7.24	7.23	0.01	7.36	7.23	0.01	3.95
2	H	4-OH	6.70	6.86	−0.16	6.83	6.86	−0.16	3.39
3	H	4-NH_2	6.12	6.07	0.05	5.86	6.07	0.05	2.83
4	H	4-NH_2,5-Me	6.91	6.86	0.05	7.29	6.86	0.05	3.39
5	3-$CH_2CH{=}CH_2$	4-OH,5-Me	6.65	6.70	−0.05	6.96	6.70	−0.05	5.05
6	3-Me	4-OH,5-Me	7.25	7.16	0.09	7.48	7.16	0.09	4.51

[a] Activity in MOLT4 cells. [b] Activity in PBL cells. [c] Activity in MT4 cells. [d] Activity in CEM cells.

to anti-HIV activity, as the activities of the compounds in four different cell systems (MOLT4, PBL, MT4, and CEM) were found to be correlated as

$$\log(1/C)_{\mathrm{MOLT4}} = 0.83(\pm0.61)\mathrm{Clog}\,P + 4.18(\pm2.24)$$
$$n = 5,\ r^2 = 0.860,\ s = 0.25,\ q^2 = 0.771 \qquad (25)$$
Outlier 1

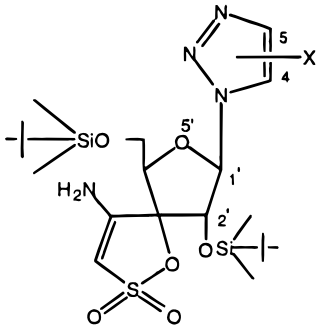
$$\log(1/C)_{\mathrm{PBL}} = 6.46(\pm3.13)\mathrm{Clog}\,P - 0.75(\pm0.39)(\mathrm{Clog}\,P)^2 + 6.47(\pm6.05)$$
$$n = 6,\ r^2 = 0.964,\ s = 0.15,\ q^2 = 0.860 \qquad (26)$$
$$\log P_0 = 4.32\ (4.13{-}4.78)$$

$$\log(1/C)_{\mathrm{MT4}} = 5.66(\pm2.48)\mathrm{Clog}\,P - 0.68(\pm0.31)(\mathrm{Clog}\,P)^2 + 4.50(\pm4.79)$$
$$n = 6,\ r^2 = 0.956,\ s = 0.12,\ q^2 = 0.681 \qquad (27)$$
$$\log P_0 = 4.15\ (4.00{-}4.40)$$

$$\log(1/C)_{\mathrm{CEM}} = 7.25(\pm4.70)\mathrm{Clog}\,P - 0.86(\pm0.59)(\mathrm{Clog}\,P)^2 + 7.71(\pm9.08)$$
$$n = 6,\ r^2 = 0.919,\ s = 0.22,\ q^2 = 0.745 \qquad (28)$$
$$\log P_0 = 4.21\ (3.99{-}4.87)$$

Though these correlations are based on too small a number of data points, they suggest the importance of hydrophobicity. It is of interest that the optimum log *P* values are in reasonable agreement with those of eqs 10, 11, 13, and 18.

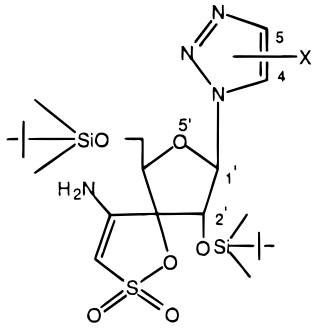
*(iv) EC_{50} Activity of TASO-Tz Derivatives (**27**) in MT-4 Cells (Table 18)[87] and CEM Cells (Table 19).*[87]



27

Alvarez et al.[87] synthesized a series of TSAO−Tz analogues (**27**), where Tz refers to a 1,2,3-triazole ring, and evaluated the anti-HIV activity. For MT-4 cells, we could correlate the activity EC_{50} (50% effective concentration required to reduce the number of viable HIV-1 infected MT-4 cells) as shown by eq 29 and for CEM cells (50% inhibition of virus induced giant cell formation in HIV-1 infected CEM cells) as

Table 18. EC_{50} Activity of TSAO−Tz Derivatives (27)[87]



		log 1/*C*						
no.	X	obsd	calcd (eq 29)	Δ	$\sigma_{I,X}$	$B5_{SUM}$	I_N	Clog *P*
1	4-COOMe	6.22	6.05	0.17	0.32	4.36	0.00	3.80
2	4-$COOC_2H_5$	6.10	5.67	0.43	0.21	5.41	0.00	4.31
3	4-COMe	5.82	5.79	0.04	0.30	4.13	0.00	3.25
4	4-$CH(OCH_2CH_3)_2$	4.14	4.25	−0.12	−0.17	6.40	0.00	4.38
5	4-CH_2OMe[a]	4.57	5.29	−0.72	0.11	4.40	0.00	3.59
6	4-CH_2CH_2Me	6.04	5.76	0.28	−0.01	4.49	0.00	5.35
7	4-$CH_2CH_2CH_2Me$	5.59	5.66	−0.07	−0.04	5.54	0.00	5.93
8	5-COOMe	6.05	6.05	−0.01	0.32	4.36	0.00	3.79
9	5-$COOC_2H_5$	5.50	5.67	−0.18	0.21	5.41	0.00	4.31
10	5-COMe	5.64	5.79	−0.15	0.30	4.13	0.00	3.25
11	5-CH_2OMe	5.62	5.29	0.33	0.11	4.40	0.00	3.59
12	5-CH_2CH_2Me	5.50	5.76	−0.26	−0.01	4.49	0.00	5.35
13	5-$CH_2CH_2CH_2Me$	5.80	5.66	0.14	−0.04	5.54	0.00	5.93
14	H	5.42	5.72	−0.30	0.00	2.00	0.00	3.80
15	4,5-$(COOMe)_2$	6.32	6.38	−0.06	0.64	6.72	0.00	3.78
16	4,5-$(COOC_2H_5)_2$	5.42	5.63	−0.20	0.42	8.82	0.00	4.82
17	4-$CONH_2$[a]	5.92	6.49	−0.57	0.28	4.07	1.00	2.31
18	5-$CONH_2$	6.28	6.49	−0.21	0.28	4.07	1.00	2.31
19	4-CONHMe	6.39	6.57	−0.19	0.28	4.16	1.00	2.53
20	5-CONHMe	6.80	6.57	0.22	0.28	4.16	1.00	2.53
21	4-$CONMe_2$	6.38	6.21	0.17	0.28	5.04	1.00	2.29
22	5-$CONMe_2$[a]	7.22	6.21	1.01	0.28	5.04	1.00	2.29
23	4-Me	5.55	5.58	−0.03	−0.04	3.04	0.00	4.36

[a] Data points not used in deriving equation.

shown by eq 30. In these equations, σ_I is the field/inductive electronic parameter.

$$\log 1/C = 3.10(\pm 0.91)\sigma_{I,X} - 0.28(\pm 0.11)B5_{SUM} + 1.21(\pm 0.44)I_N + 0.49(\pm 0.22)\ \text{Clog}\ P + 4.41(\pm 0.85)$$

$$n = 20,\ r^2 = 0.857,\ s = 0.24,\ q^2 = 0.777 \quad (29)$$

Outliers 3

$$\log 1/C = 0.76(\pm 0.58)\sigma_{I,X} - 0.39(\pm 0.12)\pi_{X,5} + 0.19(\pm 0.08)(\pi_{X,5})^2 + 5.74(\pm 0.16)$$

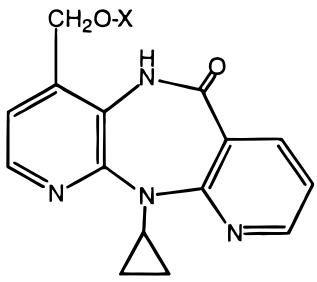
$$n = 20,\ r^2 = 0.835,\ s = 0.20,\ q^2 = 0.748 \quad (30)$$

$(\pi_5)_0 = 1.03$ (0.65−1.84), Outliers 3

Equation 29, therefore, indicates a dominant electronic effect of X-substituents on the activity in MT-4 cells but simultaneously a significant steric effect too because of their width. In QSAR 29, indicator variable I_N was used with a value of unity for 4- or/and 5-$CONR_2$ substituents and its positive coefficient shows that the presence of this group is advantageous to the activity. In CEM cells, however, the electronic effect of X-substituents does not appear to be so dominant (eq 30) but it is conducive to activity. The 5-substituents with low hydrophobic character are shown to be disadvantageous but with a value of $\pi > 1.03$ can be beneficial.

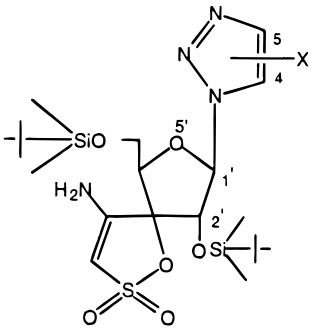
d. Nevirapine Derivatives. Mutant viruses that are resistant to HIV-1 RT inhibitors such as nevirapine have emerged in both cell culture and clinical settings. Nevirapine derivatives have been studied in detail for their anti-HIV activity by some authors.

*(i) IC_{50} Activity of Nevirapine Analogues (**28**) against Wild-Type HIV-1 Reverse Transcriptase (Table 20).*[88] For a small series of 4-substituted nevirapine



(**28**) studied by Kelly et al.,[88] we derived eq 31, which is not a good QSAR with 3 outliers and a low q^2. There is high collinearity ($r^2 = 0.72$) between CMR, log *P*, and MgVol, so it is difficult to say which parameter is best. We choose Clog *P* simply because

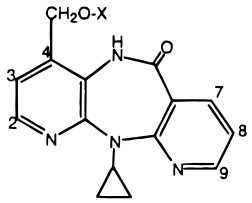
Table 19. EC_{50} Activity of TSAO−Tz Derivatives (27)[87]



		log 1/C				
no.	X	obsd	calcd (eq 30)	Δ	$\sigma_{I,X}$	$\pi_{5,X}$
1	4-COOMe[a]	6.74	5.99	0.76	0.32	0.00
2	4-$COOC_2H_5$	6.27	5.90	0.37	0.21	0.00
3	4-COMe	5.96	5.97	−0.01	0.30	0.00
4	4-CH_2OMe	5.64	5.83	−0.19	0.11	0.00
5	4-CH_2CH_2Me	5.89	5.74	0.15	−0.01	0.00
6	4-$CH_2CH_2CH_2Me$	5.82	5.71	0.11	−0.04	0.00
7	4-$CH_2CH_2CH_2Cl$[a]	6.18	5.76	0.42	0.02	0.00
8	5-COOMe[a]	6.89	5.99	0.90	0.32	−0.01
9	5-$COOC_2H_5$	5.89	5.75	0.13	0.21	0.51
10	5-COMe	5.96	6.24	−0.28	0.30	−0.55
11	5-CH_2OMe	6.05	6.24	−0.20	0.11	−0.78
12	5-CH_2CH_2Me	5.41	5.59	−0.18	−0.01	1.55
13	5-$CH_2CH_2CH_2Me$	5.82	5.74	0.09	−0.04	2.13
14	H	5.47	5.74	−0.28	0.00	0.00
15	4,5-$(COOMe)_2$	6.28	6.23	0.04	0.64	−0.01
16	4,5-$(COOC_2H_5)_2$	5.80	5.91	−0.12	0.42	0.51
17	4-$CONH_2$	6.05	5.96	0.09	0.28	0.00
18	5-$CONH_2$	6.85	6.95	−0.10	0.28	−1.49
19	4-CONHMe	5.92	5.96	−0.04	0.28	0.00
20	5-CONHMe	7.10	6.75	0.35	0.28	−1.27
21	4-$CONMe_2$	5.92	5.96	−0.04	0.28	0.00
22	5-$CONMe_2$	6.92	6.97	−0.05	0.28	−1.51
23	4-Me	5.85	5.71	0.14	−0.04	0.00

[a] Data points not used in deriving equation.

Table 20. IC_{50} Activity of Nevirapine Analogues (28)[88]



		log 1/C				
no.	X	obsd	calcd (eq 31)	Δ	Clog P	MR_X
1	H	5.52	5.79	−0.27	1.00	0.10
2	C_2H_5	5.89	5.49	0.40	2.34	1.03
3	$CH_2CH{=}CH_2$	5.85	5.66	0.20	2.59	1.45
4	C_6H_5[a]	6.92	5.70	1.22	3.69	2.54
5	$CH_2C_6H_5$	6.57	6.56	0.01	3.16	3.00
6	$CH_2CH_2C_6H_5$	6.38	6.37	0.01	3.88	3.47
7	C_6H_4-2-Me[a]	6.54	5.69	0.85	4.19	3.00
8	C_6H_4-3-Me	5.61	5.69	−0.08	4.19	3.00
9	C_6H_4-4-Me[a]	6.24	5.69	0.55	4.19	3.00
10	C_6H_4-2-OH	6.34	6.48	−0.14	2.96	2.72
11	C_6H_4-2-Cl	5.52	5.62	−0.11	4.31	3.04
12	C_6H_4-4-NH_2	7.00	6.98	0.02	2.63	2.98
13	C_6H_4-4-NHC_2H_5	7.10	6.66	0.44	3.90	3.81
14	C_6H_4-4-OMe	5.98	6.19	−0.21	3.78	3.17
15	C_6H_4-4-CN	6.39	6.38	0.01	3.45	3.07
16	C_6H_4-4-NO_2	5.95	6.23	−0.28	3.73	3.17

[a] Data points not used in deriving equation.

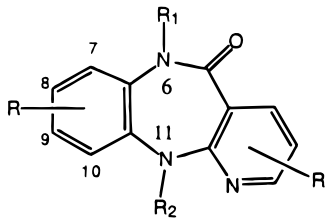
it gives the best correlation. No significant correlation was observed if we use only one of these parameters.

$$\log 1/C = -0.84(\pm 0.36)\text{Clog } P + 0.89(\pm 0.31)MR_X + 6.54(\pm 0.62)$$

$$n = 13,\ r^2 = 0.808,\ s = 0.25,\ q^2 = 0.595 \quad (31)$$

Outliers 3

*(ii) IC_{50} Activity of Pyridobenzodiazepin-5-ones (**29**) against HIV-1 Reverse Transcriptase (Table 21).*[89]



29

Hargrave et al.[89] reported anti-HIV data on 6,11-dihydro-5*H*-pyridobenzodiazepin-5-ones derivatives, from which we developed eq 32.

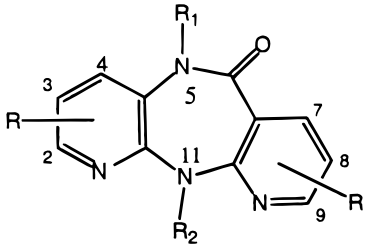
$$\log 1/C = 0.94(\pm 0.19)\text{Clog } P - 0.82(\pm 0.21)B5_{11} - 1.09(\pm 0.41)I + 6.32(\pm 0.67)$$

$$n = 28,\ r^2 = 0.847,\ s = 0.31,\ q^2 = 0.765 \quad (32)$$

Outliers 5

Equation 32 shows that there are hydrophobic interactions between the receptor and ligand. Steric bulk at the 11-position seems to be detrimental to activity as indicated by B5, the Verloop sterimol parameter. In eq 32, indicator variable *I* was used with a value of unity for 2- and 3-substituted compounds; its negative coefficient suggests that these substituents have detrimental steric effects on the activity.

*(iii) IC_{50} Activity of Dipyridobenzodiazepin-6-ones (**30**) against HIV-1 Reverse Transcriptase (Table 22).*[89]



30

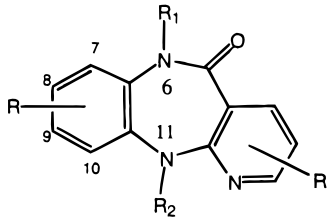
Other data by Hargrave et al.[89] on 5,11-dihydro-6*H*-dipyridobenzodiazepin-6-ones gave eq 33.

$$\log 1/C = 0.38(\pm 0.11)\text{Clog } P - 1.54(\pm 0.52)B1_{11} - 0.51(\pm 0.36)I_3 + 7.84(\pm 0.86)$$

$$n = 20,\ r^2 = 0.840,\ s = 0.26,\ q^2 = 0.789 \quad (33)$$

Outliers 4

Table 21. IC_{50} Activity of Pyridobenzodiazepin-5-ones (29)[89]

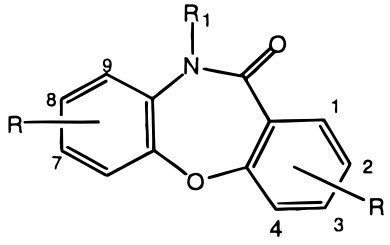


		log 1/*C*					
no.	substituents	obsd	calcd (eq 32)	Δ	Clog *P*	$B5_{11}$	*I*
1	11- C_2H_5	6.42	6.67	−0.25	3.12	3.17	0.00
2	6-Me,11-$CH_2C_6H_5$	5.05	5.03	0.02	3.85	6.02	0.00
3	6-allyl[a]	5.85	8.12	−2.27	2.78	1.00	0.00
4	6-Me,11-C_2H_5	6.46	6.19	0.26	2.61	3.17	0.00
5	6,11-C_2H_5	6.42	6.69	−0.27	3.14	3.17	0.00
6	6-Me,11-C_3H_7	6.64	6.43	0.21	3.14	3.49	0.00
7	6-Me,11-$CHMe_2$	6.70	6.48	0.22	2.72	3.17	0.00
8	6-Me,11-C_4H_9	6.00	6.07	−0.07	3.67	4.54	0.00
9	6-Me,11-CH_2SMe	5.18	5.61	−0.43	2.30	3.53	0.00
10	6-Me,11-$CH_2COOC_2H_5$[a]	4.80	2.94	1.86	2.17	6.64	0.00
11	6-Me,11-CH_2CH_2F	6.00	5.93	0.07	2.33	3.17	0.00
12	6-Me,11-acetylene	6.19	6.27	−0.09	1.33	1.60	0.00
13	6,11-di-C_2H_5,7-Me[a]	5.10	7.16	−2.07	3.64	3.17	0.00
14	6,7-di-Me,11-C_2H_5	6.62	6.66	−0.04	3.11	3.17	0.00
15	11-C_2H_5,7-Me	7.42	7.14	0.28	3.62	3.17	0.00
16	6,8-di-Me,11-C_2H_5	6.72	6.66	0.06	3.11	3.17	0.00
17	6-Me,8-OMe,11-C_2H_5	5.68	6.13	−0.46	2.55	3.17	0.00
18	3-Cl,6-Me,11-C_2H_5	5.12	5.80	−0.68	3.35	3.17	1.00
19	6-Me,9-CF_3,11-C_2H_5	6.64	6.96	−0.32	3.43	3.17	0.00
20	2-F,6-Me,11-C_2H_5	5.54	5.26	0.28	2.78	3.17	1.00
21	6-Me,9-NO_2,11-C_2H_5	6.40	5.92	0.48	2.32	3.17	0.00
22	6-Me,9-NH_2,11-C_2H_5	5.35	5.04	0.31	1.38	3.17	0.00
23	6-Me,9-N_3,11-C_2H_5	6.85	6.61	0.24	3.05	3.17	0.00
24	6-Me,8-COOMe,11-C_2H_5	5.96	6.10	−0.14	2.52	3.17	0.00
25	6-Me,8-$CONH_2$,11-C_2H_5	4.57	4.82	−0.25	1.16	3.17	0.00
26	6,8,9-tri-Me,11-C_2H_5	7.40	7.09	0.31	3.56	3.17	0.00
27	6,8,9-tri-Me,11-COMe[a]	6.38	4.60	1.78	0.88	3.13	0.00
28	6,7,8-tri-Me,11-C_2H_5	7.44	7.09	0.36	3.56	3.17	0.00
29	6-Me,8,9-di-Cl,11-C_2H_5	7.21	7.46	−0.25	3.95	3.17	0.00
30	6-Me,7-Cl,9-CF_3,11-C_2H_5[a]	6.54	7.64	−1.10	4.15	3.17	0.00
31	3-Br,6-Me,9-NO_2,11-C_2H_5	6.05	5.65	0.40	3.19	3.17	1.00
32	6-Me,8-Cl,11-C_2H_5	6.77	6.89	−0.12	3.35	3.17	0.00
33	6-Me,9-Cl,11-C_2H_5	6.77	6.89	−0.12	3.35	3.17	0.00

[a] Data points not used in deriving equation.

Equation 33 shows a hydrophobic effect of the molecule and a detrimental steric effect of 11-position substituents. Indicator variable I_3 was used with a value of 1.0 for 3-Cl and 3-Me; its negative coefficient shows the negative effect of these substituents.

*(iv) IC_{50} Activity of Benzoxazepinones (**31**) against HIV-1 Reverse Transcriptase (Table 23).*[90] For a series of dibenz[*b*,*f*]oxazepin-11(10*H*)-one (**31**) analogues of **26**, studied by Klunder et al.,[90] we formulated eq 34, which exhibits that while 2-, 7-, and 9-substituents may be advantageous because of their involvement in the dispersion interaction with the enzyme, 8- and 10-substituents may be detrimental due to some steric effects. It is not clear why there is no hydrophobic term.



31

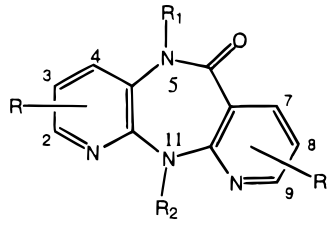
$$\log 1/C = 0.55(\pm 0.25)MR_{2,7,9} - 0.52(\pm 0.23)MR_{8,10} + 6.94(\pm 0.46)$$

$$n = 19,\ r^2 = 0.821,\ s = 0.22,\ q^2 = 0.727 \quad (34)$$

Outliers 3

*(v) IC_{50} Activity of Pyridobenzoxazepinones (**32**) against HIV-1 Reverse Transcriptase (Table 24).*[90] When we performed QSAR on another series of pyrido[2,3-*b*[1,5]benzoxazepin-5(6*H*)-ones (**32**) re-

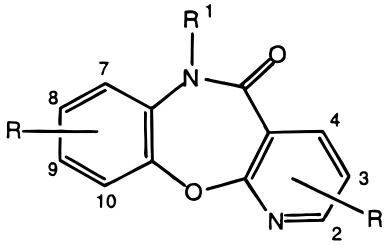
Table 22. IC_{50} Activity of Dipyridobenzodiazepin-6-ones (30)[89]



		log 1/*C*					
no.	substituents	obsd	calcd (eq 33)	Δ	Clog *P*	$B1_{11}$	I_3
1	5-Me,11-C_2H_5	6.90	6.38	0.53	2.30	1.52	0.00
2	11-C_2H_5	6.36	6.63	−0.27	2.95	1.52	0.00
3	11-Cy-C_3H_5	6.35	6.48	−0.13	2.69	1.55	0.00
4	11-Cy-C_4H_5	6.37	6.35	0.01	3.24	1.77	0.00
5	5-Me,11-C_3H_7	6.35	6.58	−0.23	2.82	1.52	0.00
6	11-$CHMe_2$	6.23	6.16	0.07	3.26	1.90	0.00
7	5-Me,11-CMe_3	4.96	4.98	−0.02	3.00	2.60	0.00
8	5-Me,11-$CH_2CH{=}CH_2$	6.07	6.47	−0.40	2.54	1.52	0.00
9	5-Me,11-COMe	4.82	4.67	0.15	−1.86	1.60	0.00
10	5-Me,11-CH_2CH_2F[a]	5.54	6.27	−0.73	2.02	1.52	0.00
11	5-Me,11-CH_2SMe	6.07	6.26	−0.19	1.99	1.52	0.00
12	2,5-di-Me,11-C_2H_5	6.77	6.57	0.20	2.79	1.52	0.00
13	2-Cl,5-Me,11-C_2H_5	6.82	6.62	0.21	2.92	1.52	0.00
14	3,5-di-Me,11-C_2H_5	6.12	6.06	0.07	2.79	1.52	1.00
15	4-Me,11-C_2H_5[a]	7.46	6.57	0.89	2.80	1.52	0.00
16	4,5-di-Me,11-C_2H_5[a]	5.72	6.41	−0.69	2.39	1.52	0.00
17	4-Me,11-Cy-C_3H_5[a]	7.08	6.42	0.65	2.53	1.55	0.00
18	4-Cl,11-C_2H_5	7.02	6.59	0.44	2.85	1.52	0.00
19	4,11-C_2H_5	6.96	6.77	0.19	3.33	1.52	0.00
20	4-CH_2OH,11-Cy-C_3H_5	5.52	5.83	−0.31	1.00	1.55	0.00
21	4-CN,11-Cy-C_3H_5	5.90	6.02	−0.12	1.49	1.55	0.00
22	2,3-Me,11-C_2H_5	6.39	6.48	−0.09	3.90	1.52	1.00
23	2,3,5-Me,11-C_2H_5	6.62	6.74	−0.12	3.24	1.52	1.00
24	2-NH_2,3-Cl,5-Me,11-C_2H_5	6.07	6.05	0.03	2.77	1.52	1.00

[a] Data points not used in deriving equation.

ported by Klunder et al.,[90] eq 35 was derived that shows that not only the overall hydrophobicity of the



32

molecule, but also certain specific properties of some substituents may be favorable to activity. According to eq 35, 7-position substituents could have a positive steric effect and 9-position substituents can also increase the potency by electron withdrawal.

$$\log 1/C = 0.67(\pm 0.31)\text{Clog } P + 1.36(\pm 0.68)B1_7 + 0.75(\pm 0.40)\sigma_7 + 3.59(\pm 1.44)$$

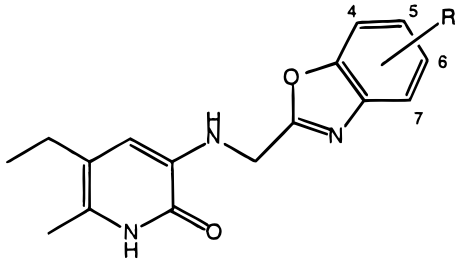
$$n = 13,\ r^2 = 0.862,\ s = 0.23,\ q^2 = 0.608 \quad (35)$$

Outliers 2

In deriving eqs 31−35 some compounds, as indicated in the respective tables, were not included as they were found to be misfit in the correlations. No specific reasons could be assigned in any case to these aberrations.

e. Pyridinone Derivatives. An extensive SAR study was made on RT inhibition data of pyridinone analogues reported by Hoffman et al.[91−94] QSAR studies were performed by us on four different series as represented by **33**−**36**.

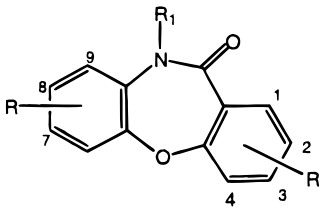
*(i) IC_{50} Activity of Pyridinone Derivatives (**33**) against HIV-1 Reverse Transcriptase (Table 25).*[92] For



33

the series of **33** where there were variations only in the R-substituents of the benzoxazole moiety, a highly significant correlation between the activity and hydrophobic, electronic, and steric properties of substituents was obtained, suggesting a positive role of hydrophobic 4- and 7-position R-substituents. Electron-releasing R-substituents that are marginal

Table 23. IC_{50} Activity of Benzoxazepinones (31)[90]



		log 1/*C*				
no.	substituents	obsd	calcd (eq 34)	Δ	$MR_{2,7,9}$	$MR_{8,10}$
1	10-C_2H_5	6.25	6.52	−0.27	0.31	1.13
2	10-$CH_2CH{=}CH_2$	6.33	6.30	0.03	0.31	1.55
3	10-C_3H_7	6.43	6.28	0.15	0.31	1.60
4	10-$CHMe_2$	6.47	6.28	0.19	0.31	1.60
5	2-NH_2,10-Me	6.75	7.00	−0.26	0.75	0.67
6	2-NH_2,10-C_2H_5	6.84	6.76	0.08	0.75	1.13
7	7,10-di-Me[a]	6.22	7.01	−0.79	0.77	0.67
8	2-NH_2,7,10-di-Me	7.52	7.25	0.27	1.21	0.67
9	2-NH_2,7-Me,10-C_3H_7	6.89	6.77	0.12	1.21	1.60
10	2-NH_2,8-Me,10-C_2H_5	6.62	6.52	0.10	0.75	1.60
11	2-NH_2,9-Me,10-C_2H_5	7.22	7.01	0.21	1.21	1.13
12	2-NH_2,9-Me,10-C_3H_7	6.59	6.77	−0.18	1.21	1.60
13	2-NH_2,7,8-di-Me,10-C_3H_7	6.78	6.53	0.26	1.21	2.06
14	2-NH_2,8,9-di-Me,10-C_3H_7	6.17	6.53	−0.35	1.21	2.06
15	2-NH_2,6,8-di-Me,10-C_3H_7	6.05	6.27	−0.23	0.75	2.06
16	2-NH_2,7,9,10-tri-Me	7.70	7.51	0.19	1.67	0.67
17	2-NH_2,7,9-di-Me,10-C_2H_5	7.37	7.26	0.10	1.67	1.13
18	2-OH,7,10-di-Me[a]	6.19	7.11	−0.92	0.95	0.67
19	2-CN,7,10-Me[a]	6.55	7.30	−0.75	1.30	0.67
20	7,9,10-tri-Me	7.24	7.27	−0.02	1.23	0.67
21	2-CN,7,9,10-tri-Me	7.30	7.55	−0.25	1.76	0.67
22	2-OH,7,9,10-tri-Me	7.20	7.36	−0.16	1.42	0.67

[a] Data points not used in deriving equation.

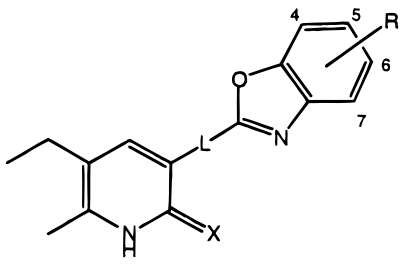
at best seem to favor activity, while a negative MR term indicates the detrimental steric effect of R-substituents.

$$\log 1/C = 1.85(\pm 0.32)\pi_{4,7} - 1.46(\pm 0.55)MR_R - 0.62(\pm 0.52)\sigma_R + 7.12(\pm 0.43)$$

$$n = 18,\ r^2 = 0.918,\ s = 0.30,\ q^2 = 0.885 \quad (36)$$

Outliers 2

*(ii) IC_{50} Activity of Pyridinone Derivatives (**34**) against HIV-1 Reverse Transcriptase (Table 26).*[93] For



34

the series of **34** where there were variations at the pyridine ring, too, at its 2-position (X = O or S) and the alterations were made in the linker chain L, the correlation obtained was

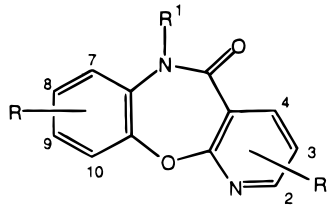
$$\log 1/C = 23.20(\pm 3.94)L_L - 2.96(\pm 0.51)(L_L)^2 - 1.28(\pm 0.42)I_{5,6} - 2.28(0.85)\sigma_{I,L} - 37.78(\pm 7.59)$$

$$n = 29,\ r^2 = 0.884,\ s = 0.33,\ q^2 = 0.671 \quad (37)$$

$(L_L)_0 = 3.92$ (3.85−3.99), Outliers 4

Equation 37 indicates that the length of the whole linker chain affects activity, and its parameter being parabolic in nature suggests that a chain length up to an optimum value of 3.92 will promote activity. The R-substituents at 5- and 6-positions may have steric effects, as an indicator variable $I_{5,6}$ used for them with a value of unity has a significant negative coefficient. Also the $\sigma_{I,L}$ parameter shows the linker chain to have an inductive effect on activity. Electronegative groups of the linker chain were considered attached to the pyridine ring. There is no hydrophobic term in this equation. In this series the variation in the activity was mainly due to variation in the linker chain (see Table 26), which might be twisting the benzoxazole ring out of plane so that the molecule is unable to bind in the hydrophobic pocket. It might also be hitting the receptor wall as it becomes longer.

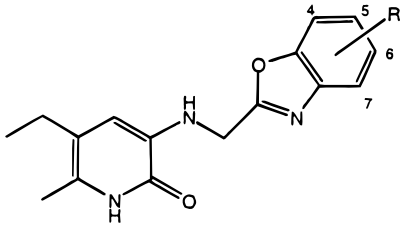
Table 24. IC_{50} Activity of Pyridobenzoxazepinones (32)[90]



no.	substituents	log 1/*C* obsd	log 1/*C* calcd (eq 35)	Δ	Clog *P*	$B1_7$	σ_7
1	6-C_2H_5,7-NO_2[a]	6.47	7.34	−0.88	1.28	1.70	0.78
2	6,9-di-Me,3-NH_2	6.60	6.42	0.18	2.20	1.00	0.00
3	6,8,9-tri-Me,3-NH_2	6.49	6.72	−0.23	2.65	1.00	0.00
4	6-C_2H_5,8,9-di-Me,3-NH_2	7.15	7.08	0.07	3.18	1.00	0.00
5	6-$CHMe_2$,8,9-di-Me,3-NH_2	7.35	7.28	0.07	3.49	1.00	0.00
6	6,7,9-tri-Me,3-NH_2	7.34	7.07	0.27	2.30	1.52	−0.17
7	6-C_2H_5,7,9-di-Me,3-NH_2	7.57	7.42	0.15	2.83	1.52	−0.17
8	6-C_2H_5,7,9-di-Me	7.54	7.60	−0.06	3.10	1.52	−0.17
9	6-C_2H_5,7-NO_2,9-Me	7.51	7.67	−0.16	1.78	1.70	0.78
10	6-C_2H_5,7-NH_2,9-Me[a]	6.88	6.61	0.27	2.52	1.35	−0.66
11	6-C_2H_5,7-CN,9-Me	7.72	7.85	−0.13	2.39	1.60	0.66
12	6-Me,7-CN,9-Me	7.26	7.50	−0.24	1.86	1.60	0.66
13	6-Me,7-NO_2,9-Me	7.64	7.32	0.32	1.25	1.70	0.78
14	6-Me,7-NH_2,9-Me	6.00	6.26	−0.26	1.99	1.35	−0.66
15	6-Me,7-CO_2Me,9-Me	7.66	7.63	0.03	2.21	1.64	0.45

[a] Data points not used in deriving equation.

Table 25. IC_{50} Activity of Pyridinone Derivatives (33)[92]

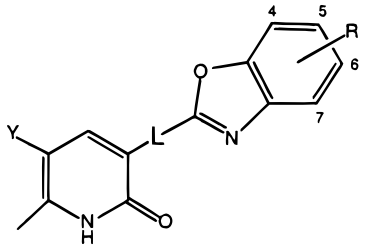


no.	R	log 1/*C* obsd	log 1/*C* calcd (eq 36)	Δ	$\pi_{4,7}$	MR_R	σ_R
1	H	6.68	6.52	0.16	0.00	0.41	0.00
2	4-Me	6.92	6.99	−0.07	0.56	0.87	−0.17
3	5-Me	5.90	5.95	−0.05	0.00	0.87	−0.17
4	6-Me	5.78	5.95	−0.17	0.00	0.87	−0.17
5	7-Me	7.26	6.99	0.27	0.56	0.87	−0.17
6	7-C_2H_5	6.59	7.15	−0.57	1.02	1.34	−0.15
7	4,7-di-Me	7.70	7.46	0.24	1.12	1.34	−0.34
8	4-Cl	6.82	6.96	−0.14	0.71	0.91	0.23
9	7-Cl	7.19	6.96	0.22	0.71	0.91	0.23
10	4,7-di-Cl	7.72	7.41	0.31	1.42	1.41	0.46
11	4-F	6.96	6.76	0.20	0.14	0.40	0.06
12	5-F	6.33	6.50	−0.17	0.00	0.40	0.06
13	6-F	5.90	6.50	−0.60	0.00	0.40	0.06
14	7-F	7.04	6.76	0.28	0.14	0.40	0.06
15	4-F,7-Cl	6.98	7.20	−0.23	0.85	0.90	0.29
16	4,7-di-F	7.16	7.00	0.16	0.28	0.39	0.12
17	4-OMe[a]	6.75	5.65	1.09	−0.02	1.10	−0.27
18	4-OH[a]	6.36	5.24	1.12	−0.67	0.59	−0.37
19	4-NO_2	4.61	4.60	0.01	−0.28	1.05	0.78
20	4-NH_2	4.17	4.01	0.17	−1.23	0.85	−0.66

[a] Data points not used in deriving equation.

*(iii) IC_{50} Activity of Pyridinone Derivatives (**35**) against HIV-1 Reverse Transcriptase (Table 27).*[91,92]

Hoffman et al. reported IC_{50} (minimum concentration of compound producing 50% inhibition of HIV-1



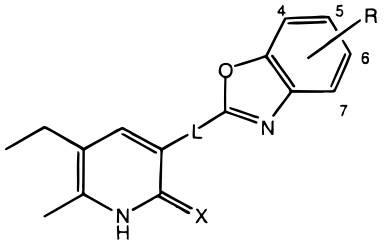
35

reverse transcriptase) data for the series of **35,** where the substituent effect was studied at the 5-position, R was only 4,7-di-Cl or H, and linker chain L was either $-CH_2CH_2-$ or $-NHCH_2-$. The best equation obtained was eq 38.

$$\log 1/C = -0.99(\pm 0.36)MR_Y + 1.98(\pm 0.90)B1_Y + 0.69(\pm 0.12)\text{Clog } P - 3.34(\pm 1.38)$$
$$n = 23,\ r^2 = 0.902,\ s = 0.33,\ q^2 = 0.858 \quad (38)$$
Outlier 1

Equation 38 shows that the hydrophobicity of compounds is important for activity. However, the negative MR_Y term and the positive $B1_Y$ term indicate steric effects of Y-substituents. The correlation analysis did not reveal any preference between the two groups of alkyl linker chains.

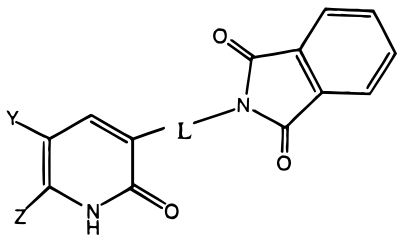
Table 26. IC_{50} Activity of Pyridinone Derivatives (34)[93]



	substituents			log 1/*C*					
no.	R	X	L	obsd	calcd (eq 37)	Δ	L_L	$I_{5,6}$	$\sigma_{I,L}$
1	H	S	CH_2CH_2	7.62	7.57	0.05	4.11	0.00	−0.01
2	4,7-di-Me	S	CH_2CH_2	7.28	7.57	−0.29	4.11	0.00	−0.01
3	4,7-di-Cl	S	CH_2CH_2	7.52	7.57	−0.05	4.11	0.00	−0.01
4	4,7-di-F	S	CH_2CH_2	7.82	7.57	0.25	4.11	0.00	−0.01
5	4-F	S	CH_2CH_2	7.89	7.57	0.32	4.11	0.00	−0.01
6	7-F	S	CH_2CH_2	7.37	7.57	−0.20	4.11	0.00	−0.01
7	4-Cl	S	CH_2CH_2	7.52	7.57	−0.05	4.11	0.00	−0.01
8	7-Cl	S	CH_2CH_2	7.54	7.57	−0.03	4.11	0.00	−0.01
9	H	O	CH_2CH_2	7.64	7.57	0.07	4.11	0.00	−0.01
10	4-Me	O	CH_2CH_2	7.48	7.57	−0.09	4.11	0.00	−0.01
11	4-Cl	O	CH_2CH_2	7.21	7.57	−0.36	4.11	0.00	−0.01
12	4-F	O	CH_2CH_2	7.82	7.57	0.25	4.11	0.00	−0.01
13	7-Me	O	CH_2CH_2	7.40	7.57	−0.17	4.11	0.00	−0.01
14	7-Cl	O	CH_2CH_2	7.41	7.57	−0.16	4.11	0.00	−0.01
15	7-F	O	CH_2CH_2	7.43	7.57	−0.14	4.11	0.00	−0.01
16	4,7-di-Me	O	CH_2CH_2	7.55	7.57	−0.02	4.11	0.00	−0.01
17	4,7-di-Cl	O	CH_2CH_2	7.85	7.57	0.28	4.11	0.00	−0.01
18	4,7-di-F	O	CH_2CH_2	7.85	7.57	0.28	4.11	0.00	−0.01
19	6-Me	O	CH_2CH_2	6.76	6.29	0.47	4.11	1.00	−0.01
20	6-F	O	CH_2CH_2	6.35	6.29	0.06	4.11	1.00	−0.01
21	5-F	O	CH_2CH_2	5.77	6.29	−0.52	4.11	1.00	−0.01
22	H	O	OCH_2	6.72	7.03	−0.31	3.98	0.00	0.27
23	4,7-di-Cl	O	OCH_2	7.06	7.03	0.03	3.98	0.00	0.27
24	4,7-di-Cl	O	SCH_2[a]	7.96	6.66	1.31	4.30	0.00	0.25
25	4,7-di-Cl	O	$S(O)CH_2$	5.85	6.43	−0.58	4.11	0.00	0.49
26	4,7-di-Cl	O	SO_2CH_2	6.69	6.21	0.49	4.11	0.00	0.59
27	H	O	$NHCH_2$	6.68	6.92	−0.24	3.53	0.00	0.13
28	4,7-di-Cl	O	$NHCH_2$	7.70	6.92	0.78	3.53	0.00	0.13
29	H	O	CH_2NH[a]	5.90	7.63	−1.73	4.02	0.00	0.00
30	H	O	CH═CH(*trans*)[a]	5.23	6.70	−1.77	4.29	0.00	0.11
31	H	O	CH═CH(*cis*)[a]	5.52	6.70	−1.48	4.29	0.00	0.11
32	H	O	CH_2	4.35	4.51	−0.16	2.87	0.00	−0.04
33	H	O	$(CH_2)_3$	4.80	4.70	0.10	4.92	0.00	−0.01

[a] Data points not used in deriving equation.

*(iv) IC_{50} Activity of Pyridinone Derivatives (**36**) against HIV-1 Reverse Transcriptase (Table 28).*[91]



36

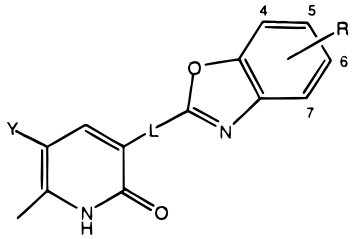
Hoffman et al.[91] reported another series of pyridinone derivatives (**36**) in which the chain links the pyridinone ring to a phthalimide ring and not to the benzoxazole ring (as in **33**−**35**), and the substitution effects were studied only at the pyridinone ring and not at the phthalimide. We obtained the following correlation for the data, but we did not observe any hydrophobic term here as observed in eqs 36 and 38. This is a small set with no variation in the phenyl ring R-substituent. It seems that variation in R is required to bring out the hydrophobic interaction. Garg et al.[95] proposed hydrophobic interactions between the R-substituent at phenyl ring and receptor.

$$\log 1/C = 0.92(\pm 0.77)MR_Y - 2.08(\pm 0.77)L_L + 13.01(\pm 3.00)$$

$$n = 11,\ r^2 = 0.839,\ s = 0.58,\ q^2 = 0.671 \quad (39)$$

Outliers 2

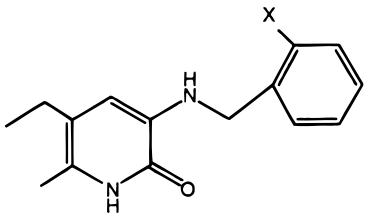
Table 27. IC_{50} Activity of Pyridinone Derivatives (35)[91,92]



	substituents			log 1/*C*					
no.	R	Y	L	obsd	calcd (eq 38)	Δ	MR_Y	$B1_Y$	Clog *P*
1	H	Me	CH_2CH_2	6.77	7.07	−0.30	0.57	1.52	1.87
2	H	C_2H_5	CH_2CH_2	7.64	6.98	0.66	1.03	1.52	2.39
3	H	C_3H_7	CH_2CH_2	7.08	6.89	0.19	1.50	1.52	2.92
4	H	$CHMe_2$	CH_2CH_2	7.32	7.55	−0.23	1.50	1.90	2.79
5	H	C_6H_5	CH_2CH_2	5.96	6.29	−0.33	2.54	1.71	3.00
6	H	CN	CH_2CH_2	6.51	6.87	−0.36	0.63	1.60	1.44
7	H	NHCOMe	CH_2CH_2	4.38	4.54	−0.16	1.49	1.35	0.03
8	H	CH_2OH	CH_2CH_2	5.47	5.85	−0.38	0.72	1.52	0.33
9	H	CH_2OMe	CH_2CH_2	5.68	5.94	−0.26	1.21	1.52	1.14
10	H	NMe_2	CH_2CH_2	5.73	5.77	−0.04	1.56	1.35	1.89
11	4,7-di-Cl	Me	CH_2CH_2[a]	7.01	8.12	−1.11	0.57	1.52	3.37
12	4,7-di-Cl	C_2H_5	CH_2CH_2	7.85	8.03	−0.18	1.03	1.52	3.90
13	4,7-di-Cl	C_3H_7	CH_2CH_2	7.47	7.93	−0.46	1.50	1.52	4.43
14	4,7-di-Cl	SMe	$NHCH_2$	7.37	7.24	0.13	1.38	1.70	2.75
15	4,7-di-Cl	C_2H_5	$NHCH_2$	7.72	7.52	0.20	1.03	1.52	3.17
16	4,7-di-Cl	$CH{=}CH_2$	$NHCH_2$	7.64	7.43	0.21	1.10	1.60	2.92
17	4,7-di-Cl	OMe	$NHCH_2$	6.94	6.93	0.01	0.79	1.35	2.46
18	4,7-di-Cl	OCOMe	$NHCH_2$	6.52	6.08	0.44	1.25	1.35	1.89
19	H	C_2H_5	$NHCH_2$	6.68	6.47	0.21	1.03	1.52	1.67
20	H	SMe	$NHCH_2$	6.72	6.19	0.53	1.38	1.70	1.25
21	H	SC_2H_5	$NHCH_2$	6.37	6.10	0.27	1.84	1.70	1.78
22	H	SO_2Me	$NHCH_2$	5.94	5.87	0.07	1.35	2.03	−0.20
23	H	$CO_2C_2H_5$	$NHCH_2$	5.76	5.86	−0.10	1.75	1.64	1.47
24	H	S(O)Me	$NHCH_2$	4.50	4.61	−0.11	1.37	1.40	−0.19

[a] Data point not used in deriving equation.

*(v) IC_{50} Activity of Pyridinone Derivatives (**37**) against HIV-1 Reverse Transcriptase (Table 29).*[94]



37

Wai et al.[94] reported data on benzylamino pyridinone derivatives that have various substituents at the 2-position of the phenyl ring. Our analysis gave eq 40, emphasizing the role of hydrophobicity in anti-HIV RT inhibition activity of pyridinone derivatives. The X-position substituents also show steric effects on the activity.

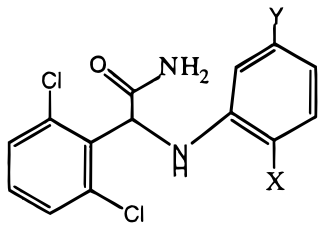
$$\log 1/C = 5.51(\pm 1.80)\text{Clog } P - 1.19(\pm 0.38)(\text{Clog } P)^2 + 0.60(\pm 0.52)MR_X - 0.51(\pm 2.01)$$

$$n = 12,\ r^2 = 0.877,\ s = 0.26,\ q^2 = 0.681 \quad (40)$$

$$\log P_0 = 2.31\ (2.19-2.43),\ \text{Outliers } 2$$

Garg and Gupta [95] also reported QSAR studies on 2-pyridinone[91−93] derivatives.

f. α-APA Derivatives. *(i) IC_{50} and CC_{50} Activities of α-Anilinophenylacetamide Derivatives (38) in MT-4 Cells (Table 30).*[96] Pauwels et al.[96] reported anti-HIV



38

data on a small series of α-anilinophenylacetamides (**38**). We obtained the correlations for anti-HIV (inhibitory concentration of the compound required to achieve 50% protection of MT-4 cells against cytopathicity of HIV) and cytotoxic (compound dose required to reduce the viability of mock-infected MT-4 cells by 50%) activities as shown by eqs 41 and 42, respectively.

$$\log 1/C = -2.16(\pm 1.78)\sigma_{X,Y} - 2.47(\pm 0.58)I + 8.90(\pm 1.05)$$

$n = 8$, $r^2 = 0.96$, $s = 0.26$, $q^2 = 0.889$ (41)

Outliers 2

$$\log 1/C = -0.68(\pm 0.40)\sigma_{X,Y} - 1.22\ (\pm 0.29)\text{CMR} + 14.99(\pm 2.61)$$

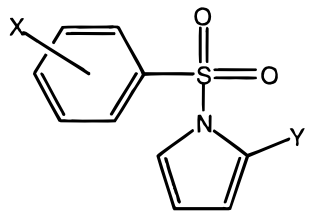
$n = 9$, $r^2 = 0.951$, $s = 0.13$, $q^2 = 0.885$ (42)

Outliers 1

In eq 41, I is an indicator variable used with a value of unity for dextrorotatory isomers. Its negative coefficient indicates that such isomers would be less effective than anti-HIV agents. However, a comparison between eqs 41 and 42 shows that while steric bulk will inhibit cytotoxicity, an electron-releasing substituent will increase the anti-HIV and cytotoxic effect. The absence of a hydrophobic term indicates that these substituents are not able to reach the hydrophobic pocket of the receptor.

g. Miscellaneous. In addition to the above-discussed important classes of RT inhibitors, some miscellaneous kinds of RT inhibitors have also been reported.

*(i) CC_{50} Activity of Pyrrole Derivatives (**39**) in MT-4 Cells (Table 31).*[97] A series of pyrrole derivatives was



39

reported by Artico et al.[97] from which we derived eq 43.

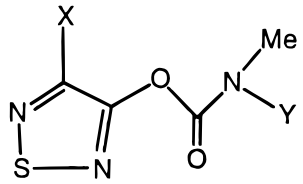
$$\log 1/C = 0.99(\pm 0.19)\sigma_{X,SUM} - 0.45(\pm 0.37)\text{MgVol} + 4.95(\pm 0.80)$$

$n = 28$, $r^2 = 0.830$, $s = 0.31$, $q^2 = 0.789$ (43)

Outliers 4

This equation suggests that by decreasing the size of the molecule and substituting the phenyl ring with electron-attracting substituents, the toxicity of the compound can be decreased. It is surprising that no electronic effect is seen for the nitrogen ring. It is not clear to us why we do not see any hydrophobic term here.

*(ii) EC_{50} and CC_{50} Activities of Thiadiazole Derivatives (**40**) in MT-4 Cells (Table 32).*[98] Hanasaki et al.[98]



40

reported a series of thiadiazole derivatives. For this series of compounds, we found both the anti-HIV activity (EC_{50}) and the cytotoxic effect (CC_{50}) to be significantly correlated with the hydrophobicity of molecule (eqs 44 and 45, respectively).

$$\log 1/C = 5.20(\pm 2.62)\text{Clog } P - 0.54(\pm 0.30)(\text{Clog } P)^2 - 4.79(\pm 5.37)$$

$n = 7$, $r^2 = 0.921$, $s = 0.37$, $q^2 = 0.835$ (44)

$\log P_0 = 4.83$ (4.50−5.64)

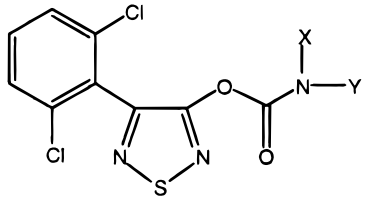
$$\log 1/C = 0.27(\pm 0.07)\text{Clog } P + 3.04(\pm 0.25)$$

$n = 8$, $r^2 = 0.939$, $s = 0.14$, $q^2 = 0.892$ (45)

Outlier 1

These correlations suggest that both the activities can increase with an increase in hydrophobic character of the molecule. However, while there would be an optimum value of Clog $P = 4.83$ for anti-HIV activity, no such optimization is exhibited for the cytotoxic effect (eq 45). This is normal for nonspecific toxicity.

*(iii) EC_{50} Activitiy of Thiadiazole Derivatives (**41**) in MT-4 Cells (Table 33).*[99] For another series of



41

dialkyl carbamate thiadiazole derivatives reported by Ijichi et al., [99] where only the side carbamate chain was substituted with alkyl groups at the nitrogen, we could correlate the anti-HIV activity with only Taft's electronic parameter σ^*, which suggested that these side chain substituents may be involved in some electronic interactions with the receptor. No hydrophobic effect is seen here, and too few data points limit the use of more parameters.

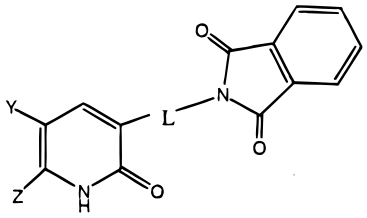
$$\log 1/C = 8.50(\pm 4.06)\sigma^*_{X,Y} + 8.40(\pm 0.70)$$

$n = 5$, $r^2 = 0.937$, $s = 0.13$, $q^2 = 0.800$ (46)

Outliers 2

(iv) IC_{50} Activity of Phosphonomethoxy Ethylgua-

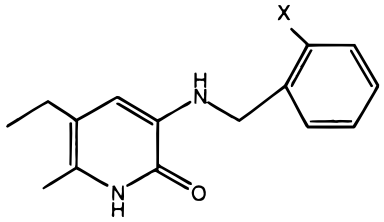
Table 28. IC_{50} Activity of Pyridinone Derivatives (36)[91]



no.	substituents Y	Z	L	log 1/C obsd	calcd (eq 39)	Δ	MR_Y	L_L
1	H	H	$NHCH_2$	5.78	5.75	0.03	0.10	3.53
2	H	Me	$NHCH_2$	5.09	5.75	−0.66	0.10	3.53
3	Me	Me	$NHCH_2$	6.66	6.18	0.48	0.57	3.53
4	C_2H_5	Me	$NHCH_2$	7.52	6.60	0.92	1.03	3.53
5	C_3H_7	Me	$NHCH_2$	7.24	7.03	0.21	1.50	3.53
6	C_4H_9	Me	$NHCH_2$	6.89	7.45	−0.56	1.96	3.53
7	Me	C_2H_5	$NHCH_2$	6.05	6.18	−0.13	0.57	3.53
8	C_2H_5	Me	$NH(CH_2)_2$	4.35	3.90	0.45	1.03	4.83
9	C_2H_5	Me	$NH(CH_2)_3$[a]	3.72	1.32	2.40	1.03	6.07
10	C_2H_5	Me	CH_2[a]	4.58	8.00	−3.40	1.03	2.87
11	C_2H_5	Me	$(CH_2)_2$	5.43	5.40	0.04	1.03	4.11
12	C_2H_5	Me	$CH_2)_3$	3.68	3.71	−0.03	1.03	4.92
13	C_2H_5	Me	CH=CH(*trans*)	4.27	5.02	−0.75	1.03	4.29

[a] Data points not used in deriving equation.

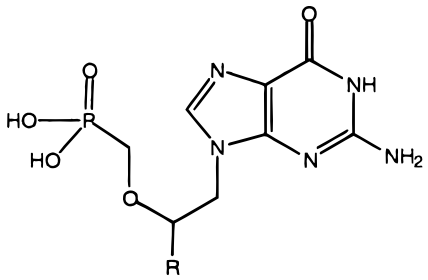
Table 29. IC_{50} Activity of Pyridinone Derivatives (37)[94]



no.	X	log 1/C obsd	calcd (eq 40)	Δ	Clog *P*	MR_X
1	2-OMe	6.58	6.33	0.25	2.31	0.79
2	H[a]	5.28	5.91	−0.63	2.39	0.10
3	$2\text{-}OC_2H_5$	6.46	6.27	0.19	2.84	1.25
4	$2\text{-}NO_2$	6.28	6.22	0.06	2.06	0.74
5	2-CN	6.10	6.09	0.01	1.97	0.63
6	2-F	5.85	5.85	0.01	2.54	0.09
7	2-Cl	5.84	5.46	0.38	3.11	0.60
8	2-SMe	5.82	6.20	−0.37	2.95	1.38
9	2-Me	5.66	5.86	−0.20	2.84	0.57
10	2-OH	5.34	5.62	−0.28	1.73	0.29
11	$2\text{-}C_2H_5$	5.16	5.13	0.03	3.37	1.03
12	$2\text{-}CF_3$	4.90	5.05	−0.14	3.28	0.50
13	$2\text{-}NH_2$	4.55	4.48	0.08	1.12	0.54
14	$2\text{-}CH_2OMe$[a]	4.45	6.54	−2.09	2.12	1.21

[a] Data points not used in deriving equation.

*nines (**42**) in CEM Cells (Table 34).*[100] Yu et al.[100] reported data for a small series of 2′-substituted phosphonomethoxy ethyl guanines (**42**). We derived eq 47 for the data.



42

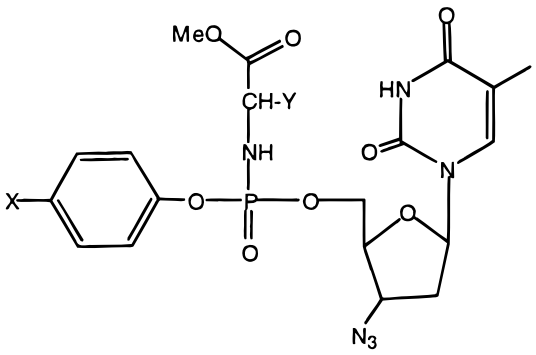
$$\log 1/C = -1.58(\pm 0.46)\mathrm{L} + 0.75(\pm 0.85)\mathrm{Clog}\,P + 12.33(\pm 3.34)$$

$$n = 10,\ r^2 = 0.910,\ s = 0.36,\ q^2 = 0.840 \quad (47)$$

Outliers 2

Equation 47 shows that the side chain substituent R produces a negative steric effect; the very weak Clog *P* term (note confidence limits) indicates that some possible hydrophobic interactions are also involved. The R-substituents were mostly substituted or unsubstituted alkyl groups.

*(v) EC_{50} Activity of Phosphonomethoxy Azidothymidines (**43a**, **43b**) in MT-4 cells (Table 35).*[101a] McGuigan et al.[101a] reported anti-HIV-1 data for a series of phosphonomethoxy azidothymidines (**43a**).



43a

We developed eq 48a.

$$\log 1/C = -1.92(\pm 0.82)\mathrm{B1}_X + 1.21(\pm 0.43)\mathrm{I}_Y + 8.89(\pm 0.85)$$

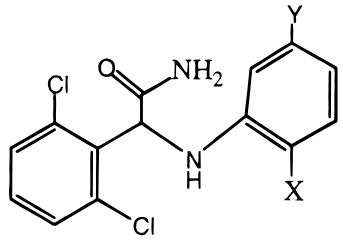
$$n = 9,\ r^2 = 0.890,\ s = 0.16,\ q^2 = 0.644 \quad (48a)$$

Outlier 1

This equation shows the strong steric effect of X-substituents. Indicator variable I_Y was used with a value of unity for Y = Me; its positive coefficient shows that this group is beneficial to the activity. However, we did not find any hydrophobic term.

Recently, Siddiqui et al.[101b] reported anti-HIV data for another series of aryl phosphoramidate (**43b**)

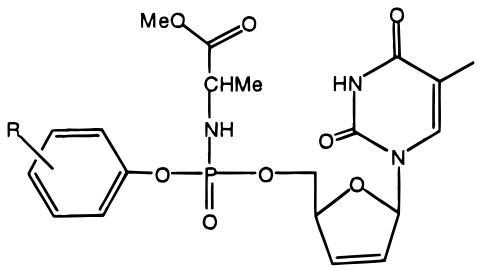
Table 30. IC_{50} and CC_{50} Activities of α-Anilinophenylacetamide Derivatives (38)[96]



	substituents		log 1/*C*								
no.	X	Y	obsd (eq 41)	calcd (eq 41)	Δ	obsd (eq 42)	calcd (eq 42)	Δ	CMR	*I*	$\sigma_{X,Y}$
1	2-OMe	H	6.22[a]	9.49	−3.27	4.82	2.07	2.76	8.50	0.00	−0.27
2	2-NO_2	H	7.06	7.22	−0.16	4.10	4.12	−0.02	8.50	0.00	0.78
3	2-NO_2(+)	H	5.77[a]	4.75	1.02	4.09	4.12	−0.04	8.50	1.00	0.78
4	2-NO_2(−)	H	7.48	7.22	0.26	4.18	4.12	0.06	8.50	0.00	0.78
5	2-COMe	H	7.59	7.82	−0.24	3.89	3.90	−0.02	8.85	0.00	0.50
6	2-COMe(+)	H	5.19	5.36	−0.17	3.92	3.90	0.02	8.85	1.00	0.50
7	2-COMe(−)	H	7.72	7.82	−0.10	4.28[a]	3.90	0.38	8.85	0.00	0.50
8	2-COMe	5-Me	7.89	7.97	−0.09	3.15	3.37	−0.22	9.31	0.00	0.43
9	2-COMe(+)	5-Me	5.68	5.51	0.17	3.57	3.37	0.20	9.31	1.00	0.43
10	2-COMe(−)	5-Me	8.30	7.97	0.33	3.38	3.37	0.01	9.31	0.00	0.43

[a] Data points not used in deriving respective equations.

derivatives of anti-HIV drug d4T. Correlation analy-



43b

sis yielded eqs 48b−e that show hydrophobic interactions between the receptor and ligand. The coefficient with Clog *P* is in the normal range. Siddiqui et al. have suggested that enhancing lipophilicity may serve to increase the cellular uptake of the prodrug by passive diffusion, leading to the expression of antiviral potency at reduced prodrug concentrations.

EC_{50} activity against HIV-1 in CEM/0 cells:

$$\log 1/C = 1.16(\pm 0.40)\text{Clog } P + 5.85(\pm 0.52)$$
$$n = 8,\ r^2 = 0.896,\ s = 0.16,\ q^2 = 0.763 \quad \text{(48b)}$$
Outliers 2

EC_{50} activity against HIV-2 in CEM/0 cells:

$$\log 1/C = 1.03(\pm 0.40)\text{Clog } P + 6.00(\pm 0.51)$$
$$n = 8,\ r^2 = 0.871,\ s = 0.16,\ q^2 = 0.774 \quad \text{(48c)}$$
Outliers 2

EC_{50} activity against HIV-1 in CEM/TK cells:

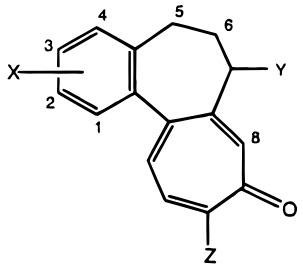
$$\log 1/C = 1.26(\pm 0.41)\text{Clog } P + 5.90\ (\pm 0.58)$$
$$n = 8,\ r^2 = 0.883,\ s = 0.14,\ q^2 = 0.801 \quad \text{(48d)}$$
Outliers 2

CC_{50} activity against HIV-2 in CEM/0 cells:

$$\log 1/C = 0.71(\pm 0.30)\text{Clog } P + 0.17(\pm 0.13)L + 2.28(\pm 0.69)$$
$$n = 8,\ r^2 = 0.893,\ s = 0.11,\ q^2 = 0.700 \quad \text{(48e)}$$
Outliers 2

Siddique et al.[101b] also correlated anti-HIV data with hydrophobicity using their experimentally determined log *P* values. Clog *P* (calculated log *P*) used by us correlated well with these log *P* values (r^2 = 0.861).

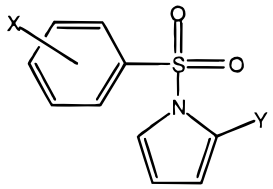
*(vi) IC_{50} and EC_{50} Activities of Cholchicines (**44**) against HIV Replication in H9 Lymphocytes (Table 36).*[102] For a series of cholchicines, anti-HIV data



44

were reported by Tatematsue et al.;[102] we observed

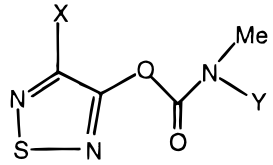
Table 31. CC_{50} Activity of Pyrrole Derivatives (39)[97]



	substituents		log 1/*C*				
no.	X	Y	obsd	calcd (eq 43)	Δ	$\sigma_{X,SUM}$	MgVol
1	2-NO_2	H[a]	4.44	4.99	−0.55	0.78	1.64
2	2-NH_2	H	3.64	3.60	0.05	−0.66	1.57
3	2-Cl	$CO_2C_2H_5$	3.85	4.25	−0.40	0.23	2.09
4	2-F	$CO_2C_2H_5$	3.88	4.12	−0.25	0.06	1.98
5	2-F	CONH-Cy-C_6H_{11}	3.52	3.90	−0.38	0.06	2.48
6	2-F	CONH-Cy-C_3H_5	3.52	4.09	−0.57	0.06	2.05
7	2-F	CONHNHCHO[a]	3.52	4.12	−0.59	0.06	2.00
8	2-NO_2	COMe[a]	4.00	4.86	−0.86	0.78	1.94
9	2-NO_2,4-Cl	H	4.82	5.17	−0.34	1.01	1.76
10	2-NH_2-4-Cl	H	4.13	3.77	0.36	−0.43	1.69
11	2-NH_2-4-Cl	CO_2Me	3.61	3.61	0.00	−0.43	2.04
12	2-NH_2-4-Cl	$CO_2C_2H_5$	3.62	3.55	0.08	−0.43	2.19
13	2-NO_2-4,5-di-Cl	$CO_2C_2H_5$	5.30	5.26	0.04	1.38	2.38
14	2-NO_2-5-Cl	H	5.70	5.31	0.39	1.15	1.76
15	2-NH_2-5-Cl	H	3.78	3.91	−0.13	−0.29	1.69
16	2-NO_2-5-Cl	CO_2Me	4.92	5.15	−0.23	1.15	2.12
17	2-NO_2-5-Cl	$CO_2C_2H_5$	4.92	5.08	−0.16	1.15	2.26
18	2-NO_2-5-Cl	$CO_2C_3H_7$	5.00	5.02	−0.02	1.15	2.40
19	2-NH_2-5-Cl	$CO_2C_3H_7$	3.96	3.62	0.34	−0.29	2.33
20	2-NO_2-5-Cl	$CO_2CH_2CH{=}CH_2$	5.37	5.04	0.33	1.15	2.36
21	2-NH_2-5-Cl	$CO_2CH_2CH{=}CH_2$	4.00	3.64	0.36	−0.29	2.28
22	2-NO_2-5-Cl	$CO_2CH_2C_6H_5$[a]	3.52	4.87	−1.35	1.15	2.73
23	2-NO_2-5-Cl	$CO_2CH_2CH_2N(C_2H_5)_2$	4.52	4.79	−0.26	1.15	2.92
24	2-NH_2-5-Cl	$CO_2CH_2CH_2N(C_2H_5)_2$	3.79	3.39	0.40	−0.29	2.85
25	2-NHMe-5-Cl	$CO_2C_2H_5$	3.52	3.58	−0.06	−0.33	2.33
26	2-NHC_2H_5-5-Cl	$CO_2C_2H_5$	3.52	3.61	−0.09	−0.24	2.47
27	2-Cl-5-NO_2	H	5.70	5.10	0.60	0.94	1.76
28	2-Cl-5-NH_2	H	4.04	4.27	−0.23	0.07	1.69
29	2-Cl-5-NO_2	CO_2Me	4.92	4.94	−0.02	0.94	2.12
30	2-Cl-5-NH_2	CO_2Me	3.82	4.11	−0.28	0.07	2.04
31	2-Cl-5-NO_2	$CO_2C_2H_5$	5.22	4.87	0.35	0.94	2.26
32	2-Cl-5-NH_2	$CO_2C_2H_5$	4.14	4.04	0.10	0.07	2.19

[a] Data points not used in deriving equation.

Table 32. EC_{50} and CC_{50} Activities of Thiadiazole Derivatives (40)[98]

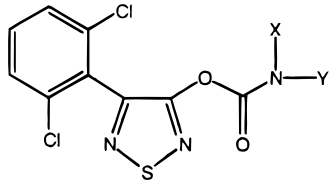


	substituents		log 1/*C*						
no.	X	Y	obsd (eq 44)	calcd (eq 44)	Δ (eq 44)	obsd (eq 45)	calcd (eq 45)	Δ (eq 45)	Clog *P*
1	H	Me				3.25	3.16	0.10	0.44
2	Me	Me				3.27	3.22	0.05	0.69
3	C_6H_5	Me	4.64	4.94	−0.31	3.74	3.72	0.02	2.54
4	2-Cl-C_6H_4	Me	6.50	5.98	0.51	3.80	3.84	−0.05	3.01
5	C_6H_4-2,6-di-Cl	Me	6.62	6.78	−0.16	3.84	3.97	−0.13	3.47
6	C_6H_4-2,6-di-Cl	C_2H_5	7.41	7.40	0.01	3.87	4.11	−0.24	4.00
7	C_6H_4-2,6-di-Cl	C_3H_7	7.89	7.72	0.16	3.88[a]	4.25	−0.37	4.53
8	C_6H_4-2,6-di-Cl	C_4H_9	7.41	7.75	−0.34	4.54	4.39	0.14	5.06
9	C_6H_4-2,6-di-Cl	C_6H_{13}	7.00	6.88	0.12	4.80	4.68	0.12	6.12

[a] Data point not used in deriving eq 45.

that the ring substituents have a dominant electronic effect (eqs 49 and 50). Equation 49 is for the potency of the compounds to inhibit the HIV-1 replication in H9 cells (IC_{50}) (Table 36a), and eq 50 is for their

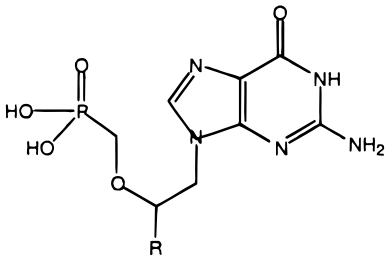
Table 33. EC_{50} Activity of Thiadiazoles (41)[99]



	substituents		log 1/*C*			
no.	X	Y	obsd	calcd (eq 46)	Δ	$\sigma^*_{X,Y}$
1	Me	Me[a]	6.69	8.40	−1.71	0.00
2	Me	C_2H_5	7.44	7.55	−0.11	−0.10
3	Me	C_3H_7[a]	7.90	7.38	0.53	−0.12
4	Me	C_4H_9	7.47	7.29	0.18	−0.13
5	Me	C_6H_{13}	6.94	6.95	−0.01	−0.17
6	C_2H_5	C_2H_5	6.61	6.70	−0.09	−0.20
7	C_2H_5	C_4H_9	6.48	6.44	0.03	−0.23

[a] Data points not used in deriving equation.

Table 34. EC_{50} Activity of Phosphonomethoxy Ethylguanines (42)[100]



		log 1/*C*				
no.	R	obsd	calcd (eq 47)	Δ	*L*	Clog *P*
1	H	6.70	6.84	−0.14	2.06	−2.96
2	Me(R)	6.00	5.79	0.21	2.87	−2.65
3	Me(S)[a]	4.92	5.79	−0.87	2.87	−2.65
4	CH_2OH(R)	3.30	3.54	−0.24	3.97	−3.33
5	CH_2F(R)	5.10	5.02	0.08	3.30	−2.78
6	CH_2F(S)	5.16	5.02	0.14	3.30	−2.78
7	CH_2Cl(R)	4.35	4.34	0.01	3.89	−2.44
8	CH_2Cl(S)	4.30	4.34	−0.04	3.89	−2.44
9	CH=CH_2(R)[a]	4.89	3.73	1.16	4.29	−2.41
10	CH=CH_2(S)	4.31	3.73	0.58	4.29	−2.41
11	C_2H_5(R)	4.27	4.23	0.04	4.11	−2.12
12	C_2H_5(S)	3.60	4.23	−0.63	4.11	−2.12

[a] Data points not used in deriving equation.

ability to protect the cells from the infection of HIV-1 replication (EC_{50}) (Table 36b).

$$\log 1/C = 4.11(\pm 1.68)\sigma_{I,Y} + 1.01(\pm 0.60)MR_X + 3.81(\pm 1.39)$$

$$n = 11,\ r^2 = 0.828,\ s = 0.38,\ q^2 = 0.757 \quad (49)$$

Outliers 3

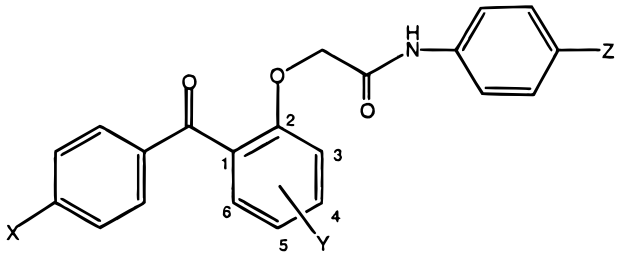
$$\log 1/C = 3.89(\pm 1.79)\sigma_{I,Y} + 1.02(\pm 0.65)MR_X + 3.85(\pm 1.48)$$

$$n = 11,\ r^2 = 0.797,\ s = 0.40,\ q^2 = 0.718 \quad (50)$$

Outliers 3

A positive high coefficient of MR might be indicative of hydrophobic rather than steric effects of substituents.

*(vii) IC_{50} Activity of Benzophenones (**45**) against HIV-1 Reverse Transcriptase (Table 37).*[103] We de-



45

rived eq 51 for the data reported by Wyatt et al.[103] for a small series of benzophenones.

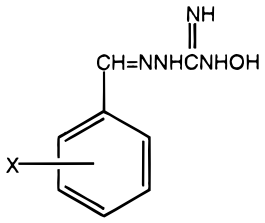
$$\log 1/C = 3.23(\pm 2.14)\sigma_{Y,5} - 0.52(\pm 0.24)CMR + 14.82(\pm 3.27)$$

$$n = 7,\ r^2 = 0.975,\ s = 0.27,\ q^2 = 0.940 \quad (51)$$

Outlier 1

The dominance of the electronic effect of Y-substituents was observed. However, why lowering the electron density on the ring is good is not obvious. The overall bulk of the molecule inhibits activity. Clog *P* and CMR are highly collinear ($r^2 = 0.73$). We believe there is a steric effect as a negative CMR term is observed.

*(viii) TC_{50} Activity of N-Hydroxy-N^1-Aminoguanidines (**46**) in Human T-Lympocytes Cells against Infection by HIV-1 (Table 38).*[104] We developed eq 52



46

for the TC_{50} data (concentration of a compound causing a 50% reduction in cell growth) reported by Doubell and Oliver[104] that shows that the hydrophobicity of the molecule and electron-donating X-substituent can significantly affect the cytotoxicity of the compounds. Also, para X-substituents have a steric effect on the activity.

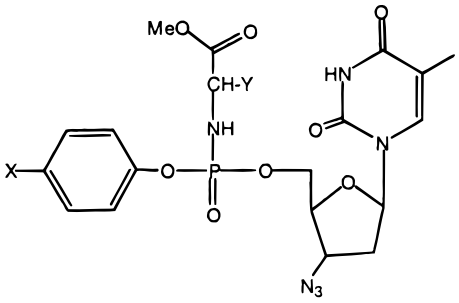
$$\log 1/C = 0.37(\pm 0.21)\text{Clog } P + 0.67(\pm 0.27)\sigma_X + 0.49(\pm 0.31)MR_{X,4} + 6.76(\pm 0.23)$$

$$n = 11,\ r^2 = 0.895,\ s = 0.12,\ q^2 = 0.760 \quad (52)$$

Outliers 2

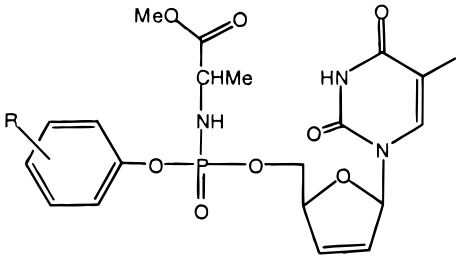
Table 35. EC_{50} Activity of Phosphonomethoxyazidothymidine (43a)[101a]

(a) Phosphonomethoxyazidothymidine



	substituents		log 1/*C*				
no.	X	Y	obsd	calcd (eq 48a)	Δ	$B1_X$	I_Y
1	H	Me	8.22	8.18	0.04	1.00	1.00
2	Me	Me	7.29	7.18	0.11	1.52	1.00
3	C_2H_5	Me	7.17	7.18	−0.02	1.52	1.00
4	C_3H_7	Me	7.37	7.18	0.18	1.52	1.00
5	C_5H_{11}	Me	6.99	7.18	−0.20	1.52	1.00
6	MeO	Me[a]	7.15	7.51	−0.36	1.35	1.00
7	F	Me	7.39	7.51	−0.12	1.35	1.00
8	H	H	6.80	6.97	−0.17	1.00	0.00
9	H	CH_2CHMe_2	7.00	6.97	0.04	1.00	0.00
10	H	$CH_2C_6H_5$	7.10	6.97	0.13	1.00	0.00

(b) 2′,3′-Didehydrophosphonomethoxythymidines

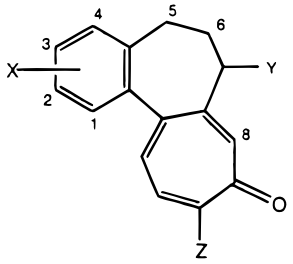


		log 1/*C*						
no.	X	obsd (eq 48b)	calcd (eq 48b)	Δ (eq 48b)	obsd (eq 48c)	calcd (eq 48c)	Δ (eq 48c)	Clog *P*
1	4-NO_2[a]	6.89	7.25	−0.37	6.72	7.25	−0.53	1.21
2	4-CN	6.82	6.92	−0.10	6.89	6.96	−0.07	0.92
3	4-CO_2Me	7.60	7.46	0.14	7.66	7.44	0.22	1.39
4	3-COMe	6.96	6.90	0.06	7.00	6.94	0.06	0.90
5	4-Cl	8.30	8.20	0.11	8.16	8.09	0.07	2.02
6	4-F	7.28	7.53	−0.26	7.22	7.50	−0.28	1.45
7	4-Me[a]	7.40	7.78	−0.39	7.30	7.72	−0.42	1.66
8	4-Me	7.24	7.31	−0.06	7.28	7.30	−0.02	1.25
9	4-COMe	7.10	6.90	0.20	7.05	6.94	−0.11	0.90
10	H	7.13	7.21	−0.08	7.13	7.20	−0.08	1.16

		log 1/*C*							
no.	X	obsd (eq 48d)	calcd (eq 48d)	Δ (eq 48d)	obsd (eq 48e)	calcd (eq 48e)	Δ (eq 48e)	Clog *P*	*L*
1	4-NO_2[a]	7.05	7.26	−0.21	4.14	4.28	−0.14	1.21	3.44
2	4-CN	7.05	6.94	0.11	4.22	4.21	0.01	0.92	4.23
3	4-CO_2Me	7.60	7.46	0.14	4.51	4.63	−0.12	1.39	4.73
4	3-COMe	7.30	6.92	0.39	4.40	3.82	0.58	0.90	2.06
5	4-Cl	8.22	8.17	0.06	4.96	4.86	0.10	2.02	3.52
6	4-F	7.66	7.53	0.13	4.59	4.31	0.28	1.45	2.65
7	4-Me[a]	7.60	7.77	−0.17	4.47	4.50	−0.03	1.66	2.87
8	4-Me	7.32	7.31	0.02	4.46	4.40	0.05	1.25	3.92
9	4-COMe	7.35	6.92	0.43	4.30	4.17	0.13	0.90	4.06
10	H	7.13	7.21	−0.08	4.00	4.01	−0.01	1.16	2.06

[a] Data points not used in deriving equations.

Table 36. IC_{50} Activity of Cholchicine Derivatives (44)[102]

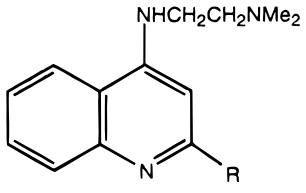


	substituents			log 1/*C*				
no.	X	Y	Z	obsd	calcd (eq 49)	Δ	$\sigma_{I,Y}$	MR_X
1	1,2,3-$(OMe)_3$	NHCOMe	OMe	7.75	7.41	0.34	0.30	2.36
2	1,2,3-$(OMe)_3$	NH_2	OMe	6.70	6.17	0.53	0.00	2.36
3	1,3-$(OMe)_2$,2-OH	NHCOMe	OMe	6.59	6.90	−0.32	0.30	1.86
4	1,3-$(OMe)_2$,2-OH	NHCOMe	OH	6.57	6.90	−0.33	0.30	1.86
5	1,2,3-$(OMe)_3$	NH_2	OH	5.92	6.17	−0.25	0.00	2.36
6	1,2,3-$(OH)_3$	NHCOMe	OH	5.92	5.89	0.03	0.30	0.86
7	1,2,3-$(OMe)_3$	$NHCOCF_3$	OMe	8.18	8.02	0.16	0.45	2.36
8	1,2,3-$(OMe)_3$	$NHCOCF_3$	OH[a]	6.70	8.02	−1.32	0.45	2.36
9	1,2,3-$(OMe)_3$	NHCOMe	SMe[a]	8.77	7.41	1.36	0.30	2.36
10	1,2,3-$(OMe)_3$	NH_2	SMe[a]	7.89	6.17	1.71	0.00	2.36
11	1,3-$(OMe)_2$,2-OH	NHCOMe	SMe	7.28	6.90	0.38	0.30	1.86
12	1,3-$(OMe)_2$,2-OH	NH_2	SMe	5.85	5.67	0.19	0.00	1.86
13	1,2,3-$(OMe)_3$	NHCOMe	NMe_2	7.07	7.41	−0.34	0.30	2.36
14	1,2,3-$(OMe)_3$	NH_2	NMe_2	5.80	6.17	−0.38	0.00	2.36

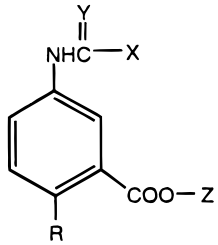
	substituents			log 1/*C*				
no.	X	Y	Z	obsd	calcd (eq 50)	Δ	$\sigma_{I,Y}$	MR_X
1	1,2,3-$(OMe)_3$	NHCOMe	OMe	8.00	7.44	0.56	0.30	2.36
2	1,2,3-$(OMe)_3$	NH_2	OMe	6.70	6.27	0.43	0.00	2.36
3	1,3-$(OMe)_2$,2-OH	NHCOMe	OMe	6.59	6.92	−0.34	0.30	1.86
4	1,3-$(OMe)_2$,2-OH	NHCOMe	OH	6.57	6.92	−0.35	0.30	1.86
5	1,2,3-$(OMe)_3$	NH_2	OH	5.89	6.27	−0.38	0 .00	2.36
6	1,2,3-$(OH)_3$	NHCOMe	OH	5.92	5.90	0.03	0.30	0.86
7	1,2,3-$(OMe)_3$	$NHCOCF_3$	OMe	8.06	8.02	0.04	0.45	2.36
8	1,2,3-$(OMe)_3$	$NHCOCF_3$[a]	OH	6.24	8.02	−1.78	0.45	2.36
9	1,2,3-$(OMe)_3$	NHCOMe[a]	SMe	8.75	7.44	1.31	0.30	2.36
10	1,2,3-$(OMe)_3$	NH_2[a]	SMe	7.89	6.27	1.62	0.00	2.36
11	1,3-$(OMe)_2$,2-OH	NHCOMe	SMe	7.33	6.92	0.41	0.30	1.86
12	1,3-$(OMe)_2$,2-OH	NH_2	SMe	5.96	5.75	0.21	0.00	1.86
13	1,2,3-$(OMe)_3$	NHCOMe	NMe_2	7.07	7.43	−0.37	0.30	2.36
14	1,2,3-$(OMe)_3$	NH_2	NMe_2	6.03	6.27	−0.24	0.00	2.36

[a] Data points not used in deriving equations.

*(ix) EC_{50} Activity of Quinoline Derivatives (**48**) in Human Lymphocyte CEM Cells (Table 39).*[105,106] For a small series of 2-(aryl or heteroaryl)-quinolin-4-amines (**47**), Strekowski et al.[105] reported eq 53 for the anti-HIV data of the compounds, which shows that the anti-HIV activity depends on the electronic character of the ring substituents. We were unable to reproduce this equation or any suitable equation.



47



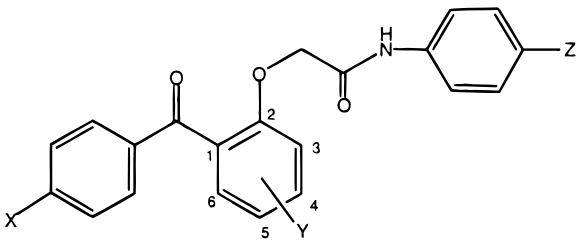
48

$$\log 1/C = -2.83(\pm 0.62)\sigma_I + 1.06$$
$$n = 10,\ r^2 = 0.721,\ s = 0.42 \qquad (53)$$

However, we derived eq 54 for the EC_{50} data (compound concentration reducing HIV-1 induced giant cell formation in CEM cells by 50%) reported by Balzarini et al.[106] on another series of quinolines (**48**).

$$\log 1/C = 1.36(\pm 0.34)\text{Clog } P - 7.39(\pm 1.81)\log(\beta \times 10^{\text{Clog } P} + 1) - 0.54(\pm 0.49)\sigma^*_Z - 2.94(\pm 1.81)\sigma^*_X + 7.62(\pm 3.43)$$
$$n = 27,\ r^2 = 0.836,\ s = 0.27,\ q^2 = 0.730 \qquad (54)$$
$$\log P_0 = 5.67(\pm 0.26), \beta = -6.33, \text{Outliers } 3$$

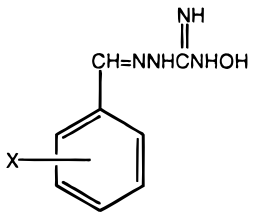
Table 37. IC_{50} Activity of Benzophenone Derivatives (45)[103]



no.	substituents X	Y	Z	log 1/C obsd	log 1/C calcd (eq 51)	Δ	CMR	$\sigma_{Y,5}$
1	OMe	H	$OCH_2CH_2N(C_2H_5)_2$	7.85	7.77	0.08	13.62	0.00
2	F	H	$OCH_2CH_2N(C_2H_5)_2$	7.67	8.08	−0.41	13.02	0.00
3	H	H	$OCH_2CH_2N(C_2H_5)_2$	8.29	8.09	0.20	13.00	0.00
4	H	4-OMe	$OCH_2CH_2N(C_2H_5)_2$	7.88	7.77	0.11	13.62	0.00
5	H	5-Cl	$OCH_2CH_2N(C_2H_5)_2$	9.03	9.02	0.01	13.49	0.37
6	H	5-Cl	$OCH_2CH_2NMe_2$[a]	9.68	9.26	0.42	13.03	0.37
7	H	5-Cl	H	10.56	10.74	−0.17	10.19	0.37
8	H	5-F	H	11.07	10.88	0.18	9.72	0.34

[a] Data point not used in deriving equation.

Table 38. TC_{50} Activity of *N*-Hydroxy-N^1-aminoguanidines (46)[104]



no.	X	log 1/C obsd	log 1/C calcd (eq 52)	Δ	Clog *P*	σ_X	$MR_{X,4}$
1	H	6.71	6.62	0.09	−0.51	0.00	0.10
2	3-F	6.93	6.90	0.03	−0.36	0.34	0.10
3	3-F,4-OMe	7.05	7.00	0.05	−0.50	0.07	0.79
4	3,4-$(OMe)_2$	6.59	6.73	−0.14	−0.85	−0.15	0.79
5	3,4-$(OMe)_2$,6-NO_2	7.20	7.26	−0.06	−0.83	0.63	0.79
6	3,4-$(OMe)_2$,6-Br	7.30	7.23	0.07	0.08	0.08	0.79
7	3,4-$(OH)_2$[a]	6.69	6.07	0.62	−1.77	−0.25	0.29
8	3-OH,4-OMe	6.71	6.55	0.16	−1.32	−0.15	0.79
9	3-OMe,4-OH	6.23	6.23	−0.00	−1.32	−0.25	0.29
10	3-OMe,4-OH,6-Cl	6.57	6.73	−0.16	−0.40	−0.02	0.29
11	2-Cl,3-OH,4-OMe[a]	6.57	6.93	−0.35	−0.73	0.08	0.79
12	2,4-$(OMe)_2$	6.49	6.59	−0.11	−0.50	−0.54	0.79
13	2,4-$(OMe)_2$,3-Me	6.80	6.73	0.07	0.00	−0.61	0.79

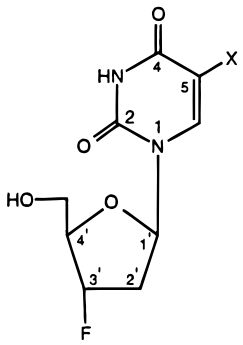
[a] Data points not used in deriving equation.

Equation 54 shows that not only the hydrophobicity of the molecule, but also the electronic character of small substituents can influence the activity. The overall hydrophobic character of the molecule also appears to play a major role. This is not a very good correlation but it does bring out the importance of log *P*. The steep negative slope of the Clog *P* term (1.36−7.02 = −5.66) suggests a steric effect for large substituents.

2. Nucleoside RT Inhibitors

The 2′,3′-dideoxypurine and pyrimidine nucleoside analogues have been well studied for their RT inhibition and anti-HIV activities, but QSARs on them are scanty. However, when we attempted QSARs on them, the activities were found to be well correlated with the physicochemical properties of the substituents.

*(i) ED_{50} Activity of Uridine Derivatives (**49**) in MT-4 Cells (Table 40).*[107] Balzarini et al.[107] reported the



49

ED_{50} data (effective dose of the compound required for 50% inhibition of HIV-1 and HIV-2 induced cytopathicity in MT-4 cells) on 5-halogen-3′-fluoro-2′,3′-dideoxyuridines (FddUrd). We derived eqs 55 and 56, respectively for them.

Against HIV-2:

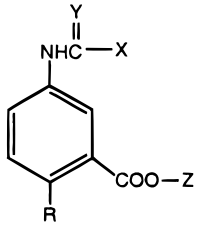
$$\log 1/C = -3.26(\pm 1.95)\sigma^+ + 6.80(\pm 0.34)$$
$$n = 5,\ r^2 = 0.904,\ s = 0.24,\ q^2 = 0.784 \qquad (55)$$

Against HIV-1:

$$\log 1/C = -5.51(\pm 1.87)\sigma^+ + 7.26(\pm 0.32)$$
$$n = 5,\ r^2 = 0.967,\ s = 0.23,\ q^2 = 0.907 \qquad (56)$$

Both equations were found to be significantly correlated with the electronic factor of the X-substituents. The negative coefficient with σ^+ refers to the electron-donating ability of the substituent via resonance. This might indicate that a high electron density on N-1 is desirable. It is a very small set so other parameters could not be included.

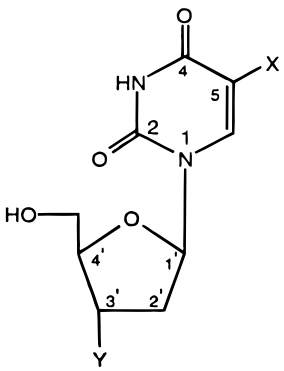
Table 39. EC_{50} Activity of Quinolines (48)[106]



	substituents				log 1/*C*					
no.	X	Y	R	Z	obsd	calcd (eq 54)	Δ	Clog *P*	σ_Z*	σ^*_X
1	$OCHMe_2$	S	4-Cl	$CHMe_2$	10.54	10.14	0.41	5.24	−0.19	1.51
2	$OCHMe_2$	O	4-Cl	$CHMe_2$	8.74	8.65	0.08	3.97	−0.19	1.51
3	OC_4H_9	S	4-Cl	$CHMe_2$	9.74	10.08	−0.34	5.99	−0.19	1.55
4	OC_2H_5	S	4-Cl	$CHMe_2$	9.58	9.34	0.23	4.93	−0.19	1.68
5	OC_5H_{11}	S	4-Cl	$CHMe_2$	8.80	9.10	−0.30	6.52	−0.19	1.52
6	OC_3H_7	S	4-Cl	$CHMe_2$	9.80	10.11	−0.31	5.46	−0.19	1.57
7	OMe	S	4-Cl	$CHMe_2$	8.18	8.45	−0.27	4.40	−0.19	1.77
8	$OCHMe_2$	S	4-Cl	C_3H_7	10.50	10.24	0.26	5.46	−0.12	1.51
9	$OCHMe_2$	S	4-Cl	C_4H_9	9.92	10.17	−0.25	5.99	−0.13	1.51
10	$OCHMe_2$	S	4-Cl	$CH(Me)C_2H_5$	10.22	10.34	−0.12	5.77	−0.21	1.51
11	$OCHMe_2$	S	4-Cl	$CH(C_2H_5)_2$	9.93	9.74	0.20	6.30	−0.23	1.51
12	$OCHMe_2$	S	4-Me	C_2H_5	9.97	9.94	0.04	5.07	−0.10	1.51
13	$OCHMe_2$	S	4-Cl	CH_2CHMe_2	10.52	10.26	0.26	5.86	−0.13	1.51
14	$OCHMe_2$	S	4-Cl	C_2H_5	10.00	9.80	0.21	4.93	−0.10	1.51
15	$OCHMe_2$	S	4-SMe	C_2H_5	9.72	9.99	−0.27	5.13	−0.10	1.51
16	$OCHMe_2$	S	4-Cl	Me	8.76	9.11	−0.35	4.40	0.00	1.51
17	$OCHMe_2$	S	4-Cl	CH_2CF_3[a]	9.01	−31.65	40.66	5.69	0.87	1.51
18	$OCHMe_2$	S	4-SMe	$CHMe_2$	9.91	10.27	−0.36	5.44	−0.19	1.51
19	$OCHMe_2$	S	4-Cl	CMe_3[a]	9.92	−30.71	40.62	5.64	−0.30	1.51
20	$OCHMe_2$	S	4-Cl	CH_2-Cy-C_3H_5	9.91	10.13	−0.21	5.38	0.01	1.51
21	$OCHMe_2$	S	4-Cl	$CH_2C_6H_5$[a]	9.02	−33.95	42.97	6.13	0.20	1.51
22	$OCHMe_2$	S	4-Cl	C_6H_4-2-Me	8.96	9.00	−0.04	6.41	0.62	1.51
23	$OCHMe_2$	S	4-Cl	$CH_2C{\equiv}CH$	9.03	8.83	0.21	4.50	0.76	1.51
24	$OCHMe_2$	S	4-Cl	$CH_2CH_2SiMe_3$	9.27	9.10	0.18	6.55	−0.33	1.51
25	$OCHMe_2$	S	4-Cl	Cy-C_5H_9	10.23	10.29	−0.06	5.88	−0.20	1.51
26	$OCHMe_2$	S	4-Cl	Cy-C_3H_5	9.90	9.63	0.27	4.76	−0.15	1.51
27	$OCHMe_2$	S	4-Cl	$CH_2CH{=}CH_2$	9.90	9.92	−0.02	5.18	0.12	1.51
28	$OC(Me)C_2H_5$	S	4-Cl	$CHMe_2$	10.22	10.01	0.21	5.77	−0.19	1.62
29	$OC(Me)C_2H_5$	S	4-Cl	$CHMe_2$	10.04	10.01	0.04	5.77	−0.19	1.62
30	$OC(Me)C_2H_5$	S	4-Cl	CMe_3	10.06	9.70	0.36	6.17	−0.30	1.62

[a] Data points not used in deriving equation.

*(ii) EC_{50} Activity of Uridine Derivatives (**50**) in MT-4 Cells (Table 41).*[108,109] For another series of



50

uridine analogues where the substituents at the 3′-position, i.e., Y, were also varied, Herdewijn et al.[108,109] reported data for inhibition of HIV-1 replication in MT-4 cells. We obtained eq 57 from their data, where *F* refers to the field/inductive effect of the Y-substituent.

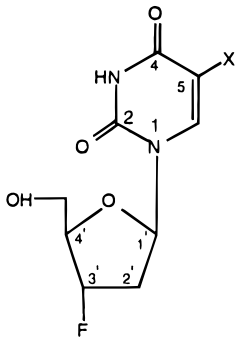
$$\log 1/C = 7.79(\pm1.79)F_Y - 5.63(\pm1.24)\sigma_X + 0.55(\pm0.41)\text{Clog } P + 4.62(\pm0.64)$$

$$n = 12,\ r^2 = 0.953,\ s = 0.39,\ q^2 = 0.798 \quad (57)$$

Outliers 2

Thus, eq 57 indicates that not only the hydrophobic character of both the X- and Y-substituents can be effective, but also the electronic characters of these substituents, particularly the electron-donating ability of the former and the inductive effect of the latter, can influence the activity, suggesting that as with eq 57 a high electron density is desirable or at least low electron density is not good. In deriving eq 57, however, compounds 10 and 11 were not included, as they exhibited aberrant behavior. The equation predicts very high activity for them as compared to their corresponding observed activity. The very low observed activities of these compounds can be attributed to their X-substituents, an ethyl group,

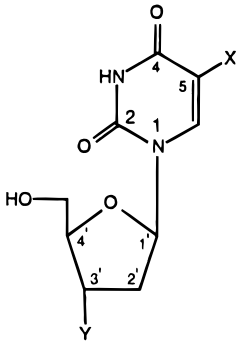
Table 40. ED_{50} Activity of Fluorodideoxyuridines (49)[107]



no.	X	log 1/C obsd[a] (eq 55)	calcd (eq 55)	Δ (eq 55)	obsd[b] (eq 56)	calcd (eq 56)	Δ (eq 56)	σ^+
1	5-Cl	6.16	6.44	−0.28	6.42	6.66	−0.24	0.11
2	5-Br	6.27	6.31	−0.04	6.39	6.44	−0.05	0.15
3	5-I	6.64	6.34	0.30	6.80	6.49	0.30	0.14
4	H	6.80	6.80	0.00	7.22	7.26	−0.04	0.00
5	5-Me	7.82	7.81	0.02	9.00	8.97	0.03	−0.31

[a] Against HIV-2. [b] Against HIV-1.

Table 41. EC_{50} of Uridine Derivatives (50)[108,109]



no.	substituents X	Y	log 1/C obsd	calcd (eq 57)	Δ	F_Y	Clog P	σ_X
1	H	H	3.68	3.93	−0.26	0.00	−1.23	0.00
2	H	F	7.40	7.44	−0.04	0.45	−1.24	0.00
3	H	N_3	6.44	6.70	−0.26	0.30	−0.46	00.0
4	Cl	F	6.42	6.72	−0.30	0.45	−0.19	0.23
5	Br	F	6.39	6.80	−0.42	0.45	−0.04	0.23
6	I	F	6.80	7.23	−0.43	0.45	0.22	0.18
7	Me	H	5.22	5.17	0.06	0.00	−0.73	−0.17
8	Me	F	9.00	8.67	0.33	0.45	−0.74	−0.17
9	Me	N_3	8.40	7.93	0.47	0.30	0.04	−0.17
10	C_2H_5	F[a]	3.48	8.85	−5.37	0.45	−0.21	−0.15
11	C_2H_5	N_3[a]	4.19	8.11	−3.92	0.30	0.57	−0.15
12	CN	F	4.28	3.84	0.45	0.45	−1.03	0.66
13	Br	N_3	6.30	6.07	0.23	0.30	0.74	0.23
14	Cl	N_3	6.14	5.98	0.16	0.30	0.59	0.23

[a] Data points not used in deriving equation.

which might be producing significant steric effects.

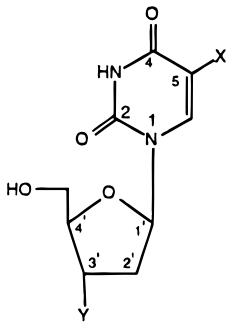
*(iii) EC_{50} Activity of Uridine Derivatives (**50**) in Human PBM Cells (Table 42).*[110] Mahmoudian[110a] reported QSAR studies on another series of uridine derivatives, for which we rederived eq 58.

$$\log 1/C = -0.95(\pm 0.52)L_X + 3.52(\pm 1.62)B1_X + 1.88(\pm 0.92)I_Y - 5.11(\pm 2.49)\sigma_{I,X} + 3.06(\pm 1.68)$$

$$n = 16,\ r^2 = 0.804,\ s = 0.63,\ q^2 = 0.549 \quad (58)$$

Outliers 4

Table 42. EC_{50} Activity of Uridine Derivatives (50)[110a]

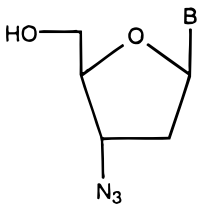


no.	substituents X	Y	log 1/C obsd	calcd (eq 58)	Δ	L_X	$B1_X$	I_Y	$\sigma_{I,,X}$
1	H	N_3	6.50	6.52	−0.02	2.06	1.00	1.00	0.00
2	Me	N_3	8.26	7.79	0.47	2.87	1.52	1.00	−0.04
3	C_2H_5	N_3	6.28	6.46	−0.18	4.11	1.52	1.00	−0.01
4	C_3H_7	N_3[a]	4.20	5.70	−1.50	4.92	1.52	1.00	−0.01
5	Br	N_3	5.98	5.95	0.03	3.82	1.95	1.00	0.44
6	I	N_3	5.94	6.52	−0.58	4.23	2.15	1.00	0.39
7	F	N_3	5.32	4.54	0.78	2.65	1.35	1.00	0.52
8	NH_2	N_3[a]	5.21	6.46	−1.25	2.78	1.35	1.00	0.12
9	OH	N_3	5.00	5.63	−0.63	2.74	1.35	1.00	0.29
10	OMe	N_3	4.15	4.56	−0.41	3.98	1.35	1.00	0.27
11	OC_2H_5	N_3	4.27	3.73	0.54	4.80	1.35	1.00	0.28
12	SCN	N_3[a]	5.29	3.81	1.49	4.08	1.70	1.00	0.64
13	H	NH_2	4.22	4.64	−0.42	2.06	1.00	0.00	0.00
14	C_2H_5	NH_2	4.26	4.58	−0.32	4.11	1.52	0.00	−0.01
15	H	I	4.92	4.64	0.28	2.06	1.00	0.00	0.00
16	Me	I[a]	4.33	5.91	−1.58	2.87	1.52	0.00	−0.04
17	C_2H_5	I	4.07	4.58	−0.51	4.11	1.52	0.00	−0.01
18	H	H	4.01	4.64	−0.63	2.06	1.00	0.00	0.00
19	Me	H	6.77	5.91	0.86	2.87	1.52	0.00	−0.04
20	C_2H_5	H	5.31	4.58	0.73	4.11	1.52	0.00	−0.01

[a] Data points not used in deriving equation.

Equation 58 shows that X-substituents produce steric and electronic effects. Indicator variable I_Y was used with a value of unity for $Y = N_3$. Its positive coefficient shows that this group is conducive to activity. Equation 58 is a poor QSAR statistically, but it does show the importance of steric and electronic effects of the substituents. It is not clear why there is no hydrophobic term.

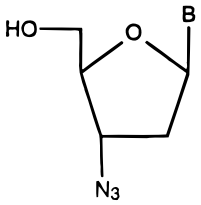
*(iv) ED_{50} Activity of Pyrimidine Derivatives (**51**) in MT-4 Cells (Table 43).*[111] Herdewijn et al.[111] reported



51

anti-HIV data (effective dose of compound achieving 50% protection of MT-4 cells) on pyrimidines analogues, where the uridine ring is replaced by other nucleic acid bases. We could derive eq 59 for the data, which shows that decreasing the size of the molecule will favor the activity. This seems a very crude prediction of the effect of volume on activity. The choice of variation in substituents limits experimen-

Table 43. ED_{50} Activity of Pyrimidine Derivatives (51)[111]



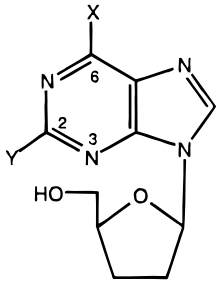
no.	substituent (B)	log 1/*C* obsd	calcd (eq 59)	Δ	MgVol
1	thymin-1-yl[a]	8.40	5.71	2.68	1.82
2	adenin-9-yl	5.30	5.57	−0.27	1.85
3	guanin-9-yl	5.55	5.30	0.25	1.91
4	uracil-1-yl	6.44	6.36	0.07	1.68
5	cytosin-1-yl	5.51	5.52	−0.01	1.86
6	N^4-methylcytosin-1-yl[a]	3.22	6.17	−2.95	1.72
7	4-(hydroxyamino)-5-methyl-1,2-dihydro-2-pyrimidin-1-yl	5.82	5.98	−0.15	1.76
8	5-methyocytosin-1-yl	5.75	5.52	0.22	1.86
9	N^4-5-dimethylcytosin-1-yl	4.76	4.87	−0.11	2.00

[a] Data points not used in deriving equation.

tation with other parameters.

$$\log 1/C = -4.62(\pm 2.18)\mathrm{MgVol} + 14.12(\pm 4.03)$$
$$n = 7,\ r^2 = 0.856,\ s = 0.21,\ q^2 = 0.747 \qquad (59)$$
Outliers 2

*(v) ED_{50} Activity of 6-Halopurine Nucleosides (**52**) in ATH8 Cells (Table 44).*[112] Murakami et al.[112]

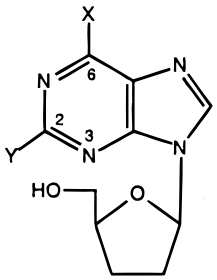


52

reported data for a small series of 6-halo-2′,3′-dideoxypurine nucleosides which were synthesized enzymatically with live *E. coli* in an effort to enhance the lipophilicity of this class of anti-HIV compounds and thereby facilitate drug delivery into the central nervous system. We derived eq 60 for their data, which shows a steric rather than hydrophobic effect of X-substituents.

$$\log 1/C = -0.68(\pm 0.22)\mathrm{B1_X} + 6.47(\pm 0.40)$$
$$n = 7,\ r^2 = 0.924,\ s = 0.07,\ q^2 = 0.835 \qquad (60)$$
Outlier 1

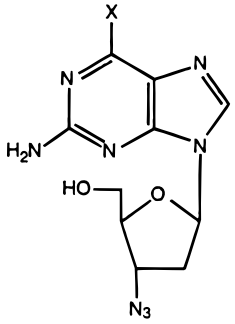
Table 44. ED_{50} Activity of 6-Halo Purine Nucleosides (52)[112]



no.	substituents X	Y	log 1/*C* obsd	calcd	Δ	$B1_X$
1	6-F	NH_2	5.62	5.56	0.06	1.35
2	6-Cl	NH_2	5.26	5.26	0.00	1.80
3	6-Br	NH_2	5.16	5.16	0.01	1.95
4	6-I	NH_2	5.10	5.02	0.08	2.15
5	6-F	H	5.55	5.56	−0.01	1.35
6	6-Cl	H	5.16	5.26	−0.10	1.80
7	6-Br	H	5.11	5.16	−0.04	1.95
8	6-I	H[a]	4.34	5.02	−0.68	2.15

[a] Data point not used in deriving equation.

Table 45. IC_{50} Activity of Purine Nucleosides (53)[113]

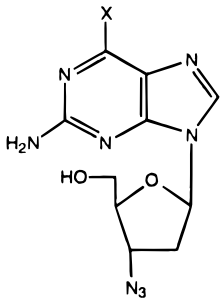


no.	X	log 1/*C* obsd	calcd (eq 61)	Δ	$B5_X$
1	OH	5.30	5.39	−0.09	1.93
2	OMe	5.05	5.06	−0.02	3.07
3	OC_2H_5[a]	4.64	4.98	−0.34	3.36
4	OC_3H_7[a]	5.05	4.67	0.38	4.42
5	$OCHMe_2$	4.66	4.76	−0.10	4.10
6	OC_4H_9	4.72	4.56	0.16	4.79
7	OC_6H_5	4.37	4.24	0.13	5.89
8	$OCH_2C_6H_5$	5.10	4.94	0.16	3.50
9	NH_2	5.70	5.38	0.32	1.97
10	NHMe[a]	4.28	5.06	−0.77	3.08
11	NHC_2H_5	4.77	4.96	−0.19	3.42
12	NHC_3H_7	4.64	4.65	−0.01	4.47
13	NH-Cy-C_3H_5	4.75	4.70	0.04	4.30
14	NHC_4H_9	4.54	4.53	0.01	4.87
15	NHC_6H_5	3.95	4.22	−0.27	5.95
16	$NHCH_2CH_2C_6H_5$[a]	4.68	3.79	0.88	7.40
17	NMe_2	4.92	5.06	−0.14	3.08
18	$N(Me)C_2H_5$	4.96	4.96	0.00	3.42
19	NC_3H_6	5.05	4.84	0.21	3.82
20	NC_4H_8[a]	5.10	4.76	0.34	4.09
21	Cl	5.22	5.43	−0.21	1.80

[a] Data points not used in deriving equation.

*(vi) IC_{50} Activity of Purine Nucleosides (**53**) in MT-4 Cells (Table 45).*[113] For a similar but larger series of purine nucleosides were a variety of substituent

inhibition data were reported by Freeman et al.,[113] we derived eq 61, which also shows a steric effect for X-substituents. The equation is not satisfactory with



53

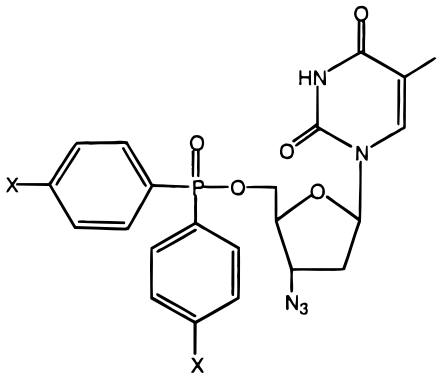
5 outliers. B5 is collinear with CMR and MgVol (r^2 = 0.76). We simply used B5 as it yields better correlation. The reason for the absence of a hydrophobic term is not clear.

$$\log 1/C = -0.29(\pm 0.07)B5_X + 5.96(\pm 0.29)$$

$$n = 16,\ r^2 = 0.837,\ s = 0.17,\ q^2 = 0.761 \quad (61)$$

Outliers 5

*(vii) EC_{50} Activity of Pyrimidine Nucleosides Phosphates (**54**) in C8166 Cells (Table 46).*[114] For some



54

pyrimidine nucleoside phosphates, McGuigan et al.[114] reported anti-HIV data. We obtained eq 62 for the data that shows that the substituents of the aryl rings attached to the phosphorus atom produce a favorable electronic effect. Electron withdrawal by X would favor nucleophilic attack on P with the possible displacement of pyrimidine moiety.

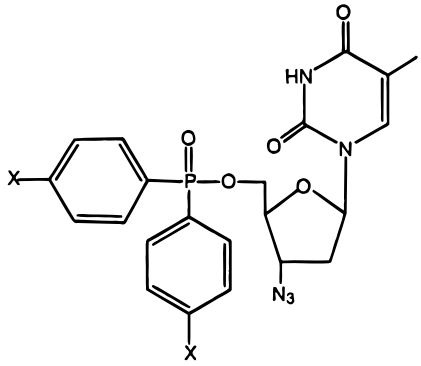
$$\log 1/C = 1.42(\pm 0.23)\sigma_X + 6.47(\pm 0.21)$$

$$n = 7,\ r^2 = 0.981,\ s = 0.17,\ q^2 = 0.961 \quad (62)$$

B. Protease (PR) Inhibitors

Since HIV-1-protease (HIV-1-PR) is an aspartic protease and since, because of that, its substrate is peptidic in nature, a number of peptide-derived compounds have been identified as HIV-1-PR inhibitors.[115] The primary sequence of an aspartic protease has two different Asp-Thr-Gly sequences, and the apostructure of it shows these two chains running

Table 46. EC_{50} Activity of Pyrimidine Nucleoside Phosphates (54)[114]



		log 1/*C*			
no.	X	obsd	calcd (eq 62)	Δ	σ_X
1	NO_2	8.50	8.68	−0.19	1.56
2	CN	8.50	8.34	0.15	1.32
3	SMe	6.40	6.47	−0.08	0.00
4	CF_3	8.20	8.00	0.19	1.08
5	I	6.80	6.98	−0.19	0.36
6	OMe	5.80	5.71	0.09	−0.54
7	H	6.50	6.47	0.02	0.00

in opposite directions with a water molecule bound between two aspartates. This water molecule is believed to be a nucleophile for the enzyme-catalyzed amide hydrolysis of the substrate.[34] The substrate possesses a scissile bond (Figure 3) which, in the substrate−enzyme interaction, is attacked by the water molecule of the enzyme and a few amino acid residues of the substrate interact with corresponding binding sites on the enzyme. This interaction is stabilized by several hydrogen bondings between the backbone of the substrate and the enzyme.[31,34] The discovery of peptide-based substrate-mimicking PR inhibitors was, therefore, directed toward the synthesis of substrate analogues in which the scissile bond was replaced by a noncleavable isostere with tetrahedral geometry that could mimic the tetrahedral transition state of the proteolytic reaction. Thus, several inhibitors with hydroxyethylene or hydroxyethylamine isostere replacement were prepared that could bind with the enzyme as shown in Figure 4. In the inhibitor−enzyme interaction, the enzyme's water molecule hydrogen bonds with both the inhibitor and the enzyme with approximately tetrahedral

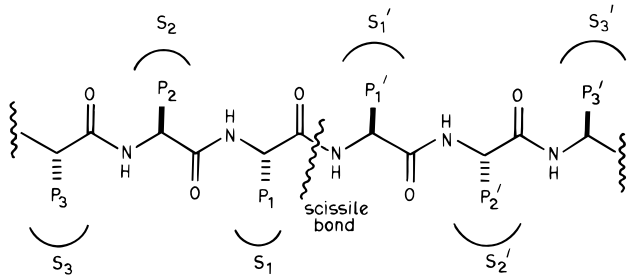


Figure 3. Peptidic substrate of aspartic proteases. The P_1, P_2, ...P_n and P_1', P_2', ...P_n' are amino acid residues, and S_1, S_2, ...S_n and S_1', S_2', ...S_n' are the corresponding binding sites at the enzyme. These nomenclatures are according to Schechter, I.; Burger, A. *Biochem. Biophys. Res. Commun.* **1967**, 27, 157. (Reprinted with permission from ref 34. Copyright 1997 American Chemical Society.)

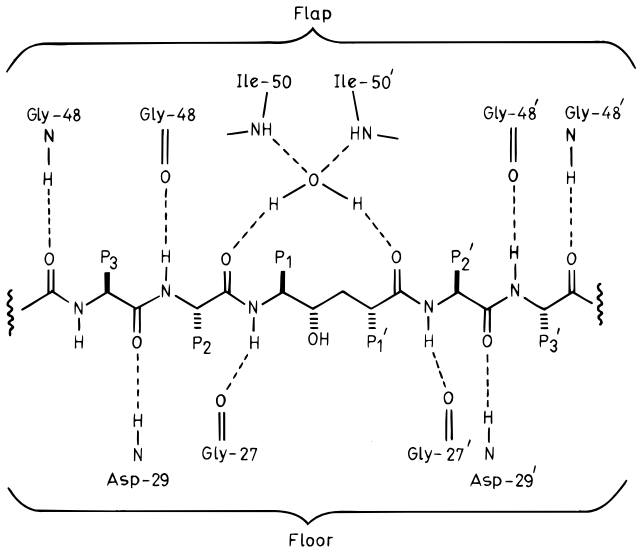


Figure 4. Model of binding of a substrate-based HIV-1-PR inhibitor with the enzyme. (Reprinted with permission from ref 34. Copyright 1997 American Chemical Society.)

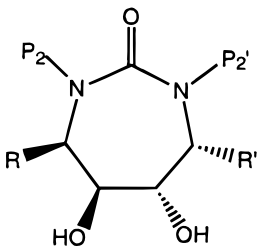
geometry. This water molecule in the complex is known as "flap" water.

However, the clinical development of peptide-derived compounds has been hindered by their poor pharmacokinetics, including low oral bioavailability and rapid excretion[116] and complex and expensive synthesis.[117] Therefore, attention has been focused on the investigation of nonpeptide inhibitors of low molecular weight that can interact with a limited number of binding sites at the enzyme, critical for the inhibition. We, however, discuss here the available or derivable QSAR studies on both nonpeptidic and peptidic inhibitors.

1. Nonpeptidic Inhibitors

a. Cyclic Urea Derivatives. Cyclic urea-based PR inhibitors (**55**) have been extensively studied. The enzyme inhibition constant (K_i) and the antiviral potency (IC_{90}, the molar concentration of the compound required to reduce the concentration of HIV viral RNA by 90% from the level measured in an infected culture) for three different series have been reported.[118,119]

*(i) K_i data of cyclic urea derivatives (**55**) against HIV-protease (Table 47).*[118] When the correlation



55

analysis was performed by us on the K_i data reported by Nugiel et al.,[118] inhibition constants were found to be correlated with McGowan volume of R/R′-substituents and two indicator variables as given in eq 63.

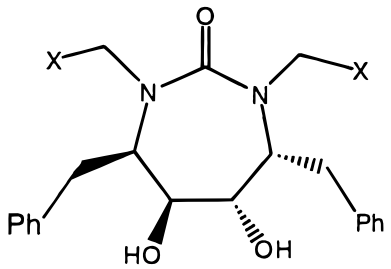
$$\log 1/K_i = 4.90(\pm3.00)\mathrm{MgVol} - 0.60(\pm0.35)(\mathrm{MgVol})^2 + 2.40(\pm0.57)I_a - 1.16(\pm0.51)I_o - 3.91(\pm6.06)$$

$$n = 33,\ r^2 = 0.858,\ s = 0.40,\ q^2 = 0.796 \quad (63)$$

$$(\mathrm{MgVol})_0 = 4.08\ (3.67-4.33),\ \text{Outliers } 5$$

Of the indicator variables, $I_a = 1$ stands for an R/R′-substituent containing an aromatic moiety and $I_o = 1$ stands for an ortho substituent in the aromatic moiety. Equation 63 suggests that while the size of the substituents up to a limited bulk tolerance and an aromatic moiety in them favor the activity, an ortho substituent in the aromatic moiety is detrimental. P2/P2′ were either benzyl or $CH_2-Cy-Pr$ groups.

*(ii) K_i and IC_{90} Data of Cyclic Urea Derivatives (**56** and **57**) against HIV-Protease in MT-2 Cells (Tables 48−51).*[119] Lam et al.[119] reported inhibition constants (K_i, inhibition of HIV-protease) and antiviral activity data (IC_{90}, concentration of inhibitor resulting in 90% inhibition of viral RNA production in HIV-1-infected MT-2 cells) of the compounds (**56**) as listed in Tables 48 and 49, where the R/R′ groups were benzyl and P2/P2′-substituents were largely varied. We derived



56

eqs 64 and 65 for the K_i and IC_{90} data respectively.

$$\log 1/K_i = 1.44(\pm0.42)\mathrm{Clog}\,P - 2.13(\pm0.64)\log(\beta \times 10^{\mathrm{Clog}\,P} + 1) + 0.68(\pm0.42)\mathrm{MR_X} - 0.64(\pm2.22)$$

$$n = 21,\ r^2 = 0.813,\ s = 0.51,\ q^2 = 0.727 \quad (64)$$

$$\log P_0 = 6.53\ (7.51-5.55),\ \log\beta = -6.13,\ \text{Outliers } 5$$

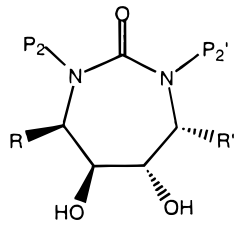
$$\log 1/C = 0.77(\pm0.25)\mathrm{Clog}\,P - 1.24(\pm0.48)\log(\beta \times 10^{\mathrm{Clog}\,P} + 1) + 1.05(\pm1.37)$$

$$n = 15,\ r^2 = 0.813,\ s = 0.33,\ q^2 = 0.665 \quad (65)$$

$$\log P_0 = 6.96\ (7.75-6.17),\ \log\beta = -6.84,\ \text{Outliers } 2$$

Equations 64 and 65 show that the hydrophobicity of the molecule is important for the activity in the bilinear model. There also seems to be a positive steric interaction of X-substituents for the inhibitory

Table 47. PR Inhibition Data of Cyclic Ureas (55)[118]

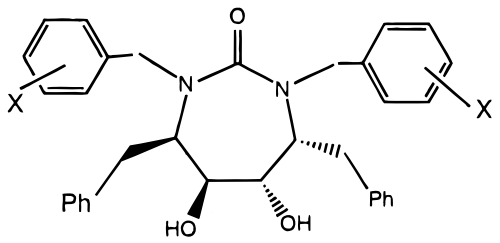


		log $1/K_i$					
no.	R/R′	obsd	calcd (eq 63)	Δ	MgVol	I_a	I_O
1	$CH_2C_6H_5$ (A)[a]	8.47	8.48	−0.01	4.03	1.00	0.00
2	Me (A)	5.30	5.13	0.18	2.82	0.00	0.00
3	$CH_2C_6H_4$-4-$CHMe_2$(A)	8.96	8.10	0.86	4.88	1.00	0.00
4	$CH_2C_6H_4$-4-$CHMe_2$(A)	8.47	8.29	0.18	4.64	1.00	0.00
5	CH_2CHMe_2(A)	5.77	5.98	−0.21	3.66	0.00	0.00
6	CH(Me)SMe(A)	5.96	6.00	−0.04	3.71	0.00	0.00
7	CH_2-3-indolyl(A)[a]	6.24	8.38	−2.14	4.49	1.00	0.00
8	CH_2−Cy−C_6H_{11}(A)[a]	7.55	6.05	1.50	4.29	0.00	0.00
9	$CH_2CH_2C_6H_5$(A)[a]	6.50	8.45	−1.95	4.31	1.00	0.00
10	CH_2-2-naphthyl(A)	8.01	8.19	−0.18	4.77	1.00	0.00
11	CH_2-3-furanyl(A)	8.08	8.38	−0.31	3.67	1.00	0.00
12	$CH_2C_6H_4$-3-SMe(A)	8.60	8.29	0.31	4.64	1.00	0.00
13	$CH_2C_6H_4$-4-SO_2Me−(A)	8.60	8.10	0.51	4.88	1.00	0.00
14	$CH_2C_6H_4$-2-OMe−(A)	7.22	7.25	−0.03	4.43	1.00	1.00
15	$CH_2C_6H_4$-2-OH(A)	7.46	7.32	0.13	4.15	1.00	1.00
16	$CH_2C_6H_4$-3-OMe(A)	8.33	8.41	−0.08	4.43	1.00	0.00
17	$CH_2C_6H_4$-4-OMe(A)	8.07	8.41	−0.34	4.43	1.00	0.00
18	$CH_2C_6H_4$-4-OH(A)	8.96	8.48	0.48	4.15	1.00	0.00
19	$CH_2C_6H_4$-3-NH_2(A)	8.55	8.47	0.09	4.23	1.00	0.00
20	$CH_2C_6H_4$-3-NMe_2(A)	8.37	8.17	0.20	4.80	1.00	0.00
21	$CH_2C_6H_4$-4-NH_2(A)	8.07	8.47	−0.40	4.23	1.00	0.00
22	C_6H_4-4-NH_2-2HCl(A)	8.15	8.47	−0.32	4.23	1.00	0.00
23	$CH_2C_6H_4$-4-NMe_2(A)	7.34	8.17	−0.83	4.80	1.00	0.00
24	CH_2-4-pyridyl(A)	7.66	8.47	−0.82	3.95	1.00	0.00
25	3-(2,5-Me-pyrolyl)-$CH_2C_6H_4$(A)	6.80	7.21	−0.41	5.53	1.00	0.00
26	$CH_2C_6H_4$-3,4-(−OCH_2O−)(A)	8.89	8.44	0.44	4.33	1.00	0.00
27	$CH_2C_6H_5$(B)[b]	8.72	8.24	0.48	3.45	1.00	0.00
28	CH_2CHMe_2(B)[c]	7.07	5.48	1.59	3.07	0.00	0.00
29	$CHMe_2$(B)[c]	6.60	5.09	1.51	2.79	0.00	0.00
30	CH(Me)SMe(B)	5.60	5.53	0.07	3.12	0.00	0.00
31	$CH_2C_6H_4$-4-F(B)	8.24	8.27	−0.03	3.48	1.00	0.00
32	$CH_2C_6H_4$-2-OMe(B)	7.19	7.29	−0.11	3.84	1.00	1.00
33	$CH_2C_6H_4$-3-OMe(B)	9.06	8.45	0.61	3.84	1.00	0.00
34	$CH_2C_6H_4$-3-OH(B)	7.89	8.32	−0.44	3.56	1.00	0.00
35	$CH_2C_6H_4$-4-OMe(B)	8.54	8.45	0.09	3.84	1.00	0.00
36	CH_2-naphthyl(B)	8.37	8.48	−0.11	4.18	1.00	0.00
37	$CH_2C_6H_3$-3,5-OMe(B)	8.57	8.47	0.10	4.24	1.00	0.00
38	CH_2-2-thienyl(B)	8.04	8.12	−0.07	3.29	1.00	0.00

[a] Compounds **1**−**26** (A) where $P_2/P_2' =$ benzyl. [b] Compounds **27**−**38** (B) where $P_2/P_2' = CH_2$−Cy−C_3H_5. [c] Data points not included in deriving equation.

activity. The presence of bilinear hydrophobic terms in eqs 64 and 65, where there are variations at the P/P′ position of **55**, indicate that hydrophobic interactions are important up to a rather high value near 7. Equation 64 is a poor correlation with a high *s*-value, but we feel that a poor correlation is better than none for comparing with others.

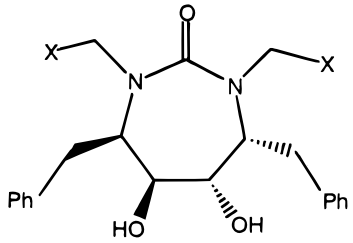
Lam et al.[119] also reported benzyl-substituted P/P′



57

analogues (**57**) (Tables 50 and 51). When we analyzed the data, not only the steric effect of the ortho substituent of the benzyl ring but also the overall lipophilicity of the molecule appears to be detrimental to both inhibition (K_i = inhibition of HIV-protease) and antiviral (IC_{90} = concentration of inhibitor resulting in 90% inhibition of viral RNA production in HIV-1-infected MT-2 cells) potencies (eqs 66 and 67). However, the enzyme inhibition potency is found to be helped by the meta substituent through a steric effect. Both eqs 66 and 67 contain a negative hydrophobic term. It is not clear to us why. When we tried to analyze the data with respect to the hydrophobicity of individual substituent positions (using π values), we again observed a negative term. Could it be a steric effect that we see here via a −Clog *P* term? It is possible that the rigidity of phenyl ring does not allow the substituents to interact well with the hydrophobic pocket. It would be of interest to study

Table 48. PR Inhibition Data of Cyclic Ureas (56)[119]



		log 1/K_i				
no.	X	obsd	calcd (eq 64)	Δ	Clog *P*	MR_X
1	H	5.24	5.01	0.24	3.88	0.10
2	Me	7.00	6.79	0.21	4.94	0.57
3	C_2H_5	8.10	8.18	−0.08	6.00	1.03
4	C_3H_7	8.85	8.46	0.39	7.05	1.50
5	C_4H_9	8.80	8.14	0.65	8.11	1.96
6	C_5H_{11}	8.34	7.74	0.60	9.17	2.42
7	C_6H_{13}	6.59	7.32	−0.74	10.23	2.89
8	CH_2OMe	6.10	6.40	−0.31	4.34	1.21
9	$CH_2OC_2H_5$[a]	5.96	−3.23	9.19	5.40	1.67
10	$CHMe_2$[a]	7.31	−4.31	11.62	6.79	1.50
11	CH_2CHMe_2	7.92	8.31	−0.39	7.85	1.96
12	$CH_2CH_2CHMe_2$	8.16	7.92	0.24	8.91	2.42
13	$CH_2CH_2CH_2CHMe_2$	7.52	7.50	0.02	9.97	2.89
14	CH_2CMe_3	7.44	8.09	−0.65	8.65	2.42
15	$CH{=}CH_2$[a]	8.28	−3.64	11.92	5.43	1.10
16	$C(Me){=}CH_2$	8.14	8.63	−0.50	6.23	1.56
17	$C{\equiv}CH$[a]	7.66	−2.81	10.47	4.08	0.96
18	C_3H_5	8.68	8.29	0.39	5.83	1.35
19	C_4H_7	8.89	8.71	0.18	6.94	1.79
20	C_5H_9	8.37	8.34	0.03	8.06	2.20
21	C_6H_{11}	7.43	7.90	−0.46	9.18	2.67
22	C_6H_5	8.52	9.08	−0.55	7.24	2.54
23	2-pyridyl	6.84	7.01	−0.18	4.24	2.30
24	4-pyridyl	7.05	7.01	0.03	4.24	2.30
25	α-naphthlyl[a]	7.07	−4.02	11.08	8.99	4.16
26	β-naphthyl	9.51	8.63	0.88	9.59	4.16

[a] Data points not used in deriving equation.

flexible aliphatic substituents.

$$\log 1/K_i = -0.39(\pm 0.17)\text{Clog } P - 3.82(\pm 1.43)MR_2 + 0.79(\pm 0.72)MR_3 + 11.73(\pm 1.32)$$

$$n = 20,\ r^2 = 0.818,\ s = 0.52,\ q^2 = 0.672 \quad (66)$$

Outliers 2

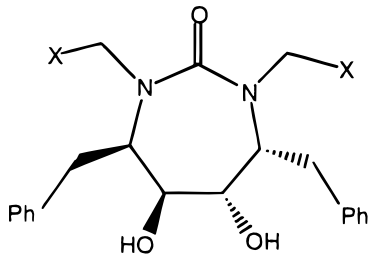
$$\log 1/C = -0.47(\pm 0.12)\text{Clog } P - 1.99(\pm 1.00)MR_2 + 9.81(\pm 0.96)$$

$$n = 20,\ r^2 = 0.833,\ s = 0.38,\ q^2 = 0.728 \quad (67)$$

Outliers 2

Gupta et al.[120] also analyzed the data[118,119] in detail. On the basis of their studies, they proposed a model of interaction of cyclic ureas with the receptor as shown in Figure 5.

Table 49. IC_{90} Data of Cyclic Ureas (56)[119]



		log 1/*C*			
no.	X	obsd	calcd (eq 65)	Δ	Clog *P*
1	C_2H_5[a]	4.27	−1.82	6.08	6.00
2	C_3H_7	6.17	5.93	0.24	7.05
3	C_4H_9	5.82	5.65	0.17	8.11
4	CH_2CHMe_2	5.50	5.76	−0.26	7.85
5	$CH_2CH_2CHMe_2$	5.10	5.30	−0.20	8.91
6	$CH{=}CH_2$	5.33	5.19	0.14	5.43
7	$C(Me){=}CH_2$	5.12	5.70	−0.58	6.23
8	$C{\equiv}CH$	4.38	4.17	0.20	4.08
9	Cy-C_3H_5	5.75	5.46	0.28	5.83
10	Cy-C_4H_7	6.00	5.93	0.07	6.94
11	Cy-C_5H_9	5.77	5.67	0.10	8.06
12	C_6H_5	6.08	5.92	0.16	7.24
13	2-pyridyl	4.31	4.30	0.01	4.24
14	3-pyridyl[a]	5.06	−0.98	6.04	4.24
15	4-pyridyl	4.01	4.30	−0.29	4.24
16	α-naphthyl	4.80	5.26	−0.47	8.99
17	β-naphthyl	5.41	4.98	0.43	9.59

[a] Data points not used in deriving equation.

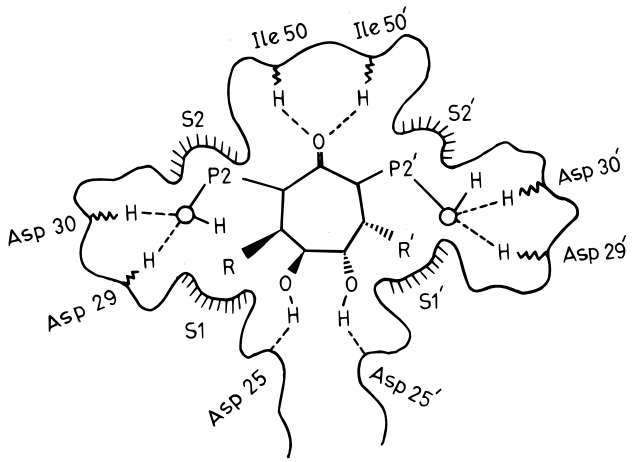
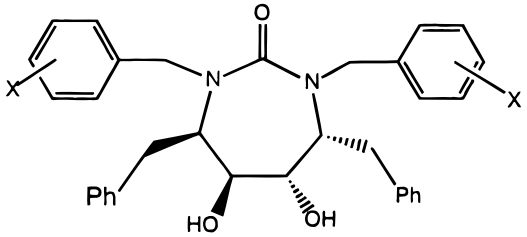


Figure 5. Schematic representation of the interaction of cyclic ureas with HIV-1 proteases. (Taken from ref 120.)

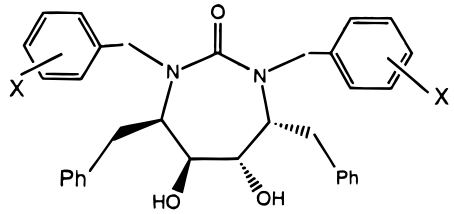
*(iii) K_i and IC_{90} Data of Cyclic Ureas (**58**) against HIV-Protease in MT-2 Cells (Tables 52 and53).*[121a]



58

Jadhav et al.[121a] reported inhibition data (K_i = inhibition of HIV-protease) and antiviral potencies

Table 50. PR Inhibition Data of Cyclic Ureas (57)[119]



		log 1/K_i					
no.	X	obsd	calcd (eq 66)	Δ	Clog *P*	MR_2	MR_3
1	2-F[a]	7.47	8.55	−1.08	7.52	0.09	0.10
2	3-F	8.52	8.50	0.02	7.52	0.10	0.09
3	4-F	8.85	8.51	0.35	7.52	0.10	0.10
4	2-Cl	6.62	6.16	0.46	8.66	0.60	0.10
5	3-Cl	9.05	8.46	0.59	8.66	0.10	0.60
6	4-Cl	8.28	8.07	0.22	8.66	0.10	0.10
7	3-Br	8.85	8.57	0.28	8.96	0.10	0.89
8	4-Br	7.57	7.95	−0.38	8.96	0.10	0.10
9	3-Me	8.16	8.60	−0.44	8.24	0.10	0.57
10	4-Me	8.24	8.23	0.01	8.24	0.10	0.10
11	3-CF_3	7.66	8.25	−0.59	9.00	0.10	0.50
12	4-CF_3	7.29	7.94	−0.64	9.00	0.10	0.10
13	2-OMe	5.73	6.07	−0.34	7.08	0.79	0.10
14	3-OMe	8.80	9.22	−0.43	7.08	0.10	0.79
15	4-OMe[a]	6.80	8.68	−1.88	7.08	0.10	0.10
16	3-NO_2	8.55	9.32	−0.77	6.72	0.10	0.74
17	3-I	9.38	8.77	0.60	9.48	0.10	1.39
18	4-CH_2OH	9.47	9.42	0.05	5.16	0.10	0.10
19	3-CH_2OH	9.85	9.91	−0.06	5.16	0.10	0.72
20	4-OH	9.92	9.13	0.79	5.90	0.10	0.10
21	3-OH	9.92	9.28	0.64	5.90	0.10	0.29
22	3-NH_2	9.55	9.92	−0.36	4.78	0.10	0.54

[a] Data points not included in deriving equation.

(IC_{90} is the concentration of inhibitor resulting in 90% inhibition of viral RNA production in HIV-1 infected MT-2 cells after 3 days) of some cyclic urea derivatives. QSAR analyses performed by us gave eqs 68 and 69, which revealed the electronic effect of substituents. We observe here again the same negative Clog *P* term as seen in eqs 66 and 67. It is of interest here to note that Clog *P* is collinear with MgVol and CMR ($r^2 = 0.80$). We believe one should be careful with a negative Clog *P* term as it might be a steric effect that is operative in such examples. Equations 68 and 69 show that electron-releasing substituents would favor the enzyme inhibition and antiviral potencies; from another point of view electron-attracting substituents are bad.

$$\log 1/K_i = -1.29(\pm 0.99)\sigma - 0.61(\pm 0.20)\text{Clog } P + 12.79(\pm 1.44)$$

$$n = 12,\ r^2 = 0.850,\ s = 0.57,\ q^2 = 0.733 \quad (68)$$

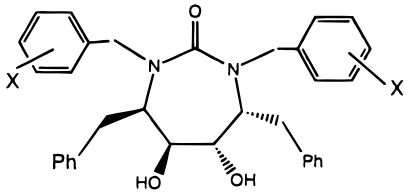
Outlier 1

$$\log 1/C = -1.34(\pm 0.82)\sigma - 0.56(\pm 0.17)\text{Clog } P + 10.04(\pm 1.19)$$

$$n = 12,\ r^2 = 0.879,\ s = 0.47,\ q^2 = 0.786 \quad (69)$$

Outlier 1

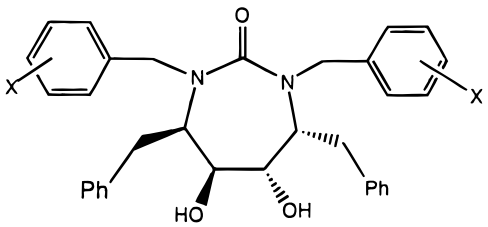
Table 51. IC_{90} Data of Cyclic Ureas (57)[119]



		log 1/*C*				
no.	X	obsd	calcd (eq 67)	Δ	Clog *P*	MR_2
1	2-F[a]	5.26	6.10	−0.85	7.52	0.09
2	3-F	6.15	6.08	0.07	7.52	0.10
3	4-F	6.22	6.08	0.14	7.52	0.10
4	2-Cl	4.95	4.55	0.40	8.66	0.60
5	3-Cl	5.89	5.55	0.34	8.66	0.10
6	4-Cl	5.35	5.55	−0.20	8.66	0.10
7	3-Br	5.92	5.41	0.51	8.96	0.10
8	4-Br	5.09	5.41	−0.32	8.96	0.10
9	3-Me	5.62	5.75	−0.13	8.24	0.10
10	4-Me	5.37	5.75	−0.38	8.24	0.10
11	3-CF_3	5.11	5.39	−0.28	9.00	0.10
12	4-CF_3	5.14	5.39	−0.25	9.00	0.10
13	2-OMe	4.64	4.93	−0.29	7.08	0.79
14	3-OMe	5.89	6.29	−0.41	7.08	0.10
15	4-OMe[a]	5.12	6.29	−1.17	7.08	0.10
16	3-NO_2	6.01	6.46	−0.44	6.72	0.10
17	3-I	5.52	5.16	0.36	9.48	0.10
18	4-CH_2OH	7.24	7.19	0.06	5.16	0.10
19	3-CH_2OH	7.42	7.19	0.23	5.16	0.10
20	4-OH	7.50	6.84	0.65	5.90	0.10
21	3-OH	7.27	6.84	0.43	5.90	0.10
22	3-NH_2	6.89	7.37	−0.48	4.78	0.10

[a] Data points not included in deriving equation.

Table 52. PR Inhibition Data of Cyclic Ureas (58)[121a]

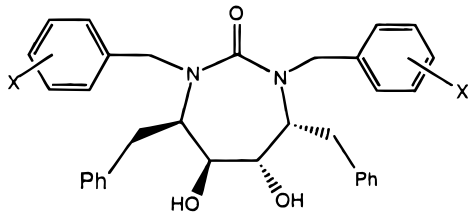


		log 1/K_i				
no.	X	obsd	calcd (eq 68)	Δ	σ	Clog *P*
1	H	8.52	8.43	0.10	0.00	7.24
2	3-NO_2	8.55	7.82	0.73	0.71	6.72
3	4-NO_2	7.50	7.73	−0.24	0.78	6.72
4	3-NH_2	9.55	10.11	−0.56	−0.16	4.78
5	4-NH_2[a]	8.96	10.76	−1.80	−0.66	4.78
6	3-CN	8.52	8.39	0.14	0.56	6.10
7	4-CN	7.28	8.26	−0.97	0.66	6.10
8	3-OH	9.92	9.08	0.85	0.12	5.90
9	4-OH	9.92	9.71	0.21	−0.37	5.90
10	3-$OCH_2C_6H_5$	6.47	6.26	0.21	0.10	10.61
11	4-$OCH_2C_6H_5$	6.27	6.69	−0.42	−0.23	10.61
12	3-CH_2OH	9.85	9.68	0.18	0.00	5.16
13	4-CH_2OH	9.47	9.68	−0.21	0.00	5.16

[a] Data point not included in deriving equation.

*(iv) K_i and IC_{90} Data of Bisbenzamide Cyclic Ureas (**59**) against HIV-Protease in MT-2 Cells (Tables 54*

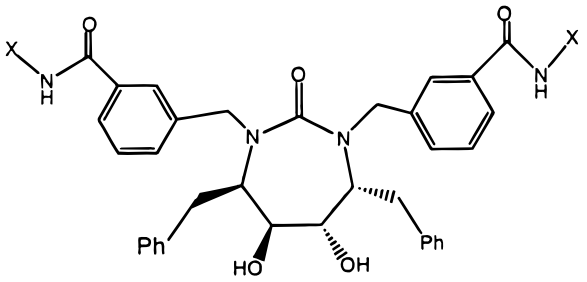
Table 53. IC_{90} Data of Cyclic Ureas (58)[121a]



		log 1/*C*				
no.	X	obsd	calcd (eq 69)	Δ	σ	Clog *P*
1	H	6.08	6.03	0.05	0.00	7.24
2	3-NO_2	6.01	5.37	0.65	0.71	6.72
3	4-NO_2	5.08	5.27	−0.20	0.78	6.72
4	3-NH_2	6.89	7.60	−0.72	−0.16	4.78
5	4-NH_2[a]	6.96	8.27	−1.31	−0.66	4.78
6	3-CN	5.66	5.91	−0.25	0.56	6.10
7	4-CN	5.24	5.78	−0.54	0.66	6.10
8	3-OH	7.27	6.61	0.66	0.12	5.90
9	4-OH	7.50	7.26	0.23	−0.37	5.90
10	3-$OCH_2C_6H_5$	4.16	4.03	0.13	0.10	10.61
11	4-$OCH_2C_6H_5$	4.16	4.47	−0.31	−0.23	10.61
12	3-CH_2OH	7.42	7.18	0.24	0.00	5.16
13	4-CH_2OH	7.24	7.18	0.06	0.00	5.16

[a] Data point not included in deriving equation.

and 55).[122] Wilkerson et al.[122] reported data for the PR inhibition (K_i = inhibition of HIV-protease) and



59

antiviral (IC_{90} = concentration of inhibitor resulting in 90% inhibition of viral RNA production in HIV-1 infected MT-2 cells) activities of a series of bisbenzamide cyclic ureas. We reanalyzed their data, which resulted in eqs 70 and 71.

$$\log 1/K_i = -1.31(\pm 0.40)\text{MgVol} + 1.58(\pm 0.40)I - 1.44(\pm 1.01)\text{IP} - 30.14(\pm 9.92)$$

$$n = 26,\ r^2 = 0.792,\ s = 0.32,\ q^2 = 0.720 \quad (70)$$

Outliers 3

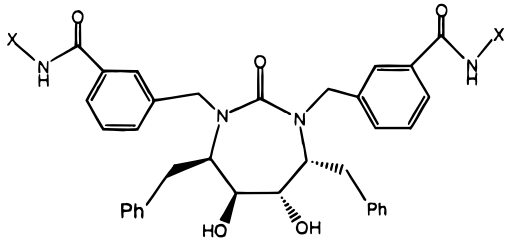
$$\log 1/C = -0.95(\pm 0.52)\text{MgVol} + 1.86(\pm 0.40)I - 1.67(\pm 0.99)\text{IP} - 1.13(\pm 0.54)\text{HBD} + 27.34(\pm 10.33)$$

$$n = 25,\ r^2 = 0.903,\ s = 0.29,\ q^2 = 0.849 \quad (71)$$

Outliers 4

Clog *P* and MgVol showed a high mutual correlation (r^2 = 0.76), thus making the role of both ambiguous. Equation 70 also has indicator variables

Table 54. PR Inhibition Data of Bis-Benzamide Cyclic Ureas (59)[122]



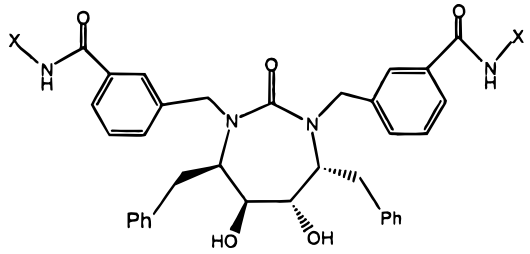
		log 1/K_i					
no.	X	obsd	calcd (eq 70)	Δ	MgVol	I	IP
1	H	10.41	10.61	−0.20	4.54	0.00	9.42
2	NH_2	10.75	10.51	0.24	4.74	0.00	9.30
3	OH	10.70	10.59	0.11	4.66	0.00	9.32
4	OMe	10.35	10.23	0.12	4.94	0.00	9.31
5	Me	10.18	10.27	−0.09	4.83	0.00	9.39
6	C_2H_5	9.68	9.92	−0.24	5.11	0.00	9.38
7	$CHMe_2$	9.24	9.60	−0.32	5.39	0.00	9.37
8	C_3H_7	9.45	9.55	−0.11	5.39	0.00	9.38
9	C_4H_9	9.37	9.18	0.19	5.67	0.00	9.38
10	CMe_3	8.62	9.21	−0.59	5.67	0.00	9.36
11	CH_2-Cy-C_3H_5	9.13	9.47	−0.34	5.46	0.00	9.38
12	CH_2CF_3	9.68	9.65	0.03	5.22	0.00	9.47
13	CH_2CN	10.20	9.71	0.49	5.14	0.00	9.50
14	C_6H_5	9.37	9.45	−0.09	5.76	0.00	9.11
15	4-Py[a]	9.39	8.97	0.42	5.68	0.00	9.52
16	3-Py	9.54	9.15	0.39	5.68	0.00	9.39
17	2-Py	10.37	10.97	−0.60	5.68	1.00	9.23
18	2-Py-3-Me−[d]	9.59	10.78	−1.20	5.96	1.00	9.11
19	2-Py-4-Me	10.57	10.67	−0.10	5.96	1.00	9.18
20	2-Py-5-Me	10.96	10.85	0.11	5.96	1.00	9.06
21	2-Py-6-Me	10.70	10.78	−0.08	5.96	1.00	9.11
22	2-Py-4,6-Me_2	10.80	10.48	0.32	6.24	1.00	9.06
23	2-Py-5-Cl	10.92	10.50	0.42	5.92	1.00	9.33
24	2-Py-3,5-Cl_2[d]	9.61	10.01	−0.40	6.17	1.00	9.45
25	2-Py-5-Br	10.46	10.21	0.25	6.03	1.00	9.44
26	2-Pyrim[b]-4-Me	9.94	10.50	−0.56	5.88	1.00	9.38
27	2-Py-5-CF_3	10.07	10.04	0.03	6.07	1.00	9.53
28	2-Pyraz[c]	10.74	10.74	0.00	5.60	1.00	9.47
29	2-Pyrim[d]	9.82	10.85	−1.03	5.60	1.00	9.39

[a] Pyridyl. [b] Pyrimidinyl. [c] Pyrazinyl. [d] Data points not included in deriving equation.

I with a value of unity for the 2-pyridyl substituent and 0.0 for others, and IP is the calculated ionization potential of the molecule. Through these variables, the equation suggests that a molecule having a low ionization potential and bearing a 2-pyridyl group would have better inhibition potency. The advantageous role of the 2-pyridyl group can be attributed to its nitrogen atom, which can participate in hydrogen bonding with the receptor (Figure 6).[122] The IP is the least important parameter. MgVol and *I* yield a QSAR with r^2 = 0.725. Adding IP, r^2 = 0.777.

Equation 71 was derived to show that the antiviral potency is a function of the enzyme inhibition and that a larger molecule bearing a substituent that may not be able to act as a hydrogen-bond donor would prove a better antiviral agent. In the equation, HBD is a hydrogen-bond donor parameter, which takes a value of 1 for a hydrogen-bond donor substituent. The MgVol term is highly collinear with Clog *P* (r^2 =

Table 55. IC_{90} Data of Bis-Benzamide Cyclic Ureas (59)[122]



		log 1/C						
no.	X	obsd	calcd (eq 71)	Δ	MgVol	I	IP	HBD
1	H	6.15	6.17	−0.12	4.54	0.00	9.42	1.00
2	NH_2	6.05	6.17	−0.11	4.74	0.00	9.30	1.00
3	OH	6.35	6.22	0.13	4.66	0.00	9.32	1.00
4	OMe	6.67	7.09	−0.42	4.94	0.00	9.31	0.00
5	Me	7.10	7.07	0.02	4.83	0.00	9.39	0.00
6	C_2H_5	6.86	6.82	0.03	5.11	0.00	9.38	0.00
7	$CHMe_2$	6.58	6.57	0.01	5.39	0.00	9.37	0.00
8	C_3H_7	6.59	6.56	0.03	5.39	0.00	9.38	0.00
9	C_4H_9	6.59	6.29	0.31	5.67	0.00	9.38	0.00
10	CMe_3	6.14	6.32	−0.17	5.67	0.00	9.36	0.00
11	CH_2-Cy-C_3H_5	6.31	6.49	−0.18	5.46	0.00	9.38	0.00
12	CH_2CF_3	7.03	6.57	0.46	5.22	0.00	9.47	0.00
13	CH_2CN	6.23	6.60	−0.38	5.14	0.00	9.50	0.00
14	C_6H_5	6.29	6.65	−0.36	5.76	0.00	9.11	0.00
15	4-Py[a]	7.03	6.05	0.98	5.68	0.00	9.52	0.00
16	3-Py	6.91	6.26	0.65	5.68	0.00	9.39	0.00
17	2-Py	8.55	8.38	0.17	5.68	1.00	9.23	0.00
18	2-Py_3-Me− [a]	7.30	8.33	−1.03	5.96	1.00	9.11	0.00
19	2-Py-4-Me	8.11	8.20	−0.08	5.96	1.00	9.18	0.00
20	2-Py-5-Me	8.51	8.41	0.10	5.96	1.00	9.06	0.00
21	2-Py-6-Me	8.50	8.32	0.17	5.96	1.00	9.11	0.00
22	2-Py-4,6-Me_2	8.18	8.13	0.05	6.24	1.00	9.06	0.00
23	2-Py-5-Cl	7.83	7.98	−0.15	5.92	1.00	9.33	0.00
24	2-Py-3,5-Cl_2	7.37	7.55	−0.19	6.17	1.00	9.45	0.00
25	2-Py-5-Br	7.55	7.70	−0.15	6.03	1.00	9.44	0.00
26	2-Pyrim[b]-4-Me[a]	6.91	7.95	−1.05	5.88	1.00	9.38	0.00
27	2-Py-5-CF_3	7.20	7.52	−0.32	6.07	1.00	9.53	0.00
28	2-Pyraz[c]	8.46	8.07	0.39	5.60	1.00	9.47	0.00
29	2-Pyrim[a]	6.66	8.20	−1.54	5.60	1.00	9.39	0.00

[a] Pyridyl. [b] Pyrimidinyl. [c] Pyrazinyl. [d] Data points not included in deriving equation.

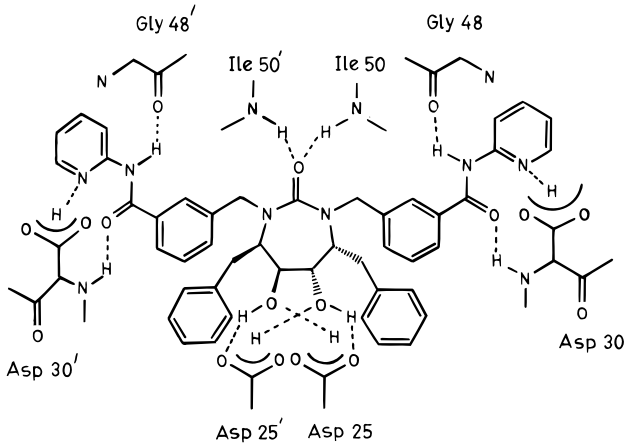
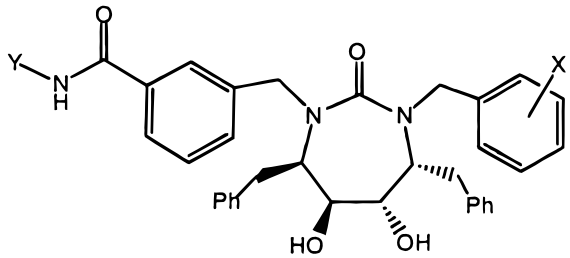


Figure 6. Model proposed for the interaction of a 2-pyridyl-containing cyclic urea derivative. (Reprinted with permission from ref 122. Copyright 1996 American Chemical Society.)

0.83). We choose MgVol simply because it gives a better correlation; also, the Clog *P* term observed was negative.

*(v) K_i Data of Nonsymmetrical Cyclic Ureas (**60**) against HIV-Protease (Table 56).*[123] Another series of



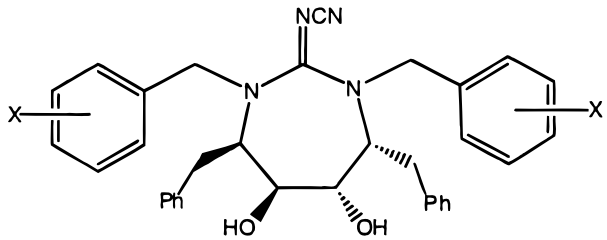
60

nonsymmetrically substituted cyclic urea derivatives was reported by Wilkerson et al.[123] When we reanalyzed their data (K_i = inhibition of recombinant single-chain dimeric HIV protease), we derived eq 72. We were unable to reproduce the QSAR reported by Wilkerson et al.[123]

$$\log 1/K_i = -1.69(\pm 0.98)\text{Clog } P + 2.67(\pm 1.63)\log(\beta \times 10^{\text{Clog } P} + 1) - 6.03(\pm 2.41)\text{MgVol} + 48.39(\pm 15.26)$$
$$n = 13,\ r^2 = 0.871,\ s = 0.35,\ q^2 = 0.757 \quad (72)$$
$$\log P_0 = 6.00\ (\pm 0.65),\quad \log \beta = -5.38,\ \text{Outliers } 2$$

This is a poor correlation since there are four variables and only 13 data points. A negative bilinear term was observed that we find difficult to explain. The initial negative slope followed by strong positive slope is very unusual.

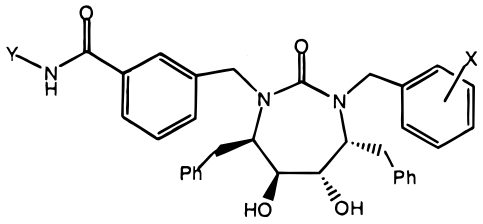
*(vi) K_i and IC_{90} Data of Cyclic Cyanoguanidines (**61**) against HIV-Protease in MT-2 Cells (Table 57).*[121a] Jadhav et al.[121a] also reported inhibition



61

potencies (K_i = inhibition of HIV-protease) and antiviral activity (IC_{90} = concentration of inhibitor resulting in 90% inhibition of viral RNA production in HIV-1-infected MT-2 cells) of nonpeptidic cyclic cyanoguanidines for which we derived eqs 73 and 74, respectively.

Table 56. PR Inhibition Data of Nonsymmetric Cyclic Ureas (60)[123]



	substituents		log 1/*C*				
no.	X	Y	obsd	calcd (eq 72)	Δ	Clog *P*	MgVol
1	3,5-di-OMe	2-Pyraz[a]	8.60	8.95	−0.35	6.38	5.21
2	3,5-di-OMe	5-Me-2-Py[b]	8.72	8.93	−0.21	7.58	5.40
3	3,5-di-OMe	6-Me-2-Py	9.16	8.93	0.22	7.58	5.40
4	3,5-di-OMe	2-Py	9.07	9.31	−0.24	7.08	5.25
5	3-OMe	2-Pyraz	10.42	10.09	0.33	6.29	5.01
6	3-OMe	5-Me-2-Py	10.16	10.05	0.11	7.49	5.20
7	3-OMe	6-Me-2-Py	10.28	10.05	0.23	7.49	5.20
8	3-OMe	2-Py	10.33	10.43	−0.10	7.00	5.06
9	3-NH_2	5-Me,2-Py	10.12	9.64	0.48	6.35	5.10
10	3-NO_2	2-Py	10.02	10.42	−0.40	6.82	5.03
11	3-NH_2	CMe_3[e]	9.39	24.37	−14.98	5.97	4.95
12	3-NH_2	2-Pyraz	10.80	10.61	0.19	5.14	4.91
13	3-NH_2	2-Benz[c,e]	10.64	23.75	−13.11	6.77	5.19
14	3-NH_2	2-Imid[d]	10.92	11.18	−0.25	5.18	4.82
15	3-CONH-CH_2CO_2H	CH_2CN	10.62	10.61	0.01	3.84	5.20

[a] Pyrazinyl. [b] Pyridinyl. [c] Benzimidazolyl. [d] Imidazolyl. [e] Data points not included in deriving equation.

$$\log 1/K_i = -1.77(\pm 0.65)\text{MgVol} - 1.26(\pm 0.87)\sigma_{sum} + 16.21\ (\pm 3.14)$$

$$n = 11,\ r^2 = 0.844,\ s = 0.47,\ q^2 = 0.747 \quad (73)$$

Outliers 2

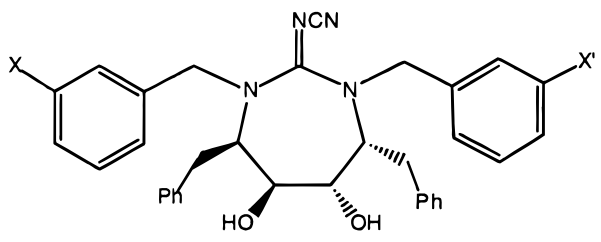
$$\log 1/C = -1.36(\pm 0.69)\sigma_{sum} - 1.53(\pm 0.51)\text{MgVol} + 13.01(\pm 2.49)$$

$$n = 11,\ r^2 = 0.877,\ s = 0.38,\ q^2 = 0.783 \quad (74)$$

Outliers 2

Equations 73 and 74 show that electron-releasing substituents seem to enhance antiviral activity while enzyme inhibition activity seems to depend on the steric effect of X-substituents. Gupta and Babu[121b] recently reported QSAR study on the data of Jadhav et al.

*(vii) K_i and IC_{90} Data of Cyclic Ureas (**62**) against HIV-Protease in MT-2 Cells (Table 58).*[124] Further



62

studies on HIV protease inhibitors were reported by Han et al.[124] on cyclic ureas, for which we found eqs 75 and 76. In eq 76, the indicator variable *I* was used with a value of 1.0 for C(=O)−R (compounds **1**−**6**) and 0.0 for C(=NOH)−R (compounds **7**−**11**).

$$\log 1/K_i = -1.44(\pm 0.61)I - 0.49(\pm 0.28)\text{CMR} + 19.86(\pm 5.29)$$

$$n = 9,\ r^2 = 0.873,\ s = 0.36,\ q^2 = 0.651 \quad (75)$$

Outlier 2

$$\log 1/C = -2.18(\pm 0.77)I - 0.66(\pm 0.30)\text{CMR} + 20.85(\pm 5.93)$$

$$n = 8,\ r^2 = 0.914,\ s = 0.29,\ q^2 = 0.735 \quad (76)$$

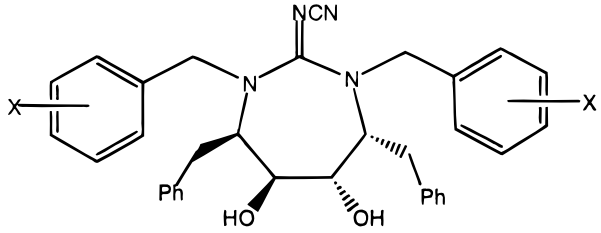
Outliers 2

The negative CMR term in both equations indicates a steric effect on enzyme inhibition and antiviral activities. The indicator variable *I* with its negative coefficient also shows that oximes are preferred over ketones.

In comparing eqs 63−76, it is worth mentioning that only eqs 64 and 65 contain a positive bilinear hydrophobic term in which the P/P′ position of **55** is substituted mostly by alkyl groups (Tables 48 and 49). All other series contain X-benzyl substituents at that position. Could it be that the rigid phenyl ring does not allow the substituent to reach hydrophobic space and interact? Also, can the phenyl ring locate substituents beyond hydrophobic space? However, a significant steric interaction seems to be involved in almost all of them.

b. Cycloalkylpyranones. The lead structures 4-hydroxycoumarin (**63**) and cycloalkylpyranone (**64**) have been extensively studied for PR-inhibition activity. In a series of successive studies, Romines et al.[125−128] reported the effects on the inhibition constant (K_i) of various modifications in the structure

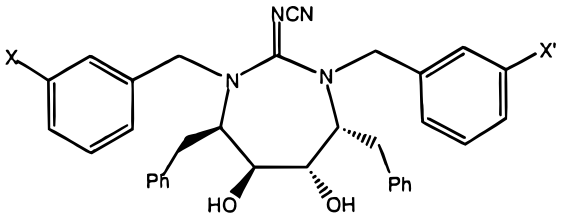
Table 57. PR Inhibition and IC_{90} Data of Cyclic Cyanoguanidines (61)[121a]



		log 1/K_i			log 1/C				
no.	X	obsd (eq 73)	calcd (eq 73)	Δ (eq 73)	obsd (eq 74)	calcd (eq 74)	Δ (eq 74)	MgVol	σ_{SUM}
1	H[a]	7.70	8.74	−1.04	5.42	6.52	−1.10	4.23	0.00
2	3-NO_2	7.05	7.23	−0.18	4.75	5.02	−0.27	4.58	0.71
3	4-NO_2	7.17	7.14	0.04	4.71	4.93	−0.22	4.58	0.78
4	3-NH_2	8.13	8.59	−0.46	6.30	6.44	−0.13	4.43	−0.16
5	4-NH_2[a]	7.60	9.22	−1.62	5.64	7.11	−1.48	4.43	−0.66
6	3-CN	7.57	7.49	0.08	5.51	5.29	0.22	4.54	0.56
7	4-CN	6.89	7.36	−0.47	5.11	5.15	−0.04	4.54	0.66
8	3-OH	9.14	8.38	0.76	6.89	6.18	0.71	4.35	0.12
9	4-OH	8.59	9.00	−0.42	6.60	6.85	−0.24	4.35	−0.37
10	3-$OCH_2C_6H_5$	5.86	5.76	0.10	4.17	3.91	0.26	5.84	0.10
11	4-$OCH_2C_6H_5$	6.05	6.18	−0.13	4.17	4.36	−0.19	5.84	−0.23
12	3-CH_2OH	8.77	8.03	0.74	6.23	5.91	0.32	4.63	0.00
13	4-CH_2OH	7.96	8.03	−0.08	5.50	5.91	−0.42	4.63	0.00

[a] Data points not included in deriving both equations.

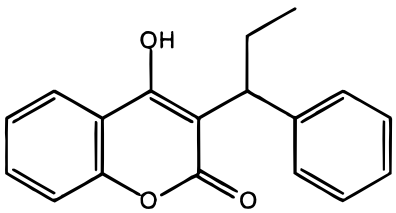
Table 58. PR Inhibition and IC_{90} Data of Cyclic Cyanoguanidines (62)[124]



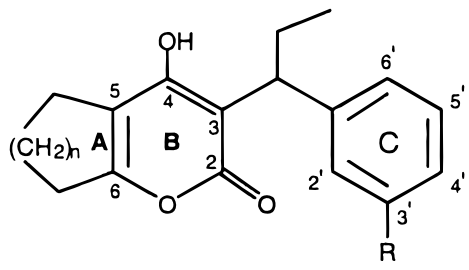
		log 1/K_i			log 1/C				
no.	X/X′	obsd (eq 75)	calcd (eq 75)	Δ (eq 75)	obsd (eq 76)	calcd (eq 76)	Δ (eq 76)	I	CMR
1	C(=O)H	9.36[a]	10.52	−1.17	7.44	7.88	−0.43	1.00	16.30
2	C(=O)Me	10.22	10.08	0.15	7.41	7.26	0.15	1.00	17.23
3	C(=O)C_2H_5	9.68	9.63	0.05	6.85	6.65	0.20	1.00	18.15
4	C(=O)C_3H_7	8.85	9.18	−0.32	5.84	6.03	−0.19	1.00	19.08
5	C(=O)CF_3	10.43	10.03	0.40	7.48	7.20	0.28	1.00	17.32
6	C(=O)CMe_3	8.44	8.73	−0.28	8.30[a]	9.25	−0.94	1.00	20.01
7	C(=NOH)H	11.00	11.37	−0.37	8.70	8.63	0.07	0.00	17.52
8	C(=NOH)Me	10.75	10.92	−0.17	8.16	8.02	0.14	0.00	18.45
9	C(=NOH)C_2H_5	10.51	10.47	0.04	7.19	7.40	−0.21	0.00	19.38
10	C(=NOH)C_3H_7	10.51	10.02	0.49	6.15[a]	8.57	−2.42	0.00	20.31
11	C(=NOH)CF_3	8.41[a]	10.87	−2.46				0.00	18.54

[a] Data points not included in deriving respective equations.

(**64**), including the change in the size of the alkyl ring A.



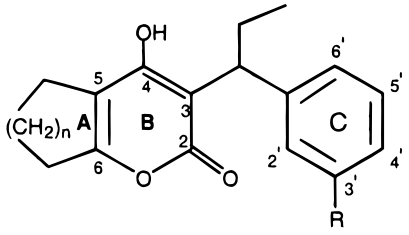
63



64

*(i) K_i Data of Cycloalkyldihydropyranones (**64**) against HIV-Protease (Table 59).*[127] For a series of compounds in which variation in the cycloalkyl ring along with other substituent effects were studied by Romines et al.,[127] we formulated eq 77, which shows a positive correlation with the CMR term, indicating

Table 59. PR Inhibition Data of Cycloalkylpyranones (64)[127]



				log $1/K_i$			
no.	R	n^a	ring type[b]	obsd	calcd (eq 77)	Δ	CMR
1	H	1	A	6.16	5.54	0.62	7.75
2	H	2	A	5.96	6.00	−0.04	8.21
3	H	3	A	6.32	6.45	−0.13	8.67
4	H	4	A	7.13	6.80	0.32	9.03
5	H[d]	6	A	5.70	7.72	−2.02	9.96
6	$NHSO_2$-(5-CN-2-Py)[c]	1	B	10.13	9.49	0.63	11.76
7	$NHSO_2$-(5-CN-2-Py)	2	B	10.19	10.40	−0.22	12.69
8	$NHSO_2$-(5-CN-2-Py)	3	B	10.75	10.40	0.34	12.69
9	$NHSO_2$-(5-CN-2-Py)	4	B	11.16	10.76	0.40	13.05
10	$NHSO_2$-(5-CN-2-Py)[d]	6	B	9.60	11.67	−2.07	13.98
11	$NHSO_2$-(5-CN-2-Py)	1	C	9.83	9.49	0.34	11.76
12	$NHSO_2$-(5-CN-2-Py)	2	C	10.30	9.95	0.35	12.23
13	$NHSO_2$-(5-CN-2-Py)	3	C	9.72	10.41	−0.68	12.69
14	$NHSO_2$-(5-CN-2-Py)	4	C	10.13	10.76	−0.63	13.05
15	H	1	Δ	5.70	5.54	0.16	7.75
16	H	2	Δ	6.17	6.00	0.18	8.21
17	H	3	Δ	5.70	6.45	−0.75	8.67
18	H	4	Δ	5.92	6.81	−0.88	9.03

[a] n = number of CH_2 units in cycloalkyl ring. [b] Ring type A & B = 6,7-dihydrocycloalkylpyrone; C & D = 5,6,7,7a-tetrahydrocycloalkylpyrone. [c] Py = pyridyl. [d] Data point not used in deriving equation.

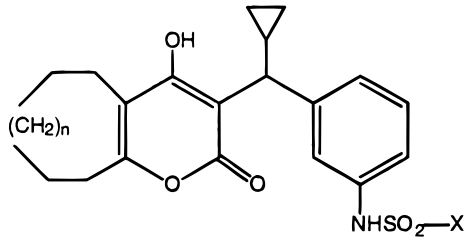
that a bulky molecule would favor the inhibitory activity.

$$\log 1/K_i = 0.98(\pm 0.13)\mathrm{CMR} + 2.08(\pm 1.41)$$

$$n = 16,\ r^2 = 0.948,\ s = 0.52,\ q^2 = 0.933 \qquad (77)$$

Outlier 2

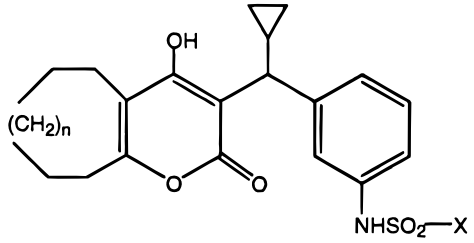
*(ii) K_i Data of Cycloalkylpyranones (**65**) against HIV-Protease (Table 60).*[128] For another series of



65

sulfonamide-substituted cycloalkylpyranones reported by Skulnick et al.,[128] eq 78 was obtained by us. In this equation indicator variable I is 1.0 for $n = 1$ and 0.0 for $n = 2$. Its negative coefficient shows that the cyclooctyl ring is preferred over the cycloheptyl ring for enzyme inhibition. The positive coefficient for the substituent parameter MR indicates that sterically bulky substituents would enhance the inhibitory

Table 60. PR Inhibition Data of Cycloalkylpyranones (65)[128]



			log $1/K_i$				
no.	X	n	obsd	calcd (eq 78)	Δ	MR_X	I
1	Me	1	6.92	6.99	−0.07	0.57	1.00
2	C_2H_5	1	7.20	7.13	0.07	1.03	1.00
3	Me	2	7.96	7.87	0.09	0.57	0.00
4	C_2H_5	2	8.00	8.01	−0.01	1.03	0.00
5	$CH{=}CH_2$	2	7.75	8.03	−0.29	1.10	0.00
6	C_3H_7	2	8.40	8.16	0.24	1.50	0.00
7	$CHMe_2$[a]	2	7.48	8.16	−0.68	1.50	0.00
8	C_4H_9	2	8.34	8.30	0.04	1.96	0.00
9	C_8H_{17}[a]	2	8.22	8.88	−0.66	3.82	0.00
10	C_6H_5	2	8.50	8.48	0.01	2.54	0.00
11	C_6H_{11}	2	8.44	8.52	−0.08	2.67	0.00

[a] Data points not used in deriving equation.

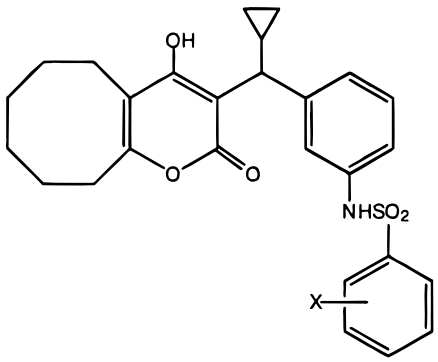
activity. No hydrophobic term was found to be significant. The small number of data points limits the use of more parameters.

$$\log 1/K_i = 0.31(\pm 0.21)\mathrm{MR_X} - 0.88(\pm 0.37)I + 7.69(\pm 0.37)$$

$$n = 9,\ r^2 = 0.935,\ s = 0.17,\ q^2 = 0.883 \qquad (78)$$

Outliers 2

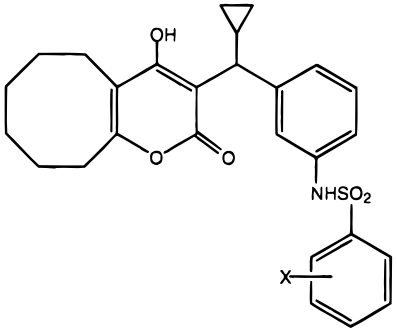
*(iii) K_i Data of Cyclooctylpyranones (**66**) for Inhibition of HIV-Protease (Tables 61 and 62).*[128] For



66

another series of cyclooctylpyranones reported by Skulnick et al.,[128] eqs 79 and 80 were derived by us. In eq 79 the inhibition potency seems to be governed by steric effects of substituents at ortho and meta positions of the phenyl ring. While large substituents at the ortho position are not favorable to the inhibition activity, it seems that the activity can be

Table 61. PR Inhibition Data of Cyclooctylpyranones (66)[128]



		log $1/K_i$					
no.	X	obsd	calcd (eq 79)	Δ	$B5_2$	L_3	$B5_3$
1	H	8.50	8.30	0.20	1.00	2.06	1.00
2	2-Me	8.03	7.93	0.10	2.04	2.06	1.00
3	2-F	7.96	8.17	−0.21	1.35	2.06	1.00
4	2-Cl	8.03	8.01	0.02	1.80	2.06	1.00
5	2-CF_3	7.68	7.73	−0.05	2.61	2.06	1.00
6	2-CN	8.13	8.08	0.05	1.60	2.06	1.00
7	3-Me	8.28	8.31	−0.04	1.00	2.87	2.04
8	3-Cl	8.57	8.68	−0.11	1.00	3.52	1.80
9	3-Br	8.59	8.76	−0.17	1.00	3.82	1.95
10	3-CF_3	8.52	8.32	0.21	1.00	3.30	2.61
11	3-NO_2	8.33	8.43	−0.10	1.00	3.44	2.44
12	3-COOH	8.35	8.57	−0.22	1.00	3.91	2.66
13	3-COOMe	8.85	8.70	0.16	1.00	4.73	3.36
14	3-NH_2[a]	8.64	8.30	0.34	1.00	2.78	1.97
15	3-CN	9.22	9.05	0.17	1.00	4.23	1.60

[a] Data point not used in deriving equation.

optimized by selecting a meta substituent of optimum length and small width.

$$\log 1/K_i = -0.35(\pm 0.28)B5_2 + 0.44(\pm 0.22)L_3 - 0.33(\pm 0.26)B5_3 + 8.07(\pm 0.75)$$

$$n = 14,\ r^2 = 0.853,\ s = 0.17,\ q^2 = 0.627 \quad (79)$$

Outlier 1

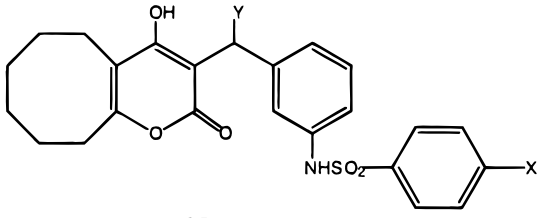
Equation 80 was derived for the para-substituted derivatives only, which shows that the substituents at this position also have steric interactions but, on the other hand, electron-donating substituents seem to favor inhibition activity through resonance.

$$\log 1/K_i = -0.57(\pm 0.28)B1 - 0.23(\pm 0.19)B5 + 0.55(\pm 0.26)\sigma^- + 9.87(\pm 0.60)$$

$$n = 17,\ r^2 = 0.824,\ s = 0.18,\ q^2 = 0.706 \quad (80)$$

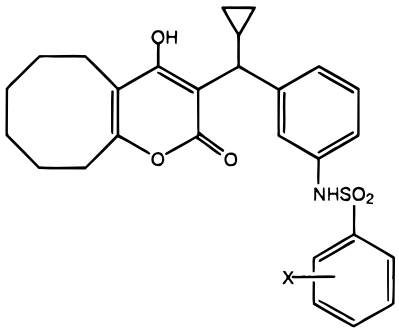
Outliers 3

*(iv) K_i Data of Cycloalkylpyranones (**67**) for Inhibition of HIV-Protease (Table 63).*[128] Skulnick et al.[128]



67

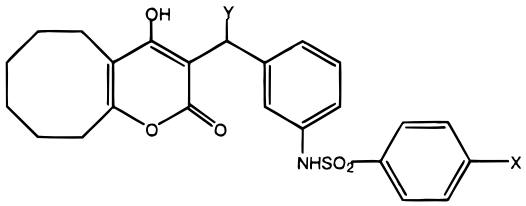
Table 62. PR Inhibition Data of Cyclooctylpyranones (66)[128]



		log $1/K_i$						
no.	X	obsd	calcd (eq 62)	Δ	B1	B5	σ^-
1	4-Me	8.48	8.44	0.04	1.52	2.04	−0.17
2	4-C_2H_5	8.30	8.16	0.14	1.52	3.17	−0.19
3	4-C_3H_7	7.96	8.16	−0.20	1.52	3.49	−0.06
4	4-$CHMe_2$	7.89	7.97	−0.08	1.90	3.17	−0.16
5	4-CMe_3	7.50	7.59	−0.09	2.60	3.17	−0.13
6	4-F	8.51	8.77	−0.26	1.35	1.35	−0.03
7	4-Cl	8.60	8.53	0.07	1.80	1.80	0.19
8	4-Br	8.68	8.45	0.23	1.95	1.95	0.25
9	4-I	8.51	8.30	0.21	2.15	2.15	0.27
10	4-CF_3	8.21	8.49	−0.28	1.99	2.61	0.65
11	4-CN	9.10	9.14	−0.04	1.60	1.60	1.00
12	4-NO_2[a]	8.57	9.03	−0.47	1.70	2.44	1.27
13	4-COOH[a]	7.96	8.76	−0.81	1.60	2.66	0.77
14	4-$CONH_2$	8.72	8.64	0.08	1.50	3.07	0.61
15	4-OMe	8.41	8.25	0.16	1.35	3.07	−0.26
16	4-OC_4H_9	7.75	7.81	−0.07	1.35	4.79	−0.32
17	4-OCF_3	8.40	8.41	−0.01	1.35	3.61	0.27
18	4-NH_2	8.16	8.30	−0.14	1.35	1.97	−0.63
19	4-NMe_2	8.55	8.32	0.23	1.35	3.08	−0.12
20	4-N_3[a]	8.59	8.11	0.48	1.50	4.18	0.11

[a] Data points not used in deriving equation.

Table 63. PR Inhibition Data of Cycloalkylpyranones (67)[128]

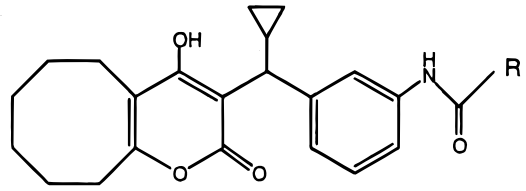


			log $1/K_i$					
no.	X	Y	obsd	calcd (eq 63)	Δ	σ^-_X	L_Y	$B1_Y$
1	4-Cl	Cy-C_3H_5	8.60	8.61	−0.00	0.19	4.14	1.55
2	4-Cl	C_2H_5	8.46	8.67	−0.22	0.19	4.11	1.52
3	4-Cl	C_3H_7	8.40	8.55	−0.15	0.19	4.92	1.52
4	4-Cl	$CHMe_2$	7.75	7.90	−0.16	0.19	4.11	1.90
5	4-CN	Cy-C_2H_5	9.10	8.94	0.16	1.00	4.14	1.55
6	4-CN	C_2H_5[a]	8.51	9.00	−0.49	1.00	4.11	1.52
7	4-CN	C_3H_7	8.85	8.88	−0.02	1.00	4.92	1.52
8	4-CN	C_4H_9	8.68	8.68	0.00	1.00	6.17	1.52
9	4-CN	$CHMe_2$	8.23	8.23	0.00	1.00	4.11	1.90
10	4-CN	CH_2CHMe_2	8.82	8.84	−0.02	1.00	5.14	1.52
11	4-F	Cy-C_3H_5	8.51	8.52	−0.01	−0.03	4.14	1.55
12	4-F	C_3H_7	8.68	8.46	0.22	−0.03	4.92	1.52
13	4-F	C_4H_9	8.22	8.26	−0.05	−0.03	6.17	1.52
14	4-F	$CHMe_2$	7.96	7.81	0.15	−0.03	4.11	1.90
15	4-F	CH_2CHMe_2	8.55	8.46	0.10	−0.03	4.92	1.52

[a] Data point not used in deriving equation.

also studied the PR-inhibition activity of cyclooctylpyranones (**67**), where X-substituents were either

Table 64. PR Inhibition Data of Cycloalkylpyranones (68)[126]



no.	X	log 1/K_i obsd	log 1/K_i calcd	Δ	Clog P
1	C_6H_5[a]	7.38	7.97	−0.59	5.30
2	4-F-C_6H_5	7.26	7.56	−0.30	5.64
3	CH=CH−C_6H_5	6.33	6.35	−0.02	6.31
4	CH_2−$NHCO_2CMe_3$	8.26	8.33	−0.07	4.72
5	CH_2CH_2−$NHCO_2CMe_3$	8.40	8.17	0.23	5.05
6	$CH_2CH_2CH_2$−$NHCO_2CMe_3$	7.85	7.88	−0.02	5.38
7	CH(Me)−$NHCO_2CMe_3$(S)	8.16	8.19	−0.02	5.03
8	CH(Me)−$NHCO_2CMe_3$(R)[a]	7.50	8.19	−0.69	5.03
9	*N*-(CO_2CMe_3)-pyrrolidine-2-yl(S)	7.37	7.15	0.21	5.90
10	CH(CH_2-1*H*-imidazol-4-yl)-$NHCO_2CMe_3$(S)	8.52	8.24	0.28	4.05
11	CH(CH_2-1*H*-imidazol-4-yl)-$NHCO_2CMe_3$(R)	7.96	8.24	−0.29	4.05

[a] Data points not used in deriving equation.

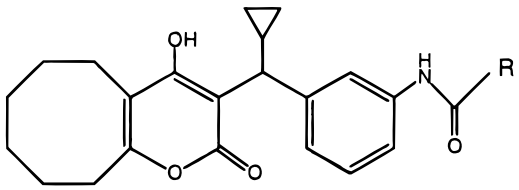
4-Cl, 4-CN, or 4-F and Y substituents were mostly alkyl groups. We obtained eq 81 for these data, which shows the same electronic effect as in eq 80. The negative coefficient of L and B1 parameters for Y-substituents show that sterically large Y-groups are not suitable for enzyme inhibition activity. The positive σ^- term suggests that a more acidic H on N is significant.

$$\log 1/K_i = 0.41(\pm 0.19)\sigma_X^- - 0.16(\pm 0.14)L_Y - 2.03(\pm 0.64)B1_Y + 12.31(\pm 1.45)$$

$$n = 14,\ r^2 = 0.881,\ s = 0.14,\ q^2 = 0.780 \quad (81)$$

Outlier 1

*(v) K_i Data of Cycloalkylpyranones (**68**) for Inhibition of HIV-Protease (Table 64).*[126] In place of sulfonamide groups, some carboxamide groups were also tried (**68**) by Romines et al.,[126] for which we derived eq 82, which shows that the inhibitory activity is totally governed by the hydrophobicity of molecule. NH here is not acidic enough to benefit the activity.



68

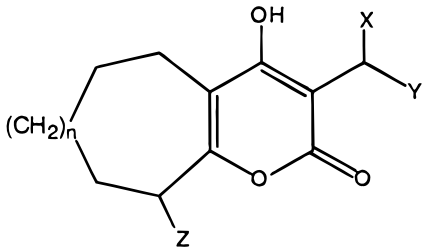
$$\log 1/K_i = 5.44(\pm 3.81)\text{Clog } P - 0.61(\pm 0.37)(\text{Clog } P)^2 - 3.84(\pm 9.58)$$

$$n = 9,\ r^2 = 0.910,\ s = 0.24,\ q^2 = 0.770 \quad (82)$$

$$\log P_0 = 4.49\ (3.46-4.78),\ \text{Outliers } 2$$

In comparing eq 81 and 82, it is of note that eq 81 shows a positive σ_X^- term for substituents attached to sulfonamide groups, indicating that electron-attracting groups would favor activity, while eq 82 derived for amide derivatives contains no electronic term.

*(vi) K_i Data of Cycloalkylpyranones (**69**) (Table 65).*[126] Romines et al.[126] also studied the effects of



69

changes in both the substituents on the α-carbon and also the effect of substitution at the cycloalkyl ring. Equation 83 was derived by us for the data, which appears to rest on the steric effect alone. In eq 83, indicator variables I_X and I_Y were used with a value of unity for X = C_2H_5 and Y = C_6H_5. In this series, however, it was found to matter little whether *n* (number of CH_2 unit in the cycloalkyl ring) is 1 or 2.

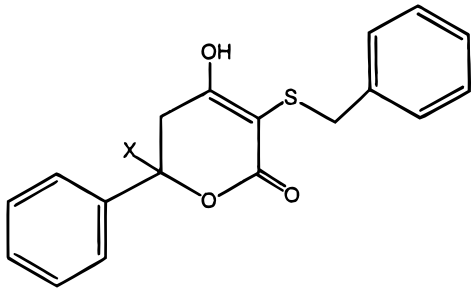
$$\log 1/K_i = -0.93(\pm 0.17)I_X + 0.88(\pm 0.20)I_Y - 1.09(\pm 0.33)B1_Z + 8.29(\pm 0.45)$$

$$n = 14,\ r^2 = 0.954,\ s = 0.11,\ q^2 = 0.927 \quad (83)$$

Outlier 1

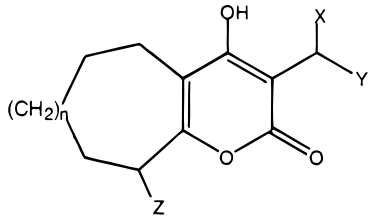
Gupta et al.[129] also reported detailed QSAR studies on the data reported by Romines et al.[125−128]

*(vii) IC_{50} Data of Dihydropyranones (**70**) for Inhibition of HIV-Protease (Table 66).*[130] Tait et al.[130] reported another series of 4-hydroxy-5,6-dihydropyrone protease inhibitors (**70**), where the X groups were mostly alkyl. Analysis by us revealed eq 84, which shows that while the overall hydrophobicity of a molecule favors antiviral activity, sterically large



70

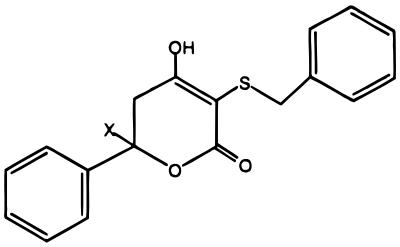
Table 65. PR Inhibition Data of Cycloalkylpyranones (69)[126]



	substituents			log 1/*C*					
no.	X	Y	Z[a]	obsd	calcd (eq 83)	Δ	I_X	I_Y	$B1_Z$
1	C_2H_5	C_6H_5	H(A)	7.23	7.14	0.09	1.00	1.00	1.00
2	C_2H_5	C_6H_5	Me(A)	6.48	6.58	−0.09	1.00	1.00	1.52
3	C_2H_5	C_6H_5	$CH_2C_6H_5$(A)	6.58	6.58	0.00	1.00	1.00	1.52
4	Cy−C_3H_5	Cy−C_3H_5	H(A)	7.24	7.20	0.05	0.00	0.00	1.00
5	Cy−C_3H_5	Cy−C_3H_5	$CH_2C_6H_5$(A)	6.59	6.63	−0.05	0.00	0.00	1.52
6	Cy−C_3H_5	C_6H_5	H(A)	7.93	8.07	−0.14	0.00	1.00	1.00
7	Cy−C_3H_5	C_6H_5	C_2H_5(A)	7.48	7.51	−0.03	0.00	1.00	1.52
8	Cy−C_3H_5	C_6H_5	C_3H_7(A)	7.59	7.51	0.08	0.00	1.00	1.52
9	Cy−C_3H_5	C_6H_5	C_4H_9(A)	7.51	7.51	0.00	0.00	1.00	1.52
10	Cy−C_3H_5	C_6H_5	CH_2CHMe_2(A)	7.52	7.51	0.02	0.00	1.00	1.52
11	Cy−C_3H_5	C_6H_5	CH_2−Cy−C_3H_5(A)	7.62	7.51	0.11	0.00	1.00	1.52
12	Cy−C_3H_5	C_6H_5	$CH_2CH_2CHMe_2$(A)	7.30	7.51	−0.21	0.00	1.00	1.52
13	Cy−C_3H_5	C_6H_5	H(B)[b]	7.02	8.07	−1.05	0.00	1.00	1.00
14	Cy−C_3H_5	C_6H_5	CH_2CH_2OMe(B)	7.55	7.51	0.05	0.00	1.00	1.52
15	Cy−C_3H_5	C_6H_5	CH_2−Cy−C_3H_5(B)	7.62	7.51	0.11	0.00	1.00	1.52

[a] Cyclooctyl ring(A) where $n = 2$; cyclohepyl ring(B) where $n = 1$. [b] Data point not used in deriving equation.

Table 66. IC_{50} Data of Dihydropyranones (70)[130]



		log 1/*C*				
no.	X	obsd	calcd (eq 84)	Δ	Clog *P*	$B5_X$
1	H[a]	5.05	6.33	−1.28	3.32	1.00
2	C_3H_7	6.54	6.53	0.01	4.90	3.49
3	C_4H_9	6.57	6.51	0.06	5.43	4.54
4	C_5H_{11}	6.91	6.77	0.14	5.96	4.94
5	C_6H_{13}	6.75	6.75	−0.01	6.49	5.96
6	CH_2CHMe_2	6.36	6.44	−0.08	5.30	4.45
7	$CH_2CH_2CHMe_2$	7.01	6.83	0.18	5.83	4.54
8	$CH_2CH_2CH_2CHMe_2$	6.74	6.81	−0.07	6.36	5.59
9	CH_2−C_6H_{13}	6.86	6.99	−0.13	6.49	5.42
10	C_6H_5	6.58	6.68	−0.10	4.88	3.11
11	$CH_2CH_2C_6H_5$	7.22	7.22	0.00	5.79	3.58

[a] Data point not used in deriving equation.

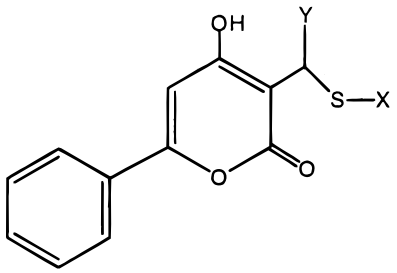
X-substituents would be detrimental to activity. More hydrophobic derivatives should be tested to establish log P_0.

$$\log 1/C = 0.82(\pm 0.32)\text{Clog } P - 0.44(\pm 0.20)B5_X + 4.05(\pm 1.11)$$

$$n = 10,\ r^2 = 0.841,\ s = 0.12,\ q^2 = 0.742 \quad (84)$$

Outlier 1

*(viii) IC_{50} Data of Pyranones (**71**) for Inhibition of HIV-Protease (Table 67).*[131] Inhibition data of the



71

pyranones were reported by Vara Prasad et al.[131] on pyranones derivatives in which the X-groups were mostly phenyl or Cy−alkyl and Y was mostly alkyl and Cy−alkyl. We analyzed the data and obtained eq 85, which shows that sterically bulky X- and Y-substituents would favor antiviral activity.

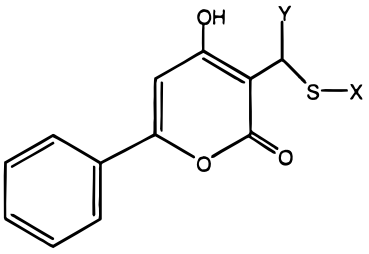
$$\log 1/C = 2.42(\pm 1.66)B1_X + 0.37(\pm 0.25)B5_X + 0.58(\pm 0.20)B5_Y - 1.67(\pm 3.56)$$

$$n = 17,\ r^2 = 0.848,\ s = 0.32,\ q^2 = 0.741 \quad (85)$$

Outliers 2

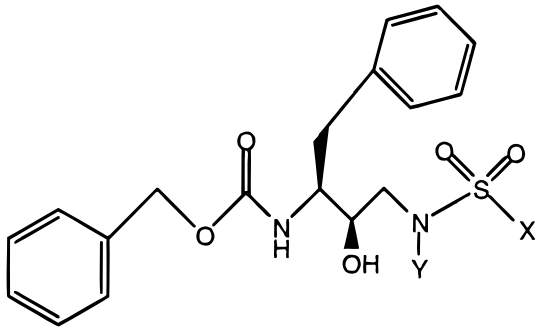
Similar to cyclic ureas, we do not observe a hydrophobic term in eqs 77−85 on cycloalkylpyranones protease inhibitors, except in eqs 82 and 84. It seems to us that due to some spatial restrictions these molecules are not able to bind in hydrophobic space as the protease receptor does have hydrophobic binding sites. Gupta et al.[137] reported extensive QSAR studies on cyclopyranones and proposed hydrophobic interaction between receptor and ligand.

Table 67. IC_{50} Data of Pyranones (71)[131]



	substituent		log 1/*C*					
no.	X	Y	obsd	calcd (eq 85)	Δ	$B1_X$	$B5_X$	$B5_Y$
1	C_6H_5	H	4.07	4.19	−0.12	1.71	3.11	1.00
2	C_6H_5	C_6H_5	6.11	5.41	0.70	1.71	3.11	3.11
3	C_6H_5	2-naphthyl	5.11	6.10	−0.99	1.71	3.11	4.31
4	C_6H_5	$Cy-C_6H_{11}$	5.61	5.62	−0.01	1.71	3.11	3.49
5	C_6H_5	CH_2CHMe_2	6.39	6.18	0.21	1.71	3.11	4.45
6	C_6H_5	$CH_2CH_2CHMe_2$	6.41	6.23	0.18	1.71	3.11	4.54
7	2-naphthyl	C_6H_5	5.61	5.84	−0.23	1.71	4.31	3.11
8	$CH_2C_6H_5$	C_6H_5	6.32	6.01	0.31	1.52	6.02	3.11
9	$CH_2C_6H_5$	CH_2CHMe_2	6.59	6.78	−0.19	1.52	6.02	4.45
10	$CH_2C_6H_5$	$CH_2-Cy-C_3H_5$	7.08	6.73	0.35	1.52	6.02	4.36
11	$Cy-C_6H_{11}$	C_6H_5	6.32	6.03	0.29	1.91	3.49	3.11
12	$Cy-C_6H_{11}$	CH_2CHMe_2	6.50	7.80	−0.31	1.91	3.49	4.45
13	$Cy-C_6H_{11}$	$CH_2-Cy-C_3H_5$	6.83	6.75	0.09	1.91	3.49	4.36
14	$Cy-C_6H_{11}$	$CH_2-Cy-C_3H_5$	6.27	6.75	−0.48	1.91	3.49	4.36
15	$Cy-C_6H_{11}$	CH_2CMe_3	6.52	6.64	−0.12	1.91	3.49	4.18
16	$CH_2-Cy-C_6H_{11}$	CH_2CHMe_2	6.07	6.60	−0.49	1.52	5.42	4.45
17	$Cy-C_5H_9$	$Cy-C_5H_9$	6.65	6.79	−0.14	1.90	4.09	4.09
18	$Cy-C_5H_9$	CH_2CHMe_2	7.24	6.80	0.44	1.91	3.49	4.45
19	$Cy-C_5H_9$	$CH_2-Cy-C_3H_5$	7.16	6.94	0.22	1.90	4.09	4.36

c. Miscellaneous. *IC_{50} and K_i Data of Isostere (**72**) for Inhibition of Recombinant HIV-Protease in CEM Cells (Table 68).*[132] Study of (*R*)-hydroxyethyl sulfonamide isostere derivatives reported by Vazquez et al.[132] gave eqs 86 and 87; they bring out steric effects of X-substituents. Surprisingly, Y did not appear to have any effect. There was little variation in Y-substituents.



72

$$\log 1/C = 11.12(\pm 5.82)B1_X - 11.55(\pm 9.72)$$

$$n = 7,\ r^2 = 0.829,\ s = 0.59,\ q^2 = 0.763 \qquad (86)$$

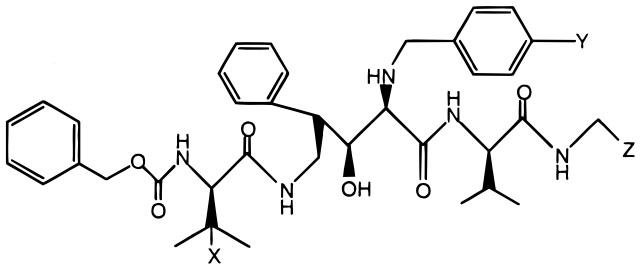
Outliers 2

$$\log 1/K_i = 0.62(\pm 0.30)L_X + 4.06(\pm 1.75)$$

$$n = 6,\ r^2 = 0.894,\ s = 0.40,\ q^2 = 0.819 \qquad (87)$$

Outliers 2

*(ii) K_i and IC_{50} Data of Isosteres (**73**) for Inhibition of HIV-1-Proteinase (Tables 69 and 70).*[133] However,



73

for a series of compounds studied by Billich et al.,[133] the binding data ($K_{i)}$ (Table 69) gave eq 88.

$$\log 1/K_i = -0.65(\pm 0.34)\sigma_Y + 0.34(\pm 0.16)I_Z + 7.93(\pm 0.12)$$

$$n = 9,\ r^2 = 0.899,\ s = 0.10,\ q^2 = 0.779 \qquad (88)$$

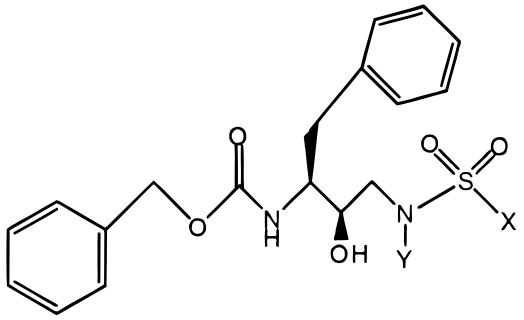
Outliers 2

We also derived eq 89 for the IC_{50} activity (concentration to reduce P24 antigen level in the supernatant of infected cell cultures by 50% HIV) for the same data (Table 70a).[133]

$$\log 1/C = -0.35(\pm 0.40)\sigma^+_Y + 1.02(\pm 0.31)I_Z + 6.80(\pm 0.23)$$

$$n = 11,\ r^2 = 0.891,\ s = 0.22,\ q^2 = 0.791 \qquad (89)$$

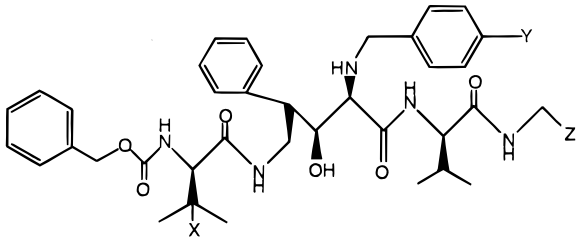
Table 68. IC_{50} and PR Inhibition Data of Isostere Derivatives (72)[132]



	substituents		log 1/*C*			log 1/K_i				
no.	X	Y	obsd (eq 86)	calcd (eq 86)	Δ (eq 86)	obsd (eq 87)	calcd (eq 87)	Δ (eq 87)	$B1_X$	L_X
1	Me	$CH_2CH_2CHMe_2$	5.50	5.36	0.14	5.73	5.84	−0.11	1.52	2.87
2	C_3H_7	$CH_2CH_2CHMe_2$	6.68[a]	5.36	1.32	6.89	7.11	−0.22	1.52	4.92
3	C_6H_5	$CH_2CH_2CHMe_2$	7.80	7.47	0.32	8.00	7.95	0.05	1.71	6.28
4	C_6H_5	$CH_2CH_2CHMe_2$(S)	6.35[a]	7.47	−1.13	6.53[a]	7.95	−1.42	1.71	6.28
5	$CH_2C_6H_5$	$CH_2CH_2CHMe_2$	5.00	5.36	−0.36				1.52	
6	C_6H_4-4-Cl	$CH_2CH_2CHMe_2$	8.00	8.48	−0.48	8.39	8.86	−0.47	1.80	7.74
7	C_6H_5	CH_2CHMe_2	8.22	7.47	0.75	8.50	7.95	0.54	1.71	6.28
8	C_6H_5	$CH_2C_6H_5$	6.72	7.47	−0.75	6.90[a]	7.95	−1.05	1.71	6.28
9	C_6H_5	CH_2−Cy−C_6H_{11}	7.85	7.47	0.38	8.17	7.95	0.22	1.71	6.28

[a] Data points not used in deriving respective equations.

Table 69. PR Inhibition Data of Isostere Derivatives (73)[133]



	substituents			log 1/K_i				
					calcd			
no.	X	Y	Z	obsd	(eq 88)	Δ	σ_Y	I_Z
1	H	H	C_6H_5[b]	8.21	7.93	0.29	0.00	0.00
2	Me	H	C_6H_5	8.04	7.93	0.11	0.00	0.00
3	H	OMe	C_6H_5	8.13	8.11	0.03	−0.26	0.00
4	H	Cl	C_6H_5	7.77	7.80	−0.01	0.19	0.00
5	H	Br	C_6H_5	7.64	7.75	−0.14	0.25	0.00
6	H	H	benz[a,b]	7.89	8.26	−0.38	0.00	1.00
7	H	OMe	benz	8.47	8.45	0.03	−0.26	1.00
8	H	Cl	benz	8.22	8.13	0.10	0.19	1.00
9	H	Br	benz	8.10	8.09	−0.02	0.25	1.00
10	Me	OMe	benz	8.34	8.45	−0.11	0.26	1.00
11	Me	Cl	benz	8.13	8.13	0.01	0.19	1.00

[a] 2-Benzimidazolyl. [b] Data point not used in deriving equation.

For the same data (Table 70b),[133] IC_{50} data (concentration to reduce virus-induced cytopathic effect by 50% in MT-4 cells) gave eq 90. The σ^+_Y term is of marginal value.

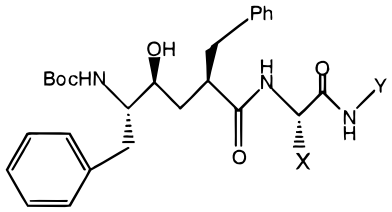
$$\log 1/C = -0.77(\pm 0.55)\sigma^+_Y + 0.69(\pm 0.39)I_Z + 6.45(\pm 0.28)$$

$$n = 10,\ r^2 = 0.866,\ s = 0.24,\ q^2 = 0.727 \quad (90)$$

Outlier 1

Equations 88−90 suggest that an electron-donating 4-substituent (Y-substituent) at the aryl ring may have an advantageous effect on both the antiviral and the protease inhibition activities. These equations also indicated that of the two Z-substituents, 2-benzimidazolyl and phenyl groups, tried at the terminal of the chain, the former would be superior to the latter, as the indicator variable I_Z used in the equation takes a value of 1 for the former and 0 for the latter. The X-substituents, being either H or CH_3, were not found to make any difference. The superiority of the benzimidazolyl group can be attributed to its nitrogens that can participate in the hydrogen bonding with the receptor.

(iii) IC_{50} Data of Isostere Derivatives ***(74)*** *for Inhibition of HIV-Protease in H9 Human T-Lymphocyte Cells (Table 71).*[134] Studies reported by Desolms



74

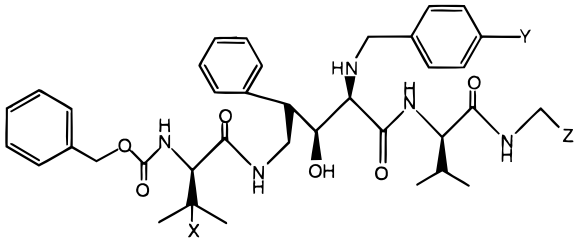
et al.[134] on **74** in which variations of the carboxy terminus of the HIV protease inhibitor L-682,679 were studied on inhibition data gave eq 91, which shows that overall bulky molecules favor activity while there is a negative steric effect of large X-substituents.

$$\log 1/C = 2.56(\pm 1.23)B1_X - 0.32(\pm 0.29)B5_X + 1.36(\pm 0.64)\mathrm{MgVol} - 1.00(\pm 2.81)$$

$$n = 31,\ r^2 = 0.707,\ s = 0.67,\ q^2 = 0.623 \quad (91)$$

Outlier 1

Table 70. IC_{50} of Isostere Derivatives (73)[133]



(a) Concentrated To Reduce P24 Antigen Level in the Supernatant of Infected Cell Cultures by 50% HIV

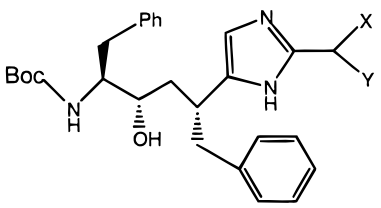
	substituents			log 1/*C*				
no.	X	Y	Z	obsd (eq 89)	calcd (eq 89)	Δ (eq 89)	σ^+_Y	I_Z
1	H	H	C_6H_5	6.50	6.80	−0.30	0.00	0.00
2	Me	H	C_6H_5	7.02	6.80	0.22	0.00	0.00
3	H	OMe	C_6H_5	6.88	7.08	−0.20	−0.78	0.00
4	H	Cl	C_6H_5	6.96	6.76	0.20	0.11	0.00
5	H	Br	C_6H_5	6.83	6.75	0.08	0.15	0.00
6	H	H	benz[a]	7.88	7.83	0.05	0.00	1.00
7	H	OMe	benz	8.14	8.10	0.04	−0.78	1.00
8	H	Cl	benz	7.46	7.79	−0.33	0.11	1.00
9	H	Br	benz	7.69	7.77	−0.08	0.15	1.00
10	Me	OMe	benz	8.27	8.10	0.17	−0.78	1.00
11	Me	Cl	benz	7.95	7.79	0.16	0.11	1.00

(b) Concentrated To Reduce Virus-Induced Cytopathic Effect by 50%

	substituents			log 1/*C*				
no.	X	Y	Z	obsd	calcd (eq 90)	Δ	σ^+_Y	I_Z
1	H	H	C_6H_5	6.24	6.45	−0.21	0.00	0.00
2	Me	H	C_6H_5	6.50	6.45	0.05	0.00	0.00
3	H	OMe	C_6H_5[b]	6.50	7.04	−0.55	−0.78	0.00
4	H	Cl	C_6H_5	6.40	6.36	0.04	0.11	0.00
5	H	Br	C_6H_5	6.46	6.33	0.13	0.15	0.00
6	H	H	benz[a]	6.85	7.14	−0.28	.00	1.00
7	H	OMe	benz	7.60	7.73	−0.13	−0.78	1.00
8	H	Cl	benz	6.92	7.05	−0.13	0.11	1.00
9	H	Br	benz	6.96	7.02	−0.06	0.15	1.00
10	Me	OMe	benz	7.92	7.73	0.19	−0.78	1.00
11	Me	Cl	benz	7.47	7.05	0.42	0.11	1.00

[a] 2-Benzimidazolyl. [b] Data point not used in deriving equation.

*(iv) K_i Data of Isostere Derivatives (**75**) for Inhibition of HIV-Protease (Table 72).*[135] Thompson et al.[135]



75

studied another series of hydroxyethylene-based HIV-1 protease inhibitors containing heterocyclic $P'_1-P'_2$ amide bond isostere (**75**) derivatives, where X-substituents were either a keto (=O) or hydroxyl (OH) group while Y-substituents were mainly alkyl.

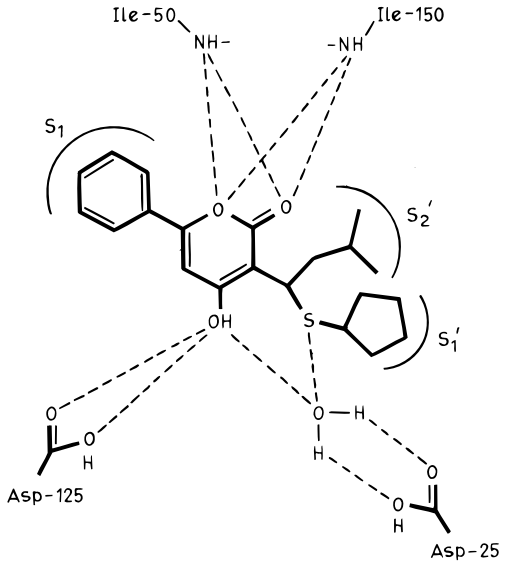


Figure 7. Model of binding of arylthiomethanes with HIV-1 protease based on X-ray crystallographic studies. S_1, S_1', S_2', etc., can be van der Waals or hydrophobic sites. (Reprinted with permission from ref 131. Copyright 1995 American Chemical Society.)

QSAR studies on K_i data by us gave eq 92.

$$\log 1/K_i = 1.26(\pm 0.70)I + 0.49(\pm 0.34)B5_Y + 3.94(\pm 1.08)$$
$$n = 8,\ r^2 = 0.885,\ s = 0.33,\ q^2 = 0.683 \qquad (92)$$

In eq 92, indicator variable $I = 1.0$ for a X = keto group and 0.0 for a hydroxy group. Its large positive coefficient shows that a keto group is favored over OH. The positive B5 parameter for Y-substituents indicates steric interactions.

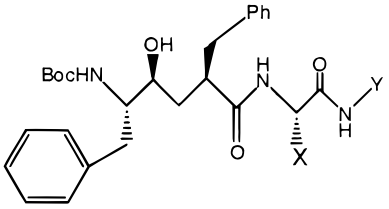
Equations 86−92 have no hydrophobic terms, while there are certainly hydrophobic binding sites at the receptor. Could it be that most of the equations are based on a small number of data points. The variation in the substituents also does not allow for much choice in the use of different physicochemical parameters. A minimum of five data points with good variation in substituents per parameter is required to derive a reliable and meaningful QSAR. Only the data supporting eq 91 could be studied in detail and it does bring out the specific binding properties of substituents.

2. Peptidic Inhibitors

Figure 4 shows how a peptidic inhibitor, i.e., a substrate-based HIV-1-PR inhibitor, can interact with the receptor. A substrate-based inhibitor can be obtained by replacing the scissile P_1-P_1' amide bond of the substrate (Figure 3) by a nonhydrolyzable isostere with tetrahedral geometry. A number of such isosteres that have been studied are shown in Figure 8.

The natural product pepstatin A (Val-Val-Val-Sta-Ala-Sta),[138] in which the isostere Statine occurs twice, is considered to be a typecasting inhibitor of aspartic acid proteases. It was also demonstrated that it could block the proteolytic action of several retroviral proteases[139] and inhibit hydrolysis of both polyprotein

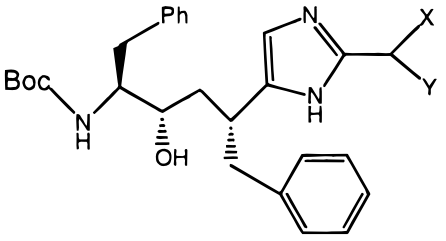
Table 71. IC_{50} Data of L-682679 Isostere Derivatives (74)[134]



	substituents		log 1/*C*					
no.	X	Y	obsd	calcd (eq 91)	Δ	$B1_X$	$B5_X$	MgVol
1	$CHMe_2$	H	8.96	8.53	0.43	1.90	3.15	4.18
2	H	H	6.26	6.34	−0.08	1.00	1.00	3.76
3	Me	H	8.16	7.53	0.62	1.52	2.04	3.90
4	CH_2CHMe_2	H	8.00	7.33	0.67	1.52	4.45	4.32
5	$CH(Me)C_2H_5$	H	8.57	8.61	−0.05	1.90	3.49	4.32
6	C_6H_5	H	9.05	8.31	0.74	1.71	3.11	4.37
7	$Cy-C_6H_{11}$	H	8.28	8.87	−0.59	1.91	3.49	4.50
8	$CH_2C_6H_5$	H	7.54	7.08	0.46	1.52	6.02	4.51
9	$CH_2-Cy-C_6H_{11}$	H	7.19	7.45	−0.25	1.52	5.42	4.64
10	CH_2OH	H	6.22	7.40	−1.18	1.52	2.70	3.96
11	d-CH_2CH-Me_2	H	5.51	7.33	−1.82	1.52	4.45	4.32
12	CH_2CHMe_2	Phe-NH_2	9.22	8.89	0.34	1.52	4.45	5.47
13	$CH(Me)C_2H_5$	Phe-NH_2	9.60	10.17	−0.57	1.90	3.49	5.47
14	CH_2CHMe_2	$CH_2C_6H_5$	8.85	8.35	0.51	1.52	4.45	5.07
15	$CH(Me)C_2H_5$	$CH_2C_6H_5$	9.15	9.63	−0.48	1.90	3.49	5.07
16	$CHMe_2$	$CH_2C_6H_5$	9.62	9.54	0.08	1.90	3.17	4.93
17	C_6H_5	$CH_2C_6H_5$[d]	8.24	9.32	−1.09	1.71	3.11	5.12
18	CH_2CHMe_2	$CH_2CH_2C_6H_5$	7.82	8.54	−0.71	1.52	4.45	5.21
19	CH_2CHMe_2	$CH(CH_2OH)CH_2C_6H_5$	8.82	8.81	0.02	1.52	4.45	5.41
20	CH_2CHMe_2	CH_2CH_2OH	8.68	7.79	0.88	1.52	4.45	4.67
21	$CH(Me)-C_2H_5$	CH_2CH_2OH	9.34	9.08	0.26	1.90	3.49	4.67
22	$CHMe_2$	CH_2CH_2OH	9.36	8.99	0.37	1.90	3.17	4.52
23	$CH(Me)-C_2H_5$	$CH_2CH(OH)CH_2OH$	9.82	9.35	0.48	1.90	3.49	4.86
24	$CHMe_2$	$CH_2CH(OH)CH_2OH$	10.30	9.26	1.04	1.90	3.17	4.72
25	$CH(Me)C_2H_5$	CH_2-2-Py[a]	8.89	9.57	−0.69	1.90	3.49	5.03
26	$CHMe_2$	CH_2-2-Py	9.13	9.49	−0.36	1.90	3.17	4.89
27	$CHMe_2$	CH_2-3-Py	9.16	9.49	−0.32	1.90	3.17	4.89
28	$CHMe_2$	CH_2-4-Py	9.21	9.49	−0.28	1.90	3.17	4.89
29	$CH(Me)C_2H_5$	CH_2-2-Imid[b]	9.06	9.39	−0.33	1.90	3.49	4.89
30	$CHMe_2$	CH_2-2-Benz[c]	9.70	9.80	−0.10	1.90	3.17	5.12
31	C_6H_5	CH_2-2-Benz	10.16	9.58	0.58	1.71	3.11	5.31
32	$CH(Me)C_2H_5$	CH_2-2-Benz	10.22	9.89	0.34	1.90	3.49	5.26

[a] Pyridine. [b] Imidazoline. [c] Benzimidazoline. [d] Data point not used in deriving equation.

Table 72. PR Inhibition Data of L-682679 Isosteres Derivatives (75)[135]



	substituents		log 1/*C*				
no.	X	Y	obsd	calcd (eq 92)	Δ	I	$B5_Y$
1	=O	H	5.46	5.69	−0.24	1.00	1.00
2	=O	Me	6.43	6.20	0.23	1.00	2.04
3	=O	C_2H_5	7.04	6.76	0.28	1.00	3.17
4	=O	C_3H_7	6.82	6.92	−0.10	1.00	3.49
5	=O	$CHMe_2$	7.08	6.76	0.32	1.00	3.17
6	=O	$CMe_2CH=CH_2$	6.57	7.06	−0.49	1.00	3.78
7	OH	Me	4.88	4.94	−0.07	0.00	2.04
8	OH	$CHMe_2$	5.57	5.50	0.07	0.00	3.17

and oligopeptide substrates by HIV-1 protease.[140−143] Acetyl pepstatin (Ac-Val-Val-Sta-Ala-Sta) was, however, found to be more potent against HIV-1 PR than pepstatin A.[141] The crystal structures of enzyme−inhibitor complexes have provided deeper insight into the mechanism of the PR inhibition. Structures of the enzyme complexed with four structurally different peptide isosteres[144−147] exhibited that all four inhibitors were bound in an extended conformation, spanning from P4 to P3′. An extensive network of hydrogen bonds could be illustrated between the enzyme and the polar atoms in the inhibitor. These postulated hydrogen bonds are formed primarily with backbone atoms of the floor and flap regions of HIV proteases (Figure 4). One striking feature of all four inhibitor complexes is that a tightly bound water molecule bridges the two enzyme flaps to the inhibitor through hydrogen bonds formed by the Ile50 and Ile50′ amide hydrogens and P_2 and P_1' carbonyl oxygens of the inhibitors. The binding pockets discernible from P_2 to P_2' are comprised almost entirely of hydrophobic residues in the enzyme.[144−146] According to Huff,[148] the inhibitor−enzyme binding is dominated by hydrophobic interactions.

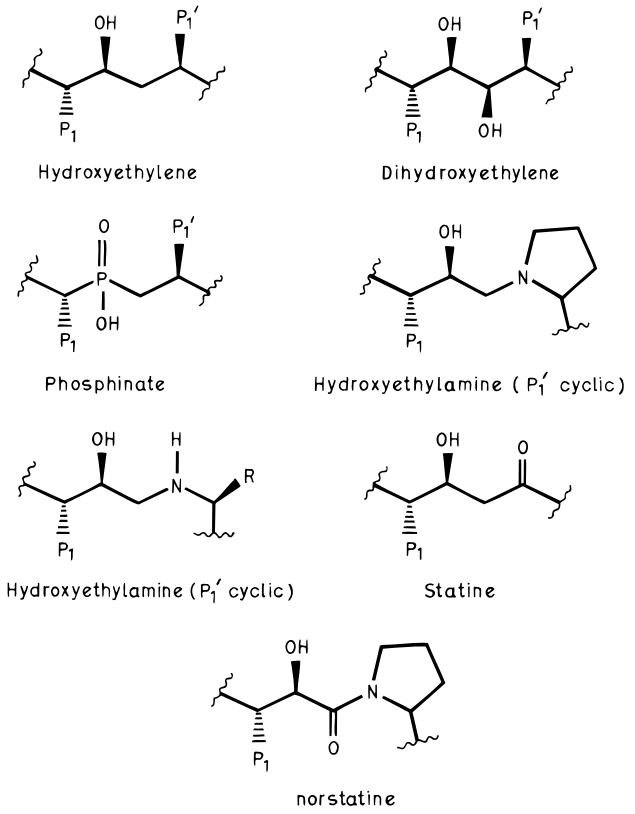
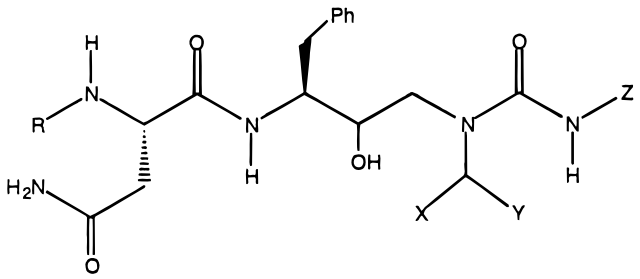


Figure 8. Some nonhydrolyzable transition-state isosteres employed to replace the P_1-P_1' amide bond of the substrate for the design of HIV-1-PR inhibitors.

*(i) IC_{50} Data of Urea Isostere Derivatives (**76**) for Inhibition of Recombinant HIV-Protease (Table 73).*[149]



76

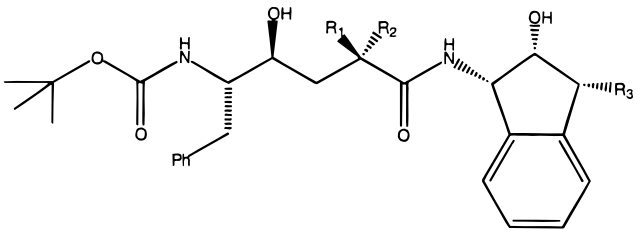
In a series of (*R*)-hydroxyethylurea isosteres (**76**) (Table 73), a representative compound (**5**) was observed to interact with the hydrophobic regions of the receptor as shown in Figure 9.[149] However, for this series of compounds, we correlated the inhibition activity as given in eq 93

$$\log 1/C = 2.27(\pm 0.50)I_X - 0.75(\pm 0.34)I_R + 1.26(\pm 0.33)B1_Z + 2.70(\pm 1.09)$$

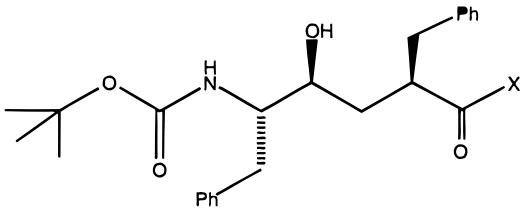
$$n = 18,\ r^2 = 0.941,\ s = 0.29,\ q^2 = 0.912 \quad (93)$$

where $I_X = 1$ stands for X = Me and $I_R = 1$ stands for R = Cbz (carbobenzyloxy). This equation illustrates that while the width of the Z-substituent interacting with the S_2' site of the receptor would be beneficial, a CHMeY group interacting with the S_1' site will have a negative effect relative to the CH_2Y group and the Cbz group, as an R-substituent, interacting with the S_3 site will be detrimental to the activity. There seem to be some hydrophobic interactions also, as adding a Clog *P* term improves the correlation slightly ($r^2 = 0.96$), but it was marginal, having a coefficient of 0.24 (±0.17).

*(ii) IC_{50} Data of Protease Inhibitors (**77**, **78**) for Inhibition of HIV-Protease (Table 74).*[150] For a combined series of isoestere derivatives of **77** and **78**, Holloway et al.[150] observed a high correlation between the intermolecular interaction energy (E_{int}) calculated for HIV−PR inhibitor complexes and the observed in vitro enzyme inhibition activity. The native and the



77



78

acetylpepstatin and L-689,502 inhibited HIV-1 protease. X-ray coordinates and the force field technique were employed in the calculation of E_{int} (intermolecular interaction energy) and the correlations obtained were

Native

$$\log 1/C = -0.15435,\ E_{int} - 8.069$$

$$n = 33,\ r^2 = 0.728,\ r_{cv} = 0.831 \quad (94)$$

Acetylpepstatin inhibited

$$\log 1/C = -0.17302,\ E_{int} - 14.90$$

$$n = 33,\ r^2 = 0.581,\ r_{cv} = 0.724 \quad (95)$$

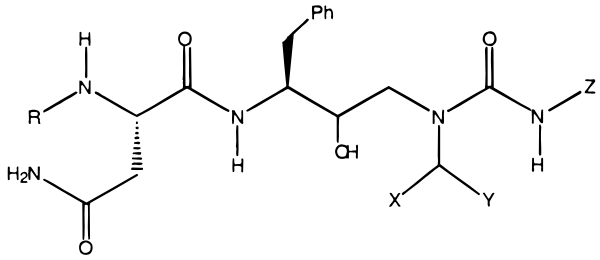
L-689,502 inhibited

$$\log 1/C = -0.16946,\ E_{int} - 15.707$$

$$n = 33,\ r^2 = 0.783,\ r_{cv} = 0.869 \quad (96)$$

In eqs 94−96, r_{CV} is the cross-validated *r*. For the same series of compounds (Table 74), we correlated

Table 73. IC_{50} Data of Urea Isosteres (76)[149]



	substituents				log 1/*C*					
no.	R	X	Y	Z	obsd	calcd (eq 93)	Δ	I_R	I_X	$B1_Z$
1	Cbz[a]	H	$CHMe_2$	Me	5.82	6.15	−0.33	1.00	1.00	1.52
2	Cbz	H	$CHMe_2$	n-Bu	6.03	6.15	−0.12	1.00	1.00	1.52
3	Qua[b]	H	$CHMe_2$	n-Bu	6.90	6.90	0.01	0.00	1.00	1.52
4	Cbz	H	$CHMe_2$	n-Pr	6.29	6.15	0.14	1.00	1.00	1.52
5	Cbz	H	$CHMe_2$	Et	6.48	6.15	0.33	1.00	1.00	1.52
6	Cbz	H	$CHMe_2$	i-Pr	6.59	6.63	−0.04	1.00	1.00	1.90
7	Cbz	H	$CHMe_2$	t-Bu	7.46	7.51	−0.06	1.00	1.00	2.60
8	Qua	H	$CHMe_2$	t-Bu	8.22	8.26	−0.04	0.00	1.00	2.60
9	Cbz	H	CH_2CHMe_2	t-Bu	7.89	7.51	0.37	1.00	1.00	2.60
10	Qua	H	CH_2CHMe_2	t-Bu	8.52	8.26	0.26	0.00	1.00	2.60
11	Cbz	H	C_6H_{11}	t-Bu	7.54	7.51	0.03	1.00	1.00	2.60
12	Qua	H	C_6H_{11}	t-Bu	8.30	8.26	0.04	0.00	1.00	2.60
13	Cbz	H	C_6H_5	t-Bu	7.72	7.51	0.21	1.00	1.00	2.60
14	Qua	H	C_6H_5	t-Bu	8.52	8.26	0.26	0.00	1.00	2.60
15[c]	Cbz	Me	C_6H_5	t-Bu	5.19	5.24	−0.05	1.00	0.00	2.60
16[d]	Cbz	Me	C_6H_5	t-Bu	5.29	5.24	0.05	1.00	0.00	2.60
17	Cbz	H	4-Py	t-Bu	6.98	7.51	−0.53	1.00	1.00	2.60
18	Qua	H	4-Pr	t-Bu	7.72	8.26	−0.54	0.00	1.00	2.60

[a] Carbobenzyloxy. [b] Quinolinyl-2-carboxamide. [c] CHXY in *R*-configuration. [d] CHXY in *S*-configuration.

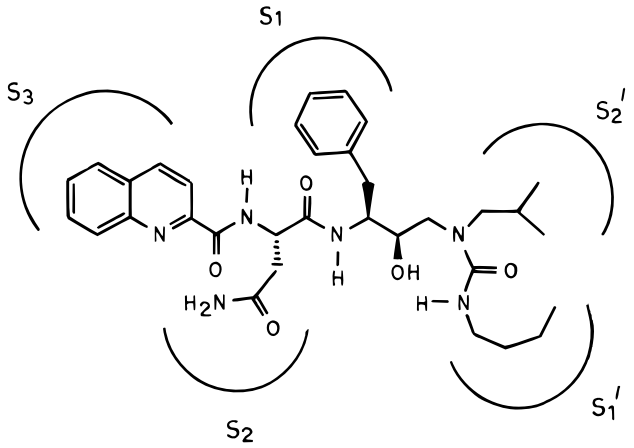


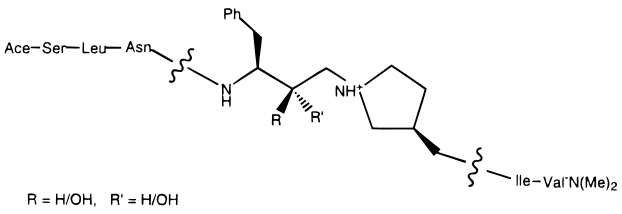
Figure 9. Schematic representation of binding of a representative compound (**5**, Table 73) with hydrophobic regions of HIV-1 protease based on observations. (Reprinted with permission from ref 149. Copyright 1993 American Chemical Society.)

the enzyme inhibition data with molar refractivity as

$$\log 1/C = 1.46(\pm 0.53)\mathrm{CMR} - 1.56(\pm 1.19)\log(\beta \times 10^{\mathrm{CMR}} + 1) + 2.38(\pm 0.64)I + 15.64(\pm 7.98)$$
$$n = 30,\ r^2 = 0.816,\ s = 0.69,\ q^2 = 0.759 \quad (97)$$
$$(\mathrm{CMR})_o = 16.87,\ \log \beta = -15.72,\ \text{Outliers } 3$$

suggesting that the molecule may be involved in dispersion interactions with the receptor but very bulky molecules may be unfavorable. In this correlation, $I = 1$ stands for an X moiety which has an OH group cis to its NH group as shown in **77**. A high positive coefficient with this variable suggests that the presence of OH cis to NH would increase the inhibition potency, possibly due to being in the proper orientation to form a hydrogen bond with the receptor.

The determination of the relative free energies for the binding of peptide inhibitors of the type **79** with

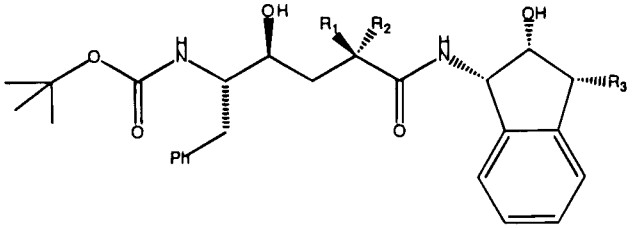


79

HIV-1 protease led Ferguson et al.[151] to suggest that the ethylamine hydroxyl group produces marked stabilization of the enzyme−inhibitor complex due to hydrogen bonding with aspartyl residues. This interaction was shown to induce a conformational change in the *R* diasteromer that resulted in a decrease in binding affinity. The addition of a second hydroxyl group to the inhibitor might help avoid the conformational requirements for binding that depend on the configuration of the inhibitor. It has been shown that the binding of symmetric glycol-containing inhibitors is not dependent on the configuration of the two carbon centers with which the hydroxyl groups are attached, and it was suggested that this

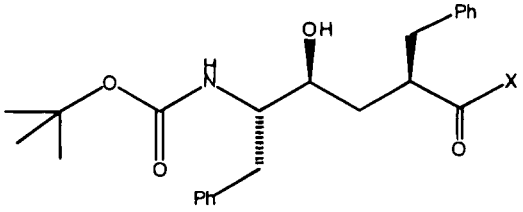
Table 74. Substituents, IC_{50} Data, and Physicochemical Parameters[150] for Protease Inhibitors (77 and 78)

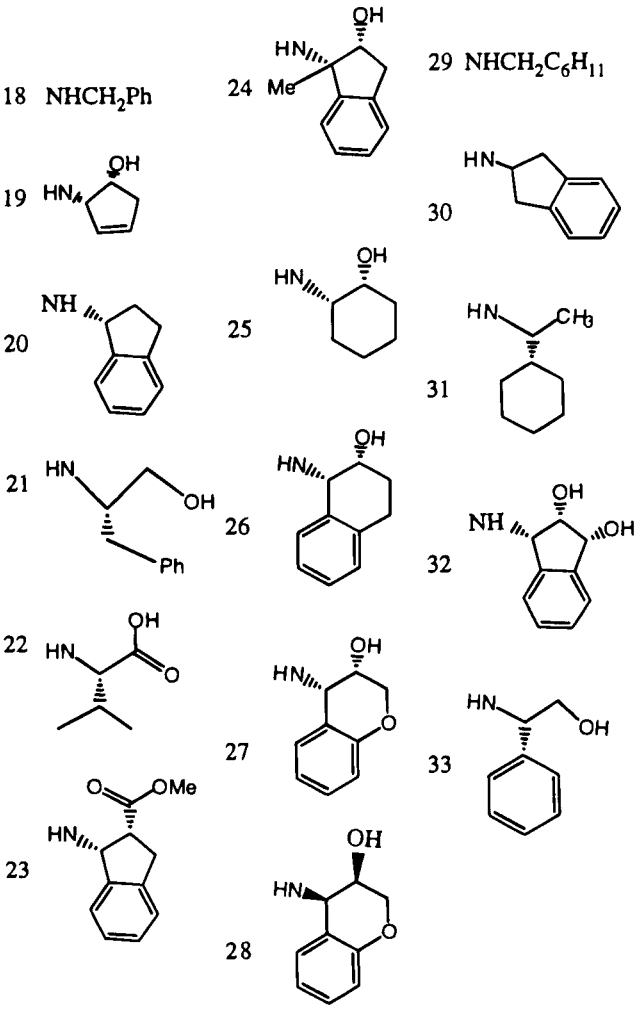
(a) Substituents R_1, R_2, and R_3 for Protease Inhibitors (**77**) (Compounds **1**−**17** Used for Deriving eq 97)[150]



no.	R_1	R_2	R_3
1	CH_2Ph	H	H
2	CH_2Ph	Me	H
3	CH_2CH_2Ph	H	OH
4	CH_2-4-CF_3Ph	H	H
5	(E)CH_2CH=$CHPh$	H	H
6	$CH_2C_6F_5$	H	H
7	CH_2-4-CH_3Ph	H	H
8	CH_2-4-NH_2Ph	H	H
9	CH_2-4-NO_2Ph	H	H
10	H	H	H
11	CH_2-4-OHPh	H	H
12	CH_2CH=CH_2	H	H
13	CH_2-4-IPh	H	H
14	$CH_2C(O)Ph$	H	H
15	CH_2-4-pyridyl	H	H
16	CH_2SPh	H	H
17	CH_2-4-CMe_3Ph	H	H

(b) Substituent X for Protease Inhibitors (**78**) (Compounds **18**−**33** Used for Deriving eq 97)[150]





(c) IC_{50} Data and Physicochemical Parameters

	log 1/*C*				
no.	obsd	calcd (eq 97)	Δ	CMR	I
1	9.60	9.20	0.40	15.76	1.00
2	8.11	9.41	−1.29	16.22	1.00
3	9.72	9.48	0.24	16.84	1.00
4	9.59	9.42	0.17	16.27	1.00
5	9.64	9.48	0.16	16.74	1.00
6	9.22	9.25	−0.02	15.84	1.00
7	9.54	9.41	0.13	16.22	1.00
8	9.51	9.38	0.13	16.13	1.00
9	9.57	9.44	0.13	16.37	1.00
10	5.53	5.37	0.17	12.78	1.00
11	9.80	9.29	0.51	15.91	1.00
12	7.56	7.34	0.22	14.15	1.00
13	9.14	9.47	−0.33	17.07	1.00
14	8.27	9.42	−1.15	16.26	1.00
15	9.28	9.04	0.23	15.55	1.00
16	9.60	9.47	0.14	16.57	1.00
17	9.77	9.44	0.33	17.61	1.00
18	6.94	5.92	1.03	14.86	0.00
19	8.02	7.34	0.68	14.15	1.00
20	7.47	6.71	0.75	15.61	0.00
21	6.16	7.06	−0.90	16.40	0.00
22	6.79	7.67	−0.88	14.39	1.00
23	7.18	7.10	0.08	16.72	0.00
24	6.67[a]	−14.91	21.59	16.22	1.00
25	6.91	8.02	−1.10	14.64	1.00
26	9.16	9.20	−0.04	15.76	1.00
27	9.75	9.29	0.46	15.91	1.00
28	7.39[a]	−14.88	22.27	15.91	1.00
29	4.52	6.04	−1.51	14.95	0.00
30	6.89	6.71	0.18	15.61	0.00
31	6.84	6.45	0.38	15.32	0.00
32	10.00	9.29	0.71	15.91	1.00
33	7.41[a]	−14.84	22.25	15.47	1.00

[a] Data points not used in deriving equation.

may be due to flexibility of the inhibitors at the hydroxyl-carbon centers or a result of the availability of multiple binding modes for the diastereomers, afforded by the presence of a second hydroxyl group

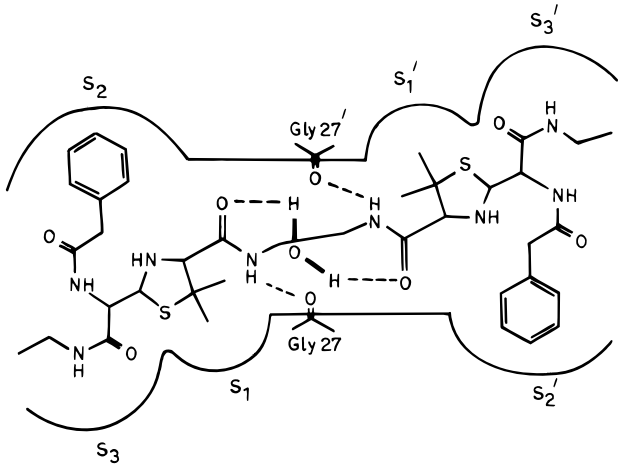
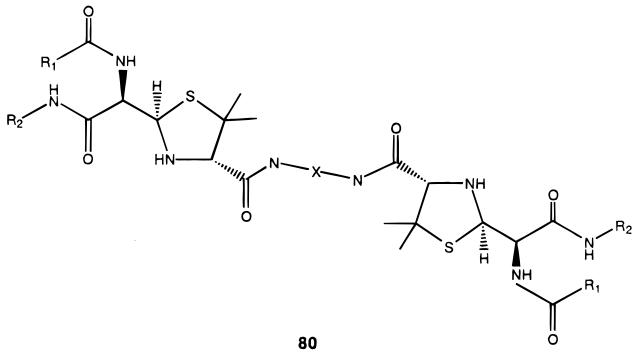


Figure 10. Model proposed for the binding of a penicillin-derived C_2 symmetric dimer inhibitor with HIV protease. (Reprinted with permission from ref 153. Copyright 1993 American Chemical Society.)

in the active site.[152] However, one must always question the hydrogen-bonding effect since the free energy change in the OH binding to water and to the hydrogen bonding component in the receptor would probably be small. Multiple H-bonding would, of course, be more significant.

Studies on the binding modes of a series of penicillin-based C_2-symmetric dimer inhibitors (**80**) led



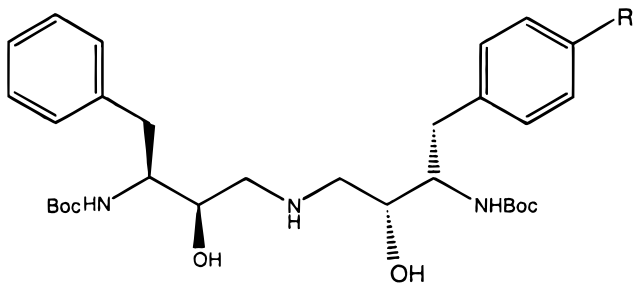
80

some authors[153,154] to suggest that these inhibitors would bind in a symmetrical fashion, tracing an S-shaped course through the active site, with good hydrophobic interactions in the S_1/S_1' and S_2/S_2' pockets and hydrogen bonding of inhibitor amide groups (Figure 10).[153] Interactions with the catalytic aspartates were found to be poor and the protein conformation to be very similar to that observed in complexes with peptidomimetics, despite the major differences in ligand structure.[153]

Results of some three-dimensional QSAR studies,[155−157,180,181] using the CoMFA (comparative molecular field analysis) model, on different kinds of peptide isosteres and cyclic ureas were found to support almost a common mode of binding and to stress the involvement of steric and electrostatic interactions.

In protease−inhibitor binding, it has been found, however, that the protease is singly protonated.[151,158] Though hydrogen bonding plays a crucial role in the stabilization of protease−inhibitor complexes, adequate treatment of the enzyme active site protonation state is important for their accurate molecular simulations. Calculations have shown that in HIV-1-Pr−inhibitor complexes, only one catalytic aspartic acid residue is protonated.[151,158]

For a series of C_2-symmetric aminodiols (**81**), however, a high degree of correlation was observed between lipophilicity, measured by reverse-phase HPLC constant k', and the cytotoxicity (CC_{50}) of the compounds (eq 98).[159]

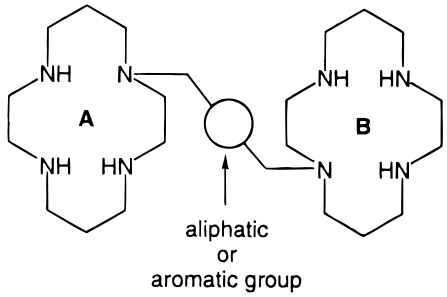


81

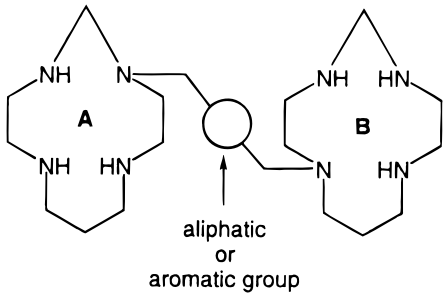
$$\log 1/C = 1.1886 \log k' - 2.7466$$
$$n = 174,\ r^2 = 0.702 \qquad (98)$$

C. Virus Uncoating Inhibitors

A series of bis-tetraazamacrocyclic compounds, consisting of two cyclam units linked in the way shown in **82** or **83** via an aliphatic linker or a linker containing an aromatic moiety, were evaluated for their anti-HIV and cytotoxic effects.[160] A partial least-



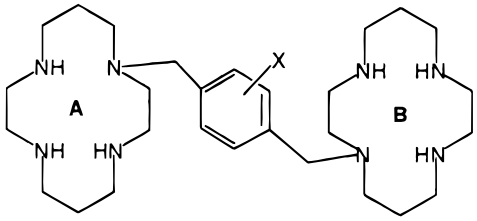
82



83

squares analysis was then performed on these com-

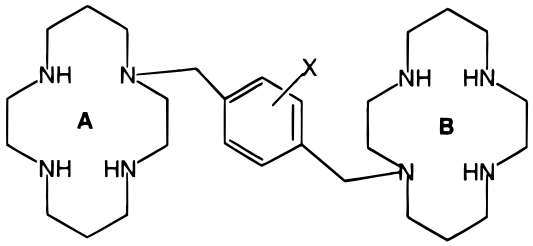
Table 75. EC_{50} Data of Virus Uncoating Inhibitors (84)[160] against HIV-ROD



		log 1/*C*				
no.	X[a]	obsd (eq 99)	calcd (eq 99)	Δ (eq 99)	I	MR_2
1	2,5-diMe(A)	8.96	8.45	0.51	1.00	0.57
2	2,5-di-Cl(A)	8.60	8.43	0.18	1.00	0.60
3	2-Br(A)	8.46	8.25	0.21	1.00	0.89
4	2-C_6H_5(A)	7.10	7.21	−0.11	1.00	2.54
5	2-NO_2(A)[b]	7.14	8.34	−1.21	1.00	0.74
6	2,5-di-OMe(A)	8.18	8.31	−0.13	1.00	0.79
7	2,3,5,6-tetra-F(A)	8.10	8.75	−0.65	1.00	0.09
8	5-C_6H_5(B)	7.61	7.27	0.34	0.00	0.10
9	2-Br(B)	6.61	6.78	−0.17	0.00	0.89
10	5-Br(B)	7.27	7.27	0.00	0.00	0.10
11	5-NO_2(B)	7.25	7.27	−0.03	0.00	0.10
12	2-F(B)	7.13	7.28	−0.14	0.00	0.09

[a] A = 1,4-phenyl ring. B = 1,3-phenyl ring. [b] Data point not used in deriving equation.

Table 76. EC_{50} Data of Virus Uncoating Inhibitors (84)[160] against HIV-1 IIIB



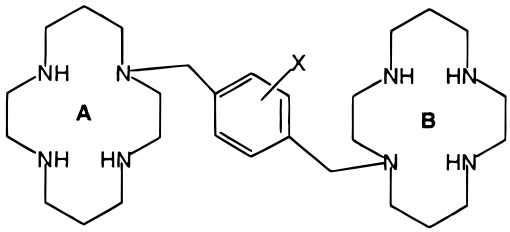
		log 1/*C*				
no.	X[a]	obsd (eq 100)	calcd (eq 100)	Δ (eq 100)	I	CMR
1	2,5-di-Me(A)	8.19	8.01	0.18	1.00	16.42
2	2,5-di-Cl(A)	7.97	7.99	−0.02	1.00	16.47
3	2-Br(A)	8.22	8.07	0.15	1.00	16.26
4	2-C_6H_5(A)	6.97	7.38	−0.41	1.00	18.00
5	2-NO_2(A)[b]	7.19	8.14	−0.95	1.00	16.10
6	2,5-di-OMe(A)	8.24	7.89	0.35	1.00	16.72
7	2,3,5,6-tetra-F(A)	8.10	8.35	−0.25	1.00	15.55
8	5-C_6H_5(B)	6.69	6.47	0.22	0.00	18.00
9	2-Br(B)	6.86	7.16	−0.30	0.00	16.26
10	5-Br(B)	7.07	7.16	−0.09	0.00	16.26
11	5-NO_2(B)	7.39	7.22	0.17	0.00	16.10
12	2-F(B)	7.46	7.46	0.00	0.00	15.50

[a] A = 1,4-phenyl ring. B = 1,3-phenyl ring. [b] Data point not used in deriving equation.

pounds,[161] resulting in models with high predictive abilities for antiviral activities ($r_{cv}^2 = 0.76$ against HIV-1 and $r_{cv}^2 = 0.70$ against HIV-2). The best descriptors deduced from the analysis were the metal affinities for both rings A and B, the metal−metal distance in the complex, ring size (atoms in the ring), and the angle and torsion between the planes defined to represent the face of each macrocyclic ring. Since the high binding of azamacrocycles for transition metals is well established, the metal coordination based mechanism for antiviral action of biscyclams cannot be ruled out and hence the complex related parameters were introduced. An extensive study of the antiviral properties of a variety of cyclam derivatives and their metal complexes has been recently reported.[162]

The above analysis by Joao et al.[161] was initially performed on 37 compounds. But for an extended series of 80 compounds (additional compounds to be reported), these authors found, using the same descriptors, $r_{cv}^2 = 0.79$ for anti-HIV-1 activity.[161] Thus, on the basis of their analysis, Joao et al. proposed the following structural requirements for antiviral activity of bismacrocycles. (a) Molecules require two chelating macrocyclic rings for high activity; (b) Distance between metal-binding centers must be 9.5−11.5 Å; (c) Plane torsions of −60° to −30° and 120° to 140° are allowed; (d) Plane angles of 40° to 70° and 110° to 140° are allowed; (e) Maximize metal affinity for each macrocyclic ring; (f) Optimum ring size for cyclam rings is 14 atoms.

Earlier, Bridger et al.[160] reported that the activity of bicyclam analogues was insensitive to the electronic properties of substituents introduced at the aryl ring in an aryl-containing linker but could be markedly reduced by sterically hindering groups such as phenyl. Bridge et al. data (EC_{50}, effective concentration of the compound required to protect 50% of the virus-infected MT-4 cells against viral cytopathicity) (Tables 75 and 76)[160] on phenylene bis(methylene)-linked bis-tetraazamacrocycles (**84**) that inhibit HIV replication gave eqs 99 and 100, respectively.



84

EC_{50} data for HIV−ROD:

$$\log 1/C = 1.47(\pm 0.54)I - 0.63(\pm 0.39)MR_2 + 7.34(\pm 0.37)$$

$$n = 11,\ r^2 = 0.832,\ s = 0.34,\ q^2 = 0.619 \quad (99)$$

Outlier 1

EC_{50} data for HIV-1 IIIB:

$$\log 1/C = -0.40(\pm 0.24)CMR + 0.91(\pm 0.38)I + 13.60(\pm 3.94)$$

$$n = 11,\ r^2 = 0.841,\ s = 0.27,\ q^2 = 0.604 \quad (100)$$

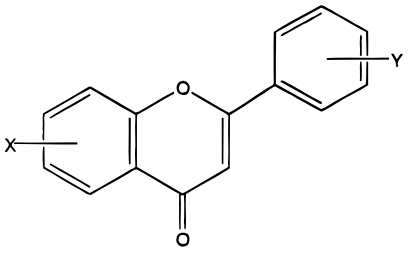
Outlier 1

In eqs 99 and 100, indicator variable *I* is used with a value of 1.0 for 1,4-phenyl (A) and 0.0 for 1,3-phenyl (B) derivatives (see Tables 75 and 76). Its positive coefficient indicates that the 1,4-phenyl linkage enhances inhibitory activity. The negative MR_2 and CMR term in both the equation shows that sterically

large substituents would be detrimental to inhibitory and antiviral activities.

D. Integrase Inhibitors

Many SAR studies are available on integrase inhibitors. Only Raghavan et al.[163] have reported a CoMFA analysis on a set of flavone analogues (**85**) that were found to inhibit HIV-1-integrase-mediated cleavage and integration in vitro. The results have



85

shown a strong correlation between the inhibitory activity of these flavones and the steric and electrostatic fields around them. For 14 compounds r_{cv}^2 was found to be 0.8 with considerably high predictive ability, and the fields were found to contribute steric = 20.5% and electrostatic = 79.5%.

VI. Overview

It seems appropriate to summarize the results in light of the information available about the structure of the receptor with which a particular type of anti-HIV drug interacts. In this respect, the two enzymes, HIV-1 reverse transcriptase and HIV-1 protease, are fairly well studied and only these two enzymes have been, so far, the prime targets for the development of anti-HIV chemotherapy.

A high-resolution electron density map of HIV-1 RT complexed with nevirapine has revealed its structure as an asymmetric dimer.[164] The enzyme is processed initially from the pol gene product as 66-kD (kilodalton) polypeptide that has both a pol and an RNase H (ribonuclease H) domain. But a subsequent proteolytic cleavage of a homodimer of the 66-kD subunits removes the RNase H domain from one subunit, leaving a heterodimer containing one 66-kD subunit (p66) and one 51-kD subunit (p51). The p66−p51 heterodimer appears to have only one pol active site, one RNase H active site, one tRNA binding site, and one nevirapine binding site.[164] The p66 domain possesses a large cleft analogous to that of the Klenow fragment of *E. coli* DNA polymerase, but the p51 domain of an identical sequence has no such domain.

The p66 subunit is folded into five separate subdomains, the four pol domains and one RNase H domain (Figure 11).[164] Its anatomical resemblance to a right-hand has led to naming the subdomains as fingers, palm, and thumb. The fourth pol subdomain lies between the rest of the pol domain and the RNase H domain, leading it to be called "connection" subdomain.

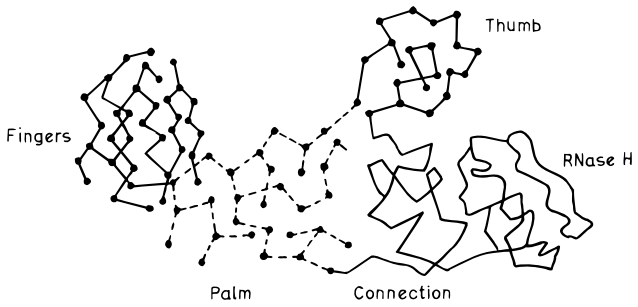


Figure 11. Schematic diagram of poll66 subunit. (Reprinted with permission from ref 164. Copyright 1992 American Association for the Advancement of Science.)

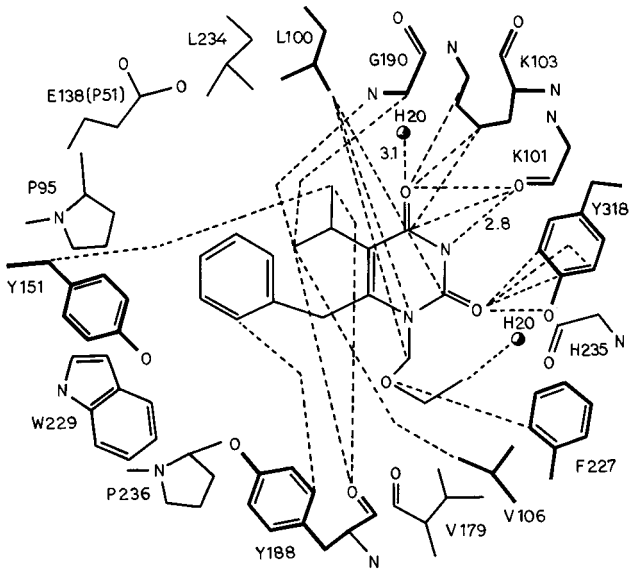


Figure 12. Schematic diagram of binding of a non-nucleoside RT inhibition with RT. (Reprinted with permission from ref 167. Copyright 1996 American Chemical Society.)

While the ddN analogues, which act as competitive inhibitors or alternate substrates of RT, interact at the substrate binding site of the enzyme characterized by its catalytic triad (D110, D185, D186), the NNRTIs bind to a site entirely distinct from it. For them, the binding site is located in a pocket lying in the p66 palm domain.[165] This site is some 10 Å away from the polymerase catalytic triad. Its internal surface is composed mainly of hydrophobic residues with few hydrophilic residues in the vicinity. It is thought that there is a common mechanism of inhibition for all NNRTIs: the displacement of a catalytic triad on binding.[166]

On the basis of their study on the crystal structures of HIV-1 RT complexed with some HEPT analogues, Hopkins et al.[167] represented the binding of an inhibitor (MKC-442) of this class as shown in Figure 12. The hydrophobic nature of the NNRTIs pocket provides relatively few possibilities for polar interactions and hydrogen bonding. However, the geometry of interaction for HEPT analogues appears to be constrained by a strong hydrogen bond from the 3-NH of the pyrimidine ring to the carbonyl oxygen of Lys 101. Additionally, a water molecule present in the NNRTI binding pocket near a channel to the bulk solvent can form a triad of hydrogen bonds between the 4-carbonyl oxygen of the HEPT ana-

logues, the main chain nitrogen of Lys 101, and a carbonyl oxygen of Glu 138 in the p51 chain (Figure 12).

All non-nucleoside inhibitors appear to work by a common inhibitory mechanism.[166] The pocket in which all NNRTIs bind is created, presumably on NNRTI entry, by a conformational switch of key residues to mimic the inactive polymerase site in the p51 chain.[166] While the Tyr181 switch appears to be the major determinant of NNRTI activity, there is another part of the protein structure that changes depending on the NNRTI: the Pro236 loop. This flexible loop forms few contacts with the rest of the protein, but in the presence of an NNRTI, it is drawn toward the inhibitor, apparently due to the hydrophobic nature of the residues in the loop.

Now it is clear from the above discussions that NNRTIs involve in their binding with the receptor mainly the hydrophobic interactions and only a few polar interactions or hydrogen bonding. In this light, we find that most of the QSAR equations reported for NNRTIs involve significant hydrophobic terms.

TIBO Derivatives. Equations 1−3 all contain a very significant hydrophobic term indicating hydrophobic interactions with the receptor. However, the presence of a bilinear term in eq 3 with an optimum value of Clog P = 5.55 indicates that there is a limitation to the size of the hydrophobic pocket. The presence of the I_Z parameter in eqs 1 and 2 suggests a better effect of 2-S than 2-O. Since S is more polarizable than O, this difference in their effects can be attributed to their involvement in polar interactions with the receptor. Besides, there seems to be some specific positive steric interaction at the 8-position of the A-ring.

HEPT Derivatives. The hydrophobic effect of substituents or of the whole molecule has been found to be very significant for the activity in all the equations (4−21) except eqs 5 and 17, which we believe may be due to testing an insufficient number of compounds with good variation in substituents. It is of interest to note that the eqs 10, 11, 13, and 18 show a parabolic correlation with Clog P with an optimum Clog P value ranging from 3.81 to 4.82, emphasizing that there is a limitation to the size of the hydrophobic pocket where HEPT derivatives bind at the reverse transcriptase receptor. The appearance of STERIMOL parameters with positive coefficients in some equations or the presence of volume terms indicate steric interactions. In this light, the presence of the parameters I_6 or I_6' or both in eqs 10−12 and 14−16 for 6-substituents (Table 10) also indicates the steric interaction of these substituents with the enzyme.

TSAO Derivatives. Similarly in the case of TSAO derivatives (eqs 23−30), hydrophobic interactions are involved. The presence of a parabolic Clog P term in eqs 26−28 with an optimum Clog P = 4.21 conforms with previous results. The occurrence of the electronic parameters in several equations does indicate the presence of polar interactions. Also, there are some steric interactions at the N-3 position of the pyrimidine ring (eqs 23b, 24a−c, and 29).

Nevirapine Derivatives. The binding of nevirapine with RT has been studied in detail by electron spectroscopy,[164,168] and it has been observed that it binds in a deep hydrophobic pocket which is formed by the central three β-strands of a five-stranded β-sheet in the palm subdomain and β-meander at the base of the thumb subdomain and is thus in contact with 38 protein atoms of these two subdomains. Thus, the binding involves mainly hydrophobic interactions. In conformity to the above discussion, a very significant hydrophobic term appears in eqs 32, 33, and 35, indicating hydrophobic interactions. In eq 31, the negative Clog P is collinear with MR and MgVol (r^2 = 0.72). However, the electronic effect has been scarcely displayed. Only eq 35 exhibits the involvement of the 7-substituent in some electronic interactions. Equations 31−34 indicate steric effects of ring substituents.

Pyridinone Derivatives. Equations 36, 38, and 40 indicate hydrophobic interactions. Again, a parabolic Clog P term with an optimum value of Clog P = 2.31 in eq 40 conforms with the previous results as to the limitation to the size of the hydrophobic pocket at the receptor. Besides, the inhibitory activity of these derivatives is also governed by steric and electronic effects of the substituents. The length and nature of the linker chain seem to be very important for the activity.

α-APA Derivatives. It is surprising to note the absence of any hydrophobic interactions for these derivatives. Probably the substituents are not able to fit properly in to the hydrophobic pocket. Also, few data points limit our perspective.

ddN Analogues. These analogues bind at the substrate binding site, and the steric and electronic properties of some substituents have been found to govern the activity. Surprisingly, the hydrophobic term was not found to be significant, and that may be due to the increase in their bulk, as the latter is also shown to be unfavorable (eq 59). Equations 55 and 56 which exhibit the dominance of only an electronic parameter cannot be very reliable, as they are based on very small numbers of data points (five only in each case). It is of interest to note that among all the QSAR reported here on ddN derivatives, only eq 57 contains a hydrophobic term.

Several other QSARs (eqs 43−54) have also been reported for different classes of RT inhibitors. Their activity seems to be governed by different physicochemical parameters. Dominance of the hydrophobic term is apparent in eqs 44, 45, 47, 48b−e, 52, and 54, among which eq 44 and 54 show parabolic and bilinear correlation, respectively. Significant electronic and steric effects are also observed in several QSARs.

Although many physicochemical properties of nonnucleoside RT inhibitors can differ, many theoretical studies seem to establish that these inhibitors possess a common three-dimensional feature which has a rigid butterfly-like configuration that fits well into a sizable internal cavity of the allosteric area of the enzyme.[169−173] A large, highly hydrophilic and constrained Ω-loop (Figure 13)[174] was dissected from the allosteric area of HIV-1 RT (segment Tyr181−

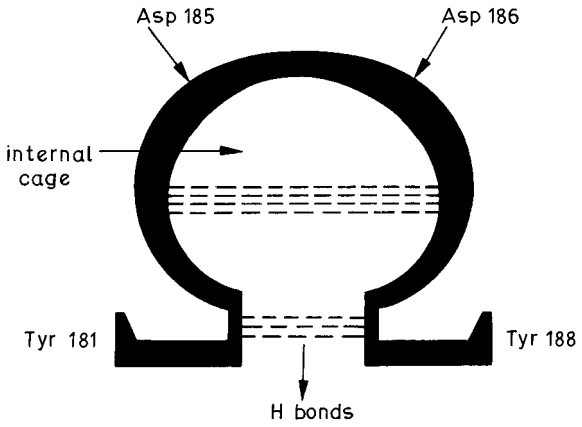


Figure 13. Schematic representation of the Ω-loop. (Reprinted with permission from ref 174. Copyright 1996 Overseas Publisher Association.)

Tyr188).[174,175] The loop contains two amino acids (Asp185 and Asp186) of catalytic aspartyl triad and two amino acids (Tyr181 and Tyr188) of NNRTI binding sites and is stabilized by the hydrogen bonding between CO of Tyr181 and the peptide NH of Tyr188. A butterfly-like configuration of an NNRTI (say nevirapine) can be represented as shown in Figure 14.[170] According to Schäfer et al.,[170] in this configuration an aromatic ring and a second extended π-system should be arranged in a roof-like orientation (or as the wings of butterfly[173]). The distance between the midpoints of the two systems should be 4.5−5 Å, and the angle between the two planes should be around 110°. Additionally, the molecule should contain a carbonyl or thiocarbonyl group, an extra lipophilic site, and a methyl group. The degree of butterfly-like configuration depends on the overall-shape parameters, the polarizability, and the lipophilicity of the molecule.[173] It was found that the butterfly-like shape fits well into a sizable internal cavity of the allosteric area of the enzyme. Structurally diverse NNRTIs interact with different amino acid residues of the allosteric pocket, the number of amino acid residues interacting with an inhibitor is correlated with the degree of butterfly-like configuration of the inhibitor, and thus the drug affinity for the enzyme and the probability of drug-resistance development will be closely correlated with the degree of butterfly-like shape of the ring.

As far as protease inhibition is concerned, it has been discussed already that HIV-1 protease is capable of forming multiple hydrogen bonds with the inhibitors. Studies on HIV-1 protease complexed with diverse inhibitors of different potencies pointed out one common feature: that most of the potent compounds make electrostatic interactions with the catalytic aspartates and displace the water molecule.[34] Potent aspartic protease inhibitors result from a combination of favorable electrostatic interactions with the catalytic aspartates (displacing the water) and favorable protein−ligand interactions in the flanking subsites with minimal strain in the linking groups. Wang et al.,[176] however, observed that at least two additional factors are important in the binding of a compound to HIV-1 PR. The first is the conformational flexibility of the inhibitor molecule, and the second is the hydrophobic interactions between an inhibitor and the enzyme. The HIV-1 PR has four hydrophobic pockets near its active sites, and it has been shown that favorable hydrophobic interactions with these pockets are desirable for an inhibitor to achieve nanomolar potency.[177]

Although both nonspecific hydrophobic interactions and specific hydrogen bondings were found to be important for an inhibitor to achieve good binding affinity, Wang et al.[176] suggested that one should focus, at least initially, on the hydrogen bonding sites when developing a pharmacophore for a 3D database. Hydrophobicity and other considerations, such as conformational energy and chemical novelty, may be more useful in the second stage in which compounds are selected for testing in a bioassay.

Cyclic Ureas. In comparing eqs 63−76, it is worth mentioning that only eqs 64 and 65 contain bilinear hydrophobic terms in which the P/P′ position of **55** is substituted mostly by alkyl groups (Tables 48 and 49). All other series contain X-benzyl substituents at that position. It may be that the rigid phenyl ring blocks the substituents from interacting with hydrophobic space. However, a significant steric interaction with the receptor seems to be involved in almost all of them. But hydrophobic and hydrogen bond interactions both are shown to be equally significant by some other QSAR studies (Figure 5).[120] It has also been proposed that there are possibilities of multiple hydrogen bondings with protease receptors (Figure 6).[122] The hydrophobic interactions have been found to be comparatively less important. Equations 68, 69, 73, and 74 contain a negative σ term for X-substituents on the benzene ring of P/P′ benzyl, which indicates that electron-releasing substituents at that position would enhance inhibitory activity.

Cycloalkylpyranones. Similar to cyclic ureas, we do not observe many hydrophobic terms in eqs 77−85, except in eqs 82 and 84. It seem to us that due to some spatial restrictions these molecules are not able to bind in hydrophobic space as the protease receptor does have hydrophobic binding sites. The findings by several researchers that the optimum size of the cycloalkyl ring should be eight-membered indicate maximal hydrophobic interaction of the ring with the receptor.[128]

Miscellaneous. Equations 86−92 do not show hydrophobic terms, while there are certainly hydrophobic binding sites at receptor. Why it is not clear to us. Could it be that most of the equations are based on a small number of data points? The variation in the substituents also does not allow for much choice in the use of different physicochemical parameters. A minimum of five data points with good variation in substituents is required per parameter to derive a sensible and meaningful QSAR. Only eq 91 could be studied in detail, and it does bring out the specific binding properties of substituents. Only electronic and hydrogen-bond interactions appear to be important, but in a series of arylthiomethanes, hydrophobic interactions have also been postulated (Figure 7).[131]

Peptidic Inhibitors. They have been found to involve predominantly hydrogen bonding. An extensive network of hydrogen bonds has been observed in the structures of the enzyme complexed with some

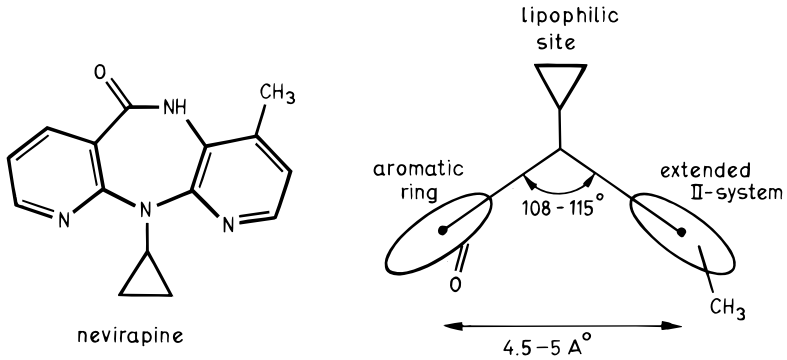


Figure 14. Schematic representation of a butterfly-like configuration of an NNRTI (nevirapine) in accordance with Schafer et al. (Reprinted with permission from ref 170. Copyright 1993 American Chemical Society.)

diverse peptide isosteres,[144−147] and Holloway et al.[150] found that the enzyme inhibition activity of peptide isosteres can be correlated with their interaction energy (eqs 94−96). Equation 97, however, suggests that there can be not only hydrogen bonding but also dispersion interactions in the binding of some isosteres with the enzyme. For a large series of C_2-symmetric aminodiols, however, the cytotoxic property was shown to depend on the lipophilicity of the compounds (eq 98).

To summarize, it is of interest to compare the optimum Clog *P* values (log P_0) observed in various correlation equations and then compare them with the Clog *P* of those anti-HIV drugs that are in the market. Clog *P* is the calculated log *P* obtained by using BioByte software[37] based on Leo's methodology.[66b]

Clog *P* values of anti-HIV drugs in the market[178] for reverse transcriptase inhibitors

a. Nucleoside Inhibitors

Name	Clog *P*
1. Zidovudine (AZT, Retrovir)	0.04
2. Zalcitabine (ddC, Hivid)	−1.29
3. Didanosine (ddI, Videx)	−1.18*
4. Stavudine (d4T, Zerit)	−0.49
5. Lamivudine (3TC, Epivir)	−1.50
6. Abacavir succinate (1592U89 succinate)	1.20*

The values marked with an asterisk (*) were calculated from new version of Clog *P* to be released shortly.

It has been shown that AZT is phosphorylated by a kinase within the cell (a process known as anabolic phosphorylation) to form AZT-5′-triphosphate.[2] It appears that this substance effectively competes with natural substrate thymidine triphosphate and acts as a chain terminator in the synthesis of DNA since it lacks the 3′-hydroxy group necessary for the formation of phosphodiester linkages.[2b,3]

Eight equations (eqs 55−62) have been developed with nucleoside inhibitors, and only one (eq 57) has a Clog *P* term. So, the hydrophobic effect of this class of compounds is not apparent from the equations. The low log *P* drug, abacavir succinate, recently approved by the FDA,[178] belongs to this class. However, one still must face the problem of the importance of hydrophobicity in the whole animal (or person) in the random movement of drugs through out the system. Testing on cells or enzymes overlooks this problem.

b. Nonnucleoside Inhibitors. Unlike nucleoside analogues, these drugs do not need to be phosphorylated to be active; rather they bind to the catalytic site of RT and inactivate the enzyme in a noncompetitive fashion.[183] The majority of the equations (1−54 out of 62) developed using compounds against RT belong to the non-nucleoside (NNRTIs) class. Out of these 54 equations, 42 have either Clog *P* or π terms, of which 11 have either parabolic or bilinear Clog *P* terms. Log P_0 values range from 2.31 to 5.72. There is a surprising uniformity in these values except for eq 40.

Log P_0 observed in QSAR equations

eq no.	log P_0 (95% confidence limits)
3	5.55 (4.8−12.7)
10	4.23 (4.0−4.6)
11	4.60 (4.2−6.4)
13	4.82 (4.3−7.6)
18	3.81 (3.5−4.2)
26	4.32 (4.1−4.8)
27	4.15 (4.0−4.4)
28	4.21 (4.0−4.9)
40	2.31 (2.2−2.4)
44	4.83 (4.5−5.6)
54	5.67 (5.4−5.9)

According to a Clog *P* calculation by Leo, the Clog *P* values of three FDA approved non-nucleoside RT inhibitors (NNRTI) are as follows.

1. Nevirapine (Viramune)	2.35
2. Delavirdine (Rescriptor)	1.59
3. Efavirenz (Sustiva)	4.95

It is interesting to note that these three drugs more or less cover the ideal log *P* values identified by the equations. So, there are drugs with high log *P* coming out as potent HIV-1 inhibitors, although our results indicate one might obtain potency by making more hydrophobic analogues.

Figure 15A shows the stereodiagram of 9-chloro-TIBO bound with HIV-1 reverse transcriptase. The figure was based on the X-ray crystal structure (1tvr[184]) available from the Protein Data Bank at the Research Collaboratory for Structural Bioinformatics (RCSB).[185] The X-ray structure clearly reveals that most of the amino acid residues surrounding the inhibitors are hydrophobic and five of them contain

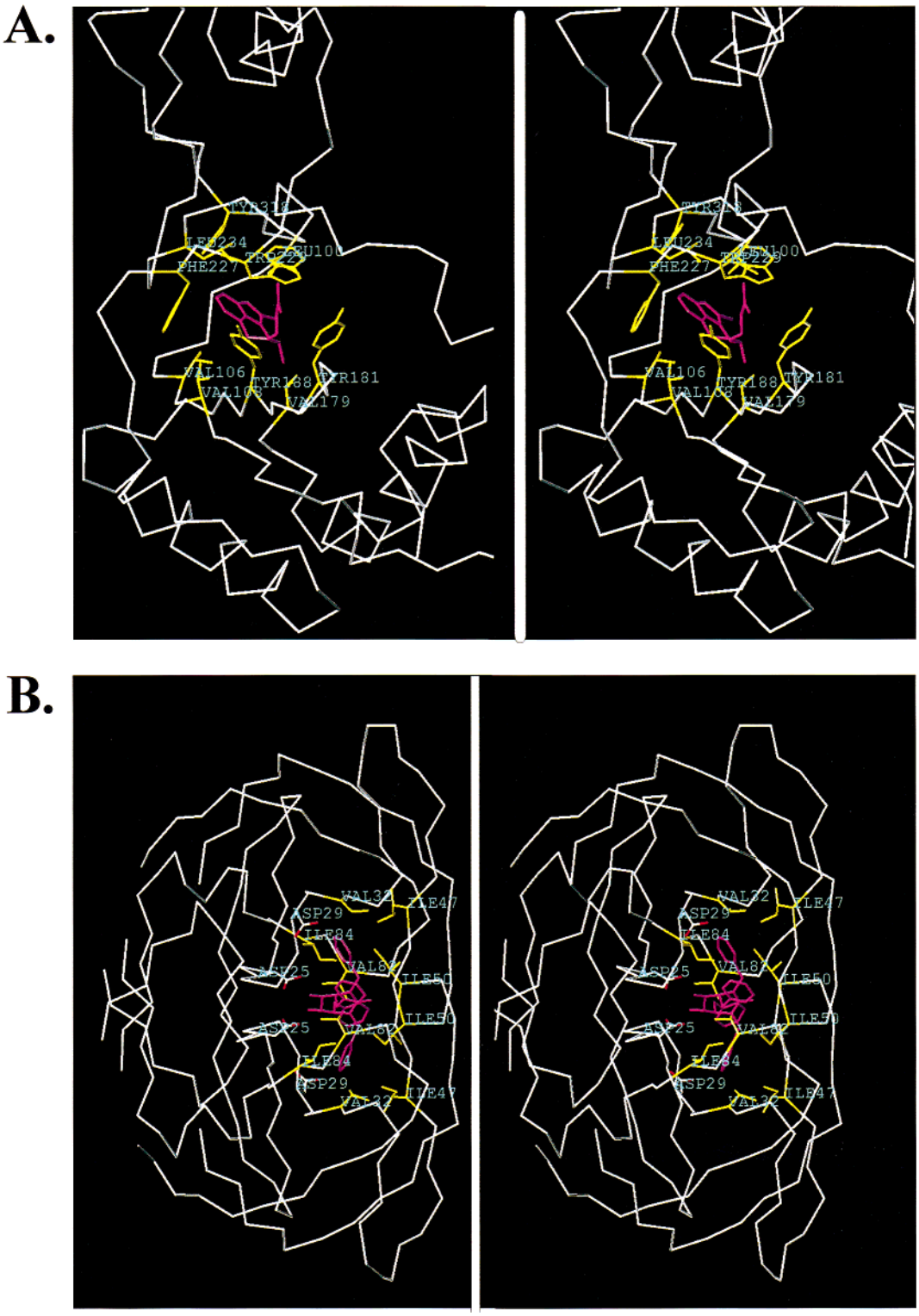


Figure 15. (A) Stereodiagram of 9-chloro-TIBO (Magenta) bound in the HIV-1 RT pocket. The binding site residues that can interact hydrophobically are indicated by yellow. The figure was generated by the Sybyl6.5 program based on the X-ray crystal structure (1tvr) obtained from the Brookhaven Protein Data Bank. (B) Stereodiagram of compound 22 of Table 48 bound in the HIV-1 PR pocket. The inhibitor and the residues interacting hydrophobically with the inhibitor have the same color annotations as in part A. The Sybyl6.5 program was used to generate the figure based on the X-ray crystal structure (1dmp) obtained from The Brookhaven Protein Data Bank.

aromatic rings (Tyr181, Tyr188, Phe227, Trp229, and Tyr318).[184] The dimethylallyl group of 9-chloro-TIBO can effectively interact with hydrophobic groups Tyr181, Tyr188, and Trp229, whereas the benzodiazepinone group interacts with several hydrophobic groups, namely, Val106, Phe227, and Tyr318. The chloro group of 9-chloro-TIBO interacts with Phe227 and Tyr318. Overall, the importance of hydrophobic groups in the 9-chloro-TIBO molecule is evident from these interaction patterns.

Clog *P* values of anti-HIV drugs in the market[178] for protease inhibitors

1. Saquinavir (Invirase)	4.73
2. Ritanovir (Norvir)	4.94
3. Indinavir (Crixivan)	3.68
4. Nelfinavir (Viracept)	5.84

log P_0 observed in QSAR equations

eq no.	log P_0 (95% confidence limits)
64	6.53 (5.5−7.5)
65	6.96 (6.2−7.8)
82	4.49 (3.5−4.8)

All protease inhibitors bind to the PR binding site pocket that has a considerable number of hydrophobic residues. There are 36 equations (63−98) for PR inhibitors, four of which (eqs 64, 65, 82, and 84) have positive Clog *P* terms and three of them (eqs 64, 65, and 82) have either parabolic or bilinear Clog *P* terms. Log P_0 values range from 4.49 to 6.96. The Clog *P* values of the FDA approved PR inhibitors correspond well with log P_0 values.

Figure 15B shows the stereodiagram of binding of a cyclic urea protease inhibitor (compound **22** of Table 48) in HIV-1 PR pocket. The figure was based on the DMP450 X-ray crystal structure (1dmp)[186] obtained from the RCSB.[185] The X-ray structure indicates that the S2/S2′ pockets in HIV-1 PR are essentially hydrophobic.[119] The residues that make up these pockets are Val32, Ile47, Ile50, and Ile84 in each monomer. The phenyl groups in the cyclic urea inhibitors make hydrophobic interactions with the hydrophobic residues. It has been shown that when Val82 and Ile84 are mutated, a substantial contact between these two residues and a potent cyclic urea derivative (DMP323) is lost.[187] A number of resistant strains against the cyclic urea inhibitors (DMP323 and DMP450) have been shown to target the hydrophobic residues for mutation. The importance of hydrophobic residues in the binding pocket reconfirms the contribution of the hydrophobicity of inhibitors on anti-HIV activity.

Surprisingly, most of the nucleoside drugs in use are hydrophilic. One wonders if this could be due to the toxicity of more hydrophobic compounds. However, this would seem to conform to our principle of minimal hydrophobicity in drug design[179] that one wants to make drugs as hydrophilic as possible *commensurate* with efficacy. In view of our findings, a combination of two or more drugs, one hydrophobic with log P_0 of 2.5−3.0 that can reach into the more or less hydrophobic compartments of body, and another hydrophilic with low log P_0 of 0.05−0.10, might prove effective in the hydrophilic regions. Considering the mutating potential of the HIV, one might want two or more drugs for each of the extremes in compartments. It seems to be necessary to irradicate every virus to cure the patients, since a patient can seem to be essentially free of the disease but it can soon return after stopping current drugs.

It must be kept in mind that Clog *P* attempts to do two jobs: account for hydrophobic interactions between ligand and receptor and for the random walk process in movement about the organism from site of injection to sites of action. Also, Clog *P* is for the *neutral form* of acids and bases that may be partially ionized. If the degree of ionization is about the same for a set of congeners, one can neglect the ionization factor. If not, using electronic terms one can often obtain good correlations where, for example, σ is associated with the degree of ionization and hence its effect on log *P*. The receptor cleft or pocket may not be completely homogeneous (hydrophobic), so that log *P* does not do a very good job for a large molecule with multiple positions of substitution. In such cases, π (the hydrophobicity of substituent) is more appropriate. However, the published π constants are from the benzene system. For substituents with lone-pair electrons, π for groups varies with electron effects of other substituents (e.g., π for OMe on benzene is not the same as OMe on pyridine or nitrobenzene).

This problem can be met by calculating Clog *P* for a derivative and subtracting Clog *P* for the parent compound. This becomes laborious when multiple substitution is involved. Also, if there is not *good* variation of substituents at each position, one has difficulty establishing a role for π. However, in serious drug research these problems can be easily overcome.

The majority of HIV research is done with cells, and these studies tend to overestimate log P_0 for animal systems. From a study of our database (omitting QSAR based on charged molecules), we have found on average that log P_0 for cells is about 1 log unit higher than for whole organisms. Thus, values above 4 may be too high. Since some of the P450 enzymes appear to attack hydrophobic molecules, this introduces another problem.

Possibly the most difficult problem in QSAR is the problem of 'congeners' that do not fit the 'final' equation. We see five major problems that are the root cause. (1) The mathematical form of the equation may be off the mark. Of course, this can always be solved by the use of more parameters using exponential and cross-product terms as the neural net approach does. However, as we have seen in this report, what one gains in keeping a large set one loses in understanding what the major forces are. (2) The parameters may not be the best. Sometimes, parameters obtained experimentally are better than those calculated and vice versa. (3) The quality of the experimental data. Often researchers do not take enough care with their test systems. We have experienced this in our laboratory. (4) Outliers due to what seem to be 'congeners' but in fact are not. This arises from trying to lump too many more or less similar compounds into a single QSAR. The resulting side reactions, we believe, are the most serious problem of all. (5) Different rates of metabolism of the members of a set.

Consider the problem of side reactions with something as simple as a single cell. Here we have in dynamic concert thousands of necessary chemical and physical reactions. The DNA codes for 50000−100000 proteins. These are formed into scores of enzymes, mitochondria, nucleus, nucleolus, peroxisomes, endoplasmic reticulum, golgi vesicles, lysomes, centrioles, and cell membranes. There are hormones mov-

ing about with messages for certain receptors. These are the complex processes of endocytosis and exocytosis and the elegant process of cell division. Why would anyone be so foolish as to attempt to understand the perturbations of say 40 chemicals acting on such a complex machine? In the first place, we are driven by a huge highly profitable drug industry still rapidly expanding that is able to treat the enormously more complex human machine to obtain spectacular results. Second, we have evidence from, at present, thousands of QSAR that give informative information. In a few instances, the researchers have obtained the QSAR from the similar chemicals acting on receptors in cells and finally in whole animals that shows surprisingly good agreement among themselves and with organic reaction mechanisms.[36,40,45,66]

Over the eons cells would seem to have evolved to protect themselves from xenobiotics. This can be done by building specialized receptors that simply ignore chemicals that are not good fits. Also, metabolism can be used to eliminate chemicals. Finally, repair mechanisms are available (e.g., for DNA). Thus, at low drug concentration, one may be able to avoid the protective mechanisms to primarily effect a specific process. However, the uncertainty that exists in describing all that is occurring with testing of set of 30−40 'congeners' in even a simple cell culture is enormous.

Someone once said that if you could not derive a good correlation equation, it was a reflection on your library. There are now such an enormous variety of parameters that almost anything, including sets of random numbers, can be correlated. Only by obtaining lateral support from a variety of points of view can one place confidence in a newly derived QSAR.

The problem of the quality of the parameters is serious, particularly that of the log *P* and steric parameters. Because there are at least a half a dozen commercial programs, of varying quality, for calculation of log *P*, few indeed take the time to experimentally determine new values. Of course, in analyzing data from the literature, the molecules are not normally available and, hence, experimental log *P* values are out of question. From our experience, we believe that while Clog *P* may be in error in terms of absolute values, the numbers are surprisingly good in relative terms.[66b] Hence, the error would largely be delegated to the intercept. Still it is always of value to determine a few experimental values to be sure that the calculated values are not unrealistic. The general quality of our Clog *P* calculations is shown in the following equation:

$$\text{Mlog } P = 0.975(\pm 0.003)\text{Clog } P + 0.049(0.007)$$

$$n = 9{,}000,\ r^2 = 0.984,\ s = 0.205$$

Steric parameters are much more difficult to define. Two approaches have been tried. We have attempted to use measured or calculated values for the various substituents. The problem with this approach is that not knowing the shape of a receptor site (except in rare examples[188,189]), the only guide is the empirical quality of the QSAR. A more elegant approach is that employed by CoMFA. Here one attempts, by trial and error, to place the members of a data set in a proper conformation from which steric interactions are estimated by exploring the outer surfaces of the set of 'congeners'. Since the positions of ligands are fixed, the 'give' in the system has to be the receptor wall. However, we know from many examples that steric effects are often a linear function of the empirical parameters. That is, as substituent size increases, activity gradually falls or rises. This more or less uniform decline or increase in activity could be due to two effects: the receptor wall could give to some degree, or the position of the ligands could gradually move, or possibly both mechanisms could operate. In our approach, we do not assume a perfectly uniform mode of binding. Thus, coefficients with steric terms may reflect the complex process of displacement of the ligand and/or the receptor wall.

Somehow all of these processes that regulate cellular metabolism, DNA replication, repair, etc., have "learned" how to avoid being seriously disrupted by xenobiotics, at least for enough time for experimental purposes. However, we cannot escape a kind of biological "uncertainty principle" that makes Heisenberg's principle seem simple by comparison. As yet we have little idea of how a foreign chemical perturbing a process affects the myriad other processes, and it would seem to be forever beyond our reach to attain any kind of complete understanding. Still QSAR allows us to light a small candle rather than curse the darkness. At low concentrations, about ($\sim 10^{-9}$ M) for a short span of time (hours), one would expect fewer visible perturbations. At a concentration of one thousand times greater (10^{-6} M) for a drug that must be taken for years or decades the problems are more difficult to address.

Hence, one has to expect outliers that must not be forgotten for they are the leads to new understanding. To cover them up by including them in a QSAR, at the cost of a lower r^2, can be more confusing than helpful. The problem is how good an r^2 can one expect all things considered. Most of our database of 6200 QSAR have r^2 of 0.85 ± 0.05.

Finally, considerable work has been done using the 3D-CoMFA methodology.[155−157,180,181] It is not at all easy to compare this work with that reviewed by us. In the first place, CoMFA, although termed QSAR, does not qualify, as it is generally used, as quantitative SAR. It is qualitative or at best semiquantitative. Authors rarely attempt to discuss their results in terms of numbers but use 3D pictures. These pictures are not precise enough to be compared with other results, in part because the terms used to formulate a regression-based model are based on principle components. Such terms will have different composition from dataset to dataset so that comparison is only possible via pictures that are not easy to understand. This is not to say that CoMFA cannot provide insight, but not the kind we are interested in for mechanistic comparisons.

At present, there are still very few academic laboratories doing serious QSAR studies and, of course, the drug companies cannot afford to publish their most interesting results. Still we believe that

QSAR will prevail and in the end will develop into a science of chemical−biological interactions.

VII. Acknowledgments

This research was supported by grants from the NIEHS 1R01 ES0759501 (C.H.).

VIII. References

(1) Barre-Sinoussi, F.; Chermann, J.-C.; Rey, F.; Nugeyre, M. T.; Chamaret, S.; Gruest, J.; Dauguet, C.; Axler-Blin, C.; Vezinet-Brun, F.; Rouzioux, C.; Rozenbaum, W.; Montagnier, L. *Science* **1983**, *220*, 868.

(2) (a) Popovic, M.; Sarngadharan, M. G.; Read, E.; Gallo, R. C. *Science* **1984**, *224*, 497. (b) Gallo, R. C.; Salahuddin, S. Z.; Popovic, M.; Shearer, G. M.; Kaplan, M.; Haynes, B. F.; Palker, T. J.; Redfield, R.; Oleske, J.; Safai, B.; White, G.; Foster, P.; Markham, P. D. *Science* **1984**, *224*, 500.

(3) Gallo, R. C.; Montagnier, L. *Sci. Am.* **1988**, *259*, 40.

(4) Coffin, J.; Hasse, A.; Levy, J. A.; Montagnier, L.; Oroszlan, S.; Teich, N.; Temin, H.; Toyoshmia, K.; Varmus, H.; Vogt, P.; Weiss, R. *Science* **1986**, *232*, 697.

(5) Levy, J. A. *JAMA* **1989**, *261*, 2997 and references therein.

(6) Clavel, F.; Guyader, M.; Guetard, D.; Salle, M.; Montagnier, L.; Alizon, M. *Nature (London)* **1986**, *324*, 691.

(7) Guyader, M.; Emerman, M.; Sonigo, P.; Clavel, F.; Montagnier, L.; Alizon, M. *Nature (London)* **1987**, *326*, 662.

(8) Bowen, D. L.; Lane, H. C.; Fauci, A. S. *Ann. Intern. Med.* **1985**, *103*, 704.

(9) Embretson, J.; Zupancic, M.; Ribas, J. L.; Burke, L.; Racz, P.; Tenner-Racz, K.; Haase, A. T. *Nature (London)* **1993**, *362*, 359.

(10) (a) Isselbacher, K. J.; Eugene, B.; Wilson, J.; Martin, J. B.; Fauci, A. S.; Dennis, L. K. In *Harrison's Principles of Internal Medicine*, 13th ed.; McGraw-Hill: New York, 1994; Vol. 1. (b) Gallo, R. C. *Sci. Am.* **1987**, 46.

(11) Fauci, A. S. *Science* **1988**, *239*, 617.

(12) De Clercq, E. *J. Med. Chem.* **1995**, *38*, 2491.

(13) Hussey, R. E.; Richardson, N. E.; Kowalski, M.; Brown, N. R.; Chang, H.-C.; Siliciano, R. F.; Dorfman, T.; Walker, B.; Sodroski, J.; Reinherz, E. L. *Nature (London)* **1988**, *331*, 78.

(14) Broder, S.; Yarchoan, R.; Collins, J. M. *Lancet* **1985**, *2*, 627.

(15) Lorentsen, K. J.; Hendrix, C. W.; Collins, J. M.; Komhauser, D. M.; Petty, B. G.; Klecker, R. W.; Flexner, C.; Eckel, R. H.; Lietman, P. S. *Ann. Intern. Med.* **1989**, *111*, 561.

(16) Balzarini, J.; Neyts, J.; Schols, D.; Hosoya, M.; Van Damme, E.; Peumans, W.; De. Clercq, E. *Antiviral Res.* **1992**, *18*, 191.

(17) Balzarini, J.; Schols, D.; Neyts, J.; Van Damme, E.; Peumans, W.; De Clercq, E. *Antimicro. Agents Chemother.* **1991**, *35*, 410.

(18) Nakashima, H.; Masuda, M.; Murakami, T.; Koyanagi, Y.; Matsumoto, A.; Fujii, N.; Yamamoto, N. *Antimicrob. Agents Chemother.* **1992**, *36*, 1249.

(19) Jansen, R. W.; Molema, G.; Pauwels, R.; Schols, D.; De Clercq, E.; Meijer, D. K. F. *Mol. Pharmacol.* **1991**, *39*, 818.

(20) Jansen, R. W.; Schols, D.; Pauwels, R.; De Clercq, E.; Meijer, D. K. F. *Mol. Pharmacol.* **1993**, *44*, 1003.

(21) (a) Mayaux, J. F.; Bousseau, A.; Pauwels, R.; Huet, T.; Henin, Y.; Dereu, N.; Evers, M.; Soler, F.; Poujade, C.; De Clercq, E.; Pecq, J.-B. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3564. (b) Debnath, A. K.; Radigam, L.; Jiang, S. *J. Med. Chem.* **1999**, *42*, 3202.

(22) Rossamann, M. G. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4625.

(23) De Clercq, E. *Med. Res. Rev.* **1993**, *13*, 229.

(24) De Clercq, E. *Exp. Opin. Invest. Drugs* **1994**, *3*, 253.

(25) Fesen, M. R.; Kohn, K. W.; Leteurtre, F.; Pommier, Y. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 2399.

(26) Fesen, M. R.; Pommier, Y.; Leteurtre, F.; Hiroguchi, S.; Yung, J.; Kohn, K. W. *Biochem. Pharmacol.* **1994**, *48*, 595.

(27) Mayaux, J. F.; Huet, T.; Pernelle, C.; Becquart, J.; Gueguen, J. C.; Tahraui, L.; Evers, M.; Henin, Y.; Bousseau, A.; Dereu, N. Abstracts of the NIH Conference on Retroviral Integrase, Bethesda, MD, January 19−20, 1995.

(28) Stein, C. A.; Cheng, Y.-C. *Science* **1993**, *261*, 1004.

(29) Kohl, N. E.; Emini, N. A.; Schleif, W. A.; Davis, L. J.; Heimbach, J. C.; Dixon, R. A. F.; Scolnick, E. M.; Sigal, I. S. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4686.

(30) Peng, C.; Ho, B. K.; Chang, T. W.; Chang, N. T. *J. Virol.* **1989**, *63*, 2550.

(31) Wlodawer, A.; Ericson, J. W. *Annu. Rev. Biochem.* **1993**, *62*, 543.

(32) Appelt, K. *Perspect. Drug Discovery Des.* **1993**, *1*, 23.

(33) Boehme, R. E.; Borthwick, A. D.; Wyatt, P. G. *Annu. Rep. Med. Chem.* **1995**, *30*, 139.

(34) Babine, R. E.; Bender, S. L. *Chem. Rev.* **1997**, *97*, 1359.

(35) Jacobsen, H.; Yasargil, K.; Winslow, D. L.; Craig, J. C.; Krohn, A.; Duncan, I. B.; Mous, J. *Virology* **1995**, *206*, 527.

(36) Ratner, L.; vander Heyden, N. *AIDS Res. Human Retrov.* **1993**, *9*, 291.

(37) BioByte Corporation, C-QSAR database, 201W Fourth St., Suite no. 204, Claremont, CA 91711-4707.

(38) Hansch, C.; Gao, H.; Hoekman, D. In *Comparative QSAR*; Devillers, J., Ed.; Taylor & Francis: Washington, DC, 1998.

(39) Hansch, C.; Leo, A.; Hoekman, D. *Exploring QSAR: Hydrophobic, Electronic and Steric Constants*; American Chemical Society: Washington, DC, 1995.

(40) Hansch, C.; Leo, A. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, DC, 1995.

(41) Verloop, A. *The STERIMOL Approach to Drug Design*; Marcel Dekker: New York, 1987.

(42) Huang, P.; Farquhar, D.; Plunkett, W. *J. Biol. Chem.* **1990**, *265*, 11914.

(43) De Clercq, E. *AIDS Res. Human Retrov.* **1992**, *8*, 119.

(44) Althaus, I. W.; Chou, J. J.; Gonzales, A. J.; Deibel, M. R.; Chou, K. C.; Kezdy, F. J.; Romero, D. L.; Thomas, R. C.; Aristoff, P. A.; Tarpley, W. G.; Reusser, F. *Biochem. Pharmacol.* **1994**, *47*, 2017.

(45) Hansch, C. *Acc. Chem. Res.* **1993**, *26*, 147.

(46) Jacobo-Molina, A.; Ding, J.; Nanni, R. G.; Clark, A. D., Jr.; Lu, X.; Tantillo, C.; Williams, R. L.; Kamer, G.; Ferris, A. L.; Clark, P.; Hizi, A.; Hughes, S. H.; Arnold, E. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6320.

(47) Nanni, R. G.; Ding, J.; Jacobo, M. A.; Hughes, S. H.; Arnold, E. *Perspect. Drug Discovery Des.* **1993**, *1*, 129.

(48) Pauwels, R.; Andries, K.; Desmyter, J.; Schols, D.; Kukla, M. J.; Breslin, H. J.; Raeymaeckers, A.; Van Gelder, J.; Woestenborghs, R.; Heykants, J.; Schellekens, K.; Janssen, M. A. C.; De Clercq, E.; Janssen, P. A. J. *Natute (London)* **1990**, *343*, 470.

(49) Debyser, Z.; Pauwels, R.; Andries, K.; Desmyter, J.; Kukla, M. J.; Janssen, P. A. J.; De Clercq, E. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 1451.

(50) Miyasaka, T.; Tanaka, H.; Baba, M.; Hayakawa, H.; Walker, R. T.; Balzarini, J.; De Clercq, E. *J. Med. Chem.* **1989**, *32*, 2507.

(51) Baba, M.; Tanaka, H.; De Clercq, E.; Pauwels, R.; Balzarini, J.; Schols, D.; Nakashima, H.; Perno, C. F.; Walker, R. T.; Miyasaka, T. *Biochem. Biophys. Res. Commun.* **1989**, *165*, 1375.

(52) Merluzzi, V. J.; Hargrave, K. D.; Labadia, M.; Grozinger, K.; Skoog, M.; Wu, J. C.; Shih, C.-K.; Eckner, K.; Hattox, S.; Adams, J. Rosehthal, A. S.; Faanes, R.; Eckner, R. J.; Karp, R. A.; Sullivan, J. L. *Science* **1990**, *250*, 1411.

(53) Goldman, M. E.; Nunberg, J. H.; O'Brien, J. A.; Quintero, J. C.; Schleif, W. A.; Freund, K. F.; Gaul, S. L.; Saari, W. S.; Wai, J. S.; Hoffman, J. M.; Anderson, P. S.; Hupe, D. J.; Emini, E. A.; Stern, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 6863.

(54) Romero, D. L.; Busso, M.; Tan, C. K.; Reusser, F.; Palmer, J. R.; Poppe, S. M.; Aristoff, P. A.; Downey, K. M.; So, A. G.; Resnick, L.; Tarpley, W. G. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 8806.

(55) Balzarini, J.; Pérez-Pérez, M.-J.; San-Felix, A.; Schols, D.; Perno, C. F.; Vandamme, A. M.; Camarasa, M. J.; De Clercq, E. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4392.

(56) Pauwels, R.; Andries, K.; Debyser, Z.; Van Daele, P.; Schols, D.; Stoffels, P.; De Vreese, K.; Woestenborghs, R.; Vandamme, A. M.; Janssen, C. G. M.; Anne, J.; Cauwenbergh, G.; Desmyter, J.; Heykants, J.; Janseen, M. A. C.; De Clercq, E.; Janssen, P. A. J. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1711.

(57) Balzarini, J.; Karlsson, A.; Pérez-Pérez, M.-J.; Camarasa, M. J.; De Clercq, E. *Virology* **1993**, *196*, 576.

(58) Vasudevachari, M. B.; Battista, C.; Lane, H. C.; Psallidopoulos, M. C.; Zhao, B.; Cook, J.; Palmer, J. R.; Romero, D. L.; Tarpley, W. G.; Salzman, N. P. *Virology* **1992**, *190*, 269.

(59) Smith, M. S.; Brian, E. L.; Pagano, J. S. *J. Virol.* **1987**, *61*, 3769.

(60) Kukla, M. J.; Breslin, H. J.; Pauwels, R.; Fedde, C. L.; Miranda, M.; Scott, M. K.; Sherrill, R. G.; Raeymackers, A.; Van Gelder, J.; Andries, K.; Janssen, M. A. C.; De Clercq, E.; Janssen, P. A. J. *J. Med. Chem.* **1991**, *34*, 746.

(61) Kukla, M. J.; Breslin, H. J.; Diamond, C. J.; Grous, P. P.; Ho, C. Y.; Miranda, M.; Rodgers, J. D.; Sherril, R. G.; De Clercq, E.; Pauwels, R.; Andries, K.; Moens, L. J.; Janssen, M. A. C.; Janssen, P. A. J. *J. Med. Chem.* **1991**, *34*, 3187.

(62) Breslin, H. J.; Kukla, M. J.; Ludovici, D. W.; Mohrbacher, R.; Ho, W.; Miranda, M.; Rodgers, J. D.; Hitchens, T. K.; Leo, G.; Gauthier, D. A.; Ho, C. Y.; Scott, M. K.; De Clercq, E.; Pauwels, R.; Andries, K.; Janssen, M. A. C.; Janssen, P. A. J. *J. Med. Chem.* **1995**, *38*, 771.

(63) (a) 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. *J. Med. Chem.* **1995**, *38*, 794. (b) Cramer, R. D., III; Bunce, J. D.; Patterson, D. E.; Frank, I. E. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18.

(64) Gupta, S. P.; Garg, R. *J. Enzyme Inhib.* **1996**, 11, 23.

(65) Pauwels, R.; Andries, K.; Debyser, Z.; Kukla, M. J.; Schols, D.; Breslin, H. J.; Woestenborghs, R.; Desmyter, J.; Janssen, M. A. C.; De Clercq, E.; Janssen, P. A. J. *Antimicrob. Agents Chemother.* **1994**, *38*, 2863.

(66) (a) Hansch, C.; Hoekman, D.; Gao, H. *Chem Rev.* **1996**, 96, 1045. (b) Leo, A. *Chem. Rev.* **1993**, 93, 1281.
(67) (a) Tanaka, H.; Takashima, H.; Ubasawa, M.; Sekiya, K.; Nitta, I.; Baba, M.; Shigeta, S.; Walker, R. T.; De Clercq, E.; Miyasaka, T. *J. Med. Chem.* **1992**, *35*, 337. (b) Hansch C.; Zhang, L. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 1165.
(68) Tanaka, H.; Takashima, H.; Ubasawa, M.; Sekiya, K.; Nitta, I.; Baba, M.; Shigeta, S.; Walker, R. T.; De Clercq, E.; Miyasaka, T. *J. Med. Chem.* **1992**, *35*, 4713.
(69) Tanaka, H.; Takashima, H.; Ubasawa, M.; Sekiya, K.; Inouye, N.; Baba, M.; Shigeta, S.; Walker, R.T.; De Clercq, E.; Miyasaka, T. *J. Med. Chem.* **1995**, *38*, 2860.
(70) Garg, R.; Kurup, A.; Gupta, S. P. *Quant. Struct.-Act. Relat.* **1997**, *16*, 20.
(71) Baba, M.; Shigeta, S.; Tanaka, H.; Miyasaka, T.; Ubasawa, M.; Umezu, K.; Walker, R. T.; Pauwels, R.; De Clercq, E. *Antiviral Res.* **1992**, *17*, 245.
(72) Balzarini, J.; Baba, M.; De Clercq, E. *Antimicrob. Agents Chemother.* **1995**, *39*, 998.
(73) Tanaka, H.; Baba, M.; Hayakawa, H.; Sakamaki, T.; Miyasaka, T.; Ubasawa, M.; Takashima, H.; Sekiya, K.; Nitta, I.; Shigeta, S.; Walker, R. T.; Balzarini, J.; De Clercq, E. *J. Med. Chem.* **1991**, *34*, 349.
(74) Pontkis, R.; Benhida, R.; Aubertin, A.-M.; Grierson, D. S.; Monneret, C. *J. Med. Chem.* **1997**, *40*, 1845.
(75) Kim, D.-K.; Gam, J.; Kim, Y.-W.; Lim, J.; Kim, H.-T.; Kim, K. H. *J. Med. Chem.* **1997**, *40*, 2363.
(76) (a) Miyasaka, T.; Tanaka, H.; Baba, M.; Hayakawa, H.; Walker, R. T.; Balzarini, J.; De Clercq, E. *J. Med. Chem.* **1989**, *32*, 2507. (b) Baba, M.; Tanaka, H.; De Clercq, E.; Pauwels, R.; Balzarini, J.; Schols, D.; Nakashima, H.; Perno, C.-F.; Walker, R. T.; Miyasaka, T. *Biochem. Biophys. Res. Commun.* **1989**, *165*, 1375. (c) Tanaka, M.; Baba, M.; Ubasawa, M.; Takashima, H.; Sekiya, K.; Nitta, I.; Shigeta, S.; Walker, R. T.; De Clercq, E.; Miyasaka, T. *J. Med. Chem.* **1991**, *34*, 1394, (d) Tanaka, H.; Baba, M.; Saito, S.; Miyasaka, T.; Takashima, H.; Sekiya, K.; Ubasawa, M.; Nitta, I.; Walker, R. T.; Nakashima, H.; De Clercq, E. *J. Med. Chem.* **1991**, *34*, 1508. (e) Tanaka, H.; Baba, M.; Ubasawa, M.; Takashima, H.; Sekiya, K.; Miyasaka, T.; Nitta, I.; Walker, R. T.; De Clercq, E. *Collec. Czech. Chem. Commun.* **1990**, *55*, 89. (f) Tanaka, H.; Baba, M.; Hayakawa, H.; Haraguchi, K.; Miyasaka, T.; Ubasawa, M.; Takashima, H.; Sekiya, K.; Nitta, I.; Walker, R. T.; De Clercq, E. *Nucleosides Nucleotides* **1991**, *10*, 397.
(77) Luco, J. M.; Ferretti, F. H. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 392.
(78) Kier, L. B.; Hall, L. H. In *Molecular Connectivity in Chemistry and Drug Research*; Bawden, D., Ed.; Research Studies Press Ltd.: Letchworth, Hertfordshire, England, 1986; Chapter 3, p 43.
(79) (a) Abraham, M. H.; McGowan, J. C. *Chromotographicia* **1987**, *23*, 243. (b) Abraham, M. H. *Chem. Soc. Rev.* **1993**, *22*, 73.
(80) Kier, L. B.; Hall, L. H. *Quant. Struct.-Act. Relat.* **1991**, *10*, 134.
(81) Wold, S. *Technometrics* **1978**, *20*, 397.
(82) Hopkins, A. L.; Ren, J.; Esnouf, R. M.; Willcox, B. E.; Jones, E. Y.; Ross, C.; Miyasaka, T.; Walker, R. T.; Tanaka, H.; Stammers, D. K.; Stuart, D. I. *J. Med. Chem.* **1996**, *39*, 1589.
(83) Kireev, D. B.; Chrétien, J. R.; Grierson, D. S.; Monneret, C. *J. Med. Chem.* **1997**, *40*, 4257.
(84) Garg, R.; Gupta, S. P. *J. Enzyme Inhib.* **1997**, *11*, 171.
(85) (a) Camarasa, M. J.; Pérez-Pérez, M.-J.; San-Felix, A.; Balzarini, J.; De Clercq, E. *J. Med. Chem.* **1992**, *35*, 2721. (b) Pérez-Pérez, M.-J.; San-Felix, A.; Balzarini, J.; De Clercq, E.; Camarasa, M. J. *J. Med. Chem.* **1992**, *35*, 2988. (c) San-Felix, A.; Sonsoles, V.; Pérez-Pérez, M.-J.; Balzarini, J.; De Clercq, E.; Camarasa, M. J. *J. Med. Chem.* **1994**, *37*, 453.
(86) Balzarini, J.; Pérez-Pérez, M.-J.; San-Félix, A.; Schols, D.; Perno, C.-F.; Vandamme, A.-M.; Camarasa, M.-J.; De Clercq, E. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4392.
(87) Alvarez, R.; Velazquez, S.; San-Felix, A.; Aquaro, S.; De Clercq, E.; Perno, C.-F.; Karlsson, A.; Balzarini, J.; Camarasa, M. J. *J. Med. Chem.* **1994**, *37*, 4185.
(88) 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. *J. Med. Chem.* **1995**, *38*, 4839.
(89) Hargrave, K. D.; Proudfoot, J. R.; Grozinger, K. G.; Cullen, E.; Kapadia, S. R.; Patel, U. R.; Fuchs, V. U.; Mauldin, S. C.; Vitous, J.; Behnke, M. L.; Klunder, J. M.; Pal, K.; Skiles, J. W.; McNeil, D. W.; Rose, J. M.; Chow, G. C.; Skoog, M. T.; Wu, J. C.; Schmidt, G.; Engel, W. W.; Eberlein, W. G.; Saboe, T. D.; Campbell, S. J.; Rosenthal, A. S.; Adams, J. *J. Med. Chem.* **1991**, *34*, 2231.
(90) Klunder, J. M.; Hargrave, K. D.; West, M. A.; Kullen, E.; Pal, K.; Behnke, M. L.; Kapadia, S. R.; McNeil, D. W.; Wu, J. C.; Chow, G. C.; Adams, J. *J. Med. Chem.* **1992**, *35*, 1887.
(91) Hoffman, J. M.; Wai, J. S.; Thomas, C. M.; Levin, R. B.; O′Brein, J. A.; Goldman, M. E. *J. Med. Chem.* **1992**, *35*, 3784.
(92) Saari, W. S.; Wai, J. S.; Fisher, T. E.; Thomas, C. M.; Hoffman, J. M.; Rooney, C. S.; Smith, A. M.; Jones, J. H.; Bamberger, D. L.; Goldman, M. E.; O′Brein, J. A.; Nunberg, J. H.; Quintero, J. C.; Schleif, W. A.; Emini, E. A.; Anderson, P. S. *J. Med. Chem.* **1992**, *35*, 3792.
(93) Hoffman, J. M.; Smith, A. M.; Rooney, C. S.; Fisher, T. E.; Wai, J. S.; Thomas, C. M.; Bamberger, D. L.; Barnes, J. L.; Williams, T. M.; Jones, J. H.; Olson, B. D.; O′Brein, J. A.; Goldman, M. E.; Nunberg, J. H.; Quintero, J. C.; Schleif, W. A.; Emini, E. A.; Anderson, P. S. *J. Med. Chem.* **1993**, *36*, 953.
(94) Wai, J. S.; Williams, T. M.; Bamberger, D. L.; Fisher, T. E.; Hoffman, J. M.; Hudcosky, R. J.; MacTough, S. C.; Rooney, C. S.; Saari, W. S.; Thomas, C. M.; Goldman, M. E.; O′Brien, J. A.; Emini, E. A.; Nunberg, J. H.; Quintero, J. C.; Schleif, W. A.; Anderson, P. S. *J. Med. Chem.* **1993**, *36*, 249.
(95) Garg, R.; Gupta, S. P. *J. Enzyme Inhib.* **1997**, *12*, 1.
(96) Pauwels, R.; Andries, K.; Debyser, Z.; Daele, P. V.; Schols, D.; Stoffels, P.; De Vreese, K.; Woestenborghs, R.; Vandamme, A.-M.; Janssen, C. G. M.; Anne, J.; Cauwenbergh, G.; Desmyter, J.; Heykants, J.; Janssen, M. A. C.; De Clercq, E.; Janssen, P. A. J. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1711.
(97) Artico, M.; Silverstri, R.; Massa, S.; Loi, A. G.; Corrias, S.; Piras, G.; La Colla, P. *J. Med. Chem.* **1996**, *39*, 522.
(98) Hanasaki, Y.; Watanabe, H.; Katsuura, K.; Takayama, H.; Shirakawa, S.; Yamaguchi, K.; Sakai, S.-I.; Ijichi, K.; Fujiwara, M.; Konno, K.; Yokota, T.; Shigeta, S.; Baba, M. *J. Med. Chem.* **1995**, *38*, 2038.
(99) Ijichi, K.; Fujiwara, M.; Hanasaki, Y.; Watanabe, H.; Katsuura, K.; Takayama, H.; Shirakawa, S.; Sakai, S.-I.; Shigeta, S.; Konno, K.; Yokota, T.; Baba, M. *Antimicrob. Agents Chemother.* **1995**, *39*, 2337.
(100) Yu, K. -L.; Bronson, J. J.; Yang, H.; Patick, A.; Alam, M.; Brankovan, V.; Datema, R.; Hitchcock, M. J. M.; Martin, J. C. *J. Med. Chem.* **1993**, *36*, 2726.
(101) (a) McGuigan, C.; Pathirana, R. N.; Balzarini, J.; De Clercq, E. *J. Med. Chem.* **1993**, *36*, 1048. (b) Siddiqui, A. Q.; Ballatore, C.; McGuigan, C.; De Clereq, E.; Balzarini, J. *J. Med. Chem.* **1999**, *42*, 393.
(102) Tatematsu, H.; Kilkuskie, R. E.; Corrigan, A. J.; Bodner, A. J.; Lee, K.-H. *J. Nat. Prod.* **1991**, *54*, 632.
(103) Wyatt, P. G.; Bethell, R. C.; Cammack, N.; Charon, D.; Dodic, N.; Dumaitre, B.; Evans, D. N.; Green, D. V. S.; Hopewell, P. L.; Humber, D. C.; Lamont, R. B.; Orr, D. C.; Plested, S. J.; Ryan, D. M.; Sollis, S. L.; Storer, R.; Weingarten, G. G. *J. Med. Chem.* **1995**, *38*, 1657.
(104) Doubell, P. C. J.; Oliver, D. W. *Arzneim.-Forsch.* **1992**, *42*, 65.
(105) Strekowski, L.; Mokrosz, J. L.; Honkan, V. A.; Czarny, A.; Cegla, M. T.; Wydra, R. L.; Patterson, S. E.; Schinazi, R. F. *J. Med. Chem.* **1991**, *34*, 1739.
(106) Balzarini, J.; Brouwer, W. G.; Felauer, E. E.; De Clercq, E.; Karlsson, A. *Antiviral Res.* **1995**, *27*, 219.
(107) Balzarini, J.; van Aerschot, A.; Pauwels, R.; Baba, M.; Schols, D.; Herdewijn, P.; De Clercq, E. *Mol. Pharmacol.* **1989**, *35*, 571.
(108) Herdewijn, P. A. M.; van Aerschot, A.; Balzarini, J.; De Clercq, E. *Med. Chem. Res.* **1991**, *1*, 9.
(109) Balzarini, J.; Baba, M.; Pauwels, R.; Herdewijn, P.; De Clercq, E. *Biochem. Pharmacol.* **1988**, *37*, 2847.
(110) (a) Mahmoudian, M. *Pharm. Res.* **1991**, *8*, 43. (b) Chu, K. C.; Schinazi, R. F.; Ahn. M. K.; Ultas, G. V.; Gu, Z. P. *J. Med. Chem.* **1989**, *32*, 612. (c) Lin, T.-S.; Guo, J.-Y.; Schinazi, R. F.; Chu, C. K.; Xiang, J.-N.; Prusoff, W. H. *J. Med. Chem.* **1988**, *31*, 336.
(111) Herdewijn, P.; Balzarini, J.; Baba, M.; Pauwels, R.; van Aerschot, A.; Janssen, G.; De Clercq, E. *J. Med. Chem.* **1988**, *31*, 2040.
(112) Murakami, K.; Shirasaka, T.; Yoshioka, H.; Kojima, E.; Aoki, S.; Ford, H., Jr.; Driscoll, J. S.; Kelley, J. A.; Mitsuya, H. *J. Med. Chem.* **1991**, *34*, 1606.
(113) Freeman, G. A.; Shaver, S. R.; Rideout, J. L.; Short, S. A. *Bioorg. Med. Chem.* **1995**, *3*, 447.
(114) McGuigan, C.; Davies, M.; Pathirana, R.; Mahmood, N.; Hay, A. J. *Antiviral Res.* **1994**, *24*, 69.
(115) Romines, K. R.; Thaisrivongs, S. *Drugs Future* **1995**, *20*, 377.
(116) Olson, G. L.; Bolin, D. R.; Bonner, M. P.; Bos, M.; Cook, C. M.; Fry, D. C.; Graves, B. J.; Hatada, M.; Hill, D. E.; Kahn, M.; Madison, V. S.; Rusiecki, V. K.; Sarabu, R.; Sepinwall, J.; Vincet, G. P.; Voss, M. E. *J. Med. Chem.* **1993**, *36*, 3039.
(117) See, for example: Maligres, P. E.; Upadhyay, V.; Rossen, K.; Cianciosi, S. J.; Purick, R. M.; Eng, K. K.; Reamer, R. A.; Askin, D.; Volante, R. P.; Reider, P. J. *Tetrahedron Lett.* **1995**, *36*, 2195.
(118) Nugiel, D. A.; Jacobs, K.; Worley, T.; Patel, M.; Kaltenbach R. F., III; Meyer, D. T.; Jadhav, P. K.; DeLucca, G. V.; Smyser, T. E.; Klabe, R. M.; Bacheler, L. T.; Rayner, M. M.; Seitz, S. P. *J. Med. Chem.* **1996**, *39*, 2156.
(119) Lam, P. Y. S.; Ru, Y.; Jadhav, P. K.; Aldrich, P. E.; DeLucca, G. V.; Eyerman, C. J.; Chang, C.-H.; Emmett, G.; Holler, E. R.; Danekar, W. F.; Li, L.; Confalone, P. N.; McHugh, R. J.; Han, Q.; Li, R.; Markwalder, J. A.; Seitz, S.P.; Sharpe, T. R.; Bacheler, L. T.; Rayner, M. M.; Klabe, R. M.; Shum, L.; Winslow, D. L.; Kornhauser, D. M.; Jackson, D. A.; Erickson-Viitanen, S.; Hodge, C. N. *J. Med. Chem.* **1996**, *39*, 3514.
(120) Gupta, S. P.; Babu, M. S.; Garg, R.; Sowmya, S. *J. Enzyme Inhib.* **1998**, *13*, 399.

(121) (a) Jadhav, P. K.; Woerner, F. J.; Lam, P. Y. S.; Hodge, C. N.; Eyermann, C. J.; Man, H.-W.; Daneker, W. F.; Bacheler, L. T.; Rayner, M. M.; Meek, J. L.; Erickson-Viitanen, S.; Jackson, D. A.; Calabrese, J. C.; Schadt, M.; Chang, C.-H. *J. Med. Chem.* **1998**, *41*, 1446. (b) Gupta, S. P.; Babu, M. S. *Bioorg. Med. Chem.* **1999**, *7*, 1.
(122) Wilkerson, W. W.; Akamike, E.; Cheatham, W. W.; Hollis, A. Y.; Collins, R. D.; DeLucca, I.; Lam, P. Y. S.; Ru, Y. *J. Med. Chem.* **1996**, *39*, 4299.
(123) Wilkerson, W. W.; Dax, S.; Cheatham, W. W. *J. Med. Chem.* **1997**, *40*, 4079.
(124) Han, Q.; Chang, C.-H.; Li, R.; Ru, Y.; Jadhav, P. K.; Lam, P. Y. S. *J. Med. Chem.* **1998**, *41*, 2019.
(125) Romines, K. R.; Watenpaugh, K. D.; Tomick, P. K.; Howe, W. J.; Morris, J. K.; Lovasz, K. D.; Mulichak, A. M.; Finzel, B. C.; Lynn, J. C.; Horng, M.-M.; Schwende, F. J.; Ruwart, M. J.; Zipp, G. L.; Chong, K.-T.; Dolak, L. A.; Toth, L. N.; Howard, G. M.; Rush, B. D.; Wilkinson, K. F.; Possert, P. L.; Dalga, R. J.; Hinshaw, R. R. *J. Med. Chem.* **1995**, *38*, 1884.
(126) Romines, K. R.; Watenpaugh, K. D.; Howe, W. J.; Tomich, P. K.; Lovasz, K. D.; Morris, J. K.; Janakiraman, M. N.; Lynn, J. C.; Horng, M.-M.; Chog, K.-T.; Hinshaw, R. R.; Dolak, L. A. *J. Med. Chem.* **1995**, *38*, 4463.
(127) Romines, K. R.; Morris, J. K.; Howe, W. J.; Tomich, P. K.; Horng, M.-M.; Chong, K.-T.; Hinshaw, R. R.; Anderson, D. J.; Strohbach, J. W.; Turner, S. R.; Mizsak, S. A. *J. Med. Chem.* **1996**, *39*, 4125.
(128) Skulnick, H. I.; Johnson, P. D.; Aristoff, P. A.; Morris, J. K.; Lovasz, K. D.; Howe, W. J.; Watenpaugh, K. D.; Janakiraman, M. N.; Anderson, D. J.; Reischer, R. J.; Schwartz, T. M.; Banitt, L. S.; Tomich, P. K.; Lynn, J. C.; Horng, M.-M.; Chong, K.-T.; Hinshaw, R. R.; Dolak, L. A.; Seest, E. P.; Schwende, F. J.; Rush, B. D.; Howard, G. M.; Toth, L. N.; Wilkinson, K. R.; Kakuk, T. J.; Johnson, C. W.; Cole, S. L.; Zaya, R. M.; Zipp, G. L.; Possert, P. L.; Dalga, R. J.; Zhong, W.-Z.; Williams, M. G.; Romines, K. R. *J. Med. Chem.* **1997**, *40*, 1149.
(129) Gupta, S. P.; Babu, M. S.; Kaw, N. *J. Enzyme Inhib.* **1999**, *14*, 109.
(130) Tait, B. D.; Hagen, S.; Domagala, J.; Ellsworth, E. L.; Gajda, C.; Hamilton, H. W.; Vara Prasad, J. V. N.; Ferguson, D.; Graham, N.; Hupe, D.; Nouhan, C.; Tummino, P. J.; Humblet, C.; Lunney, E. A.; Pavlovsky, A.; Rubin, J.; Gracheck, S. J.; Baldwin, E. T.; Bhat, T. N.; Erickson, J. W.; Gulnik, S. V.; Liu, B. *J. Med. Chem.* **1997**, *40*, 3781.
(131) Vara Prasad, J. V. N.; Para, K. S.; Tummino, P. J.; Ferguson, D.; Mcquade, E. J.; Lunney, E. A.; Rapundalo, S. T.; Batley, B. L.; Hingorani, G.; Domagala, J. M.; Gracheck, S. J.; Bhat, T. N.; Liu, B.; Baldwin, E. T.; Erickson, J. W.; Sawyer, T. K. *J. Med. Chem.* **1995**, *38*, 898.
(132) Vazquez, M. L.; Bryant, M. L.; Clare, M.; Decrescenzo, G. A.; Doherty, E. M.; Freskos, J. N.; Getman, D. P.; Houseman, K. A.; Julien, J. A.; Kocan, G. P.; Mueller, R. A.; Shieh, H.-S.; Stallings, W. C.; Stegeman, R. A.; Talley, J. J. *J. Med. Chem.* **1995**, *38*, 581.
(133) Billich, A.; Charpiot, B.; Fricker, G.; Gstach, H.; Lehr, P.; Peichl, P.; Scholz, D.; Rosenwirth, B. *Antiviral Res.* **1994**, *25*, 215.
(134) Desolms, S. J.; Giuliani, E. A.; Guare, J. P.; Vacca, J. P.; Sanders, W. M.; Graham, S. L.; Wiggins, J. M.; Darke, P. L.; Sigal, I. S.; Zugay, J. A.; Emini, E. A.; Schleif, W. A.; Quintero, J. C.; Anderson, P. S.; Huff, J. R. *J. Med. Chem.* **1991**, *34*, 2852.
(135) Thompson, S. K.; Murthy, K. H. M.; Zhao, B.; Winborne, E.; Green, D. W.; Fisher, S. M.; Desjarlais, R. L.; Tomaszek, T. A., Jr.; Meek, T. D.; Gleason, J. G.; Abdel-Meguid, S. S. *J. Med. Chem.* **1994**, *37*, 3100.
(136) Hansch, C.; Gao, H. *Chem Rev.* **1997**, *97*, 2995.
(137) Gupta, S. P.; Babu, M. S.; Sowmya, S. *Bioorg. Med. Chem.* **1998**, *6*, 2185.
(138) Umezawa, H.; Aoyagi, T.; Morishima, H.; Matsuzaki, M.; Hamada, M.; Takeuchi, T. *J. Antibiot.* **1970**, *23*, 259.
(139) Katoh, I.; Yasunaga, T.; Ikawa, Y.; Yoshinaka, Y. *Nature (London)* **1987**, *329*, 654.
(140) Krausslich, H.G.; Schneider, H.; Zybarth, G.; Carter, C. A.; Wimmer, E. *J. Virol.* **1988**, *62*, 4393.
(141) Richards, A. D.; Roberts, R.; Dunn, B. M.; Graves, M. C.; Kay, J. *FEBS Lett.* **1989**, *247*, 113.
(142) Darke, P. L.; Leu, C. T.; Davis, L. J.; Heimbach, J. C.; Diehl, R. E.; Hill, W. S.; Dixon, R. A. F.; Sigal, I. S. *J. Biol. Chem.* **1989**, *264*, 2307.
(143) Nutt, R. F.; Brady, S. F.; Darke, P. L.; Ciccarone, T. M.; Colton, C. D.; Nutt, E. M.; Rodkey, J. A.; Bennett, C. D.; Waxman, L. H.; Sigal, I. S.; Anderson, P. S.; Veber, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7129.
(144) Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B. H.; Wlodawer, A. *Science* **1989**, *246*, 1149.
(145) Fitzgerald, P. M. D.; McKeever, B. M.; Van Middlesworth, J. F.; Springer, J. P.; Heimbach, J. C.; Leu, C.-T.; Herber, W. K.; Dixon, R. A. F.; Darke, P. L. *J. Biol. Chem.* **1990**, *265*, 14209.
(146) Erickson, J.; Neidhart, D. J.; VanDrie, J.; Kempf, D. J.; Wang, X. C.; Norbeck, D. W.; Plattner, J. J.; Rittenhouse, J. W.; Turon, M.; Wideburg, N.; Kohlbrenner, W. E.; Simmer, R.; Helfrich, R.; Paul, D. A.; Knigge, M. *Science* **1990**, *249*, 527
(147) Swain, A. L.; Miller, M. M.; Green, J.; Rich, D. H.; Schneider, J.; Kent, S. B. H.; Wlodawer, A. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8805.
(148) Huff, J. R. *J. Med. Chem.* **1991**, *34*, 2305.
(149) Getman, D. P.; DeCrescenzo, G. A.; Heintz, R. M.; Reed, K. L.; Talley, J. J.; Bryant, M. L.; Clare, M.; Houseman, K. A.; Marr, J. J.; Mueller, R. A.; Vazquez, M. L.; Shieh, H.-S.; Stallings, W. C.; Stegeman, R. A. *J. Med. Chem.* **1993**, *36*, 288.
(150) Holloway, M. K.; Wai, J. M.; Halgren, T. A.; Fitzgerald, P. M. D.; Vacca, J. P.; Dorsey, B. D.; Levin, R. B.; Thompson, W. J.; Chen, L. J.; deSolms, S. J.; Gaffin, N.; Ghosh, A. K.; Giuliani, E. A.; Graham, S. L.; Guare, J. P.; Hungate, R. W.; Lyle, T. A.; Sanders, W. M.; Tucker, T. J.; Wiggins, M.; Wiscount, C. M.; Woltersdorf, O. W.; Young, S. D.; Darke, P. L.; Zugay, J. A. *J. Med. Chem.* **1995**, *38*, 305.
(151) Ferguson, D. M.; Radmer, R. J.; Kollman, P. A. *J. Med. Chem.* **1991**, *34*, 2654.
(152) Kempf, D. J.; Norbeck, D. W.; Codacovi, L.; Wang, X. C.; Kohlbrenner, W. E.; Wideburg, N. E.; Paul, D. A.; Knigge, M. F.; Vasavanonda, S.; Graig-Kennard, A.; Saldivar, A.; Rosenbrook, W., Jr.; Clement, J. J.; Plattner, J. J.; Erickson, J. *J. Med. Chem.* **1990**, *33*, 2687.
(153) Wonacott, A.; Cooke, R.; Hayes, F. R.; Hann, M. M.; Jhoti, H.; McMeekin, P.; Mistry, A.; Murray-Rust, P.; Singh, O. M. P.; Weir, M. P. *J. Med. Chem.* **1993**, *36*, 3113.
(154) Humber, D. C.; Bamford, M. J.; Bethell, R. C.; Cammack, N.; Cobley, K.; Evans, D. N.; Gray, N. M.; Hann, M. M.; Orr, D. C.; Saunders, J.; Shenoy, B. E. V.; Storer, R.; Weingarten, G. G.; Wyatt, P. G. *J. Med. Chem.* **1993**, *36*, 3120.
(155) Waller, C. L.; Oprea, T. I.; Giolitti, A.; Marshall, G. R. *J. Med. Chem.* **1993**, *36*, 4152.
(156) Oprea, T. I.; Waller, C. L.; Marshall, G. R. *J. Med. Chem.* **1994**, *37*, 2206.
(157) Kroemer, R. T.; Ettmayer, P.; Hecht, P. *J. Med. Chem.* **1995**, *38*, 4917.
(158) Chen, X.; Tropsha, A. *J. Med. Chem.* **1995**, *38*, 42.
(159) Chen, P.; Cheng, P. T. W.; Alam, M.; Beyer, B. D.; Bisacchi, G. S.; Dejneka, T.; Evans, A. J.; Greytok, J. A.; Hermsmeier, M. A.; Humphreys, W. G.; Jacobs, G. A.; Kocy, O.; Lin, P.-F.; Lis, K. A.; Marella, M. A.; Ryono, D. E.; Sheaffer, A. K.; Spergel, S. H.; Sun, C.-Q.; Tino, J. A.; Vite, G.; Colonno, R. J.; Zahler, R.; Barrish, J. C. *J. Med. Chem.* **1996**, *39*, 1991.
(160) Bridger, G. J.; Skerlj, R. T.; Thornton, D.; Padmanabhan, S.; Martelucci, S. A.; Henson, G. W.; Abrams, M. J.; Yamamoto, N.; De Vreese, K.; Pauwels, R.; De Clercq, E. *J. Med. Chem.* **1995**, *38*, 366.
(161) Joao, H. C.; De Vreese, K.; Pauwels, R.; De Clercq, E.; Henson, G. W.; Bridger, G. J. *J. Med. Chem.* **1995**, *38*, 3865.
(162) Inouye, Y.; Kanamori, T.; Yoshida, T.; Bu, X.; Shionoya, M.; Koike, T.; Kimura, E. *Biol. Pharm. Bull.* **1994**, *17*, 243.
(163) Raghavan, K.; Buolamwini, J. K.; Fesen, M. R.; Pommier, Y.; Kohn, K. W.; Weinstein, J. N. *J. Med. Chem.* **1995**, *38*, 890.
(164) Kohlstaedt, L. A.; Wang, J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A. *Science* **1992**, *256*, 1783.
(165) Ren, J.; Esnouf, R.; Garman, E.; Somers, D.; Ross, C.; Kirby, I.; Keeling, J.; Darby, G.; Jones, Y.; Stuart, D.; Stammers, D. *Nature Struct. Biol.* **1995**, *2*, 293.
(166) Esnouf, R.; Ren, J.; Ross, C.; Jones, Y.; Stammers, D.; Stuart, D. *Nature Struct. Biol.* **1995**, *2*, 303.
(167) Hopkins, A. L.; Ren, J.; Esnouf, R. M.; Willcox, B. E.; Jones, E. Y.; Ross, C.; Miyasaka, T.; Walker, R. T.; Tanaka, H.; Stammers, D. K.; Stuart, D. I. *J. Med. Chem.* **1996**, *39*, 1589.
(168) Smerdon, S. J.; Jäger, J.; Wang, J.; Kohlstaedt, L. A.; Chirino, A. J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3911.
(169) Mui, P. W.; Jacober, S. P.; Hargrave, K. D.; Adams, J. *J. Med. Chem.* **1992**, *35*, 201.
(170) Schäfer, W.; Friebe, W.-G.; Leinert, H.; Mertens, A.; Poll, T.; von der Saal, W.; Zilch, H.; Nuber, B.; Zeigler, M. L. *J. Med. Chem.* **1993**, *36*, 726.
(171) Ding, J.; Das, K.; Moereels, H.; Koymans, L.; Andries, K.; Janssen, P. A. J.; Hughes, S. H.; Arnold, E. *Nature Struct. Biol.* **1995**, *2*, 407.
(172) Mager, P. P.; De Clercq, E.; Takashima, H.; Ubasawa, M.; Sekiya, K.; Baba, M. Walther, H. *Eur. J. Med. Chem.* **1996**, *31*, 701.
(173) Mager, P. P. *Drug Des. Discovery* **1996**, *14*, 241.
(174) Mager, P. P. *Drug Des. Discovery* **1996**, *14*, 213.
(175) Mager, P. P.; Walther, H. *Drug Des. Discovery* **1996**, *14*, 225.
(176) Wang, S.; Milne, G. W. A.; Yan, X.; Posey, I. J.; Nicklaus, M. C.; Graham, L.; Rice, W. G. *J. Med. Chem.* **1996**, *39*, 2047.
(177) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *Arch. Biochem. Biophys.* **1978**, *185*, 584.
(178) http:/www.niaid.nih.gov/daids/dtpdb/fdadrug.htm.
(179) Hansch, C.; Bjorkroth, J. P.; Leo, A. *J. Pharm. Sci.* **1987**, *76*, 663.
(180) Debnath, A. K. *J. Med. Chem.* **1999**, *42*, 249.

(181) Debnath, A. K. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 761.
(182) De Clercq, E. *Il Farmaco* **1999**, *54*, 26.
(183) Spence, R. A.; Kati, W. M.; Anderson, K. S.; Johnson, K. A. *Science* **1995**, *267*, 988.
(184) Das, K.; Ding. J.; Hsiou, Y.; Clark, A. D., Jr.; Moereels, H.; Koymans, L.; Andries, K.; Pauwels, R.; Janseen, P. A. J.; Boyer, P. L.; Clark, P.; Smith, R. H., Jr.; Smith, M. B. K.; Michejda, C. J.; Hughes, S. H.; Arnold, E. *J. Mol. Biol.* **1996**, *264*, 1085.
(185) http://www.rcsb.org/pdb/.
(186) Hodge, C. N.; Aldrich, P. E.; Bacheler, L. T.; Chang, C. H.; Eyermann, C. J.; Garber, S.; Grubb, M.; Jackson, D. A.; Jadhav, P. K.; Korant, B.; Lam, P. Y.; Maurin, M. B.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Reid, C.; Sharpe, T. R.; Shurn, L.; Winslow, D. L.; Erickson-Viitanen, S. *Chem. Biol.* **1996**, *3*, 301.
(187) Ala, P. J.; Huston, E. E.; Klabe, R. M.; McCabe, D. D.; Duke, J. L.; Rizzo, C. J.; Korant, B. D.; DeLoskey, R. J.; Lam, P. Y.; Hodge, C. N.; Chang, C. H. *Biochemistry* **1997**, *36*, 1573.
(188) Hansch, C.; Klein, T. *Acc. Chem. Res.* **1986**, *19*, 392.
(189) Selassie, C. D.; Li, R.-L.; Poe, M.; Hansch, C. *J. Med. Chem.* **1991**, *34*, 46.

CR9703358