Camptothecins: A SAR/QSAR Study

Rajeshwar P. Verma* and Corwin Hansch

Department of Chemistry, Pomona College, 645 North College Avenue, Claremont, California 91711

Received July 13, 2007

Contents

1. Introduction 213
2. Structure–Activity Relationships (SAR) 215
2.1. Quinoline (A/B) Ring 215
2.2. C/D Ring 216
2.3. E Ring 216
2.4. D/E Ring 216
3. Quantitative Structure–Activity Relationships (QSAR) 217
3.1. Review of QSAR Studies from the Literature 217
3.1.1. QSAR for the Inhibition of DNA Topoisomerase I 217
3.1.2. QSAR for the Inhibition of Various Cancer Cells 218
3.2. Evaluation of New QSAR Models 220
3.2.1. Materials and Methods 220
3.2.2. Results and Discussion 221
3.2.3. Validation of the QSAR Models 231
3.2.4. Overview 232
4. Summary of the SAR and QSAR 233
5. Conclusions 233
6. References 234

1. Introduction

Camptothecin (CPT, **I**), a unique pentacyclic quinoline alkaloid originally isolated from a native tree of Tibet and China called *Camptotheca acuminata* in latin and *Xi Shu* in Chinese, is one of the prominent lead compounds in anticancer drug development.[1−3] It has been identified from the early assessments that the importance of 20*S* chiral carbon of CPT for their activity and also pointed out a dynamic equilibrium between the close-ring lactone and open-ring carboxylic acid forms at physiological pH. Due to the extremely poor solubility of CPT in water, clinical trials were initiated using its water-soluble sodium salt (**II**; Figure 1). The results were disappointing: biological activity was weak relative to xenograph models and unexpected side effects including hemorrhagic cystitis and myelotoxicity, which resulted in suspension of the trials.[4,5] Later on, it was established that the α-hydroxy lactone ring moiety must be intact for antitumor activity and that this ring was being opened in the preparation of the sodium salt.[6]

In a conformational analysis, torsional parameters for the MM3(96) force field were obtained by Carrigan et al.[7] for the α-hydroxy lactone and CPT using ab initio calculations on representative compounds containing the critical dihedral angles. MM3(96) predicts two distinct "boat-like" conformations for the α-hydroxy lactone moiety. The low-energy lactone conformation predicted by MM3(96) is in good agreement with X-ray crystal structures of CPT iodoacetate and 7-ethyl-10-(4-piperidino)piperidinylcarbonyloxy CPT HCl as well as the ab initio structure of a CPT-like α-hydroxy lactone.

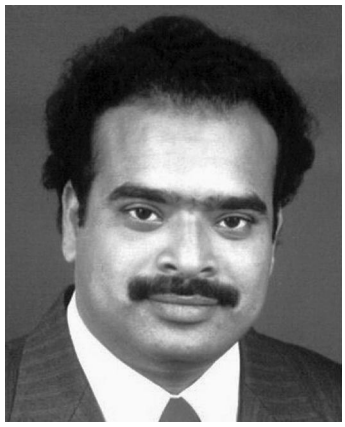
Nearly 20 years later, the discovery that the primary cellular target of CPT is DNA topoisomerase I (topo I) was the breakthrough that renewed interest in this agent and led to synthesizing more water-soluble analogues.[8−10] Two of them, topotecan (Hycamptin, **III**) for the clinical treatment of the ovarian and small-cell lung cancers,[11−14] and irinotecan (Camptosar or CPT-11, **IV**)[15,16] for the metastatic colorectal cancers have already gained approval by the Food and Drug Administration (FDA) of the U.S.A.[17,18] Irinotecan is a prodrug that is converted into their active metabolic form 10-hydroxy-7-ethylcamptothecin (SN-38, **V**; Figure 2). These two drugs (topotecan and irinotecan) and other derivatives of CPT have become a part of the multimillion dollar industry that is dedicated to finding better chemotherapeutic agents with excellent antitumor activity and less normal tissue toxicity. To achieve this goal, it is necessary to understand the details about the mechanisms of action, the targets of these drugs, and the cellular response to the drugs.

Human topoisomerase I (topo I) relaxes superhelical tension associated with DNA replication, transcription and recombination by reversibly nicking one strand of duplex DNA and forming a covalent 3′-phosphotyrosine linkage. This enzyme is the sole target of the CPT family of anticancer compounds, which acts by stabilizing the covalent protein–DNA complex and enhancing apoptosis through blocking the advancement of replication forks. Once the CPT molecule has intercalated into the topo I-DNA cleavable complex, the collision between the complex and the replication fork during S-phase is thought to result in DNA double strand breaks (DSBs) that eventually lead to cell death.[18,19] It has also been suggested that topo I cleaves DNA at multiple sites. The highest efficient sites exhibit significant sequence homology. Approximately 90% of topo I site have a tyrosine residue at position-1. However, sites of cleavage stabilized by CPT exhibit a strong preference for guanine at +1 position, while thymidine remains the preferred nucleobase at the −1 position.[20]

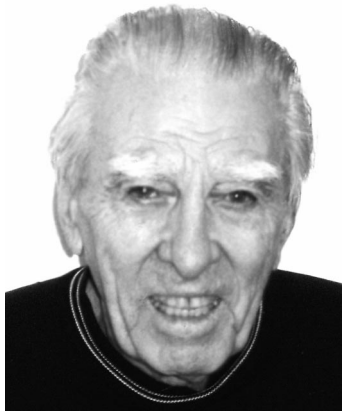
The exact mechanism by which CPT stabilizes the DNA-topo I covalent binary complex is not fully understood because the drug acts as an uncompetitive inhibitor and binds only the transient binary complex.[21] Enzymology studies have revealed that CPT does not interact with topo I alone, nor does it bind to DNA.[22] Although it has been reported that topotecan, which should be protonated at physiological pH, does bind to DNA at high concentration.[23] Despite the

* Phone: (909) 607-4249. Fax: (909) 607-7726. E-mail: rverma@pomona.edu.

10.1021/cr0780210 CCC: $71.00 © 2009 American Chemical Society
Published on Web 12/19/2008



Rajeshwar P. Verma received his M.Sc. (1988) and Ph.D. (1992) degrees in chemistry from Magadh University, Bodh-Gaya (India). After working one year at the same university as a postdoctoral fellow with Professor K. S. Sinha, he joined Roorkee University (now IIT Roorkee) as a research associate and worked with Professor S. M. Sondhi (1993−1997). He also worked as a Lecturer of Chemistry at Gurukula Kangri University, Hardwar (1994−1995). He won a Research Associateship Award in December 1994 from the Council of Scientific and Industrial Research, New Delhi (India). In 1997, he moved to Pomona College to join the renowned QSAR research group of Professor Corwin Hansch and Cynthia Selassie, working as a postdoctoral research associate. Dr. Verma's research interests include the following: isolation, characterization, and synthesis of natural products derived from medicinal plants; chemistry of isothiocyanates; synthesis of biologically important heterocyclic and phenolic compounds; quantitative structure−activity relationships (QSAR); and computer-assisted drug design (CADD).



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 Du Pont Co., 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, Switzerland, 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. He is an honorary fellow of the Royal Society of Chemistry and received the ACS Award for Computers in Chemical and Pharmaceutical Research for 1999.

apparent lack of affinity of CPT for DNA or topo I alone, the binding of CPT to the covalent binary complex is suggested to be responsible for the observed stabilization. Recently, an observation from the X-ray crystal structure of a ternary complex containing a human topo I covalently attached to a DNA duplex and bound to topotecan, suggests that topotecan intercalates at the site of DNA cleavage and is stabilized by base-stacking interactions with both the upstream (−1) and downstream (+1) base pair. The intercalation resulted in a shift of the downstream base pairs and displacement of the 5′-OH strand away from the phosphotyrosine bond thus blocking relegation. This binding occurred whether the E ring of topotecan was in the closed lactone form or the open carboxylate form; however, a higher occupancy rate (63%) was seen with the lactone form.[24]

More recently, atomic force microscope (AFM) images have been used by Argaman et al.[25] to investigate the mode of action of DNA topo I in the presence and absence of CPT. The AFM analysis revealed that the position of the enzyme in the topo I-DNA covalent complexes (in the presence of CPT) differed from its position in the absence of this drug. Topo I was attached to the double stranded relaxed DNA molecules in the absence of CPT, while in the presence of this drug the enzyme was located inside a relaxed DNA bubble.

Another possible mechanism of cell death by CPT is by blocking angiogenesis. In a study to investigate the antiangiogenic and antitumor effects of oral ST1481 (gimatecan) in human tumor xenografts, Petrangolini and co-workers[26] have suggested the possibility that the antiangiogenic properties of ST1481 contribute to its antitumor potential and that this effect might be enhanced by the continuous low-dose treatment. Recent results have shown that CPT-11 is an effective inhibitor of angiogenesis and providing strong implications for wider clinical application for colon cancer.[27]

CPT is a DNA topo I inhibitor found to be effective in treating psoriasis by inhibiting the growth of keratinocytes in vitro by inducing apoptosis. CPT inhibited keratinocyte proliferation and telomerase activity as well as the proliferation and induction of apoptosis. The inhibitory effect of CPT is correlated with its concentration, suggesting concentration dependency. Down-regulation of telomerase activity was observed not only in cells treated at concentrations able to induce apoptosis, but also in cells treated at concentration insufficient to induce apoptosis, indicating that CPT-induced apoptosis may be preceded by down-regulation of telomerase activity. The antiproliferative activity of CPT and its inhibition of keratinocyte apoptosis through down-regulation of telomerase activity may explain the drug's therapeutic mechanism in psoriasis.[28] Another study demonstrated that SN-38 produces an increase in production of pro-apoptotic factors such as p53, Bax, Bcl-xl, and p-21/WAF-1 in colon cancer.[18,29]

Resistance to chemotherapeutic agents is a major clinical complication in cancer therapy. Studies of CPT resistance using yeast and mammalian cell culture models suggest three general mechanisms of resistance: (i) reduced cellular accumulation of CPTs, (ii) alteration in the structure or location of topo I, and (iii) alterations in the cellular response to CPT−DNA−ternary complex formation. The relevance of these mechanisms to clinical drug resistance is not yet known, but evaluation of these models in clinical specimens should enhance the use of CPTs both as single agents and in combination with other anticancer drugs.[30] Recently, resistance to CPT has been attributed to enhanced drug efflux by a novel ABC transporter, the BCRP/MXR/$ABCG_2$ transporter, which is widely expressed in normal human tissues.[31,32] Mutations that impart resistance to CPT have also been identified in several regions of human topo I. Chrencik and co-workers[33] have presented the crystal structures of two CPT-resistant forms of human topo I (Phe361Ser at 2.6 Å resolution and Asn722Ser at 2.3 Å resolution) in ternary complexes with DNA and topotecan. The alteration of Asn722 to Ser leads to the elimination of a water-mediated

contact between the enzyme and topotecan. Further consideration of CPT-resistant mutations at seven additional sites in human topo I presented the structural evidence that explaining their possible impact on drug binding. These results provide better understanding toward the mechanism of cell poisoning by CPT and suggest specific modifications to the drug that may improve efficacy.

Molecular docking studies were performed by Lauria et al.[34] on a series of 24 CPT-like topo I inhibitors present in the NCI anticancer agents mechanism database, using four different topo I-DNA-inhibitor complexes, with the aim to investigate the binding modes of these derivatives. The analysis of the best docked conformations has confirmed the role of some amino acids present in the active site. These results may explain the role of some single point mutation in developing resistance to CPTs and are useful to understand the structural features required to improve the performance of CPT derivatives as topo I inhibitors.

An NMR study was undertaken by Bocian et al.[35] to evaluate the binding constant and mode of binding of the CPT family drugs CPT and TPT to the DNA octamers $d(GCGTACGC)_2$ and $d(GCGATCGC)_2$. All the results indicate binding of the TPT lactone form, and chemical shift changes show preference of TPT binding to the terminal G1 unit rather than to internal bases. The carboxylate form of TPT interacts more weakly than the lactone form with DNA. Intermolecular NOE effects between the DNA oligomer and TPT were observed at pH 5 and 3 °C. The observed cross-peaks cannot be reconciled with a single TPT/DNA complex structure, which suggests a model of a limited number of conformations in fast exchange. MD calculations on four pairs of starting structures with TPT stacked onto the G1−C8 base pair in different orientations were therefore performed. The use of selected experimental "docking" restraints yielded 10 MD trajectories covering a wide conformational space. These calculations support a combination of two major families of conformations in fast exchange. One of these is the conformation found in a crystal of a TPT/DNA/topoisomerase I ternary complex.

To improve the pharmacokinetics, drug resistance, clinical efficacy, and toxicity profiles of the original CPT molecule, scientific efforts are continued to develop new CPT analogues. At present, over a dozen new CPT derivatives are in various stages of clinical trials.[3] The historical achievements,[2,36] syntheses,[17,20,37] clinical applications,[3,38−41] and mechanism of actions[17,18,20,37,38,41−46] of CPTs have already been discussed in several excellent reviews. In the present review, we demonstrate the structure−activity relationships (SAR) and the quantitative structure−activity relationships (QSAR) of CPT derivatives to understand their chemical−biological interactions, which may provide strategies that might aid in the development of outstanding antitumor agents belonging to this family. It is important to distinguish between SAR and QSAR: SAR is qualitative in nature, often occurring in the form of structural alerts that include molecular substructures or fragment counts related to the presence or absence of biological activity; while QSAR is typically quantitative in nature, producing categorical or continuous prediction scales.[47,48]

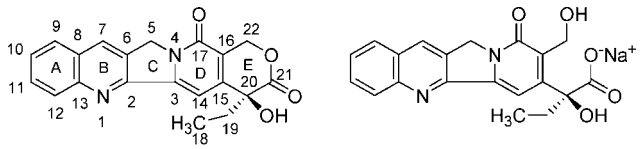
2. Structure−Activity Relationships (SAR)

The design of the novel CPT derivatives depends on the following four assumtions: (i) The essential structural features for the activity of CPT derivatives are 20(*S*)-hydroxyl, pyridone moiety (D-ring), lactone moiety (E-ring), and planarity of the pentacyclic (A, B, C, D, and E) ring system. Thus, the C-D-E rings of CPT derivatives cannot be altered.[44,49] (ii) The modifications of quinoline ring (9, 10, and 11-positions of the A-ring and 7-position of the B-ring) generally enhance the potency of the CPT derivatives in both in vivo and in vitro studies.[44,50] (iii) The development of homocamptothecins (hCPTs; a new family of CPT derivatives with seven-membered lactone E-ring) with enhanced activity necessitate a reevaluation of the E-ring lactone function.[17,51] (iv) For the targeted enzymatic activation of tumor cells, CPT derivatives must have to possess these four criteria: (a) improved water solubility, (b) stability in blood, (c) decreased cytotoxicity, and (d) susceptibility to defined enzymatic cleavage.[17,52]

The details about the structure−activity relationships (SAR) of CPT derivatives with respect to their pentacyclic (A, B, C, D, and E) ring system are as follows:

2.1. Quinoline (A/B) Ring

It has already been established that the CPT derivatives with modifications at quinoline (A/B) ring are of great interest. This is further supported by the fact that the only two CPT analogues approved for clinical use, that are, topotecan (III) and irinotecan (IV), are derivatives with substitutions within the quinoline ring. The SAR for the quinoline (A/B) ring of the CPT derivatives includes the following: (i) Monosubstitution at 9, 10, or 11 positions of the A ring by NH_2 or OH group increases the antitumor activity, whereas substitution at position 12 greatly reduces the activity.[53] (ii) Substitution at 9 and 10 positions of the A ring by halides and other electron-rich groups (e.g., NH_2, OH, etc.) generally increases the DNA topo I inhibition.[54] (iii) Substitution at 10 and 11 positions of the A ring are generally unfavorable to the biological activity.[55] Exceptions are the 10,11-methylenedioxy or 10,11-ethylenedioxy functional group at the A ring substantially increases the DNA topo I inhibition.[54,56] (iv) Small substituents at 10 position of the A ring generally increases the DNA topo I inhibition. Substitution at 10 position with a hydroxyl group contributes to the increased activity of SN-38 and topotecan.[45] (v) Substitution at the 11 position of the A ring by fluorine or cyano group also increases the DNA topo I inhibition.[57] (vi) Substitution at the 7 position of the B ring has been found to be more potent, and an increase in water solubility has been observed depending on the nature of the substituent.[54] (vii) Substitutions at 7 and 9 positions of B and A rings do not affect the DNA topo I inhibitory activity, suggesting the absence of tight interaction with the receptor site and the regions around positions 7 and 9.[45] (viii) Comparison of the stability among CPT-, SN-38-, and 10-OH-CPT-induced cleavable complexes reveals that the OH group at position 10 enhances stability of the cleavable complexes. The difference between SN-38- and 10-OH-CPT-induced cleavable complexes suggests that the ethyl group at the 7 position is also important for stabilizing the interaction between CPT derivatives and topo I-DNA complex.[58] (ix) Substitutions at 7 and 10 positions of B and A rings (where the substituent at position 10 is a hydroxyl group) result in greatly improved human blood stabilities of CPT derivatives. SN-38 is one of the best examples with 7-alkyl-10-hydroxy substitution pattern.[59] (x)The stability of the E-ring lactone in human plasma can be affected by derivatization of the quinoline ring.[20]



I II

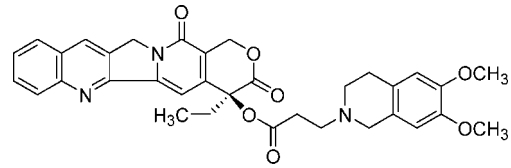
Figure 1. Structure of 20(*S*)-camptothecin (**I**) and their ring-opened carboxylic acid sodium salt (**II**).

2.2. C/D Ring

There are few reports that have focused on the modification of C and D rings of CPT (Figure 3). It has generally been suggested that either the replacement or the substitution at C/D ring would reduce the activity. Reports on C/D-ring substituents have also been limited, presumably due to the paucity of accessible carbons for substitution and more difficult synthetic routes leading to potential analogues. There are only two available sites for substituents, C-5 and C-14 in the C/D-ring.[20] The SAR for the C/D- ring of the CPT derivatives includes the following: (i) Substitution at 5 position of the C ring by acetoxy, alkoxy, amino, or hydroxyl group generally diminishes the antitumor activity.[20,60] (ii) Substitutions of alkoxy or several other groups at C-5 of the C-ring are reasonably well tolerated when accompanied by the additions of hydroxyl or nitro groups to the A-ring. 9-OH-5-OEt-CPT is one of the best examples.[42,61] Exceptions are the CPT derivatives, **VI** and **VII**, which contained no additional group within the quinoline ring and showed a good or maintain the same potency equivalent to CPT.[62,63] (iii) Substitution at C-14 of the D-ring by methyl ester group generally reduces the antitumor activity.[64] (iv) The pyridone carbonyl of CPT is an important ring for stabilizing the enzyme−DNA−CPT ternary complex. The deaza derivative of CPT (**VIII**) is the best example, which should maintain quite a similar shape and planarity relative to CPT, was found to be approximately 60-fold less efficient as a topo I inhibitor.[65] (v) 14-Azacamptothecin (**IX**) has exhibited reasonable potency as a topo I poison and topo I dependent cytotoxic agent, and stabilized enzyme-linked DNA breaks with the same sequence selectivity as CPT itself.[17,66]

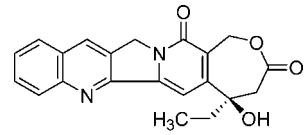
2.3. E Ring

CPT occurs in two different enantiomeric forms 20-*S* and 20-*R* indicating the particular arrangement of atoms and groups in space around the chiral center (C-20). It has already been demonstrated that 20-*R* CPT is inactive both in topo I inhibition and in vivo assays, while the 20-*S* CPT has great potency in inhibiting human colon cancer xenografts in nude mice.[67] The SAR for the E-ring of the CPT derivatives includes the following: (i)The lactone (E-ring) with 20-*S* configuration is critical for both efficient topo I inhibition and in vivo potency. (ii) Any changes in the E-ring, such as replacement of the lactone by a lactum group, reduction of the lactone, removal of the carbonyl oxygen, or removal of the 20-hydroxyl group, inactivate the molecule.[68] (iii) Under physiological conditions, the presence of α-OH group (20-*S* configuration) results in an equilibrium that favors the (inactive) open carboxylate over the (active) ring-closed lactone form.[17,20,41,59,69,70] (iv) There are two possible reasons for the importance of α-OH group (20-*S* configuration) of CPT for topo I inhibition: (a) the formation of a hydrogen bond between the hydroxyl group and the enzyme−DNA complex and (b) the presence of an intramolecular hydrogen bond with the lactone carbonyl of CPT (C-21 position). Both interactions may facilitate the possible E-ring opening reaction.[64,68] (v) Esterification of the 20-OH group, which can either eliminate the intramolecular hydrogen bonding or increase the steric hindrance of the carbonyl group of E-ring, results in the stability of the lactone ring in vivo.[17,71,72] A series of 20-*O*-linked nitrogen-based CPT esters including ester (**X**) was reported by Wang et al.,[73] which possesses the both lower cytotoxic in vitro and better antitumor activity in vivo than topotecan.

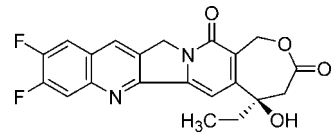


X

(vi) Replacement of the 20-OH group by amino or halogens results in the significant diminished activity.[17,20] (vii) It is important to note that homoCPT (**XI**) obtained by the replacement of α-hydroxylactone moiety with β-hydroxy-lactone exerts potent inhibition of topo I, elevated levels of cytotoxicity, and stability of the lactone after 24 h at physiological pH.[74] These demonstrations have prompted the synthesis and evaluation of numerous homoCPT derivatives.[51,74−78] Among them, the most important homoCPT derivatives is 10,11-difluoro-homoCPT (**XII**), which exhibits strong anti-proliferative activity against numerous cell lines and is currently in phase I clinical trials.[17,79]



XI



XII

2.4. D/E Ring

The SAR for the D/E ring of CPT derivatives with respect to their involvement in hydrogen bonding with topo I-DNA complex includes the following: (i) It has already been proposed that O17 (carbonyl oxygen), O20 (oxygen of the 20-OH group), O21 (lactone carbonyl), O (lactone oxygen), and the H atom of the 20-OH group of CPT would be involved in hydrogen bonding with the topo I-DNA complex.[50,80−82] (ii) On the basis of molecular electrostatic potential (MEP) values, Jena and Mishra[82] have shown that hydrogen-bond strengths are involved for the four oxygen atoms where these atoms would serve as hydrogen bond acceptors and follow the order O17 (carbonyl oxygen) > O21 (lactone carbonyl) > O (lactone oxygen) > O20 (oxygen of the 20-OH group). The H atom of the 20-OH group would be a hydrogen bond donor. The surface MEP value near to the H atom of the 20-OH group of CPT is appreciably smaller due to an intramolecular hydrogen bond between the H atom of 20-OH group and the O21 (lactone carbonyl) atom.

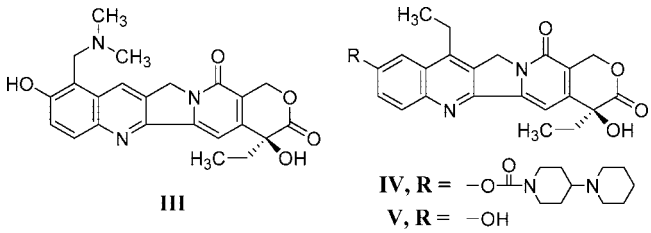


Figure 2. Structure of topotecan (**III**), Irinotecan (**IV**), and SN-38 (**V**)

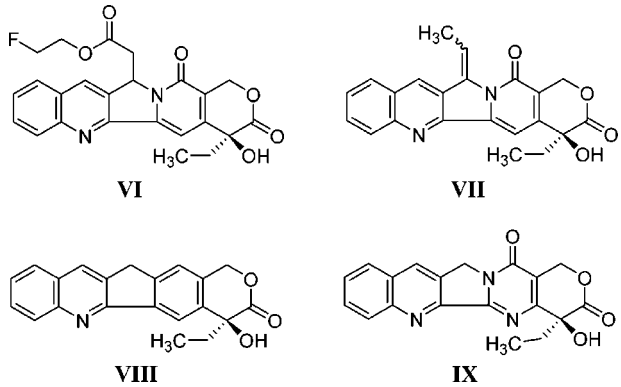


Figure 3. Structure of CPT analogues modified in C/D-ring.

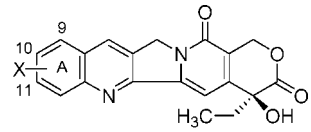
3. Quantitative Structure−Activity Relationships (QSAR)

Quantitative structure−activity relationship (QSAR) is one of the well-developed areas in the computational chemistry. The interest in the application of QSAR has steadily increased in recent decades because it has repeatedly proven itself to be a low-cost, high-return investment. In the past 45 years, the use of QSAR, since the advent of this methodology,[83] has become increasingly helpful in understanding many aspects of chemical−biological interactions in drug and pesticide research, as well as in the field of toxicology.[84,85] This method is useful in elucidating the mechanisms of chemical−biological interactions in various biomolecules, particularly enzymes as well as membranes, organelles, and cells, and in humans.[84,86−88] It has also been used for the evaluation of absorption, distribution, metabolism, and excretion (ADME) phenomena in many organisms and whole animal studies.[89,90] The QSAR approach employs extrathermodynamically derived and computational-based descriptors to correlate biological activity in isolated receptors, cellular systems, and in vivo. Four standard molecular descriptors routinely used in QSAR analysis are electronic, hydrophobic, steric, and topological indices. These descriptors are invaluable in helping to delineate a large number of receptor−ligand interactions that are critical to biological processes. The quality of a QSAR model depends strictly on the type and quality of the data, and not on the hypotheses and is valid only for the compound structure analogues to those used to build the model. QSAR models can stand alone to augment other computational approaches or can be examined in tandem with the equations of a similar mechanistic genre to establish their authenticity and reliability. Potential use of QSAR models for screening of chemical databases or virtual libraries before their synthesis appears equally attractive to chemical manufacturers, pharmaceutical companies, and government agencies. All the QSAR models reported in this review are for 20-*S*-CPT derivatives.

3.1. Review of QSAR Studies from the Literature

3.1.1. QSAR for the Inhibition of DNA Topoisomerase I

When the data of Wall et al.[56] for the inhibition of DNA topoisomerase I by X-CPTs (**XIII**) was used, eq 1 was developed.[91]



$$\log 1/C = 0.43\mathrm{C}\log P - 0.43\sigma^{+}_{X} - 0.89\mathrm{MR}_9 + 1.11I + 6.37 \quad (1)$$

where $n = 17$, $r^2 = 0.862$, $s = 0.226$, $q^2 = 0.681$, $Q = 4.108$, and $F_{4,12} = 18.739$.

In eq 1, C is the molar concentration of X-CPTs (**XIII**) that cause 50% inhibition of topo I cleavable complex formation, σ^{+}_{X} is the sum of the Hammett parameters of substituents at positions 9, 10, and 11, while MR_9 is the molar refractivity only for the substituents at positions 9. The negative coefficient with σ^{+}_{X} (−0.43) implies that highly electron releasing substituents at positions 9, 10, and 11 may strengthen the inhibitory activity of these compounds against topo I. An unfavorable steric interaction at C-9 position has been detected by the presence of a negative coefficient of MR_9 (−0.89). The indicator variable I is assigned the value of 1 and 0 for the presence and absence of X = 10-OCH_2O-11 group. Its positive coefficient suggests that the presence of a 10,11-methylenedioxy moiety at the A ring increases the activity. A positive hydrophobic effect for the whole molecule is also appeared in this equation. However, this equation does not allow any clue for an adequate distinction between the σ^{+} responsiveness of topo I inhibition for 9-, 10-, and 11-substituents. Considering these drawbacks of eq 1, QSAR **2** was developed using only 10-X-CPTs from the same data set of Wall et al.,[56] which gave a very good correlation between the inhibitory activities of 10-X-CPTs toward topo I and the hydrophobic parameters of their X-substituents.[92]

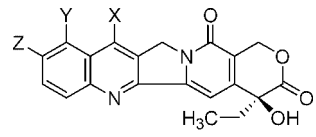
$$\log 1/C = 0.75(\pm 0.20)\pi_X + 6.24(\pm 0.10) \quad (2)$$

where $n = 8$, $r^2 = 0.936$, $s = 0.110$, $q^2 = 0.890$, $Q = 8.795$, and $F_{1,6} = 87.750$.

In the above QSAR model, π_X is the hydrophobicity of the substituents at position 10. Its positive coefficient (+0.75) suggests that the presence of highly hydrophobic substituents at position 10 increases the activity.

In a recent work,[93] two QSAR models **3** and **4** were demonstrated for the inhibition of topo I by CPT analogues **XIV** and **XV**, respectively.

Inhibition of DNA topo I by 7-X-9-Y-10-Z-CPTs (**XIV**)[93]

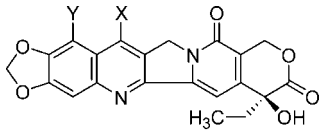


$$\log 1/C = 0.84(\pm 0.22)\text{Clog } P - 0.62(\pm 0.14)B5_X - 0.59(\pm 0.33)\text{MR}_Y - 0.45(\pm 0.36)\text{MR}_Z + 0.88(\pm 0.28)I_1 - 0.52(\pm 0.28)I_2 + 6.37(\pm 0.31) \quad (3)$$

where $n = 30$, $r^2 = 0.854$, $s = 0.258$, $q^2 = 0.738$, $Q = 3.581$, and $F_{6,23} = 22.422$.

In this equation, C represents the minimum concentration (mol/L) of CPT analogues (**XIV**) that inhibited the cleavable complex formation by 50%. Clog P is the calculated partition coefficient in n-octanol/water and is a measure of hydrophobicity for the whole molecule. Positive coefficient of Clog P suggests that the highly hydrophobic molecules will be more active. $B5_X$ is a Verloop's sterimol descriptor, which measures the maximum width of the X-substituents. MR_Y and MR_Z are the calculated molar refractivities of Y and Z substituents, respectively. A negative sign associated with $B5_X$, MR_Y, and MR_Z brings out steric effect for these substituents. The indicator variable (I_1) takes the value of 1 and 0 for the presence and absence of hydroxyl group at position 10, respectively. Similarly, I_2 takes the value of 1 and 0 for the presence and absence of n-alkyl groups at position 7, respectively. The presence of the hydroxyl group at position 10 increases the activity as evidenced by the positive coefficient of the indicator variable (I_1). The negative coefficient of the indicator variable (I_2) suggests that the presence of branched alkyl groups at position 7 is preferred over n-alkyl groups. On the basis of the QSAR model **3**, two compounds (X = H, Y = Z = Br, and X = H, Y = Z = OH) are suggested as potential synthetic targets that also fulfill the conditions of Lipinski's "rule of five" for oral bioavailability. The predicted log $1/C$ for these two compounds obtained from eq 3 is 7.29 and 7.58, respectively.

Inhibition of DNA topo I by 7-X-9-Y-10,11-methylenedioxycamptothecins (**XV**)[93]



XV

$$\log 1/C = -0.71(\pm 0.20)\pi_X - 0.88(\pm 0.39)I_3 + 0.73(\pm 0.44)I_4 + 7.19(\pm 0.22) \quad (4)$$

where $n = 16$, $r^2 = 0.873$, $s = 0.215$, $q^2 = 0.788$, $Q = 4.349$, and $F_{3,12} = 27.496$.

In this equation, C is the molar concentration of CPT analogues (**XV**) that inhibited the cleavable complex by 50%. π_X is the calculated hydrophobic parameter for X-substituents, which is the most important parameter for this series of compounds. The negative coefficient with π_X suggests that molecules (**XV**) with highly hydrophilic X-substituents would present better inhibitory activity. I_3 and I_4 are indicator variables, which pinpoint the unusual activities of X = aminomethyl groups and X = cyclic groups, respectively. The negative coefficient of I_3 suggests that the presence of X = aminomethyl groups will be detrimental to the activity. On the other hand, the presence of X = cyclic groups will promote the inhibitory activity as shown by the positive coefficient of I_4.

On the basis of the QSAR model **4**, three compounds [X = CH_2(-$NCH_2CH_2NHCH_2CH_2$-), Y = H; X = CH_2NHCH_3, Y = H; and X = cyclo-C_4H_7, Y = H] are suggested as potential synthetic targets that also fulfill the conditions of Lipinski's "rule of five" for oral bioavailability. The predicted log $1/C$ for these three compounds obtained from eq 4 is 6.50, 6.75, and 7.01, respectively.

Three-dimensional quantitative structure–activity relationship (3D QSAR) with a comparative molecular field analysis (CoMFA) was conducted by Carrigan et al.[54] on a series of 32 CPT analogues to correlate topo I inhibition with their steric and electrostatic properties. CoMFA model with the greatest predictive validity was obtained when both the R- and S-isomers were included in the data set, and semiempirical charges were calculated for MM3 minimized low-energy lactone structures. A cross-validated R^2 (r^2_{cv}) of 0.758 and a noncross-validated R^2 of 0.916 were obtained for MM3 minimized structures with PM3 ESP charges for the 32 CPT analogues.

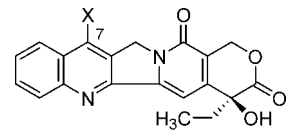
A training set of 43 structurally diverse CPT analogues for their topo I inhibition activities was used by Lu et al.[94] to construct a 3D QSAR with a comparative molecular field analysis. The CoMFA model gave a good cross-validated correlation in which q^2 was 0.495. Then, the analysis of the noncross-validated PLS model in which r^2 was 0.935 was built and permitted demonstrations of high predictability for the activities of the 10 CPT analogues in the test set. The CoMFA contour maps illustrated that a more negatively charged group at positions 9, 10, and 11 of CPT would increase activity, but excessively increasing the bulky group at position 10 would be adverse to the activity; substituents that occupy position 7 with the bulky positively charged group will enhance the inhibition activity.

A tuned 3D QSAR model was also developed by Amat et al.[95] using molecular quantum similarity measures (MQSM) to predict the topo I inhibition for 12 CPT analogues. The best regression model was obtained with $r^2 = 0.928$ and $q^2 = 0.866$. TQSAR analysis of CPT analogues using the topo I inhibition activity demonstrates that MQSM can be a useful tool to predict the activities of therapeutic agents.

3.1.2. QSAR for the Inhibition of Various Cancer Cells

Three series of CPT derivatives (**XVI–XVIII**) were used by Li and co-workers[96] to develop three QSAR models **5–7**, which represent the correlations between biological activities of these compounds and their topological molecular descriptors. They used the following topological descriptors: Ax = topological molecular descriptors Ax;[97] ${}^m\chi_v$ = Kier and Hall valence connectivity indices[98–101] ${}^mK$ = Kier shape indices[98–101] ${}^m\chi$ = Randic indices of different orders;[102] ^{k}ACIC = average complementary information content index;[103] ^{k}CIC = complementary information content index;[103] ^{k}SIC = structure information content index;[103] RPCG = relative positive charge;[104] and P = polarity.[105]

Cytotoxic activities (pIC_{50}) of 7-X-camptothecins (**XVI**) on H460 human NSCLS cell line[96]



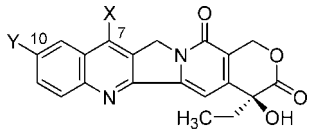
XVI

$$pIC_{50} = 3.045\,{}^1\text{ACIC} - 8.347\text{RPCG} - 0.286\,{}^2K + 7.514 \quad (5)$$

where $n = 23$, $r^2 = 0.841$, $s = 0.313$, $Q = 2.930$, and $F_{3,19} = 33.597$.

Based on this QSAR model, it had been predicted that the lower value of RPCG and the larger value of ^{1}ACIC of the X-substituents are preferred to lead to higher cytotoxic activity. This is to say that more lipophilicity leads to more potent anticancer activity.

Inhibition of HL-60 (human promyelocytic leukemic) cells by 7-X-10-Y-CPTs (**XVII**)[96]



XVII

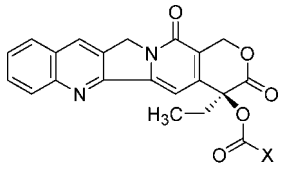
$$pIC_{50} = -0.756X_1 - 1.493X_2 + 0.669X_3 - 1.981\times10^{-2}X_4 + 8.452 \quad (6)$$

where $n = 15$, $r^2 = 0.907$, $s = 0.381$, $Q = 2.500$, and $F_{4,10} = 24.352$.

pIC_{50} is the biological activities (inhibition effects on the growth of HL-60) of CPTs (**XVII**); X_1 is ^{2}ACIC of X; X_2 is ^3K_v of X; X_3 is $^2\chi$ of X; and X_4 is $Ax3$ of Y.

According to the above QSAR model, authors[96] have predicted that the lipophilicity of the substituents at C-7 position of the ring-B play an important role to increase biological activity of CPT derivatives (**XVII**). Because X_1 represents a comparative study of lipophilicity versus topological molecular descriptors for the substituents at position 7. It's negative coefficient (−0.756) suggests that less lipophilic (hydrophobic) substituents will be needed at position 7 to increase the biological activities of the CPT derivatives (**XVII**). The steric factor of the substitutions at C-10 position of the ring-A also affects the activity. In most cases, more steric substituents at position C-10 lead to less biological activity (pIC_{50}).

Cytotoxic activities (IC_{50}) of 20-S-camptothecin alkanoic esters (**XVIII**) on BRE-MCF-7 cells (a breast cancer cell line)[96]



XVIII

$$IC_{50} = 0.215^1CIC + 36.658P'' - 1.564^0SIC - 8.354\times10^{-3} \quad (7)$$

where $n = 10$, $r^2 = 0.907$, $s = 0.542$, $Q = 1.757$, and $F_{3,6} = 19.544$.

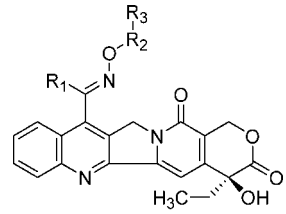
On the basis of the above QSAR equation, it has been suggested that the polarity plays an important role in increasing the biological activities of CPT esters (**XVIII**).

A QSAR study was performed by Dallavalle et al.[106] to predict cytotoxic activities (IC_{50}) of CPT derivatives (**XIX**) against H460 human NSCLC cell line, employing descriptors related to the lipophilicity. A multiple linear regression analysis afforded the following QSAR model **8**.

$$Y = -0.039(X1) - 0.020(X2) + 0.332(X3) + 6.939 \quad (8)$$

where $n = 21$, $r^2 = 0.820$, $s = 0.441$, $Q = 2.053$, and $F_{3,17} = 26.48$.

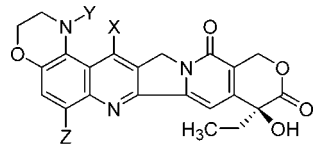
In the above equation, Y represents pIC_{50}, where IC_{50} is the cytotoxic activities of CPT derivatives (**XIX**) in molar



XIX

concentration. X1 is the molecular volume (Å^3) of R_1, X2 is the molecular volume (Å^3) of R_2, and X3 is Clog P or log D (in case of ionizable groups). According to the above equation, it has been suggested that the most important determinant for the in vitro cytotoxicity of the CPT derivatives (**XIX**) is the lipophilicity.

From the cytotoxic data of CPT derivatives (**XX**) against SKOV-3 human ovarian cancer cells,[107] QSAR eq 9 was developed.[108]



XX

$$\log 1/C = 6.90CMR - 0.33CMR^2 - 28.31 \quad (9)$$

where $n = 10$, $r^2 = 0.921$, $s = 0.278$, $q^2 = 0.723$, $Q = 3.452$, and $F_{2,7} = 40.804$ and optimum CMR = 10.50(9.65−10.90).

In this equation, C is the molar concentration of the CPT derivatives (**XX**) that causes 50% cell death (IC_{50}). QSAR **9** is a parabolic correlation in terms of CMR (calculated molar refractivity of the whole molecule), which suggests that the cytotoxic activities of CPT derivatives (**XX**) against SKOV-3 cells first increases with an increase in molar refractivity up to an optimum CMR value of 10.50 and then decreases. Clog P cannot replace CMR. Substituting log P for CMR in eq 9 gave a very poor fit (Clog P vs CMR; $r = 0.036$), indicating interaction in nonhydrophobic space.

A hierarchical clustering analysis was conducted by Fan et al.[109] to define antitumor activity patterns in a data set of 167 tested CPTs obtained from the NCI drug database. QSAR studies for 58 CPTs were carried out using the mean 50% growth inhibitory concentrations (GI_{50}) for 60 cell lines as the dependent variables. Different statistical methods, including stepwise linear regression, principal component regression (PCR), partial least-squares regression (PLS), and fully cross-validated genetic function approximation (GFA) were applied to construct quantitative structure−activity relationship models. The best model was obtained from the GFA method in terms of correlation coefficients and cross-validation analysis. The cross-validated r^2 for the final GFA model was 0.783, indicating a predictive QSAR model. The most important descriptors for the QSAR models were partial atomic charges at the 11- and 12-positions of the A-ring and three interatomic distances that define the relative spatial dispositions of three significant atoms (the hydroxyl hydrogen of the E-ring, the lactone carbonyl oxygen of the E-ring, and the carbonyl oxygen of the D-ring).

Comparative molecular field analysis (CoMFA), comparative molecular similarity index analysis (CoMSIA), and docking studies were used by Yoon et al.[110] to predict the biological activity of a 4-benzylpiperazine derivative of CPT-11 [7-ethyl-10-[4-(1-benzyl)-1-piperazino]carbonyloxycamp-

tothecin (BP-CPT)] in U373MG glioma cell lines transfected with plasmids encoding rabbit liver carboxylesterases (rCE) or human intestinal carboxylesterases (hiCE). BP-CPT has been reported to be activated more efficiently than CPT-11 by a rat serum esterase activity; however, 3D QSAR studies (CoMFA and CoMSIA) predicted that rCE would activate BP-CPT less efficiently than CPT-11. The models of CoMFA and CoMSIA were based on the relative activity values of 10 compounds (CPT-11 plus nine analogues) for the training set and 4 compounds (CPT-11 derivatives including BP-CPT) for the test set. The SN-38 portions of all the 14 molecules were superimposed. CoMFA analysis was based on steric plus electrostatic properties, PLS analysis gave a cross-validated q^2 correlation coefficient of 0.528 and conventional noncross-validated correlation coefficient of $r^2 = 1$. The steric parameter alone as the primary descriptor gave a higher cross-validated q^2 (0.583) than that observed when both steric and electrostatic properties were combined. This observation suggests that the steric descriptor predominates in this structure−activity relationship model. As similar to CoMFA analysis, a primary descriptor of steric plus electrostatic parameters gave a lower q^2 value in CoMSIA analysis than that with steric properties alone. PLS analysis of the CoMSIA values for compounds in the training set resulted in q^2 value of 0.440 and r^2 value of 0.999 for both parameters and a q^2 value of 0.656 and r^2 value of 0.997 for the steric descriptor alone. The hydrophobic descriptor also produced an excellent q^2 value of 0.678 and r^2 value of 0.998. Thus, CoMFA and CoMSIA results suggest together that steric properties may most accurately predict the biological activity of this class of prodrugs, with the hydrophobicity of the compound also a possible contributing factor. Molecular docking studies were performed to identify potential three-dimensional intermolecular interactions that determine or contribute to prodrug binding by the target enzyme rCE. Interestingly, the SN-38 moiety common to both CPT-11 and BP-CPT docked into the active site in a similar orientation. However, the carbonyl carbon of BP-CPT was closer by 2.1 Å to the Ser-221 hydroxyl group than was the carbonyl carbon of CPT-11. It was unknown whether this closer proximity of the BP-CPT to Ser-221 would be more likely to facilitate hydrolysis of the prodrug or to stabilize the enzyme/prodrug complex, thereby decreasing the efficiency of hydrolysis of the prodrug.

Irinotecan (CPT-11) is a widely used potent antitumor drug. However, overexpression of ABCG2 (BCRP/MXR/ABCP) confers cancer cells resistance to SN-38 (the active metabolite of CPT-11). A molecular modeling to circumvent cancer drug resistance associated with ABCG2 was conducted by Aida-Hyugaji et al.[111] using fourteen new SN-38 analogues, which were evaluated by molecular orbital (MO) calculations and neural network (NN) QSAR technique. Finally, it was suggested that the solvation energy (ΔG) and log P (hydrophobicity and hydrophilicity) values of compounds are critically related to determine substrates of ABCG2 and may be good indices for drug resistance.

3.2. Evaluation of New QSAR Models

3.2.1. Materials and Methods

All of the data has been collected from the literature (see individual QSAR for respective references). EC_{50} or IC_{50} is the concentration of CPT derivatives that cause 50% inhibition of topo I cleavable complex formation. Similarly, GI_{50} or IC_{50} is the concentration of CPT analogues required to produce 50% growth inhibition or 50% death of cancer cells. EC_{50} also represents the concentration of CPTs that caused 50% cytotoxicity to protozoa. In the literature, the values of EC_{50}, IC_{50}, and GI_{50} are given either in μM or nM concentration. For the comparison point of view, all the values of EC_{50}, IC_{50}, and GI_{50} have been converted into molar concentrations and explained after each equation. In QSAR studies, we often like the more effective compounds to have a higher "activity" and not a lower. This is the main reason; it is very common in QSAR development to transform the concentration of a desired effect "C" to an activity (negative logarithm of the concentration) by the following equation: $A = -\log C = \log 1/C$.

Thus, the above equation was used to transform EC_{50}, IC_{50}, and GI_{50} into their negative logarithm; that is, log $1/EC_{50}$, log $1/IC_{50}$, and log $1/GI_{50}$, respectively. log $1/EC_{50}$, log $1/IC_{50}$, and log $1/GI_{50}$ are the subsequent dependent variables, which define the biological parameter for QSAR development. Physicochemical descriptors are autoloaded, and multiregression analyses (MRA) are used to derive the QSAR by utilizing the C-QSAR program.[112] Selection of descriptors is made on the basis of permutation and correlation matrices among the descriptors in order to avoid collinearity problems. Details about the C-QSAR program, the search engine, the choice of parameters, and their use in the development of QSAR models have been discussed previously.[113,114]

The parameters used in this review have been discussed previously in detail along with their application.[84] Briefly, Clog P is the calculated partition coefficient of a compound in n-octanol/water and is a measure of its hydrophobicity, while π is the hydrophobic parameter for substituents only. σ, σ^+, and σ^- are Hammett electronic parameters that apply to substituent effects on aromatic systems. $B1$, $B5$, and L are Verloop's sterimol parameters for substituents.[115] $B1$ is a measure of the minimum width of a substituent, $B5$ defines the maximum width of the substituents, and L is the substituent length. CMR is the calculated molar refractivity for the whole molecule. Molar refractivity (MR) is calculated from the Lorentz−Lorenz equation and is described as follows: $[(n^2 - 1)/(n^2 + 2)](\mathrm{MW}/\delta)$, where n is the refractive index, MW is the molecular weight, and δ is the density of the substance. MR is dependent on volume and polarizability. It can be used for a substituent or for the whole molecule.

A new polarizability parameter, NVE, was developed, which is shown to be effective at delineating various chemico−biological interactions.[116−119] NVE represents the total number of valence electrons and is calculated by simply summing up the valence electrons in a molecule, that is, H = 1, C = 4, Si = 4, N = 5, P = 5, O = 6, S = 6, and halogens = 7. It may also be represented as: $\mathrm{NVE} = n_\sigma + n_\pi + n_n$, where n_σ is the number of electrons in σ-orbital, n_π is the number of electrons in π-orbitals, and n_n is the number of lone pair electrons. MgVol is the molar volume for the whole molecule.[120] The indicator variable I is assigned the value of 1 or 0 for special features with special effects that cannot be parametrized and has been explained wherever used.

In QSAR equations, n is the number of data points, r is the correlation coefficient between observed values of the dependent and the values calculated from the equation, r^2 is the square of the correlation coefficient and represents the goodness of fit, q^2 is the cross-validated r^2 (a measure of the quality of the QSAR model), and s is the standard

Table 1. Biological (EC_{50} or IC_{50}; mol L^{-1})[127] and Physicochemical Parameters Used To Derive QSAR Eqs 10 and 22

			log 1/EC_{50} (eq 10)			log 1/IC_{50} (eq 22)				
No.	X	Y	obsd.	Pred.	Δ	Obsd.	Pred.	Δ	π_X	π_Y
1	CH_3	H	5.97	5.70	0.27	6.93	7.03	−0.10	0.56	0.00
2	CH_2CH_3	H	5.35	5.56	−0.21	7.11	7.36	−0.25	1.02	0.00
3[a]	$(CH_2)_2CH_3$	H	5.14	5.40	−0.26	6.73	7.73	−1.00	1.55	0.00
4[a]	$(CH_2)_3CH_3$	H	5.14	5.23	−0.09	6.41	8.14	−1.73	2.13	0.00
5	H	OH	6.46	6.78	−0.32	ND[b]	7.60	ND	0.00	−0.67
6	CH_3	OH	6.89	6.62	0.27	8.28	8.00	0.28	0.56	−0.67
7	CH_2CH_3	OH	6.49	6.48	0.01	8.43	8.32	0.11	1.02	−0.67
8	$(CH_2)_2CH_3$	OH	6.22	6.32	−0.10	8.45	8.70	−0.25	1.55	−0.67
9	$(CH_2)_3CH_3$	OH	6.28	6.15	0.13	8.96	9.11	−0.15	2.13	−0.67
10	H	OCH_3	6.14	5.89	0.25	6.51	6.66	−0.15	0.00	−0.02
11	CH_3	OCH_3	5.62	5.73	−0.11	7.18	7.06	0.12	0.56	−0.02
12	CH_2CH_3	OCH_3	5.55	5.59	−0.04	7.59	7.39	0.20	1.02	−0.02
13	$(CH_2)_2CH_3$	OCH_3	5.35	5.43	−0.08	7.66	7.76	−0.10	1.55	−0.02
14	$(CH_2)_3CH_3$	OCH_3	5.55	5.26	0.29	8.47	8.17	0.30	2.13	−0.02

[a] Not included in the derivation of QSAR **22** [b] ND = not determined.

deviation. The cross-validated r^2 (q^2) is obtained by using a leave-one-out (LOO) procedure as described by Cramer et al.[121] Q is the quality factor (quality ratio), where $Q = r/s$. Chance correlation due to the excessive number of parameters (which also increases the r and s values) can, therefore, be detected by the examination of the Q value. High values of Q indicate the high predictive power of the QSAR models and the lack of "over-fitting". F represents the Fischer statistics (Fischer ratio), $F = fr^2/[(1 - r^2)m]$, where f is the number of degrees of freedom [$f = n - (m + 1)$], n is the number of data points, and m is the number of variables. The F value is actually the ratio between explained and unexplained variance for a given number of degrees of freedom. Thus, it indicates a true relationship or the significance level for MLR models. The modeling was taken to be optimal when Q reached a maximum together with F, even if slightly nonoptimal F values have normally been accepted. A significant decrease in F with the introduction of one additional variable (with increasing Q and decreasing s) could mean that the new descriptor is not as significant as expected; that is, its introduction has endangered the statistical quality of the combination. However, the statistical quality could be improved by the introduction of a more convincing descriptor.[122−124]

Compounds were deemed to be outliers on the basis of their deviation between observed and calculated activities from the equation (obsd. − pred. > 2s).[125,126] Each regression equation includes 95% confidence limits for each term in parentheses. All new QSAR models reported here are derived by us and were not formulated by the original authors.

3.2.2. Results and Discussion

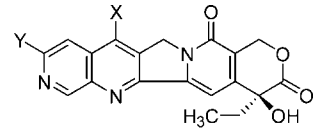
3.2.2.1. QSAR for the Inhibition of DNA Topoisomerase I. Vladu et al.[127] synthesized a series of 7-X-10-Y-CPTs (**XVII**) and evaluated their ability to trap human DNA topoisomerase I in cleavable complexes. The activity data of CPTs (**XVII**) was given in EC_{50} (μM), which required producing cleavable complexes with purified topo I in 50% of the plasmid DNA. These values were converted into log 1/EC_{50} of the analogues (**XVII**) in molar concentration and are given in Table 1. We used these data to develop eq 10 (Table 1).

$$\log 1/EC_{50} = -0.30(\pm 0.19)\pi_X - 1.37(\pm 0.42)\pi_Y + 5.87(\pm 0.28) \quad (10)$$

where $n = 14$, $r^2 = 0.859$, $s = 0.227$, $q^2 = 0.750$, $Q = 4.084$, and $F_{2,11} = 33.507$.

The hydrophobic parameters π_X and π_Y are for the substituents at positions 7 and 10, respectively. Its negative coefficients suggest that there is a need for less hydrophobic substituents at both positions.

A series of 7-X-10-Y-11-aza-CPTs (**XXI**) were synthesized and evaluated for topoisomerase I-targeting activity as well as for cytotoxicities against different cell lines by Uehling et al.[128] The topo I-mediated DNA cleavage data of these compounds (**XXI**) was given in IC_{50} (nM), where IC_{50} is the minimum drug concentration that produced 50% fragmentation of DNA in the presence of excess calf thymus topoisomerases. These data of IC_{50} were converted into log 1/IC_{50} in molar concentration and used in the development of QSAR eq 11 (Table 2).



XXI

$$\log 1/IC_{50} = -0.40(\pm 0.17)MR_Y + 1.08(\pm 0.56)I + 0.50(\pm 0.39)I_1 + 6.74(\pm 0.33) \quad (11)$$

where $n = 17$, $r^2 = 0.784$, $s = 0.334$, $q^2 = 0.652$, $Q = 2.653$, and $F_{3,13} = 15.728$; π_Y versus MR_Y; $r = 0.010$.

MR_Y is the calculated molar refractivity of Y-substituents, whereas I is an indicator variable taking the value of 1 and 0 for the presence and absence of an amidine group in the Y-substituents, respectively. Similarly, I_1 is an indicator variable taking the value of 1 for X = CH_3 and 0 for X = H or C_2H_5, respectively. The negative sign of MR_Y brings out a steric effect for the Y-substituents. π_Y (calculated hydrophobicity of Y-substituents) cannot replace MR_Y. Substituting π_Y for MR_Y in eq 11 gave a very poor fit ($r^2 = 0.388$, $q^2 = -0.046$), indicating interaction in nonhydrophobic space. The indicator variables (I and I_1) with positive coefficients suggest that the presence of an amidine group in the Y-substituents and a methyl group in the X-substituents will enhance the activity.

In an effort to improve the water solubility of CPT, Rahier et al.[129] synthesized four 20-*O*-phosphate derivatives (**XXII**). These compounds are found freely water soluble, stable at physiological pH, and stabilize the human topo I-DNA covalent binary complex with the same sequence selectivity as CPT itself. This may be possible due to either elimination of intramolecular hydrogen bonding or increasing steric hindrance of the carbonyl group of E-ring, results in the

Table 2. Biological (IC_{50}; mol L^{-1}),[128] Physicochemical, and Structural Parameters Used To Derive QSAR Eq 11

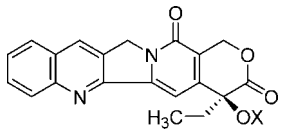
No.	X	Y	log 1/IC_{50} (eq 11) obsd.	pred.	Δ	MR_Y	I	I_1
1	H	H	6.42	6.74	−0.32	0.00	0	0
2	C_2H_5	H	6.47	6.74	−0.27	0.00	0	0
3	CH_3	Br	7.55	6.93	0.62	0.78	0	1
4	C_2H_5	Br	6.91	6.43	0.49	0.78	0	0
5	CH_3	CN	6.84	7.05	−0.21	0.48	0	1
6	C_2H_5	CN	6.57	6.55	0.02	0.48	0	0
7	CH_3	CH_2NH_2	6.63	6.91	−0.28	0.83	0	1
8	C_2H_5	CH_2NH_2	6.78	6.41	0.37	0.83	0	0
9	C_2H_5	$C(NH_2)NOH$	7.15	7.22	−0.08	1.48	1	0
10	C_2H_5	$C(NH_2)NH$	7.48	7.41	0.08	1.02	1	0
11	CH_3	$C{\equiv}CCH_2NH_2$	6.30	6.56	−0.26	1.71	0	1
12	C_2H_5	$C{\equiv}CCH_2NH_2$	5.97	6.06	−0.09	1.71	0	0
13	C_2H_5	$C{\equiv}CCH_2N(CH_3)_2$	5.88	5.69	0.19	2.64	0	0
14	CH_3	$C{\equiv}CCH_2(\text{-}NCH_2CH_2OCH_2CH_2\text{-})$	5.95	5.83	0.13	3.54	0	1
15	C_2H_5	$(CH_2)_3N(CH_3)_2$	5.83	5.66	0.16	2.69	0	0
16	C_2H_5	$COOC_2H_5$	6.02	6.11	−0.08	1.58	0	0
17	C_2H_5	$CONH(CH_2)_2N(CH_3)_2$	5.03	5.50	−0.47	3.09	0	0

Table 3. Biological (IC_{50}; mol L^{-1})[129] and Physicochemical Parameters Used To Derive QSAR Eq 12

No.	X	log 1/IC_{50} (eq 12) obsd.	pred.	Δ	π_X
1	H (CPT)	6.05	5.97	0.08	0.00
2	$P(=O)(OH)_2$	4.70	4.88	−0.18	−0.62
3	$P(=O)(OH)(OCH_3)$	4.48	4.47	0.01	−0.85
4	$P(=O)(OH)(CH_3)$	4.24	4.14	0.10	−1.04
5[a]	P(=O)(OH)(Ph)	4.43	5.51	−1.08	−0.26

[a] Not included in the derivation of QSAR **12**.

stabilization of the lactone ring.[17,71,72] Thus, it is a very good example for the comparison between α-OH (20-*S*) and α-OX (20-*S*) of CPTs in improving their water solubility and stabilizing topo I-DNA covalent binary complex. The topo I-dependent cytotoxicity data of these four compounds along with CPT (**I**) in *S. cerevisiae* was given in IC_{50} (μM), where IC_{50} is the inhibition of RS321Nph-topo I grown in minimal medium containing 3% raffinose or galactose for 2 days at 30 °C. These data of IC_{50} were converted into log 1/IC_{50} (molar concentration) and are given in Table 3. Eq 12 was derived from the data in Table 3.



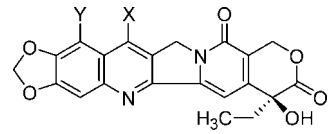
XXII

$$\log 1/IC_{50} = 1.77(\pm 0.86)\pi_X + 5.97(\pm 0.64) \quad (12)$$

where $n = 4$, $r^2 = 0.975$, $s = 0.156$, $q^2 = 0.718$, $Q = 6.327$, and $F_{1,2} = 78.000$; outlier: X = P(=O)(OH)Ph.

π_X is the calculated hydrophobicity of X-substituents, which is found to be the single most important parameter for this data set. This suggests that at all the parts where X-substituents have been entered, hydrophobic contacts have been made. The linear π_X model suggests that the molecules with highly hydrophobic X-substituents will be more active. Although this is a very small data set, it is the best model and explains 97.5% of the varience in log 1/IC_{50}. The outlier [X = P(=O)(OH)Ph] is much less active than expected by seven times the standard deviation. Possible reasons for its unusually low activity are not obvious although its bulk and/or geometry due to the presence of phenyl group (in place of OH group) may reduce the coplanarity with 20-*O*-phosphate group and minimize the topo I-dependent cytotoxicity.

Lu et al.[94] collected the inhibitory data of topo I from the literature for some series of CPT derivatives, which was determined using the cleavable complex assay. We used their data [pIC_{50} or log 1/IC_{50}, where IC_{50} was assessed by minimum concentration (mol/L) of the drug that inhibited the cleavable complex formation by 50%] for 7-X-9-Y-10,11-(methylenedioxy)camptothecins (**XXIII**) and developed eq 13 (Table 4).



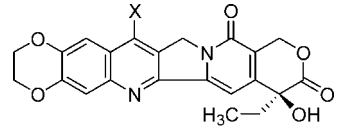
XXIII

$$\log 1/IC_{50} = -0.59(\pm 0.15)MR_X - 1.31(\pm 0.61)MR_Y - 0.71(\pm 0.47)I + 8.30(\pm 0.39) \quad (13)$$

where $n = 11$, $r^2 = 0.947$, $s = 0.117$, $q^2 = 0.788$, $Q = 8.316$, and $F_{3,7} = 41.692$; outlier: X = CH_2(−N[+]=CHCH=CHC(CH_2OH)=CH-), Y = H; π_X versus MR_X; $r = 0.627$; π_Y versus MR_Y; $r = 0.091$.

MR_X and MR_Y are the calculated molar refractivity of X- and Y-substituents, respectively. The indicator variable (I) takes the value of 1 for X = H (unsubstituted) and 0 for X = different groups (substituted), respectively. One compound in Table 4 was deemed to be an outlier on the basis of their deviation (>2s).

When the inhibition data of 7-X-10,11-(ethylenedioxy)-camptothecins (**XXIV**) against DNA topo I from the same collection of Lu et al. was used,[94] eq 14 was derived (Table 5).



XXIV

$$\log 1/IC_{50} = -0.62(\pm 0.16)MR_X - 0.44(\pm 0.28)I + 8.48(\pm 0.46) \quad (14)$$

Table 4. Biological (IC_{50}; mol L^{-1}),[94] Physicochemical, and Structural Parameters Used To Derive QSAR Eq 13

No.	X	Y	log 1/IC_{50} (eq 13) obsd.	pred.	Δ	MR_X	MR_Y	*I*
1	CH_2(-$NCH_2CH_2N(CH_3)CH_2CH_2$-)	H	6.38	6.34	0.04	3.34	0.00	0
2	CH_2Cl	H	7.82	7.74	0.08	0.96	0.00	0
3	H	H	7.57	7.59	−0.02	0.00	0.00	1
4	H	Cl	6.82	6.94	−0.12	0.00	0.49	1
5	H	NH_2	7.21	7.10	0.11	0.00	0.37	1
6	H	NO_2	6.82	6.79	0.03	0.00	0.61	1
7	CH_2(-N[+]=CHCH=CHCH=CH-)	H	6.68	6.63	0.05	2.85	0.00	0
8	CH_2(-NCH=CHN=CH-)	H	6.80	7.01	−0.21	2.19	0.00	0
9	CH_2(-NCH=CHN[+](CH_3)=CH-)	H	6.66	6.69	−0.03	2.75	0.00	0
10	CH_2(-N[+]=CHCH=NCH=CH-)	H	6.85	6.75	0.10	2.64	0.00	0
11[a]	CH_2(-N[+]=CHCH=CHC(CH_2OH)=CH-)	H	7.00	6.26	0.74	3.47	0.00	0
12	CH_2(-N[+]=CHCH=CHCH=N-)	H	6.72	6.75	−0.03	2.64	0.00	0

[a] Not included in the derivation of QSAR **13**.

Table 5. Biological (IC_{50}; mol L^{-1}),[94] Physicochemical, and Structural Parameters Used To Derive QSAR Eq 14

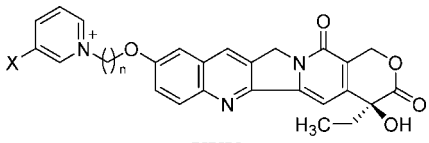
No.	X	log 1/ IC_{50} (eq 14) obsd.	pred.	Δ	MR_X	*I*
1	CH_2(−$NCH_2CH_2N(CH_3)CH_2CH_2$−)	6.52	6.39	0.13	3.34	0
2	CH_2Cl	7.96	7.88	0.08	0.96	0
3	CH_2(−N[+]=CHCH=CHCH=CH−)	6.72	6.69	0.03	2.85	0
4	CH_2(−NCH=CHN=CH−)	6.70	6.66	0.04	2.19	1
5	CH_2(−N[+]=CHCH=NCH=CH−)	6.74	6.83	−0.09	2.64	0
6	CH_2(−N[+]=CHCH=CHCH=N−)	6.60	6.83	−0.23	2.64	0
7	CH_2(−N[+]=CHCH=CHC(CH_2OH)=CH−)	6.22	6.31	−0.09	3.47	0
8	CH_2(−NCH=CHN[+](CH_3)=CH−)	6.28	6.32	−0.04	2.75	1
9	CH_2(−N[+]=CHCH=C(CH_2OH)CH=CH−)	6.47	6.31	0.16	3.47	0

where $n = 9$, $r^2 = 0.943$, $s = 0.141$, $q^2 = 0.806$, $Q = 6.887$, and $F_{2,6} = 49.632$; π_X versus MR_X; $r = 0.536$.

IC_{50} is the minimum concentration (mol/L) of the drug that inhibited the cleavable complex formation by 50%, MR_X is the calculated molar refractivity of X-substituents, whereas *I* is an indicator variable taking the value of 1 and 0 for the presence and absence of an imidazole moiety in the X-substituents. The negative sign of MR_X brings out a steric effect for the X-substituents. π_X (calculated hydrophobicity of X-substituents) cannot replace MR_X. Substituting π_X for MR_X in eq 14 gave a very poor fit ($r^2 = 0.421$, $q^2 = -0.479$), suggesting interaction in nonhydrophobic space. The negative coefficient of the indicator variable indicates that the absence of imidazole moiety in the X-substituents would be favorable for higher activity.

3.2.2.2. QSAR for the Inhibition of Various Cancer Cells. *3.2.2.2.1. Modification of Ring A.* In an effort to improve the water solubility of CPT, 16 water soluble 10-substituted quaternary ammonium salts of CPT (**XXV**) were synthesized by Zu et al.[130] and also evaluated for their antitumor activities on different three cancer cells in vitro. All of these salts possess lower cytotoxicities than CPT in comparison. The cytotoxic data was given in IC_{50} (μg/mL), which was converted into IC_{50} (mol/L). From the cytotoxicity data of these compounds against two human cancer cell lines (KB and HCT-8), we derived eqs 15 and 16.

Cytotoxic activities (IC_{50}; mol/L) of CPTs (**XXV**) to KB (human epidermoid carcinoma of the nasopharynx) cells. Data obtained from Zu et al.[130] are shown in Table 6.



XXV

$$\log 1/IC_{50} = -1.48(\pm 0.47)MR_X + 0.75(\pm 0.33)I + 9.50(\pm 0.32) \quad (15)$$

where $n = 12$, $r^2 = 0.882$, $s = 0.247$, $q^2 = 0.739$, $Q = 3.802$, and $F_{2,9} = 33.636$; outliers: X = CHO, $n = 2$; X = COOH, $n = 2$; X = F, $n = 3$; π_X versus MR_X; $r = 0.435$

A negative coefficient of MR_X (calculated molar refractivity of X-substituents) brings out a steric effect for the X-substituents. It has been observed that π_X (calculated hydrophobicity of X-substituents) cannot replace MR_X. Substituting π_X for MR_X in eq 15 gave a very poor fit ($r^2 = 0.342$, $q^2 = -0.075$), suggesting interaction in nonhydrophobic space. Indicator variable $I = 1$ and 0 for $n = 3$ and 2, respectively. The presence of $n = 3$ increases the activity as evidenced by the positive coefficient of the indicator variable (*I*). Three compounds were deemed to be outliers on the basis of their deviation (>2*s*).

Cytotoxic activities (IC_{50}; mol/L) of CPTs (**XXV**) to HCT-8 (human colon cancer) cells. Data obtained from Zu et al.[130] are shown in Table 6.

$$\log 1/IC_{50} = -1.41(\pm 0.50)MR_X + 0.88(\pm 0.35)I + 9.35(\pm 0.34) \quad (16)$$

where $n = 12$, $r^2 = 0.874$, $s = 0.265$, $q^2 = 0.697$, $Q = 3.528$, and $F_{2,9} = 31.214$; outliers: X = CHO, $n = 2$; X = COOH, $n = 2$; X = F, $n = 3$; π_X versus MR_X; $r = 0.435$.

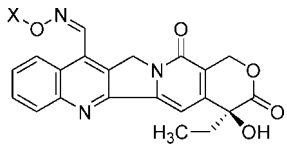
QSAR eqs 15 and 16 are very similar to each other, which suggest that CPTs (**XXV**) may target an enzyme of similar kind in human KB and HCT-8 cancer cells.

3.2.2.2.2. Modification of Ring B. Cytotoxic activities (IC_{50}; mol/L) of 7-CH=NO-X-CPTs (**XXVI**) to H460 human NSCLC cells. Data obtained from Dallavalle et al.[106] are shown in Table 7.

Table 6. Biological (IC_{50}; mol L^{-1}),[130] Physicochemical, and Structural Parameters Used To Derive QSAR Eqs 15 and 16

			log 1/ IC_{50} (eq 15)			log 1/ IC_{50} (eq 16)				
No.	X	n	obsd.	pred.	Δ	obsd.	pred.	Δ	MR_X	I
1	H	2	6.14	6.35	−0.21	6.01	6.21	−0.20	0.10	0
2	$COCH_3$	2	5.20	4.85	0.35	5.23	4.78	0.45	1.12	0
3[a]	CHO	2	6.16	5.49	0.67	6.15	5.39	0.76	0.69	0
4	CH_2OH	2	5.18	5.44	−0.26	5.13	5.34	−0.21	0.72	0
5	OH	2	6.18	6.08	0.10	6.06	5.95	0.10	0.29	0
6	CH_3	2	5.89	5.67	0.22	5.53	5.56	−0.03	0.56	0
7	F	2	6.17	6.37	−0.20	6.11	6.22	−0.12	0.09	0
8[a]	COOH	2	6.21	5.48	0.73	6.27	5.38	0.89	0.69	0
9	H	3	7.18	7.10	0.08	7.14	7.09	0.05	0.10	1
10	$COCH_3$	3	5.22	5.60	−0.38	5.22	5.66	−0.44	1.12	1
11	CHO	3	6.27	6.24	0.03	6.48	6.27	0.21	0.69	1
12	CH_2OH	3	6.21	6.19	0.02	6.18	6.22	−0.05	0.72	1
13	OH	3	7.10	6.83	0.27	7.11	6.83	0.28	0.29	1
14[a]	CH_3	3	ND[b]	6.42	ND	ND	6.44	ND	0.56	1
15[a]	F	3	5.17	7.12	−1.95	5.40	7.10	−1.71	0.09	1
16	COOH	3	6.21	6.23	−0.02	6.20	6.26	−0.06	0.69	1

[a] Not included in the derivation of QSAR eqs 15 and 16. [b] ND = not determined.



XXVI

$$\log 1/IC_{50} = 0.46(\pm 0.16)\pi_X - 0.39(\pm 0.11)MR_X + 0.86(\pm 0.38)I + 1.04(\pm 0.35)I_1 + 7.51(\pm 0.32) \quad (17)$$

where $n = 24$, $r^2 = 0.829$, $s = 0.321$, $q^2 = 0.711$, $Q = 2.835$, and $F_{4,19} = 23.028$; outliers: X = COC_6H_5; CH_2COOH; $C(CH_3)_2COOH$; $CH_2C_6H_4(4\text{-}C_6H_3)$.

In this equation, π_X and MR_X are the calculated hydrophobic and molar refractivity descriptors of the X-substituents, respectively. Positive π_X suggests that cytotoxic activities of these molecules will increase with the increase of hydrophobicity of X-substituents. On the contrary, the increase in the molar refractivity of the X-substituents (MR_X) decreases the cytotoxic activities of these compounds (negative coefficient). The indicator variable (I) takes the value of 1 and 0 for the presence and absence of the heterocycle in the X-substituents, respectively. Similarly, the indicator variable (I_1) takes the value of 1 for X = (un)substituted benzyl group and 0 for X = other than (un)substituted benzyl group, respectively. The positive coefficient of the indicator variables (I and I_1) suggests that the presence of either a heterocyclic or (un)substituted benzyl group in the X-substituents would increase the activity. Compounds 2, 11, 12, and 24 (Table 7) were omitted on the basis of their deviation (>2s).

The outlier (X = COC_6H_5) is much less active than expected by three times the standard deviation. Possible reasons for its unusually low activity are not obvious although its bulk and/or geometry due to the presence of phenyl group may reduce the coplanarity with X-group and minimize the cytotoxicity. Two other outliers [X = CH_2COOH and $C(CH_3)_2COOH$] might be due to the presence of ionizable carboxylic group, because π represents the calculated hydrophobicity of neutral X-substituents. The derivative [X = $CH_2C_6H_4(4\text{-}C_6H_5)$] was also considered to be an outlier due to being much less active than expected by 3.5 times the standard deviation. Possible reasons for its unusually low activity are also not obvious although its bulk and/or geometry due to the presence of biphenyl group may reduce the coplanarity with X- group and minimize the cytotoxicity.

Table 7. Biological (IC_{50}; mol L^{-1}),[106] Physicochemical, and Structural Parameters Used To Derive QSAR Eq 17

		log 1/ IC_{50} (eq 17)						
No.	X	obsd.	pred.	Δ	π_X	MR_X	I	I_1
1	H	7.49	7.51	−0.02	0.00	0.00	0	0
2[a]	COC_6H_5	5.75	6.71	−0.96	0.79	3.01	0	0
3	CH_3	7.40	7.26	0.14	−0.16	0.46	0	0
4	$CH_2CH{=}CH_2$	7.22	7.27	−0.05	0.62	1.37	0	0
5	$CH_2CH(CH_2)O$	7.70	7.55	0.15	−0.63	1.37	1	0
6	CMe_3	7.82	7.29	0.53	1.08	1.86	0	0
7	$C(CH_3)_2CH_2OH$	6.47	6.57	−0.10	−0.37	2.01	0	0
8	$C(CH_3)_2COOCMe_3$	6.85	6.58	0.27	1.26	3.90	0	0
9	$CH_2CH_2NH_2$	6.31	6.55	−0.24	−1.00	1.30	0	0
10	$CH_2CH_2NMe_2$	6.52	6.60	−0.08	−0.11	2.22	0	0
11[a]	CH_2COOH	5.63	6.65	−1.02	−0.94	1.12	0	0
12[a]	$C(CH_3)_2COOH$	4.32	6.57	−2.25	−0.33	2.04	0	0
13	$CH_2CONH(CH_2)_3NH_2$	5.28	5.66	−0.38	−1.43	3.09	0	0
14	$CH_2CONH(CH_2)_3NHCOOCMe_3$	5.77	5.52	0.25	0.36	5.60	0	0
15	CH_2CH_2−Morpholinyl	7.10	7.13	−0.03	−0.08	3.13	1	0
16	CH_2CH_2-3-(N-Me)Piperidinyl	7.00	7.27	−0.27	0.88	3.90	1	0
17	CH_2CH_2-1-Uracylyl	6.30	6.44	−0.14	−1.31	3.44	1	0
18	C_6H_5	6.80	7.10	−0.30	1.21	2.51	0	0
19	$CH_2C_6H_5$	7.52	7.79	−0.27	0.84	2.98	0	1
20	$CH_2C_6H_4(4\text{-}CH_3)$	7.70	7.84	−0.14	1.34	3.44	0	1
21	$CH_2C_6H_4(4\text{-}NO_2)$	7.77	7.44	0.33	0.58	3.59	0	1
22	$CH_2C_6F_5$	7.55	7.99	−0.44	1.35	3.05	0	1
23	$CH_2C_6H_4(4\text{-}NH_2)$	7.60	7.09	0.51	−0.39	3.34	0	1
24[a]	$CH_2C_6H_4(4\text{-}C_6H_5)$	6.55	7.69	−1.14	2.73	5.49	0	1
25	$C(C_6H_5)_3$	5.84	6.08	−0.24	3.63	8.00	0	0
26	CH_2-9-Anthracenyl	6.72	6.52	0.20	3.19	6.35	0	0
27	CH_2-4-Pyridyl	7.52	7.00	0.52	−0.66	2.76	1	0
28	CH_2-2-Imidazolyl	6.68	6.91	−0.23	−1.33	2.19	1	0

[a] Not included in the derivation of QSAR eq 17.

Table 8. Biological (IC_{50}; mol L^{-1})[131] and Physicochemical Parameters Used To Derive QSAR Eq 18

No.	X	log 1/ IC_{50} (eq 18) obsd.	pred.	Δ	Clog *P*
1	$CH{=}NC_6H_5$	6.89	6.66	0.23	1.84
2	$CH{=}NC_6H_{11}$	6.43	6.38	0.05	2.67
3[a]	$CH{=}N(CH_2)_2N(CH_3)_2$	6.92	5.94	0.98	−0.04
4	$CH{=}N(CH_2)_4OH$	6.44	6.16	0.28	0.23
5	$CH{=}NC_6H_4(4\text{-}NO_2)$	6.55	6.64	−0.09	1.95
6	$CH{=}NCH_2C_6H_5$	6.55	6.51	0.04	2.40
7	CH_2NH_2	5.63	5.84	−0.21	−0.15
8	$CH_2NHC_6H_5$	6.51	6.66	−0.15	1.84
9	$CH_2NHC_6H_{11}$	6.46	6.55	−0.09	2.30
10	$CH_2NH(CH_2)_2N(CH_3)_2$	6.35	6.42	−0.07	0.64
11	$CH_2NHC({=}NH)NH_2$	4.90	4.87	0.03	−0.98

[a] Not included in the derivation of QSAR eq 18.

Table 9. Biological (IC_{50}; mol L^{-1})[60] and Physicochemical Parameters Used To Derive QSAR Eq 19

No.	X	log 1/ IC_{50} (eq 19) obsd.	pred.	Δ	MR_X
1	H [20(*S*)-CPT]	7.40	7.39	0.01	0.00
2	$NHCH_2CH_2OH$	4.28	4.56	−0.28	1.54
3	$NHCH_2CH_2CH_2CH_2OH$	4.46	4.40	0.06	2.47
4	$NHCH(CH_2OH)_2$	4.30	4.41	−0.11	2.16
5	$NHCH_2CH_2N(CH_3)_2$	4.52	4.41	0.11	2.68
6	$NHCH_2CH_2N(CH_2)_4$	4.32	4.45	−0.13	3.22
7	$NHCH_2CH_2N(CH_2)_5$	4.63	4.49	0.14	3.68
8	$NHCH_2CH_2N(CH_2CH_2NHCH_2CH_2)$	4.45	4.48	−0.03	3.59
9	$NHCH_2C_6H_5$	4.33	4.45	−0.12	3.24
10	$NHCH_3$	5.33	5.15	0.18	0.92
11	NHOH	5.62	5.73	−0.11	0.61
12	$NHCH_2CH_2Cl$	4.70	4.44	0.26	1.88

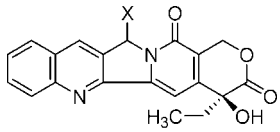
Cytotoxicity (IC_{50}; mol/L) of 7-X-CPTs (**XVI**) to H460 human NSCLC cells. Data obtained from Dallavalle et al.[131] are shown in Table 8.

$$\log 1/IC_{50} = 0.86(\pm 0.26)\text{Clog } P - 0.27(\pm 0.12)\text{Clog } P^2 + 5.97(\pm 0.22) \quad (18)$$

where $n = 10$, $r^2 = 0.924$, $s = 0.180$, $q^2 = 0.864$, $Q = 5.339$, and $F_{2,7} = 42.553$; optimum Clog $P = 1.62(1.33-2.30)$; outlier: $X = CH{=}N(CH_2)_2N(CH_3)_2$.

This is a parabolic correlation in terms of Clog *P*, which suggests that the cytotoxic activities of CPT derivatives (**XVI**) against H460 cells first increases with an increase in their hydrophobicity up to an optimum Clog *P* of 1.62 and then decreases. One derivative [$X = CH{=}N(CH_2)_2N(CH_3)_2$] was deemed to be an outlier due to being much more active than expected by five times to the standard deviation.

3.2.2.2.3. Modification of Ring C. Average cytotoxicity (IC_{50}; mol/L) of 5-X-CPTs (**XXVII**) to 56 human tumor cell lines. Data obtained from Subrahmanyam et al.[60] are shown in Table 9.



XXVII

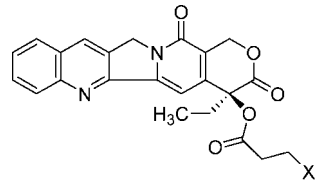
$$\log 1/IC_{50} = -3.65(\pm 0.70)MR_X + 3.74(\pm 0.85)\log(\beta \times 10^{MR_X} + 1) + 7.15(\pm 0.35) \quad (19)$$

where $n = 12$, $r^2 = 0.970$, $s = 0.182$, $q^2 = 0.952$, $Q = 5.412$, and $F_{3,8} = 86.223$; inversion point for $MR_X = 2.39$; $\log \beta = -0.81$; π_X versus MR_X; $r = 0.551$.

This is an inverted bilinear relation in terms of MR_X (calculated molar refractivity of X-substituents), which suggests that activity of these compounds first decreases linearly as the molar refractivity of the X-substituents increases, and that after a certain point (inversion point; $MR_X = 2.39$) activity begins to increase gradually. This may correspond to an allosteric reaction.[91,113,132,133]

3.2.2.2.4. Modification of Ring E. A series of 20-*O*-linked nitrogen based CPT esters (**XXVIII**) was synthesized and evaluated for their cytotoxicities against different human cancer cell lines by Wang et al.[73] From the cytotoxicity data of these compounds against two human cancer cell lines (A2780 and Bel7402), we derived eqs 20 and 21.

Cytotoxicity (IC_{50}; mol/L) of CPTs (**XXVIII**) to A2780 (human ovarian cancer) cells. Data obtained from Wang et al.[73] are shown in Table 10.



XXVIII

$$\log 1/IC_{50} = 2.05(\pm 0.96)\pi_X - 0.42(\pm 0.18)\pi_X^2 + 4.70(\pm 1.26) \quad (20)$$

where $n = 9$, $r^2 = 0.908$, $s = 0.076$, $q^2 = 0.602$, $Q = 12.539$, and $F_{2,6} = 29.609$; optimum $\pi_X = 2.46(2.22-2.59)$. Two compounds 2 and 5 (Table 10) were omitted on the basis of their deviation ($>2s$).

Cytotoxicity (IC_{50}; mol/L) of CPTs (**XXVIII**) to Bel7402 (human liver cancer) cells. Data obtained from Wang et al.[73] are shown in Table 10.

$$\log 1/IC_{50} = 12.81(\pm 3.37)MR_X - 1.25(\pm 0.33)MR_X^2 - 24.84(\pm 8.53) \quad (21)$$

where $n = 12$, $r^2 = 0.892$, $s = 0.258$, $q^2 = 0.841$, $Q = 3.659$, and $F_{2,9} = 37.167$; optimum $MR_X = 5.14(5.03-5.26)$; π_X versus MR_X; $r = 0.654$; outlier: compound 4 in Table 10.

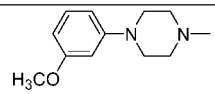
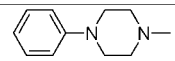
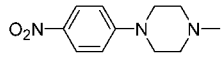
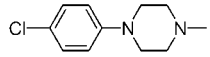
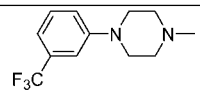
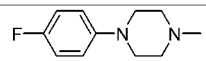
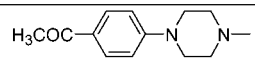
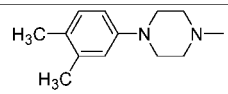
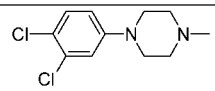
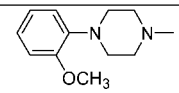
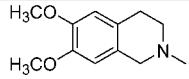
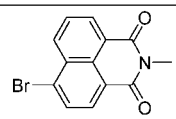
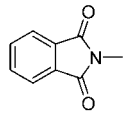
π_X and MR_X are the calculated hydrophobicity and molar refractivity of X-substituents, respectively. Equation 20 is a parabolic correlation in terms of hydrophobicity of X-substituents, whereas eq 21 is a parabolic correlation in terms of molar refractivity of X-substituents. This suggests that CPTs (**XXVIII**) may act by different mechanisms or interaction with other cellular targets in addition to nuclear topo I in each of these two human cancer (A2780 and Bel7402) cell lines.

3.2.2.2.5. Modification of A/B Rings. Growth inhibitory activity (IC_{50}; mol/L) of 7-X-10-Y-camptothecins (**XVII**) to HeLa/SF (human cervical carcinoma in serum-free media) cells. Data obtained from Vladu et al.[127] are shown in Table 1.

$$\log 1/IC_{50} = 0.71(\pm 0.26)\pi_X - 1.45(\pm 0.53)\pi_Y + 6.64(\pm 0.32) \quad (22)$$

where $n = 11$, $r^2 = 0.931$, $s = 0.231$, $q^2 = 0.847$, $Q =$

Table 10. Biological (IC_{50}; mol L^{-1})[73] and Physicochemical Parameters Used To Derive QSAR Eqs 20 and 21[c]

No.	X	log 1/ IC_{50} (Eq. **20**)			log 1/ IC_{50} (Eq. **21**)			π_X	MR_X
		Obsd	Pred.	Δ	Obsd	Pred.	Δ		
1		7.25	7.20	0.05	8.05	7.67	0.38	2.20	5.72
2[a]		7.57	7.20	0.37	8.10	8.09	0.01	2.18	5.10
3		7.17	7.22	-0.05	7.23	7.68	-0.45	2.29	5.72
4[b]		7.16	7.08	0.08	7.19	7.83	-0.64	3.06	5.60
5[a]		7.80	6.89	0.91	7.74	7.81	-0.07	3.36	5.61
6		7.25	7.23	0.02	8.15	8.09	0.06	2.49	5.12
7		7.05	7.11	-0.06	7.19	7.02	0.17	1.93	6.07
8		7.04	7.04	0.00	7.29	7.10	0.19	3.13	6.03
9		6.54	6.57	-0.03	7.11	6.98	0.13	3.72	6.09
10		7.10	7.20	-0.10	7.33	7.67	-0.34	2.20	5.72
11		7.08	7.00	0.08	8.05	7.93	0.12	1.72	5.51
12		ND	6.89	ND	6.11	6.29	-0.18	3.37	6.34
13		ND	6.64	ND	6.10	6.10	0.00	1.27	3.88

[a] Not included in the derivation of QSAR eq 20. [b] Not included in the derivation of QSAR eq 21. [c] ND = not determined.

4.177, and $F_{2,8} = 53.971$; outliers: X = $(CH_2)_2CH_3$, Y = H; X = $(CH_2)_3CH_3$, Y = H.

The hydrophobic parameters π_X and π_Y are for the substituents at the 7- and 10-position, respectively. The negative coefficients of π_Y suggests that the less hydrophobic substituent at position-10, that is, hydroxyl group will be favorable as compared to the methoxyl group for the enhanced inhibitory activity of CPTs (**XVII**) to HeLa/SF cells. Two compounds [X = $(CH_2)_2CH_3$, Y = H; X = $(CH_2)_3CH_3$, Y = H] were deemed to be outliers due to being much less active than expected by 4.3 and 7.5 times to the standard deviation, respectively.

It is important to note here that the two eqs 10 and 22 are obtained from the same data set, which are very similar to each other but differ in the nature of their π_X. Equation 10 ($-0.30\pi_X$) suggests that less hydrophobic substituents at C-7 position will increase the topo I activities of CPTs (**XVII**). On the contrary, the increase in the hydrophobicity of the X-substituents ($+0.71\pi_X$) increases the growth inhibitory activities of these compounds against HeLa/SF cells (eq 22). The reasons for these conflicting results about the hydrophobicity of the C-7 substituents are not much clear because DNA topo I is the sole target of the CPT family of anticancer compounds.

We would agree with the suggestion of Wang et al.[134] that such structural modifications may change susceptibility to cellular efflux mechanisms or to interaction with other cellular targets or loci in addition to nuclear topo I. Increasing hydrophobicity by addition of alkyl groups at the C-7 position can increase persistence of C-10 substituted CPT analogues, a result that parallels an increase in antiproliferative activity against a variety of human cancer cell lines. To provide a

Table 11. Biological (IC_{50}; mol L^{-1}),[136] Physicochemical, and Structural Parameters Used To Derive QSAR Eq 23

			log 1/ IC_{50} (eq 23)				
No.	X	Y	obsd.	pred.	Δ	π_X	I_{OH}
1	H	H	6.48	6.70	−0.22	0.00	0
2	CHO	H	6.41	6.64	−0.23	−0.09	0
3	CHO	OH	5.75	5.47	0.28	−0.09	1
4	CHO	OCH_3	6.74	6.64	0.10	−0.09	0
5	CH_2OCOCH_3	H	6.82	6.59	0.23	−0.18	0
6	CH_2OCOCH_3	OH	5.15	5.41	−0.26	−0.18	1
7[a]	CH_2OCOCH_3	OCH_3	7.40	6.59	0.81	−0.18	0
8	CN	H	5.98	6.46	−0.48	−0.39	0
9	CN	OH	5.25	5.28	−0.03	−0.39	1
10	CN	OCH_3	6.18	6.46	−0.28	−0.39	0
11	CH=CHCHO	H	6.77	6.64	0.13	−0.09	0
12	$CH{=}CHCOOC_2H_5$	H	7.40	7.22	0.18	0.85	0
13	CH=CHCN	H	6.89	6.68	0.21	−0.03	0
14	CH=C(CN)CN	H	6.52	6.26	0.26	−0.71	0
15	$CH{=}C(CN)COOC_2H_5$	H	6.60	6.80	−0.20	0.17	0
16[a]	CH=C(Br)Br	H	5.34	7.93	−2.59	2.01	0
17	$CH(OH)CH_2NO_2$	H	5.91	5.73	0.18	−1.59	0
18	$CH_2CH_2COOC_2H_5$	H	7.22	7.11	0.11	0.67	0

[a] Not included in the derivation of QSAR eq 23.

structural evidence for the observed increase in persistence for 7-alkyl modifications, Adam and colleagues[135] substituted BACPT (7-butyl-10-amino-camptothecin) in the X-ray crystal structure of the ternary topo I/DNA complex for TPT.[24] By superposing BACPT to TPT; all E-ring contacts were maintained. Modeling of the binding mode of BACPT revealed a direct hydrogen bond contact for the 10-amino to the side chain of Glu-356 of core subdomain I of topo I in addition to known contacts found for other CPTs. More important, residues 350−356 and 425−431 of core subdomain I may provide induced fit stabilization to the hydrophobic alkyl moiety at the C-7 position.

Cytotoxicity (IC_{50}; mol/L) of 7-X-10-Y-camptothecins (**XVII**) to H460 (human nonsmall-cell lung carcinoma) cells. Data obtained from Dallavalle et al.[136] are shown in Table 11.

$$\log 1/IC_{50} = 0.61(\pm 0.27)\pi_X - 1.18(\pm 0.36)I_{OH} + 6.70(\pm 0.16) \quad (23)$$

where $n = 16$, $r^2 = 0.859$, $s = 0.258$, $q^2 = 0.778$, $Q = 3.593$, and $F_{2,13} = 39.599$; outliers: X = CH_2OCOCH_3, Y = OCH_3; X = CH=C(Br)Br, Y = H.

π_X is the calculated hydrophobicity of X-substituents, whereas I_{OH} is an indicator variable taking the value of 1 and 0 for the presence and absence of a hydroxyl group in the Y-substituents, respectively. The negative coefficient of the indicator variable indicates that the presence of H-bonding polar hydroxyl group at position 10 is detrimental, a small alkoxy group, possibly an H-bond acceptor, is better tolerated. This result is in contradiction to that of eq 22, which may be due to the difference in the nature of either X-substituents of the CPTs or the cell lines. One compound [X = CH_2OCOCH_3, Y = OCH_3] was deemed to be an outlier due to being much more active than expected by 3 times to the standard deviation. On the other hand, the other derivative [X = CH=C(Br)Br, Y = H] was deemed to be an outlier due to being much less active than expected by 10 times to the standard deviation.

Cytotoxicity (IC_{50}; mol/L) of 7-X-10-Y-11-aza-camptothecins (**XXI**) to HT29 (human colon carcinoma) cells. Data obtained from Uehling et al.[128] are shown in Table 12.

Table 12. Biological (IC_{50}; mol L^{-1})[128] and Physicochemical Parameters Used To Derive QSAR Eq 24

			log 1/ IC_{50} (eq 24)			
No.	X	Y	obsd.	pred.	Δ	π_Y
1	CH_3	Br	9.00	8.57	0.43	0.78
2	C_2H_5	Br	8.15	8.57	−0.42	0.78
3	C_2H_5	CN	7.77	7.37	0.40	0.48
4	CH_3	CH_2NH_2	6.91	6.60	0.31	0.83
5	C_2H_5	CH_2NH_2	ND[b]	6.60	ND	0.83
6[a]	C_2H_5	$C(NH_2)NOH$	8.15	6.53	1.62	1.48
7	C_2H_5	$C(NH_2)NH$	6.51	6.79	−0.28	1.02
8	CH_3	$C{\equiv}CCH_2NH_2$	6.45	6.56	−0.11	1.71
9	C_2H_5	$C{\equiv}CCH_2NH_2$	6.45	6.56	−0.11	1.71
10	C_2H_5	$COOC_2H_5$	7.59	7.82	−0.23	1.58

[a] Not included in the derivation of QSAR eq 24. [b] ND = not determined.

Table 13. Biological (IC_{50}; mol L^{-1})[137] and Physicochemical Parameters Used To Derive QSAR Eq 25

			log 1/ IC_{50} (eq 25)			
No.	X	Y	obsd.	pred.	Δ	π_X
1	H	H	8.53	8.67	−0.14	0.00
2	OH	H	6.58	6.87	−0.29	−0.67
3	CH_3	H	8.86	8.97	−0.11	0.56
4	Br	H	9.03	8.68	0.35	0.86
5	Cl	H	8.67	8.86	−0.19	0.71
6	H	Br	8.54	8.67	−0.13	0.00
7	H	Cl	8.75	8.67	0.08	0.00
8	H	F	9.0	8.67	0.33	0.00
9[a]	H	OH	6.96	8.67	−1.70	0.00
10	Cl	Cl	8.80	8.86	−0.07	0.71
11[a]	NH_2	H	7.62	4.16	3.46	−1.23
12	OCH_3	F	8.94	8.64	0.31	−0.02
13	OH	F	7.06	6.87	0.19	−0.67
14	CH_3	F	8.82	8.97	−0.15	0.56
15	F	F	8.66	8.84	−0.19	0.14

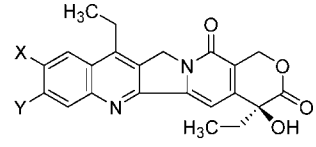
[a] Not included in the derivation of QSAR eq 25.

$$\log 1/IC_{50} = 1.02(\pm 0.39)\pi_Y + 7.66(\pm 0.33) \quad (24)$$

where $n = 8$, $r^2 = 0.874$, $s = 0.359$, $q^2 = 0.736$, $Q = 2.605$, and $F_{1,6} = 41.619$; outlier: X = C_2H_5, Y = $C(NH_2)NOH$.

π_Y is the calculated hydrophobic parameter of Y-substituents, which is the single most important parameter for this data set. The linear π_Y model suggests that the molecules with highly hydrophobic Y-substituents will be more active. One compound [X = C_2H_5, $Y = C(NH_2)NOH$] was deemed to be an outlier due to being much more active than expected by 4.5 times to the standard deviation.

Inhibitory concentration (IC_{50}; mol/L) of 7-ethyl-10-X-11-Y-camptothecins (**XXIX**) to PC-6/SN2−5H2 (SN-38-resistant) cells. Data obtained from Yoshikawa et al.[137] are shown in Table 13.



XXIX

$$\log 1/IC_{50} = 1.52(\pm 0.34)\pi_X - 1.74(\pm 0.64)\pi_X^2 + 8.67(\pm 0.21) \quad (25)$$

where $n = 13$, $r^2 = 0.914$, $s = 0.244$, $q^2 = 0.809$, $Q = 3.918$, and $F_{2,10} = 53.140$; optimum $\pi_X = 0.44(0.31-0.66)$; outlier: X = H, Y = OH; X = NH_2, Y = H.

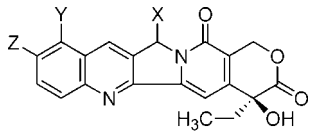
Table 15. Biological (GI_{50}; mol L^{-1}),[61] Physicochemical, and Structural Parameters Used To Derive QSAR Eqs 27 and 28

					log 1/GI_{50} (Eq. 27)			log 1/GI_{50} (Eq. 28)					
No.	R	X	Y	Z	obsd.	pred.	Δ	obsd.	pred.	Δ	Clog *P*	CMR	*I*
1	CH_2CH_2OH	H	H	OH	5.49	5.58	−0.09	5.47	5.28	0.19	0.34	10.91	0
2	CH_2CH_2F	H	H	OH	6.60	6.70	−0.10	6.30	6.35	−0.05	1.26	10.77	0
3	$CH_2CH_2OCH_3$	H	H	OH	6.17	6.36	−0.19	5.92	6.41	−0.49	0.90	11.37	0
4	CH_2CF_3	H	H	OH	7.00	6.97	0.03	7.00	7.08	−0.08	1.80	10.80	0
5	CH_2CH_2OH	H	NO_2	H	5.38	5.24	0.14	5.47	5.43	0.04	0.15	11.37	0
6	CH_2CH_2F	H	NO_2	H	6.20	6.53	−0.33	6.60	6.50	0.10	1.07	11.23	0
7	$CH_2CH_2OCH_3$	H	NO_2	H	6.36	6.13	0.23	6.37	6.56	−0.19	0.71	11.83	0
8	CH_2CF_3	H	NO_2	H	7.22	6.91	0.31	7.70	7.23	0.47	1.61	11.26	0
9	CH_2CH_2F	C_2H_5	H	OH	6.00	6.08	−0.08	5.10	5.48	−0.38	2.28	11.70	1
10[b]	$CH_2CH_2OCH_3$	C_2H_5	H	OH	6.26	6.10	0.16	6.22	5.53	0.69	1.93	12.30	1
11	CH_2CF_3	C_2H_5	H	OH	6.00	5.81	0.19	6.40	6.20	0.20	2.82	11.73	1
12[a]	CH_2CH_2OH	C_2H_5	OCH_3	H	5.30	6.01	−0.71	5.42	5.10	0.32	1.59	12.30	1
13	CH_2CH_2F	C_2H_5	OCH_3	H	6.10	6.00	0.10	5.97	6.17	−0.20	2.51	12.16	1
14	$CH_2CH_2OCH_3$	C_2H_5	OCH_3	H	5.94	6.10	−0.16	6.28	6.23	0.05	2.15	12.76	1
15[b]	CH_2CF_3	C_2H_5	OCH_3	H	5.41	5.61	−0.20	5.55	6.89	−1.34	3.05	12.19	1

[a] Not included in the derivation of QSAR eq 27. [b] Not included in the derivation of QSAR eq 28.

Parabolic dependence on π_X (calculated hydrophobic parameters of X-substituents) provides an optimum hydrophobicity of 0.44, where IC_{50} represents the drug concentration (mol L^{-1}) that required for the reduction of cell growth by 50%. One compound [X = H, Y = OH] was deemed to be an outlier due to being much less active than expected by 7 times to the standard deviation. On the other hand, the derivative [X = NH_2, Y = H] was deemed to be an outlier due to being much more active than expected by 14 times to the standard deviation.

3.2.2.2.6. Modification of A/C Rings. A series of 5-X-9-Y-10-Z-camptothecins (**XXX**) was synthesized and evaluated for their cytotoxic activities against seven cancer cell lines by Subrahmanyam et al.[62] The cytotoxicity data (GI_{50}) was given in μM, where GI_{50} stands for the concentration of the drug required to produce 50% growth inhibition of the cell under study. The cytotoxic activities of these compounds (GI_{50}) against UACC 62 (melanoma) cancer cell line was converted into molar concentration and used to develop eq 26 (Table 14).



XXX

$$\log 1/GI_{50} = 2.18(\pm 0.73)MR_X + 1.92(\pm 1.36) \quad (26)$$

where $n = 11$, $r^2 = 0.836$, $s = 0.337$, $q^2 = 0.749$, $Q = 2.712$, and $F_{1,9} = 45.878$; π_X versus MR_X; $r = 0.003$; outlier: X = CH_2COCH_3, Y = NO_2, Z = H.

MR_X is the calculated molar refractivity of X-substituents and its positive coefficient suggests that an increase in molar refractivity of X-substituents should result in stronger cytotoxicity against UACC 62 (melanoma) cancer cells. Lipophilicity of the X-substituents is not found to play a definite role as evidenced by a very poor correlation between π_X and MR_X (π_X vs MR_X; $r = 0.003$). Substituting π_X for MR_X in Eq. 26 gave a very poor fit ($r^2 = 0.012$, $q^2 = -0.562$) suggesting interaction in nonhydrophobic space. One compound [X = CH_2COCH_3, $Y = NO_2$, $Z = H$] was deemed to be an outlier due to being less active than expected by three times to the standard deviation.

3.2.2.2.7. Modification of A/B/C Rings. A series of 5-OR-7-X-9-Y-10-Z-camptothecins (**XXXI**) was also synthesized and evaluated for their cytotoxicities against seven cancer cell lines by Subrahmanyam et al.[61] The cytotoxicity data (GI_{50}) was given in μM, where GI_{50} represents the concentration of the drug required to produce 50% growth inhibition of the cell under study. The cytotoxicity data of these compounds (GI_{50}) against two cancer cell lines (DU-145 and ACHN) was converted into molar concentration to develop eqs 27, and 28.

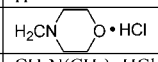
Cytotoxicity (GI_{50}; mol/L) of 5-OR-7-X-9-Y-10-Z-camptothecins (**XXXI**) to DU-145 (prostate) cancer cells. Data obtained from Subrahmanyam et al.[61] are shown in Table 15.

Table 14. Biological (GI_{50}; mol L^{-1})[62] and Physicochemical Parameters Used To Derive QSAR Eq 26

				log 1/GI_{50} (eq 26)				
No.	X	Y	Z	obsd.	pred.	Δ	MR_X	π_X
1	CH_2COOCH_3	H	H	5.19	5.56	−0.37	1.67	0.54
2	CH_2COOCH_3	H	OH	5.76	5.56	0.20	1.67	0.54
3	CH_2COOCH_3	NO_2	H	5.39	5.56	−0.17	1.67	0.54
4	CH_2COCH_3	H	H	5.45	5.22	0.23	1.52	0.24
5	CH_2COCH_3	H	OH	5.12	5.22	−0.10	1.52	0.24
6[a]	CH_2COCH_3	NO_2	H	6.52	5.22	1.30	1.52	0.24
7	$CH_2COOCH_2CH_2F$	H	H	6.70	6.60	0.10	2.15	0.79
8	$CH_2COOCH_2CH_2OH$	H	H	6.70	6.90	−0.20	2.29	−0.11
9	$CH_2COOCH_2CF_3$	H	H	7.15	6.67	0.48	2.18	0.84
10	CH_2CONH_2	H	H	4.80	5.01	−0.21	1.42	−0.63
11	CH_2CO(pyrrolidin-1-yl)	H	H	6.39	5.92	0.47	1.84	−0.67
12	CH_2CO(piperidin-1-yl)	H	H	6.48	6.93	−0.45	2.30	−0.67

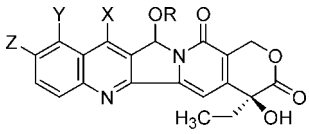
[a] Not included in the derivation of QSAR eq 26.

Table 16. Biological (IC_{50}; mol L^{-1})[107] and Physicochemical Parameters Used To Derive QSAR Eqs 29−33

No.	X	Y	Z	log 1/IC_{50} (Eq. **29**)			log 1/IC_{50} (Eq. **30**)			log 1/IC_{50} (Eq. **31**)		
				Obsd.	Pred.	Δ	Obsd.	Pred.	Δ	Obsd.	Pred.	Δ
1	H	H	H	7.64	7.81	-0.17	7.64	7.86	-0.22	7.70	7.91	-0.21
2	C_2H_5	H	H	7.80	7.33	0.47	7.85	7.34	0.51	7.70	7.33	0.37
3	H	CH_3	H	7.64	7.62	0.02	7.68	7.66	0.02	7.72	7.69	0.03
4	C_2H_5	CH_3	H	6.87	6.92	-0.05	7.07	6.89	0.18	6.82	6.83	-0.01
5	H	C_2H_5	H	7.68	7.33	0.35	7.77	7.34	0.43	7.72	7.33	0.39
6	H	CH_2CN	H	7.52	7.32	0.20	7.43	7.33	0.10	7.40	7.32	0.08
7	H	$(CH_2)_2NH_2$.AcOH	H	6.38	7.01	-0.63	6.29	7.00	-0.71	6.27	6.95	-0.68
8	H	H		5.27	5.23	0.04	5.02	5.03	-0.01	4.81	4.74	0.07
9[a]	H	H	$CH_2N(CH_3)_2$.HCl	5.50	6.51	-1.01	5.09	6.45	-1.36	5.02	6.33	-1.31
10	CPT			7.77	7.74	0.03	7.85	7.76	0.09	7.77	7.79	-0.02
11	SN-38			7.85	7.85	0.00	7.80	7.90	-0.10	8.10	7.96	0.14
12	Topotecan			7.25	7.52	-0.27	7.27	7.55	-0.28	7.42	7.57	-0.15

No.	log 1/IC_{50} (Eq. **32**)			log 1/IC_{50} (Eq. **33**)			CMR	Clog *P*
	Obsd.	Pred.	Δ	Obsd.	Pred.	Δ		
1	7.55	7.83	-0.28	7.89	8.11	-0.22	10.79	1.35
2	7.85	7.32	0.53	8.15	7.65	0.50	11.72	2.38
3	7.59	7.64	-0.05	8.00	7.96	0.04	11.26	1.79
4	6.85	6.87	-0.02	7.12	7.18	-0.06	12.19	2.82
5	7.72	7.32	0.40	7.89	7.65	0.24	11.72	2.32
6	7.49	7.31	0.18	7.85	7.63	0.22	11.74	0.94
7	6.27	6.97	-0.70	6.76	7.29	-0.53	12.09	1.05
8	4.98	4.93	0.05	5.13	5.09	0.04	13.46	1.10
9[a]	5.06	6.41	-1.35	5.07	6.69	-1.62	12.55	1.19
10	7.70	7.67	0.03	7.74	7.73	0.01	9.52	0.90
11	7.96	7.87	0.09	8.22	8.13	0.09	10.60	1.97
12	7.32	7.54	-0.22	7.51	7.86	-0.35	11.43	0.73

[a] Not included in the derivation of QSAR eqs 29−33.



XXXI

$$\log 1/GI_{50} = 2.00(\pm 0.59)\text{Clog } P - 0.49(\pm 0.19)\text{Clog } P^2 - 0.89(\pm 0.51)I + 4.96(\pm 0.42) \quad (27)$$

where $n = 14$, $r^2 = 0.874$, $s = 0.219$, $q^2 = 0.726$, $Q = 4.269$, and $F_{3,10} = 23.122$; optimum Clog $P = 2.04(1.73-2.57)$; outlier: R = CH_2CH_2OH, X = C_2H_5, Y = OCH_3, Z = H.

Cytotoxicity (GI_{50}; mol/L) of 5-OR-7-X-9-Y-10-Z-camptothecins (**XXXI**) to ACHN (renal) cancer cells. Data obtained from Subrahmanyam et al.[61] are shown in Table 15.

$$\log 1/GI_{50} = 1.29(\pm 0.44)\text{Clog } P + 0.86(\pm 0.60)\text{CMR} - 3.00(\pm 1.03)I - 6.92(\pm 4.56) \quad (28)$$

where $n = 13$, $r^2 = 0.857$, $s = 0.312$, $q^2 = 0.717$, $Q = 2.968$, and $F_{3,9} = 17.979$; Clog *P* versus CMR; $r = 0.462$; outliers: R = $CH_2CH_2OCH_3$, X = C_2H_5, Y = H, Z = OH; R = CH_2CF_3, X = C_2H_5, Y = OCH_3, Z = H.

Equation 27 is a parabolic correlation in terms of hydrophobicity with optimum Clog *P* of 2.04, whereas eq 28 is a linear correlation in terms of Clog *P* followed by CMR. Both equations have an indicator variable (*I*) that takes the value of 1 and 0 for the presence and absence of an ethyl group at position-7, respectively. The negative coefficient of the indicator variable suggests that the presence of an ethyl group at position-7 is detrimental to the activity for this data set. Equations 27 and 28 represent the cytotoxicities of a series of CPTs (**XXXI**) against two different cancer cell lines and are very different from each other, which suggests that these compounds may act by different mechanisms or interaction with other cellular targets along with topo I in each of these two human cancer (DU-145 and ACHN) cell lines. In eq 27, one compound [R = CH_2CH_2OH, X = C_2H_5, Y = OCH_3, Z = H] was deemed to be an outlier due to being less active than expected by 3 times to the standard deviation. On the other hand, two compounds [R = $CH_2CH_2OCH_3$, X = C_2H_5, Y = H, Z = OH; R = CH_2CF_3, X = C_2H_5, *Y* = OCH_3, *Z* = H] in eq 28 were deemed to be outliers due to being much more and much less active than expected by 2.2 and 4.3 times to the standard deviation, respectively.

3.2.2.2.8. Miscellaneous CPTs. It was claimed that the hexacyclic CPTs exhibited antitumor activities superior to those of the original pentacyclic ring system, probably due to the increased planarity exerted by an additional ring. Based on this fact, Kim et al.[107] synthesized a series of hexacyclic CPTs (**XX**) and evaluated their cytotoxic activities along with three known compounds (CPT, SN-38, and topotecan) against different human cancer cell lines. The cytotoxicity data (IC_{50}) was given in nM, where IC_{50} represents the concentration of the compound causing 50% cell death. The cytotoxicity data (IC_{50}) of these compounds (**XX**) against five human cancer cell lines (WiDr, A549, MKN45, SK-OV-3, and H128) was converted into molar concentration and then used in the development of eqs 29−33.

Cytotoxicity (IC_{50}; mol/L) of hexacyclic CPTs (**XX**), CPT, SN-38, and topotecan to WiDr (human colon) cancer cells. Data obtained from Kim et al.[107] are found in Table 16.

$$\log 1/IC_{50} = 5.31(\pm 3.77)\text{CMR} - 0.26(\pm 0.16)\text{CMR}^2 - 21.60(\pm 19.34) \quad (29)$$

where $n = 11$, $r^2 = 0.859$, $s = 0.334$, $q^2 = 0.775$, $Q =$

2.775, and $F_{2,8}$ = 24.369; optimum CMR = 10.25(8.03−10.86); Clog *P* versus CMR; *r* = 0.133; outlier: X = Y = H, Z = $CH_2NMe_2 \cdot HCl$.

Cytotoxicity (IC_{50}; mol/L) of hexacyclic CPTs (**XX**), CPT, SN-38, and topotecan to A549 (human lung) cancer cells. Data obtained from Kim et al.[107] are shown in Table 16.

$$\log 1/IC_{50} = 5.95(\pm 4.26)CMR - 0.29(\pm 0.19)CMR^2 - 24.36(\pm 22.69) \quad (30)$$

where $n = 11$, $r^2 = 0.851$, $s = 0.376$, $q^2 = 0.752$, $Q = 2.452$, and $F_{2,8} = 22.846$; optimum CMR = 10.29(8.06−10.89); Clog *P* versus CMR; *r* = 0.133; outlier: X = Y = H, Z = $CH_2NMe_2 \cdot HCl$.

Cytotoxicity (IC_{50}; mol/L) of hexacyclic CPTs (**XX**), CPT, SN-38, and topotecan to MKN45 (human stomach) cancer cells. Data obtained from Kim et al.[107] are shown in Table 16.

$$\log 1/IC_{50} = 6.73(\pm 3.67)CMR - 0.33(\pm 0.16)CMR^2 - 26.67(\pm 20.99) \quad (31)$$

where $n = 11$, $r^2 = 0.906$, $s = 0.324$, $q^2 = 0.812$, $Q = 2.938$, and $F_{2,8} = 38.553$; optimum CMR = 10.30(9.09−10.79); Clog *P* versus CMR; *r* = 0.133; outlier: X = Y = H, Z = $CH_2NMe_2 \cdot HCl$.

Cytotoxicity (IC_{50}; mol/L) of hexacyclic CPTs (**XX**), CPT, SN-38, and topotecan to SK-OV-3 (human ovary) cancer cells. Data obtained from Kim et al.[107] are shown in Table 16.

$$\log 1/IC_{50} = 6.41(\pm 4.20)CMR - 0.31(\pm 0.18)CMR^2 - 25.33(\pm 24.05) \quad (32)$$

where $n = 11$, $r^2 = 0.859$, $s = 0.371$, $q^2 = 0.786$, $Q = 2.499$, and $F_{2,8} = 24.369$; optimum CMR = 10.37(8.64−10.91); Clog *P* versus CMR; *r* = 0.133; outlier: X = Y = H, Z = $CH_2NMe_2 \cdot HCl$.

Cytotoxicity (IC_{50}; mol/L) of hexacyclic CPTs (**XX**), CPT, SN-38, and topotecan to H128 (human lung) cancer cells. Data obtained from Kim et al.[107] are shown in Table 16.

$$\log 1/IC_{50} = 7.70(\pm 3.64)CMR - 0.36(\pm 0.16)CMR^2 - 32.54(\pm 20.87) \quad (33)$$

where $n = 11$, $r^2 = 0.896$, $s = 0.322$, $q^2 = 0.844$, $Q = 2.941$, and $F_{2,8} = 34.462$; optimum CMR = 10.57(9.76−10.96); Clog *P* versus CMR; *r* = 0.133; outlier: X = Y = H, Z = $CH_2NMe_2 \cdot HCl$.

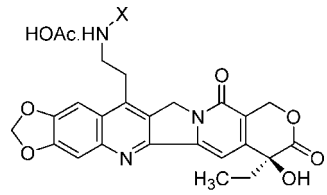
All the above five equations (eqs 29−33) are parabolic relations in terms of molar refractivity of the whole molecules with optimum CMR of 10.25−10.57. Hydrophobicity is not found to play a definite role as evidenced by a very poor correlation between Clog *P* and CMR (Clog *P* vs CMR; *r* = 0.133). One common compound (X = Y = H, Z = $CH_2NMe_2 \cdot HCl$) was deemed to be an outlier in all these five equations (eqs 29−33) due to being less active than expected by 3−5 times their standard deviations. The similarity of these equations suggests that the hexacyclic CPTs (**XX**), CPT, SN-38, and topotecan may target an enzyme of similar kind in these five human cancer (WiDr, A549, MKN45, SK-OV-3, and H128) cells.

Jew et al.[138] synthesized a series of hexacyclic CPTs (**XXXII**) and evaluated their cytotoxic activities along with two other known compounds (CPT and 7-*N*-isopropylaminoethyl-CPT) against five different human tumor cell lines. The cytotoxicity data (IC_{50}) was given in μM, where IC_{50} represents the concentration of the compound causing 50% cell death. The cytotoxicity data (IC_{50}) of these compounds against one human tumor cell line (DLD-1) was converted into molar concentration and used in the development of eq 34 (Table 17).

Table 17. Biological (IC_{50}; mol L^{-1})[138] and Physicochemical Parameters Used To Derive QSAR Eq 34

No.	X	log 1/ IC_{50} (eq 34) obsd.	pred.	Δ	Clog *P*
1	CH_2CH_3	8.21	7.70	0.51	1.38
2	$CH_2CH_2CH_3$	7.59	7.01	0.58	1.91
3	$CH(CH_3)_2$	7.05	7.30	−0.25	1.69
4	$CH_2CH(CH_3)_2$	6.19	6.48	−0.29	2.31
5[a]	$CH(CH_3)CH_2CH_3$	5.51	6.60	−1.09	2.22
6	$(CH_2)_3CH_3$	5.95	6.31	−0.36	2.44
7	$CH_2CH(CH_3)CH_2CH_3$	5.95	5.79	0.16	2.84
8	7-$(CH_2)_2NHCH(CH_3)_2$-CPT	8.14	7.98	0.16	1.17
9	CPT	7.85	8.34	−0.49	0.90

[a] Not included in the derivation of QSAR eq 34.



XXXII

$$\log 1/IC_{50} = -1.32(\pm 0.61)Clog\ P + 9.52(\pm 1.18) \quad (34)$$

where $n = 8$, $r^2 = 0.824$, $s = 0.440$, $q^2 = 0.686$, $Q = 2.064$, and $F_{1,6} = 28.091$; outlier: X = $CH(CH_3)CH_2CH_3$.

The negative Clog *P* term shows that the highly hydrophilic molecules for this data set would present better cytotoxic activities. One compound (X = $CH(CH_3)CH_2CH_3$) was deemed to be an outlier due to being less active than expected by 2.5 times to the standard deviation.

A series of hexacyclic CPTs (**XXXIII**) was also synthesized by Jew et al.[138] and evaluated their cytotoxic activities along with CPT against five different human tumor cell lines. From the cytotoxicity data (IC_{50}; mol/L) of these compounds against one human tumor cell line (HEC-1-B), we developed eq 35 (Table 18).

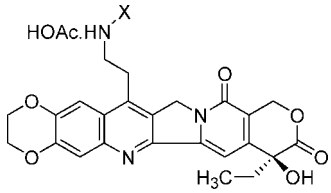
$$\log 1/IC_{50} = -1.15(\pm 0.55)Clog\ P + 8.15(\pm 1.16) \quad (35)$$

where $n = 7$, $r^2 = 0.854$, $s = 0.347$, $q^2 = 0.701$, $Q = 2.663$, and $F_{1,5} = 29.247$; outlier: X = $CH(CH_3)_2$. One compound [X = $CH(CH)_2$] was deemed to be an outlier due to being

Table 18. Biological (IC_{50}; mol L^{-1})[138] and Physicochemical Parameters Used To Derive QSAR Eq 35

No.	X	log 1/ IC_{50} (eq 35) obsd.	pred.	Δ	Clog *P*
1	CH_2CH_3	6.72	6.61	0.11	1.34
2[a]	$CH(CH_3)_2$	4.89	6.25	−1.36	1.65
3	$CH_2CH(CH_3)_2$	5.44	5.54	−0.10	2.27
4	$CH(CH_3)CH_2CH_3$	5.48	5.65	−0.17	2.18
5	$(CH_2)_3CH_3$	5.44	5.40	0.04	2.40
6	$CH_2CH(CH_3)CH_2CH_3$	5.46	4.94	0.53	2.80
7	$CH_2CH_2N(C_2H_5)_2$	4.92	5.43	−0.51	2.37
8	CPT	7.22	7.12	0.10	0.90

[a] Not included in the derivation of QSAR eq 35.



XXXIII

less active than expected by 3.9 times to the standard deviation.

3.2.2.3. QSAR for the Antiprotozoal Activity. African trypanosomes (*Trypanosoma brucei* species) are parasitic protozoa causing lethal diseases in human and cattle. It has been shown that CPT is cytotoxic to African trypanosomes and related pathogenic hemoflagellates.[139] In another study, a series of CPT derivatives (see Table 19) was tested against axenically cultured, bloodstream form, *T. brucei* by Bodley et al.[140] The cytotoxicity data (EC_{50}) was given in μM, which was converted into molar concentration and used to develop eq 36 (Table 19).

$$\log 1/EC_{50} = 0.47(\pm 0.25)\text{Clog } P - 1.25(\pm 0.29)I + 6.42(\pm 0.41) \quad (36)$$

where $n = 23$, $r^2 = 0.837$, $s = 0.319$, $q^2 = 0.770$, $Q = 2.868$, and $F_{2,20} = 51.350$; outliers: 9-NH_2-CPT; 10-NH_2−CPT. The indicator variable (I) takes the value of 1 for CPTs and 0 for EDCPT and MDCPTs.

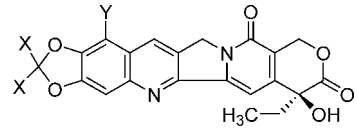
The negative coefficient of the indicator variable suggests that the presence of 10,11-ethylenedioxy or 10,11-methylenedioxy group would enhance the activity, which is further supported by the fact that it explains alone 71.4% of the varience in log $1/EC_{50}$. Two compounds (9-NH_2-CPT and 10-NH_2-CPT) were deemed to be outliers due to being more active than expected by 2.45 and 2 times, respectively, to the standard deviation.

Table 19. Biological (EC_{50}; mol L^{-1}),[140] Physicochemical, and Structural Parameters Used To Derive QSAR Eq 36[a]

No.	cmpds	log $1/EC_{50}$ (eq 36) obsd.	pred.	Δ	Clog P	I
1	CPT	5.80	5.58	0.22	0.87	1
2	7-C_2H_5-CPT	6.20	6.06	0.14	1.90	1
3	7-C_3H_7-CPT	6.10	6.30	−0.20	2.43	1
4	7-C_2H_5-9-NH_2-CPT	6.07	5.80	0.27	1.35	1
5	7-C_2H_5-9-NO_2-CPT	5.57	5.99	−0.42	1.76	1
6	7-C_2H_5-10-NH_2-CPT	6.21	5.80	0.41	1.35	1
7	7-C_2H_5-10-NO_2-CPT	6.22	5.99	0.23	1.76	1
8	9-Cl-CPT	6.09	5.93	0.16	1.63	1
9[b]	9-NH_2-CPT	6.08	5.32	0.76	0.32	1
10	9-NO_2-CPT	5.80	5.51	0.29	0.73	1
11	10-CH_3-CPT	5.64	5.81	−0.17	1.37	1
12	10-Cl-CPT	5.82	5.93	−0.11	1.63	1
13[b]	10-NH_2-CPT	5.92	5.32	0.60	0.32	1
14	10-NO_2-CPT	5.68	5.51	0.17	0.73	1
15	11-NH_2-CPT	4.74	5.32	−0.58	0.32	1
16	12-NH_2-CPT	4.92	5.32	−0.40	0.32	1
17	7-CH_3-EDCPT	7.16	7.11	0.05	1.48	0
18	MDCPT	6.80	6.88	−0.08	0.98	0
19	7-CH_3-MDCPT	7.36	7.11	0.25	1.48	0
20	7-C_2H_5-MDCPT	7.22	7.35	−0.13	2.01	0
21	7-C_2H_5-9-NH_2-MDCPT	7.24	7.11	0.13	1.48	0
22	7-C_2H_5-9-NO_2-MDCPT	6.77	7.27	−0.50	1.83	0
23	9-Cl-MDCPT	7.39	7.22	0.17	1.71	0
24	9-NH_2-MDCPT	7.13	6.63	0.50	0.45	0
25	9-NO_2-MDCPT	6.40	6.79	−0.39	0.80	0

[a] CPT = Camptothecin; MDCPT = 10,11-Methylenedioxycamptothecin; EDCPT = 10,11-Ethylenedioxycamptothecin. [b] Not included in the derivation of QSAR eq 36.

Werbovetz et al.[141] examined CPT and four of its 10,11-methylenedioxy analogues (**XXXIV**) against the pathogenic protozoan *Leishmania donavani* in vitro. The cytotoxicity data (EC_{50}) was given in μM, which was converted into molar concentration and used in the development of eq 37 (Table 20).



XXXIV

$$\log 1/EC_{50} = 8.58(\pm 2.15)\text{MgVol} - 15.61(\pm 5.60) \quad (37)$$

where $n = 4$, $r^2 = 0.993$, $s = 0.107$, $q^2 = 0.982$, $Q = 9.318$, and $F_{1,2} = 283.714$; Clog P versus MgVol; $r = 0.358$; outlier: X = Y = H.

The steric hindrance of MgVol (molar volume) was found to be the most significant variable for this data set. No role for the hydrophobic effect was found. Although this is a very small data set, it is the best model and explains 99.3% of the varience in log $1/EC_{50}$. One compound (X = Y = H) was deemed to be an outlier due to being more active than expected by six times the standard deviation.

3.2.3. Validation of the QSAR Models

QSAR model validation becomes an essential part in developing a statistically valid and predictive model, because the real utility of a QSAR model is in its ability to predict accurately the modeled property for new compounds. Criteria of validation of the QSAR models have already been discussed.[47,142−146] The following approaches have been used to validate the QSAR models **10−37**.

3.2.3.1. Fraction of the Variance. The fraction of the variance of an MRA model is expressed by r^2. It is believed that the closer the value of r^2 to unity, the better the QSAR model. The values of r^2 for these QSAR models are from 0.784 to 0.993, which suggests that these QSAR models explain 78.4−99.3% of the variance in the data set. According to the literature, the predictive QSAR model must have $r^2 > 0.6$.[145,146]

3.2.3.2. Cross-Validation Test. The values of q^2 for these QSAR models are varied from 0.602 to 0.982. The high values of q^2 validate the QSAR models. According to the literature, the predictive QSAR model must have $q^2 > 0.5$.[145,146]

3.2.3.3. Standard Deviation (s). s is the standard deviation about the regression line. The smaller the value of s the better the QSAR model. The values of s for these QSAR models are from 0.076 to 0.440.

Table 20. Biological (EC_{50}; mol L^{-1})[141] and Physicochemical Parameters Used To Derive QSAR Eq 37

No.	X	Y	log $1/EC_{50}$ (eq 37) obsd.	pred.	Δ	MgVol
1	F	H	6.80	6.82	−0.02	2.62
2[a]	H	H	7.19	6.52	0.67	2.58
3	H	Cl	7.48	7.57	−0.09	2.70
4	H	NH_2	7.49	7.38	0.11	2.68
5	CPT		5.24	5.24	0.00	2.43

[a] Not included in the derivation of QSAR eq 37.

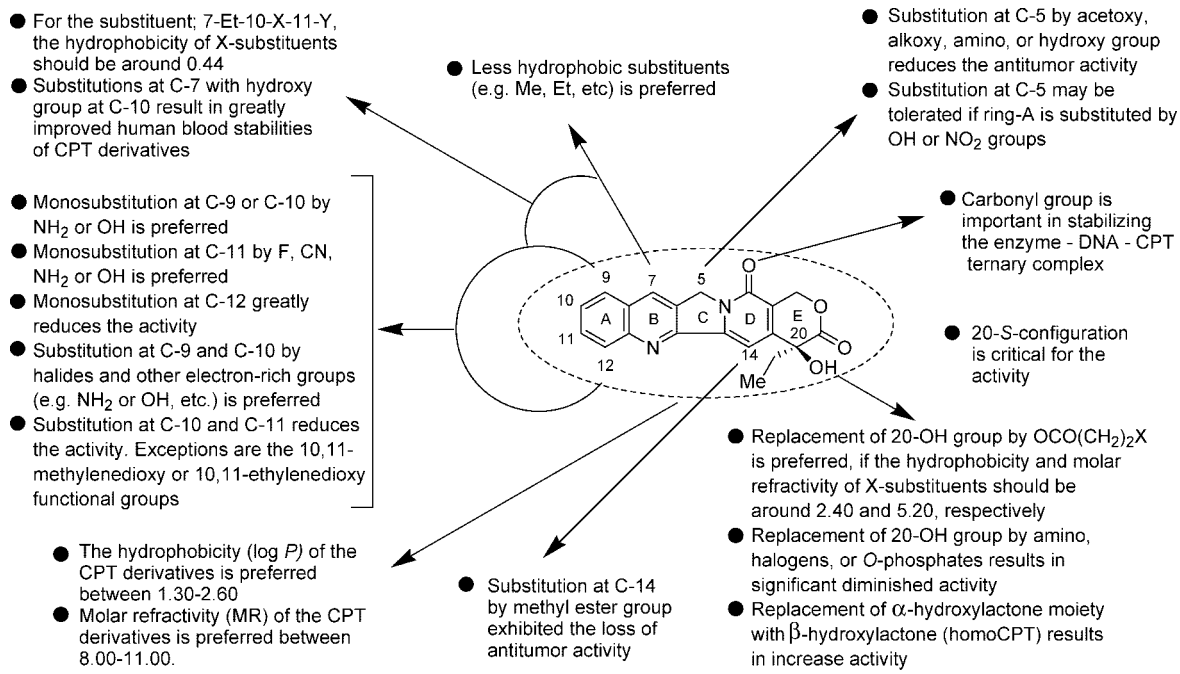


Figure 4. Brief description of SAR/QSAR of CPT on the basis of the results of SAR and QSAR models.

3.2.3.4. Quality Factor (*Q*). Chance correlation, due to the excessive number of parameter (which increases also the *r* and *s* values) is detected by the examination of *Q* value.[122−124] The high values of *Q* (2.064−12.539) for these QSAR models suggest that the high predictive power for the QSAR models as well as no over- fitting.

3.2.3.5. Fischer Statistics (*F*). Fischer statistics (*F*) is the ratio between explained and unexplained variance for a given number of degrees of freedom. The larger the value of *F* the greater the probability that the QSAR equation is significant. The *F*-values for these QSAR models are from 15.728 to 283.714, which are statistically significant at the 95% level.

All the QSAR models (except eq 13) also fulfill the thumb rule condition that (number of data points)/(number of descriptors) ≥ 4.

3.2.4. Overview

An analysis of our QSAR models (eqs 10–37) reveals two most important descriptors, for example, hydrophobicity and molar refractivity, which are valid for the activity of such a big, complicated, and flexible molecules like CPTs. Out of 28 QSAR, 14 contain a correlation between activity and hydrophobicity. A positive linear correlation is found in seven equations (eqs 12, 17, 22, 23, 24, 28, and 36). The coefficient with the hydrophobic parameter varies considerably, from a low value of 0.46 (eq 17) to the high value of 1.77 (eq 12). These data suggest that activity might be improved by increasing compound/substituents hydrophobicity. It is important to note here that the compound solubility decreases by increasing compound/substituents hydrophobicity, thus limiting bioavailability. The existence of a linear only correlation between activity and compound/substituents hydrophobicity suggests that the log P/π values were not great enough to establish the upper limit of the activity, that is, the balance between log P/π values and the compound activity. Thus, more compounds will be needed for those data sets to establish the upper limit of the log P/π values by the development of either parabolic or bilinear QSAR models. A negative linear correlation is found in 4 equations (eqs 10, 22, 34, and 35), and the coefficient range from −0.30 (eq 10) to −1.45 (eq 22). Less hydrophobic congeners in these compound families might display enhanced activity (note: eq 22 contains a positive correlation with π_X and a negative correlation with π_Y, so one should preserve a hydrophilic group at Y-position while boosting the hydrophobicity of the X-substituents). Parabolic correlations with hydrophobic parameter of the substituents are found in two equations (eqs 20 and 25), which reflect situations where activity increases with increasing hydrophobicity of the substituents up to an optimal value and then decreases. These are the encouraging examples, where the optimal values of π (hydrophobic parameter of the substituents) are well defined, which are 2.46 and 0.44, respectively. Parabolic correlations with the molecule's overall hydrophobicity are also found in two equations (eqs 18 and 27). The optimal log *P* for these equations is 1.62 and 2.04, respectively.

The second important parameter is molar refractivity of the molecules/substituents, which is present in 15 QSAR. A positive linear correlation is found in two equations (eqs 26 and 28) with coefficient 2.18 and 0.86. These data suggest that the activity might be improved by increasing compound/substituents molar refractivity/polarizability. A negative linear correlation is found in six equations (eqs 11 and 13–17) and the coefficient range is found to be from −1.48 (eq 15) to −0.39 (eq 17). Less steric congeners in these compound families might display enhanced activity. Parabolic correlation with molar refractivity of the substituents is found in one eq 21, which shows that activity is optimal for a particular value, or range of values, of MR. The optimal MR for this equation is 5.14. Parabolic correlations with the molecule's overall molar refractivity are found in five equations (eqs 29–33). The optimal molar refractivity of the whole molecules for these equations is from 10.25 to 10.57. Bilinear correlation with molar refractivity of the substituents

is also found in one eq 19, which reflects situation where activity declines with increasing molar refractivity of the substituents and then changes direction and increases. This may correspond to an allosteric reaction. Molar volume is shown in one QSAR (eq 37), but it does not seem to play as important a role as hydrophobicity and molar refractivity for the data sets that we have examined.

The presence of outliers is one of the most difficult problems in QSAR. Outliers are those compounds that have unexpected biological activities and are unable to fit in a QSAR model. It could be associated with one of the following reasons: (1) the experimental errors in the primary data, (2) the molecules may act by different mechanisms, (3) members of the data set may have different rates of metabolism, (4) the parameters used may not be the best, and (5) the intrinsic noise associated with both the original data and the methodological aspects involved in the construction of a QSAR model.[147−150]

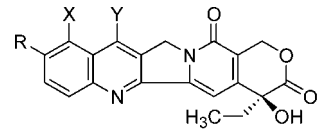
4. Summary of the SAR and QSAR

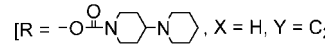
A brief summary of the structure−activity relationships (SAR) and quantitative structure−activity relationships (QSAR) is shown in Figure 4.

5. Conclusions

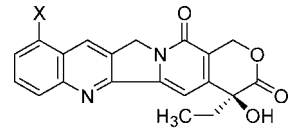
The CPT analogues are emerging as a promising group of chemotherapeutic agents. The structure−activity relationships (SAR) provide insight into the mechanism of topo I inhibition and helped in the synthesis of various CPT analogues by modifying the different rings of the original CPT molecule, giving each analogue a unique property. These modifications have resulted in various improvements in the parent molecule, including changes in bioavailability, stabilization of the lactone ring, and/or a decrease in the substrate recognition by drug-resistant proteins as well as improvements in the toxicity profile in preclinical studies. QSAR paradigm has been proved to be useful in understanding the requirements of physicochemical properties of the substituents in many key locations as well as molecules as a whole. An investigation of the QSAR results relative to the CPT analogues suggests that the two most important descriptors, for example, compound/substituent hydrophobicity and molar refractivity, are also valid for the activity of CPTs—big, complicated, and flexible molecules. According to Franke and Gruska, "The drug discovery process is of a very complex nature, effective drug design requires an entire spectrum of techniques in which QSAR methods still play an important role...".[151] Thus, the knowledge of SAR, together with the generation of QSAR, constitutes a large body of evidence that may assist in the development of new CPT with excellent antitumor activity and low toxicity.

At the end, we would like to consider the compounds that are in the advanced stage of clinical trials (**III**, **IV**, **XXXV**−**XXXXIV**). Two of them, topotecan (**III**) for the treatment of the ovarian and small-cell lung cancers and irinotecan (**IV**) for the metastatic colorectal cancers have already been approved by the FDA. The structures, Clog *P* and CMR values of these compounds (**III**, **IV**, **XXXV**−**XXXXIV**) are as follows:

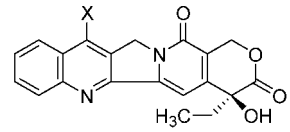


III (Topotecan)
[R = OH, X = $CH_2N(CH_3)_2$, Y = H]
(Clog *P* = 0.73; CMR = 11.43)
IV (Irinotecan)
[R =  , X = H, Y = C_2H_5]
(Clog *P* = 2.73; CMR = 16.12)

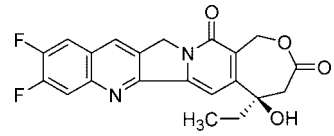
It is interesting to note that all of these compounds (except



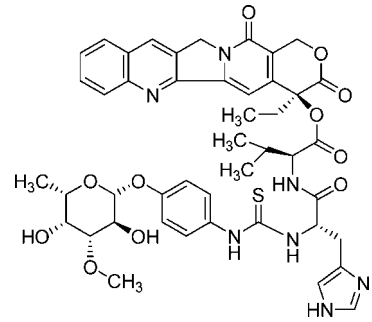
XXXV [Rubitecan; X = NO_2]
(Clog *P* = 0.76; CMR = 10.13)
XXXVI [IDEC-132; X = NH_2]
(Clog *P* = 0.35; CMR = 9.89)



XXXVII [Gimatecan; X = $CH{=}NOC(CH_3)_3$]
(Clog *P* = 2.23; CMR = 12.49)
XXXVIII [Karenitecin; X = $CH_2CH_2Si(CH_3)_3$]
(Clog *P* = 3.70; CMR = 12.73)
XXXIX [CKD-602; X = $CH_2CH_2NHCH(CH_3)_2$]
(Clog *P* = 1.17; CMR = 12.21)

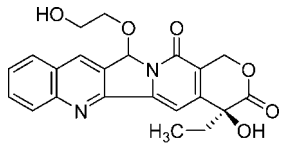


XXXX [Diflomotecan]
(Clog *P* = 1.21; CMR = 10.02)

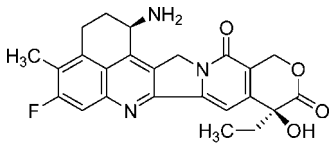


XXXXI [Bay 38-3441]
(Clog *P* = 1.54; CMR = 23.84)

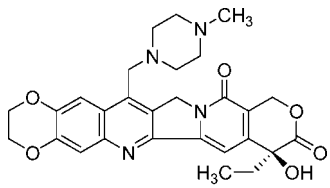
XXXVIII) have the Clog *P* values either low or very close to the predicted range (Clog *P* = 1.30−2.60; see Figure 4 and eqs 18 and 27). The CMR values of four compounds (**XXXV**, **XXXVI**, **XXXX**, and **XXXXII**) are in the predicted range (CMR = 8.00−11.00; see Figure 4 and Eqs. 29–33) and the rest have higher values (except two compounds **III** and **XXXXIII**, which have CMR values very close to the predicted range). The question arises of whether the relatively low Clog *P* is necessary or is this accidental because of the



XXXXII [DRF-1042]
(Clog *P* = 0.29; CMR = 10.76)



XXXXIII [Exatecan]
(Clog *P* = 0.95; CMR = 11.58)



XXXXIV [Lurtotecan]
(Clog *P* = 1.56; CMR = 13.92)

high CMR. Clog *P* and CMR can be collinear unless care is taken in substituent selection. Of course Clog *P* is important in bioavailability; however, usually values in the range of 1.50−3.00 suffice for this purpose.

6. References

(1) Wall, M. E.; Wani, M. C.; Cook, C. E.; Palmer, K. H.; McPhail, A. T.; Sim, G. A. *J. Am. Chem. Soc.* **1966**, *88*, 3888.
(2) Wall, M. E. *Med. Res. Rev.* **1998**, *18*, 299.
(3) Legarza, K.; Yang, L.-X. *In Vivo* **2005**, *19*, 283.
(4) Moertel, C. G.; Schutt, A. J.; Reitemeier, R. J.; Hahn, R. G. *Cancer Chemother. Rep.* **1972**, *56*, 95.
(5) Gottlieb, J. A.; Luce, J. K. *Cancer Chemother. Rep.* **1972**, *56*, 103.
(6) Wani, M. C.; Ronman, P. E.; Lindley, J. T.; Wall, M. E. *J. Med. Chem.* **1980**, *23*, 554.
(7) Carrigan, S. W.; Lii, J.-H.; Bowen, J. P. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 61.
(8) Hsiang, Y.-H.; Hertzberg, R.; Hecht, S.; Liu, L. F. *J. Biol. Chem.* **1985**, *260*, 14873.
(9) Dancey, J.; Eisenhauer, E. A. *Br. J. Cancer* **1996**, *74*, 327.
(10) Pommier, Y.; Pourquier, P.; Fan, Y.; Strumberg, D. *Biochim. Biophys. Acta* **1998**, *1400*, 83.
(11) Markman, M. *Oncologist* **2005**, *10*, 695.
(12) Ahmad, T.; Gore, M. *Expert Opin. Pharmacother.* **2004**, *5*, 2333.
(13) Stewart, D. J. *Oncologist* **2004**, *9* (Suppl. 6), 33.
(14) Ormrod, D.; Spencer, C. M. *Drugs* **1999**, *58*, 533.
(15) Fuchs, C.; Mitchell, E. P.; Hoff, P. M. *Cancer Treat. Rev.* **2006**, *32*, 491.
(16) Pavillard, V.; Charasson, V.; Laroche-Clary, A.; Soubeyran, I.; Robert, J. *Anticancer Res.* **2004**, *24*, 579.
(17) Li, Q.-Y.; Zu, Y.-G.; Shi, R.-Z.; Yao, L.-P. *Curr. Med. Chem.* **2006**, *13*, 2021.
(18) Legarza, K.; Yang, L.-X. *Anticancer Res.* **2006**, *26*, 3301.
(19) Hsiang, Y.-H.; Lihou, M. G.; Liu, L. F. *Cancer Res.* **1989**, *49*, 5077.
(20) Thomas, C. J.; Rahier, N. J.; Hecht, S. M. *Bioorg. Med. Chem.* **2004**, *12*, 1585.
(21) Wang, X.; Wang, L.-K.; Kingsbury, W. D.; Johnson, R. K.; Hecht, S. M. *Biochemistry* **1998**, *37*, 9399.
(22) Hertzberg, R. P.; Caranfa, M. J.; Hecht, S. M. *Biochemistry* **1989**, *28*, 4629.
(23) Miller, S. E.; Pilch, D. S. *Ann. N.Y. Acad. Sci.* **2000**, *922*, 309.
(24) Staker, B. L.; Hjerrild, K.; Feese, M. D.; Behnke, C. A.; Burgin, A. B., Jr.; Stewart, L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15387.
(25) Argaman, M.; Bendetz-Nezer, S.; Matlis, S.; Segal, S.; Priel, E. *Biochem. Biophys. Res. Commun.* **2003**, *301*, 789.
(26) Petrangolini, G.; Pratesi, G.; De Cesare, M.; Supino, R.; Pisano, C.; Marcellini, M.; Giordano, V.; Laccabue, D.; Lanzi, C.; Zunino, F. *Mol. Cancer Res.* **2003**, *1*, 863.
(27) Ji, Y.; Hayashi, K.; Amoh, Y.; Tsuji, K.; Yamauchi, K.; Yamamoto, N.; Tsuchiya, H.; Tomita, K.; Bouvet, M.; Hoffman, R. M. *Anticancer Res.* **2007**, *27*, 713.
(28) Liu, X.; Lin, J.; Bao, Y.; Lin, X.; An, L. *J. Dermatol. Sci.* **2006**, *42*, 262.
(29) Hayward, R. L.; Macpherson, J. S.; Cummings, J.; Monia, B. P.; Smyth, J. F.; Jodrell, D. I. *Clin. Cancer Res.* **2003**, *9*, 2856.
(30) Saleem, A.; Edwards, T. K.; Rasheed, Z.; Rubin, E. H. *Ann. N.Y. Acad. Sci.* **2000**, *922*, 46.
(31) Meng, L.-H.; Liao, Z.-Y.; Pommier, Y. *Curr. Top. Med. Chem.* **2003**, *3*, 305.
(32) Maliepaard, M.; van Gastelen, M. A.; de Jong, L. A.; Pluim, D.; van Waardenburg, R. C. A. M.; Ruevekamp-Helmers, M. C.; Floot, B. G. J.; Schellens, J. H. M. *Cancer Res.* **1999**, *59*, 4559.
(33) Chrencik, J. E.; Staker, B. L.; Burgin, A. B.; Pourquier, P.; Pommier, Y.; Stewart, L.; Redinbo, M. R. *J. Mol. Biol.* **2004**, *339*, 773.
(34) Lauria, A.; Ippolito, M.; Almerico, A. M. *J. Mol. Model.* **2007**, *13*, 393.
(35) Bocian, W.; Kawecki, R.; Bednarek, E.; Sitkowski, J.; Pietrzyk, A.; Williamson, M. P.; Hansen, P. E.; Kozerski, L. *Chem.—Eur. J.* **2004**, *10*, 5776.
(36) Oberlies, N. H.; Kroll, D. J. *J. Nat. Prod.* **2004**, *67*, 129.
(37) Driver, R. W.; Yang, L.-X. *Mini-Rev. Med. Chem.* **2005**, *5*, 425.
(38) Ulukan, H.; Swaan, P. W. *Drugs* **2002**, *62*, 2039.
(39) Garcia-Carbonero, R.; Supko, J. G. *Clin. Cancer Res.* **2002**, *8*, 641.
(40) Zunino, F.; Pratesi, G. *Expert Opin. Invest. Drugs* **2004**, *13*, 269.
(41) Hatefi, A.; Amsden, B. *Pharm. Res.* **2002**, *19*, 1389.
(42) Sriram, D.; Yogeeswari, P.; Thirumurugan, R.; Ratan Bal, T. *Nat. Prod. Res.* **2005**, *19*, 393.
(43) Adams, D. J. *Curr. Med. Chem. Anticancer Agents* **2005**, *5*, 1.
(44) Lorence, A.; Nessler, C. L. *Phytochemistry* **2004**, *65*, 2735.
(45) Kim, D.-K.; Lee, N. *Mini-Rev. Med. Chem.* **2002**, *2*, 611.
(46) Corbett, K. D.; Berger, J. M. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 95.
(47) Eriksson, L.; Jaworska, J.; Worth, A. P.; Cronin, M. T. D.; McDowell, R. M.; Gramatica, P. *Environ. Health Perspect.* **2003**, *111*, 1361.
(48) Verma, R. P.; Hansch, C. *Bioorg. Med. Chem.* **2007**, *15*, 2223.
(49) Wang, X.; Zhou, X.; Hecht, S. M. *Biochemistry* **1999**, *38*, 4374.
(50) Redinbo, M. R.; Stewart, L.; Kuhn, P.; Champoux, J. J.; Hol, W. G. J. *Science* **1998**, *279*, 1504.
(51) Bailly, C. *Crit. Rev. Oncol. Hematol.* **2003**, *45*, 91.
(52) Leu, Y.-L.; Roffler, S. R.; Chern, J.-W. *J. Med. Chem.* **1999**, *42*, 3623.
(53) Wani, M. C.; Nicholas, A. W.; Wall, M. E. *J. Med. Chem.* **1986**, *29*, 2358.
(54) Carrigan, S. W.; Fox, P. C.; Wall, M. E.; Wani, M. C.; Bowen, J. P. *J. Comput.- Aided Mol. Design* **1997**, *11*, 71.
(55) Wani, M. C.; Nicholas, A. W.; Manikumar, G.; Wall, M. E. *J. Med. Chem.* **1987**, *30*, 1774.
(56) Wall, M. E.; Wani, M. C.; Nicholas, A. W.; Manikumar, G.; Tele, C.; Moore, L.; Truesdale, A.; Leitner, P.; Besterman, J. M. *J. Med. Chem.* **1993**, *36*, 2689.
(57) Yaegashi, T.; Sawada, S.; Nagata, H.; Furuta, T.; Yokokura, T.; Miyasaka, T. *Chem. Pharm. Bull.* **1994**, *42*, 2518.
(58) Tanizawa, A.; Kohn, K. W.; Kohlhagen, G.; Leteurtre, F.; Pommier, Y. *Biochemistry* **1995**, *34*, 7200.
(59) Burke, T. G.; Mi, Z. *J. Med. Chem.* **1993**, *36*, 2580.
(60) Subrahmanyam, D.; Sarma, V. M.; Venkateswarlu, A.; Sastry, T. V. R. S.; Kulakarni, A. P.; Srinivasa Rao, D.; Krishna Reddy, K. V. S. R. *Bioorg. Med. Chem.* **1999**, *7*, 2013.
(61) Subrahmanyam, D.; Sarma, V. M.; Venkateswarlu, A.; Sastry, T. V. R. S.; Srinivas, A. S. S. V.; Krishna, C. V.; Deevi, D. S.; Kumar, S. A.; Babu, M. J.; Damodaran, N. K. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 369.
(62) Subrahmanyam, D.; Venkateswarlu, A.; Rao, K. V.; Sastry, T. V. R. S.; Vandana, G.; Kumar, S. A. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1633.
(63) Sugimori, M.; Ejima, A.; Ohsuki, S.; Matsumoto, K.; Kawato, Y.; Yasuoka, M.; Tagawa, H.; Terasawa, H. *Heterocycles* **1994**, *38*, 81.
(64) Crow, R. T.; Crothers, D. M. *J. Med. Chem.* **1992**, *35*, 4160.
(65) Nicholas, A. W.; Wani, M. C.; Manikumar, G.; Wall, M. E.; Kohn, K. W.; Pommier, Y. *J. Med. Chem.* **1990**, *33*, 972.
(66) Rahier, N. J.; Cheng, K.; Gao, R.; Eisenhauer, B. M.; Hecht, S. M. *Org. Lett.* **2005**, *7*, 835.
(67) Wani, M. C.; Nicholas, A. W.; Wall, M. E. *J. Med. Chem.* **1987**, *30*, 2317.
(68) Hertzberg, R. P.; Caranfa, M. J.; Holden, K. G.; Jakas, D. R.; Gallagher, G.; Mattern, M. R.; Mong, S.-M.; Bartus, J. O.; Johnson, R. K.; Kingsbury, W. D. *J. Med. Chem.* **1989**, *32*, 715.
(69) Mi, Z.; Burke, T. G. *Biochemistry* **1994**, *33*, 10325.
(70) Burke, T. G.; Munshi, C. B.; Mi, Z.; Jiang, Y. *J. Pharm. Sci.* **1995**, *84*, 518.
(71) Cao, Z.; Harris, N.; Kozielski, A.; Vardeman, D.; Stehlin, J. S.; Giovanella, B. *J. Med. Chem.* **1998**, *41*, 31.

(72) Cao, Z.; Pantazis, P.; Mendoza, J.; Early, J.; Kozielski, A.; Harris, N.; Vardeman, D.; Liehr, J.; Stehlin, J. S.; Giovanella, B *Ann. N.Y. Acad. Sci.* **2000**, *922*, 122.
(73) Wang, C.-Y.; Pan, X.-D.; Liu, H.-Y.; Fu, Z.-D.; Wei, X.-Y.; Yang, L.-X. *Bioorg. Med. Chem.* **2004**, *12*, 3657.
(74) Lavergne, O.; Lesueur-Ginot, L.; Rodas, F. P.; Bigg, D. C. H. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 2235.
(75) Lavergne, O.; Lesueur-Ginot, L.; Rodas, F. P.; Kasprzyk, P. G.; Pommier, J.; Demarquay, D.; Prevost, G.; Ulibarri, G.; Rolland, A.; Schiano-Liberatore, A.-M.; Harnett, J.; Pons, D.; Camara, J.; Bigg, D. C. H. *J. Med. Chem.* **1998**, *41*, 5410.
(76) Lavergne, O.; Harnett, J.; Rolland, A.; Lanco, C.; Lesueur-Ginot, L.; Demarquay, D.; Huchet, M.; Coulomb, H.; Bigg, D. C. H. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2599.
(77) Lavergne, O.; Demarquay, D.; Bailly, C.; Lanco, C.; Rolland, A.; Huchet, M.; Coulomb, H.; Muller, N.; Baroggi, N.; Camara, J.; Le Breton, C.; Manginot, E.; Cazaux, J.-B.; Bigg, D. C. H. *J. Med. Chem.* **2000**, *43*, 2285.
(78) Larsen, A. K.; Gilbert, C.; Chyzak, G.; Plisov, S. Y.; Naguibneva, I.; Lavergne, O.; Lesueur-Ginot, L.; Bigg, D. C. H. *Cancer Res.* **2001**, *61*, 2961.
(79) Lansiaux, A.; Facompré, M.; Wattez, N.; Hildebrand, M.-P.; Bal, C.; Demarquay, D.; Lavergne, O.; Bigg, D. C. H.; Bailly, C. *Mol. Pharmacol.* **2001**, *60*, 450.
(80) Kerrigan, J. E.; Pilch, D. S. *Biochemistry* **2001**, *40*, 9792.
(81) Laco, G. S.; Collins, J. R; Luke, B. T.; Kroth, H.; Sayer, J. M.; Jerina, D. M.; Pommier, Y. *Biochemistry* **2002**, *41*, 1428.
(82) Jena, N. R.; Mishra, P. C. *J. Mol. Model.* **2007**, *13*, 267.
(83) Hansch, C.; Maloney, P. P.; Fujita, T.; Muir, R. M. *Nature* **1962**, *194*, 178.
(84) Hansch, C.; Leo, A. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, D.C. 1995.
(85) Schultz, T. W.; Hewitt, M.; Netzeva, T. I.; Cronin, M. T. D. *QSAR Comb. Sci.* **2007**, *26*, 238.
(86) Selassie, C. D.; Garg, R.; Kapur, S.; Kurup, A.; Verma, R. P.; Mekapati, S. B.; Hansch, C. *Chem. Rev.* **2002**, *102*, 2585.
(87) Winkler, D. A. *Briefings Bioinf.* **2002**, *3*, 73.
(88) Hall, L. H.; Hall, L. M. *SAR QSAR Environ. Res.* **2005**, *16*, 13.
(89) Hou, T. J.; Zhang, W.; Xia, K.; Qiao, X. B.; Xu, X. J. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1585.
(90) Khan, M. T. H.; Sylte, I. *Curr. Drug Discovery Technol.* **2007**, *4*, 141.
(91) Verma, R. P. *Bioorg. Med. Chem.* **2005**, *13*, 1059.
(92) Hansch, C.; Verma, R. P. *Top. Heterocycl. Chem.* **2007**, *10*, 43.
(93) Hansch, C.; Verma, R. P. *ChemMedChem* **2007**, *2*, 1807.
(94) Lu, A.-J.; Zhang, Z.-S; Zheng, M.-Y.; Zou, H.-J.; Luo, X.-M; Jiang, H.-L. *Acta Pharmacol. Sin.* **2007**, *28*, 307.
(95) Amat, L.; Robert, D.; Besalú, E.; Carbó-Dorca, R. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 624.
(96) Li, M.-J.; Jiang, C.; Li, M.-Z.; You, T.-P. *J. Mol. Struct.: THEOCHEM* **2005**, *723*, 165.
(97) Yao, Y.-Y; Xu, L; Yang, Y.-Q.; Yuan, X.-S. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 590.
(98) Kier, L. B.; Hall, L. H. *J. Pharm. Sci.* **1981**, *70*, 583.
(99) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure-Activity Analysis*; Wiley: New York, 1986.
(100) Kier, L. B. *Quant. Struct.-Act. Relat.* **1985**, *4*, 109.
(101) Kier, L. B., Rouvray, D. H.; Eds. *Computational Chemical Graph Theory*; Nova Science Publishers: New York, 1990.
(102) Randic, M. *J. Am. Chem. Soc.* **1975**, *97*, 6609.
(103) Basak, S. C.; Harriss, D. K.; Magnuson, V. R. *J. Pharm. Sci.* **1984**, *73*, 429.
(104) Stanton, D. T.; Egolf, L. M.; Jurs, P. C.; Hicks, M. G. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 306.
(105) Osmialowski, K.; Halkiewicz, J.; Redecki, A.; Kaliszan, R. *J. Chromatogr.* **1985**, *346*, 53.
(106) Dallavalle, S.; Ferrari, A.; Biasotti, B.; Merlini, L.; Penco, S.; Gallo, G.; Marzi, M.; Tinti, M. O.; Martinelli, R.; Pisano, C.; Carminati, P.; Carenini, N.; Beretta, G.; Perego, P.; De Cesare, M.; Pratesi, G.; Zunino, F. *J. Med. Chem.* **2001**, *44*, 3264.
(107) Kim, D.-K.; Ryu, D. H.; Lee, J. Y.; Lee, N.; Kim, Y.-W.; Kim, J.-S.; Chang, K.; Im, G.-J.; Kim, T.-K.; Choi, W.-S. *J. Med. Chem.* **2001**, *44*, 1594.
(108) Verma, R. P.; Hansch, C. *Mol. Pharmaceutics* **2006**, *3*, 441.
(109) Fan, Y.; Shi, L. M.; Kohn, K. W.; Pommier, Y.; Weinstein, J. N. *J. Med. Chem.* **2001**, *44*, 3254.
(110) Yoon, K. J. P.; Krull, E. J.; Morton, C. L.; Bornmann, W. G.; Lee, R. E.; Potter, P. M.; Danks, M. K. *Mol. Cancer Ther.* **2003**, *2*, 1171.
(111) Aida-Hyugaji, S.; Nakagawa, H.; Nomura, J.; Sakurai, M.; Nagashima, U.; Ishikawa, T. *Croat. Chem. Acta.* **2007**, *80*, 277.
(112) *C-QSAR Program*; BioByte Corp.: 201 W. Fourth St., Suite 204, Claremont, CA 91711, www.biobyte.com.
(113) Hansch, C.; Hoekman, D.; Leo, A.; Weininger, D.; Selassie, C. D. *Chem. Rev.* **2002**, *102*, 783.
(114) Verma, R. P.; Hansch, C. *Nat. Protoc.* DOI: 10.1038/nprot.2007.125 (freely available at: http://www.natureprotocols.com/2007/03/05/development_of_qsar_models_usi_1.php).
(115) Verloop, A. In *The Sterimol Approach to Drug Design*; Marcel Dekker: New York, 1987.
(116) Hansch, C.; Steinmetz, W. E.; Leo, A. J.; Mekapati, S. B.; Kurup, A.; Hoekman, D. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 120.
(117) Hansch, C.; Kurup, A. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1647.
(118) Verma, R. P.; Kurup, A.; Hansch, C. *Bioorg. Med. Chem.* **2005**, *13*, 237.
(119) Verma, R. P.; Hansch, C. *Bioorg. Med. Chem.* **2005**, *13*, 2355.
(120) Abraham, M. H.; McGowan, J. C. *Chromatographia* **1987**, *23*, 243.
(121) Cramer, R. D., III; Bunce, J. D.; Patterson, D. E.; Frank, I. E. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18.
(122) Pogliani, L. *Chem. Rev.* **2000**, *100*, 3827.
(123) Pogliani, L. *J. Phys. Chem.* **1996**, *100*, 18065.
(124) Agrawal, V. K.; Singh, J.; Khadikar, P. V.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2044.
(125) Selassie, C. D.; Kapur, S.; Verma, R. P.; Rosario, M. *J. Med. Chem.* **2005**, *48*, 7234.
(126) Verma, R. P.; Hansch, C. *Virology* **2007**, *359*, 152.
(127) Vladu, B.; Woynarowski, J. M.; Manikumar, G.; Wani, M. C.; Wall, M. E.; von Hoff, D. D.; Wadkins, R. M. *Mol. Pharmacol.* **2000**, *57*, 243.
(128) Uehling, D. E.; Nanthakumar, S. S.; Croom, D.; Emerson, D. L.; Leitner, P. P.; Luzzio, M. J.; McIntyre, G.; Morton, B.; Profeta, S.; Sisco, J.; Sternbach, D. D.; Tong, W.-Q.; Vuong, A.; Besterman, J. M. *J. Med. Chem.* **1995**, *38*, 1106.
(129) Rahier, N. J; Eisenhauer, B. M.; Gao, R.; Jones, S. H.; Hecht, S. M. *Org. Lett.* **2004**, *6*, 321.
(130) Zu, Y.-G.; Li, Q.-Y.; Fu, Y.-J.; Wang, W. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4023.
(131) Dallavalle, S.; Ferrari, A.; Merlini, L.; Penco, S.; Carenini, N.; De Cesare, M.; Perego, P.; Pratesi, G.; Zunino, F. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 291.
(132) Verma, R. P. *Lett. Drug Des. Discovery* **2005**, *2*, 205.
(133) Mekapati, S. B.; Kurup, A.; Verma, R. P.; Hansch, C. *Bioorg. Med. Chem.* **2005**, *13*, 3737.
(134) Wang, X.; Wang, L.-K.; Kingsbury, W. D.; Johnson, R. K.; Hecht, S. M. *Biochemistry* **1998**, *37*, 9399.
(135) Adams, D. J.; da Silva, M. W.; Flowers, J. L.; Kohlhagen, G.; Pommier, Y.; Colvin, O. M.; Manikumar, G.; Wani, M. C. *Cancer Chemother. Pharmacol.* **2006**, *57*, 135.
(136) Dallavalle, S.; Delsoldato, T.; Ferrari, A.; Merlini, L.; Penco, S.; Carenini, N.; Perego, P.; De Cesare, M.; Pratesi, G.; Zunino, F. *J. Med. Chem.* **2000**, *43*, 3963.
(137) Yoshikawa, M.; Ikegami, Y.; Hayasaka, S.; Ishii, K.; Ito, A.; Sano, K.; Suzuki, T.; Togawa, T.; Yoshida, H.; Soda, H.; Oka, M.; Kohno, S.; Sawada, S.; Ishikawa, T.; Tanabe, S. *Int. J. Cancer* **2004**, *110*, 921.
(138) Jew, S.-S.; Kim, H.-J.; Kim, M. G.; Roh, E.-Y.; Hong, C. II; Kim, J.-K.; Lee, J.-H.; Lee, H.; Park, H.-G. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 3203.
(139) Bodley, A. L.; Shapiro, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3726.
(140) Bodley, A. L.; Wani, M. C.; Wall, M. E.; Shapiro, T. A. *Biochem. Pharmacol.* **1995**, *50*, 937.
(141) Werbovetz, K. A.; Bhattacharjee, A, K.; Brendle, J. J.; Scovill, J. P. *Bioorg. Med. Chem.* **2000**, *8*, 1741.
(142) Jaworska, J. S.; Comber, M.; Auer, C.; Van Leeuwen, C. J. *Environ. Health Perspect.* **2003**, *111*, 1358.
(143) Cronin, M. T. D.; Walker, J. D.; Jaworska, J. S.; Comber, M. H. I.; Watts, C. D.; Worth, A. P. *Environ. Health Perspect.* **2003**, *111*, 1376.
(144) Cronin, M. T. D.; Jaworska, J. S.; Walker, J. D.; Comber, M. H. I.; Watts, C. D.; Worth, A. P. *Environ. Health Perspect.* **2003**, *111*, 1391.
(145) Golbraikh, A.; Tropsha, A. *J. Mol. Graphics Modell.* **2002**, *20*, 269.
(146) Tropsha, A.; Gramatica, P.; Gombar, V. K. *QSAR Comb. Sci.* **2003**, *22*, 69.
(147) Garg, R.; Kurup, A.; Mekapati, S. B.; Hansch, C. *Chem. Rev.* **2003**, *103*, 703.
(148) Verma, R. P.; Hansch, C. *Bioorg. Med. Chem.* **2005**, *13*, 4597.
(149) Polanski, J.; Bak, A.; Gieleciak, R.; Magdziarz, T. *J. Chem. Inf. Model.* **2006**, *46*, 2310.
(150) Ekins, S.; Mestres, J.; Testa, B. *Br. J. Pharmacol.* **2007**, *152*, 9.
(151) Franke, R.; Gruska, A. *Quantitative Structure-Activity Relationhsip (QSAR) Models of Mutagens and Carcinogens*; CRC Press: Boca Raton, 2003; Chapter 1, p 1.

CR0780210