THESIS APPROVAL

GRADUATE SCHOOL, KASETSART UNIVERSITY

Doctor of Philosophy (Chemistry)
DEGREE

Chemistry
FIELD

Chemistry
DEPARTMENT

TITLE: Theoretical Investigation on the Binding of DNA-Topoisomerase I Complex With Lamellarin Derivatives and Molecular Design

NAME: Miss Poonsiri Thipnate

THIS THESIS HAS BEEN ACCEPTED BY

THESIS ADVISOR
(Associate Professor Supa Hannongbua, Dr.rer.nat.)

COMMITTEE MEMBER
(Associate Professor Supanna Techasakul, Ph.D.)

COMMITTEE MEMBER
(Mr. Chak Sangma, Ph.D.)

COMMITTEE MEMBER
(Professor Peter Wolschann, Dr.rer.nat.)

COMMITTEE MEMBER
(Mr. Poonsakdi Ploypradith, Ph.D.)

DEPARTMENT HEAD
(Assistant Professor Noojaree Prasitpan, Ph.D.)

APPROVED BY THE GRADUATE SCHOOL ON .....

DEAN
(Associate Professor Gunjana Theeragool, D.Agr.)

THESIS

THEORETICAL INVESTIGATION ON THE BINDING OF DNA-TOPOISOMERASE I COMPLEX WITH LAMELLARIN DERIVATIVES AND MOLECULAR DESIGN

POONSIRI THIPNATE

A Thesis Submitted in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy (Chemistry)
Graduate School, Kasetsart University
2009

Poonsiri Thipnate 2009: Theoretical Investigation on the Binding of DNA-Topoisomerase I Complex With Lamellarin Derivatives and Molecular Design. Doctor of Philosophy (Chemistry), Major Field: Chemistry, Department of Chemistry. Thesis Advisor: Associate Professor Supa Hannongbua, Dr.rer.nat. 104 pages.

Currently, cancer is one of leading causes of death in the world. It continued to be a serious problem because the number of death from cancer increases every year. Hence, this work focuses on anticancer compounds, especially from marine source, so called lamellarins. Molecular docking methods were applied to investigate the orientation and the binding energies of lamellarins in the binding pocket of a possible known target, topoisomerase I-DNA complex. Several lamellarins can intercalate at the site of DNA cleavage, forming base-stacking interactions with both the upstream and downstream base pairs. The hydrogen-bond interactions occurred with amino acid residues of topoisomerase I such as Asn352, Glu356, Arg364, and Asn722. The different interactions between the lamellarins containing a saturated D-ring and those with a C5-C6 double bond and common structural requirements for their cytotoxic activities against human breast cancer cell lines were determined using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) techniques. The best CoMFA and CoMSIA models for both cell lines yielded satisfactory predictive ability with r^2_{cv} values in the range of 0.659-0.728. Moreover, 4D-QSAR and 4D-fingerprint virtual screening models were built and investigated for the cytotoxicity of lamellarins against human hormone-dependent T47D breast cancer cells. 4D-QSAR models were first constructed from the exploration of eight possible receptor binding alignments for the entire training set. Since the training set is small (25 compounds), the generality of the 4D-QSAR paradigm was then exploited to devise a strategy to maximize the extraction of binding information from the training set, and to also permit virtual screening of diverse lamellarin chemistry. 4D-QSAR models were sought for only six of the most potent lamellarins of the training set as well as another subset composed of lamellarins with constrained ranges in molecular weight and lipophilicity. Overall, it was found that formation of an intermolecular hydrogen bond and hydrophobic interactions for substituents on the E ring modulate most of the cytotoxicity against T47D breast cancer cells. Hydrophobic substitutions on the F-ring can also enhance cytotoxic potency. The 4D-fingerprint QSAR model was constructed using absolute molecular similarity. This 4D-fingerprint virtual high throughput screen permits a larger range of chemistry diversity to be assayed than the 4D-QSAR models. The optimized 4D-QSAR 3D-pharmacophore model has $r^2_{cv} = 0.947$, while the optimized 4D-fingerprint virtual screening model has $r^2_{cv} = 0.719$. This work reveals that it is possible to develop significant 3D-QSAR, 4D-QSAR and virtual screening models for a small set of lamellarins showing cytotoxic behavior in breast cancer screens that can guide future drug development based upon lamellarins.

/ /

Student's signature Thesis Advisor's signature

ACKNOWLEDGEMENT

First of all, I would like to express my sincere gratitude to my supervisor, Associate Professor Dr. Supa Hannongbua. Her wide vision and knowledge in Science have been the great valuable experience for me. Her support, understanding, and encouragement have brought me many grateful opportunities in my life.

I would like to express my appreciation to my co-supervisors, Dr. Poonsakdi Ploypradith, Dr. Nopporn Thasana, and Dr. Montakarn Chittchang from Chulabhorn Research Institute and Chulabhorn Graduate Institute for their valuable advice, constructive comments, and their time throughout this work. I am extremely grateful for Professor Peter Wolschann, Professor Anton J. Hopfinger and his postdoctoral student, Dr. Jianzhong Liu, Associate Professor Dr. Supanna Techasakul, Dr. Chak Sangma, and Assistant Professor Teerasak Veerapaspong for their time, valuable comments, and suggestion. Many thanks for all my friends and my family who are sharing my hard time and my happiness. Especially, I would like to thank Ms. Sirintip Hanpradit (Rose), who brings me a wonderful time and inspiration, valuable experience, and new lovely friends for me.

This work would not have been possible without the generous supports from the Phetchaburi Rajabhat University (PBRU), the Faculty of Science, Kasetsart University, the Kasetsart University (KU) Graduate School, the Commission on Higher Education (CHE), the Postgraduate Education and Research Program in Petroleum, Petrochemical Technology and Advanced Materials, KU Research, Development Institute (KURDI), Laboratory for Computational and Applied Chemistry (LCAC) and the computing centre of KU, the Center of Nanotechnology KU, NANOTEC Center of Excellence at KU, and the Large Scale Research Laboratory of the National Electronics and Computer Technology (NECTEC).

Poonsiri Thipnate

February, 2009

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF ABBREVIATION

HBV	=	Hepatitis B virus
HPV	=	Human papilloma virus
HIV	=	Human immunodeficiency virus
DNA	=	Deoxyribonucleic acids
MDR	=	Multidrug-resistance
AutoDock	=	Automated Docking
Gold	=	Genetic Optimisation for Ligand Docking
PLS	=	Partial least squares
CoMFA	=	Comparative molecular field analysis
CoMSIA	=	Comparative molecular similarity indices analysis
3D-QSAR	=	Three-dimensional quantitative structure-activity relationship
4D-QSAR	=	Four-dimensional quantitative structure-activity relationship
P-GP	=	P-glycoprotein
CPT	=	Camptothecin
CDK	=	Cyclin-dependent kinases
GSK	=	Glycogen synthase kinase
Glu	=	Glutamic acid
Asn	=	Asparagine
Arg	=	Arginine
Asp	=	Aspartic acid
SAR	=	Structure-activity relationship
RI	=	Receptor-independent
RD	=	Receptor-dependent
4D-MS	=	Four-dimensional molecular similarity analysis
CHARMM	=	Chemistry at HARvard Molecular Mechanics
r^2_{cv}	=	Predictive ability of cross-validation

LIST OF ABBREVIATION (Continued)

RMSD	=	Root mean square deviation
PRESS	=	Prediction error sum of squares
SSY	=	Variance of the data around the mean value
S_{PRESS}	=	The uncertainty of the prediction
r^2	=	Conventional correlation coefficient
IPEs	=	Interaction pharmacophore elements
MDS	=	Molecular dynamic simulation
CEP	=	Conformational ensemble profile
GCODs	=	Grid cell occupancy descriptors
log P	=	The logarithm of the 1-octanol/water partition coefficient
MR	=	Molar refractivity
GFA	=	Genetic function approximation
MDDM	=	Main distance-dependent matrix
RMSD	=	Root mean square deviation
$G_{binding}$	=	Estimated free energies of binding
Lys	=	Lysine
Thr	=	Threonine
sp^3 C(+1)	=	Carbon sp3-hybridization with plus 1 charge probe atom
sp^3 O(-1)	=	Oxygen sp3-hybridization with minus 1 charge probe atom
H(+1)	=	Hydrogen with plus 1 charge probe atom
St	=	Steric field
El	=	Electrostatic field
Hyd	=	Hydrophobic field
Hd	=	H-bond donor field
Ha	=	H-bond acceptor field

THEORETICAL INVESTIGATION ON THE BINDING OF DNA-TOPOISOMERASE I COMPLEX WITH LAMELLARIN DERIVATIVES AND MOLECULAR DESIGN

INTRODUCTION

Nowadays, cancer is one of the leading causes of death in the world. Cancer diagnosed around the world is expected to hit 12 million persons in 2008. Furthermore, it will overtake heart disease to be the leading cause of death in the world by 2010 (Anonymous, 2008). It is a group of diseases characterized by uncontrolled growth and spread of abnormal cells. It is caused by both external factors such as tobacco, chemicals, radiation, and infection organisms such as hepatitis B virus (HBV), human papilloma virus (HPV), human immunodeficiency virus (HIV) and internal factors such as inherited mutations, hormones, immune conditions, and mutation from metabolism (American Cancer Society, 2007).

In the battle against cancer, there are many methods to treat cancer, for example, radiation therapy, surgery, biological therapy, gene therapy, and targeted cancer therapies. One of the main treatments is chemotherapy. It refers to the treatment of cancer by chemicals or anticancer drugs. The chemotherapy drugs in current use primarily target DNA. Several aspects of DNA are targeted, including the structure of individual nucleotides, the integrity of nucleotides or their bases within DNA, the main enzymes active in the synthesis phase (DNA polymerase and topoisomerases, which are active in DNA replication and in DNA unwinding, respectively), and the structures and enzymes active in the mitosis phase (Boik, 2001).

In anticancer drug development, around 60% of currently used anticancer agents are derived from natural sources, including plants, marine organisms and micro-organisms (Cragg and Newman, 2006). This work focuses on anticancer compounds from marine source, especially lamellarins which are polyaromatic pyrrole alkaloids that have been isolated from different sources, such as ascidians,

molluscs, and sponges (Andersen *et al.*, 1985; Lindquist *et al.*, 1988; Carroll *et al.*, 1993; Urban *et al.*, 1994; Urban *et al.*, 1995; Urban and Capon, 1996; Reddy *et al.*, 1997; Davis *et al.*, 1999; Ham and Kang, 2002; Krishnaiah *et al.*, 2004; Reddy *et al.*, 2005). A family of lamellarins consists of three structural groups, including unsaturated D-ring fused, saturated D-ring fused, and unfused central pyrrole ring group. To date, more than 30 lamellarins have been isolated, many of which exhibit interesting biological activities. For example, lamellarin α 20-sulfate is a potential candidate for HIV treatments as it can inhibit HIV-1 integrase both *in vitro* and *in vivo* (Reddy *et al.*, 1999; Ridley *et al.*, 2002). On the other hand, non-sulfated lamellarins appear to possess potent cytotoxic activity against cancer cells (Fan *et al.*, 2008), especially lamellarin D which has been receiving a lot more attention than all the other compounds in this series. Their cancer cytotoxicities have been attributed to the finding that lamellarin D is an effective stabilizer of human topoisomerase I-DNA covalent complexes, and thus, capable of stimulating DNA cleavage (Facompre *et al.*, 2003; Vanhuyse *et al.*, 2005). Moreover, lamellarin D has also been demonstrated to induce apoptosis, as well as disrupt the inner transmembrane potential of mitochondria, which is a novel pharmacological target for anticancer chemotherapy (Costantini *et al.*, 2000; Debatin *et al.*, 2002; Dias and Bailly, 2005; Kluza *et al.*, 2006). Additionally, nontoxic doses of some lamellarins, especially lamellarin I, can show also reverse multidrug-resistance (MDR) of cancer cells by inhibiting P-glycoprotein (P-gp)-mediated drug efflux with a 9-16 times higher MDR modulating potency than that of verapamil (Quesada *et al.*, 1996). More recently, some lamellarins inhibit several protein kinases relevant to cancer such as cyclin-dependent kinases, dual specificity tyrosine phosphorylation activated kinase 1A, casein kinase 1, glycogen synthase kinase-3 and PIM-1. These results show good correlation observed between the effects of lamellarins on protein kinases and their action on cell. Inhibition of specific kinases may contribute to the cytotoxicity of lamellarins (Baunbaek *et al.*, 2008).

Generally, lamellarins with a C5-C6 double bond are considered to be more cytotoxic than those bearing a C5-C6 single bond. However, some exceptions have been observed with lamellarin M, which is roughly equally cytotoxic to its saturated

counterpart, lamellarin K, whereas lamellarin M-triacetate is actually much less cytotoxic than its corresponding analogue in the saturated series, lamellarin K-triacetate (Quesada *et al.*, 1996; Bailly, 2004). Based on the structure-activity relationships (SAR) analysis of lamellarins with their cytotoxicities to cancer cell lines for the substituents on the lamellarin core, it appears that the hydroxyl groups at the C8 and C20 positions of lamellarin D are important structural requirements, while the hydroxyl group at C14 and the two methoxy groups at C13 and C21 are not necessary for the cytotoxic activities (Ishibashi *et al.*, 2002; Marco *et al.*, 2005). Employing a systematic approach on a large number of lamellarins and cancer cell lines, Chittchang *et al.* recent findings not only substantiate the significant contributions of the C5-C6 olefin moiety, as well as the hydroxyl groups at C8 and C20, but also demonstrate the importance of the C7-hydroxyl group for the first time (Chittchang *et al.*, 2009).

Computational chemistry has been used to verify and explain these experimental findings. Molecular models of the ternary complex formed between the topoisomerase I-DNA with a lamellarin D molecule fully intercalating into the DNA duplex have been created with the AutoDock 3.0 docking program and studied using nanosecond molecular dynamics simulations in aqueous solutions (Facompre *et al.*, 2003; Marco *et al.*, 2005). The results confirm that the C8 and C20 hydroxyl groups on the lamellarin core are the major determinants for both the topoisomerase I cleavable complex stabilization and the cytotoxic action. In contrast, there is no clear explanation for the functional role of the C5-C6 double bond, which when present in the quinoline D-ring, tends to make the compounds more cytotoxic than those having a C5-C6 single bond. Due to investigate the orientation and binding energies of all lamellarins in the binding pocket of topoisomerase I-DNA, molecular docking methods (AutoDock and Gold) were employed in this work. Moreover, quantum calculations were performed to investigate the particular interactions of unsaturated and saturated D-ring lamellarins.

In previous studies, 26 natural and non-natural lamellarins were synthesized using Ploypradith *et al.* developed synthetic routes (Ploypradith *et al.*, 2006) and

tested for their cytotoxicities against 11 cancer cell lines (Chittchang *et al.*, 2009). The prominent selectivity observed with certain lamellarins towards human hormone-dependent T47D and hormone-independent MDA-MB-231 breast cancer cells have prompted me to further investigate the possible underlying reasons. With the inclusion of two unnatural compounds (i.e., dehydrolamellarins J and Y), a total of eleven pairs of lamellarins, each of which only differ in the presence of either a saturated or an unsaturated D-ring, were used to explore their potential differences in steric, electrostatic, hydrophobic, and hydrogen bond interactions in detail. The 3D-QSAR methods were employed as they are important tools for understanding the mechanism of the interactions between ligands and an unknown receptor. In order to use these methods, the physicochemical properties of 26 lamellarin molecules were first represented in the form of molecular fields, which could then be effectively correlated with their cytotoxic activities using partial least squares (PLS) regression analysis (Cramer *et al.*, 1988; Bohm *et al.*, 1999; Klebe *et al.*, 1994). Additionally, comparative molecular field analysis (CoMFA) (Cramer *et al.*, 1988) and comparative molecular similarity indices analysis (CoMSIA) (Klebe *et al.*, 1994; Bohm *et al.*, 1999) techniques were applied to examine the different interactions between the lamellarins containing a saturated D-ring and those with a C5-C6 double bond. The common structural requirements for their cytotoxic activities against both human breast cancer cell lines were also determined.

From three conventional problems of a receptor independent of lamellarins such as the identification of the active comformations or molecular shapes of flexible compounds like lamellarins, the specification of molecular alignment, and partitioning with respect to intermolecular interaction or interacting pharmacophore of lamellarin molecules, the four dimension-quantitative structure activity relationship (4D-QSAR) analysis were applied to the lamellarin data set. The extra dimension of 4D-QSAR obtains from ensemble sampling (Hopfinger *et al.*, 1997). This analysis is a molecular modeling method which allows the construction of optimized dynamic spatial QSAR models, in the form of 3D pharmacophores, which are dependent on conformation, alignment, and pharmacophore-grouping (Hopfinger *et al.*, 1999). It has been proved to be a useful method for the construction of quantitative 3D pharmacophore models

of flexible ligand analogues when the geometry of the corresponding receptor is not known (Hopfinger *et al.*, 1997; Hopfinger *et al.*, 1999; Krasowski *et al.*, 2002). In order to develop and implement lamellarin QSAR models to provide more insight and design criterion relative to improved anti-breast cancer compounds, 4D-QSAR not only was first performed to the set of 25 lamellarins but the 4D-fingerprints were also generated as universal descriptors to build descriptive QSAR models in this study. The 4D-fingerprint are derived independence of any molecular alignment. It generates 4D-fingerprints which were performed by the 4D molecular similarity for each compound (Senese *et al.*, 2004). It was successfully applied as universal descriptors in several training sets for example the glucose analogs, the flavonoid analogs, the propofol analogs, the AHPBA and THP inhibitors of HIV-1 protease (Senese *et al.*, 2004), human serum albumin (HSA) data set (Liu *et al.*, 2005, 2006), local lymph node assay (LLNA) data base (Li *et al*., 2007a, 2007b), skin penetration enhancer data set (Iyer *et al.*, 2007), and HIV-1 integrase inhibitors (Iyer and Hopfinger, 2007).

OBJECTIVES

1. To investigate the orientation and binding energies of all lamellarins in the binding pocket of topoisomerase I-DNA by using molecular docking methods.

2. To examine the different interactions between the lamellarins containing a saturated D-ring and those with a C5-C6 double bond. The common structural requirements for their cytotoxic activities against human breast cancer cell lines were also determined by using 3D-QSAR analysis.

3. To develop and implement lamellarin QSAR models to provide more insight and design criterion relative to improve anti-breast cancer compounds by using 4D-QSAR and 4D-fingerprints analysis.

LITERATURE REVIEW

Since 1985, the first four lamellarins (A-D) were isolated and from the marine prosobranch mollusk, *Lamellaria* sp. and their structures were determined by an X-ray crystallographic and ^{1}H-NMR study (Andersen *et al.*, 1985). A family of more than 30 lamellarins was isolated from several diverse marine, including ascidians, molluscs, and sponges as shown in Table 1 (Fan *et al.*, 2008). These molecules consist of three structural groups such as unsaturated D-ring fused (Figure 1 (a)), saturated D-ring fused (Figure 1 (b)), and unfused central pyrrole ring group (Figure 1 (c)). Currently, many of them exhibit interesting biological activities against cancer cells, especially fused central pyrrole ring lamellarins (Bailly, 2004; Fan *et al.*, 2008). They have been found to be cytotoxic to a wide range of cencer cell lines such as lamellarins C and U demonstrating potent cytotoxicity against 10 human tumor cell lines with IC_{50} from 0.4-19.4 nM (Fan *et al.*, 2008), lamellarin D has potent cytotoxic activity against human prostate cancer cells (DU-145, LNCaP) and leukemia cells (K562) (Fan *et al.*, 2008), and lamellarins I, K, and L exhibited significant cytotoxicity against P388 and A549 cultured cancer cell lines (Carroll *et al.*, 1993). Additionally, lamellarin K derivative showed selectivity against the A549 human lung cancer cell line with IC_{50} around 7.6 nM (Fan *et al.*, 2008; Quesada *et al.*, 1996) and lamellarin N showed some selectivity toward both melanoma cell lines (SK-MEL-5 and UACC-62) (Reddy *et al.*, 1997). Lamellarin ζ, χ, L triacetate, and F have shown excellent activities against colorectal cancer cells (COLO-205) (Reddy *et al.*, 2005) whereas lamellarin H and α demonstrated cytotoxicity toward HeLA cells (Reddy *et al.*, 1999; Ridley, *et al.*, 2002).

By Bailly's group, lamellarin D was identified as a novel lead candidate (Facompre′ *et al.*, 2003; Bailly, 2004; Tardy *et al.*, 2004; Dias *et al.*, 2005; Marco *et al.*, 2005; Kluza *et al.*, 2006; Gallego *et al.*, 2008) because it has an effective cytotoxic activity against multidrug resistance reversal (MDR) cell lines by inhibiting P-glycoprotein (P-GP) or the multidrug transporter. It is an integral plasma membrane protein which contains 12 putative transmembrane regions and two ATP binding sites. It confers multidrug resistance by functioning as an energy-dependent drug efflux

pump that is utilized in normal cells, as well as in cancer cells, for their detoxification (Gottesman and Pastan, 1993; Licht *et al.*, 1996; Quesada *et al.*, 1996; Vanhuyse *et al.*, 2005). In 2006, Kluza *et al.* has demonstrated that lamellarin D acts on cancer cell mitochondria which are cellular organelles that perform pivotal functions essential for ATP production, homeostasis, and metabolism (Kluza *et al*., 2006; Gallego *et al.*, 2008; Toogood, 2008). There are two major alterations in mitochondria of cancer cells. First, they are often relatively resistant to the induction of mitochondrial membrane permeabilization (MMP), which is the rate-limiting step of the intrinsic pathway of apoptosis. Second, cancer cell mitochondria often exhibit a reduced oxidative phosphorylation. It means that ATP is generated through the conversion of glucose to pyruvate and excess pyruvate is then eliminated as the waste product lactate (Rustin and Kroemer, 2007). Lamellarin D has a direct mitochondrial permeability transition (MPT)-inducing effect to induce apoptosis of cancer cells (Kluza *et al.*, 2006).

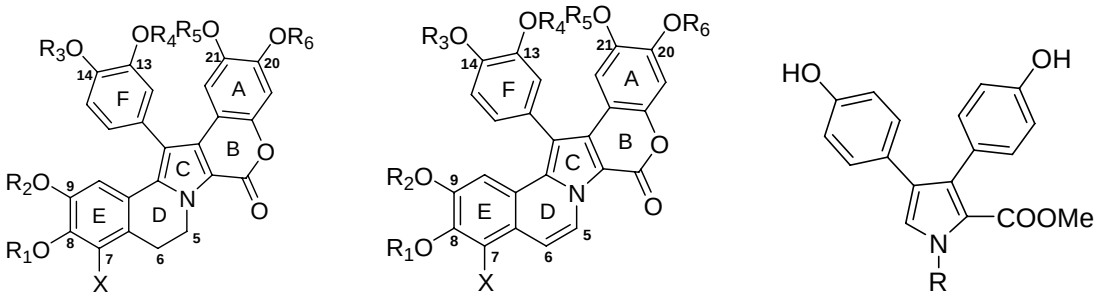


Figure 1 Three structural classification groups of lamellarins.

Moreover, Facompre′ *et al.* reported in 2003 that lamellarin D is a weak DNA binder that intercalates between base pair of the double helix. This result was supported by DNA binding measurements using absorbance, fluorescence, and electric linear dichroism spectroscopy. In the DNA relaxation assay, lamellarin D strongly promoted the conversion of supercoiled DNA into nicked DNA in the presence of topoisomerase I. It was 5 times less efficient than camptothecin (CPT), anticancer drug isolated and identified from the bark of the Chinese yew tree, *Camptotheca acuminate* (Thomas *et al.*, 2004), at stabilizing topoisomerase I–DNA

complexes. Interestingly, the two drugs exhibited slightly distinct sequence specificity profiles. Topoisomerase I-mediated DNA cleavage in the presence of lamellarin D occurred at some sites common to CPT, but a few specific sites identified with CPT but not with lamellarin D or conversely unique sites cleaved by lamellarin D but not by CPT were detected. This suggests that lamellarin D and CPT interact differently with the topoisomerase I-DNA interface (Facompre′ *et al.*, 2003). In this report, it not only showed the experimental results but also showed the molecular modeling results. A molecular modeling analysis provides structural information on the orientation of lamellarin D within the topoisomerase I–DNA covalent complex. The 2.10 Å resolution crystal structure of human topoisomerase I covalently linked to double-stranded DNA and bound to topotecan, Protein Data Bank entry 1k4t (Staker *et al.*, 2002), was used as reference structure to model the lamellarin D-mediated stabilization of topoisomerase I–DNA complex. According to this model, the phenol ring or F-ring (Figure 1 (b)) is pointing out to the protein–DNA interface and does not have any direct interaction with the protein. It may serve to help other protein partners. The C8 and C20 hydroxyl groups are at a hydrogen bond distance from the Asn722 and Glu356 residues of the enzyme, respectively while the keto group is interacting with the Arg364 residue. This result is in well agreement with the known structure–activity relationships of lamellarin D series (Ishibashi *et al.*, 2002). In 2008, Baunbaek *et al*. have been reported that lamellarins can inhibit several protein kinases of cancer cell including two cyclin-dependent kinases (CDK) CDK1/cyclin B and CDK5/p25, glycogen synthase kinase-3α/β (GSK 3α/β), up-regulated in prostate cancers (PIM-1), dualspecificity tyrosine phosphorylation activated kinase 1A (DYRK1A) and casein kinase 1 (CK1). Such a good correlation between the effects of lamellarins on protein kinases and their action on cell death indicated that inhibition of specific kinases may support the cytotoxicity of lamellarins (Baunbaek *et al.*, 2008).

The further computational analysis of lamellarin D is molecular dynamic simulation. In 2005, molecular models of the ternary complexes formed between the Topoisomerase I-DNA ensemble and lamellarin D or CPT have been generated and investigated by means of nanosecond molecular dynamics simulations in aqueous

solution. This results show that the 20-OH and 8-OH of lamellarin D can participate in H-bonding interactions with Glu356 and Asn722, respectively. The latter result suggested that leukemia CEM/C2 cells, leukemia cells which are resistant to CPT, are cross-resistant to lamellarin D. This model also reported that lamellarin D stabilizes topoisomerase I cleavage at CG base pair sites in addition to the TG base pair sites observed for CPT and rationalize SAR within the series. Moreover, the effect of replacing the 20-OH in lamellarin D with hydrogen was confirmed by using thermodynamic integration free energy simulations (Marco *et al.*, 2005).

The first reported study on SAR of lamellarins was done by Ishibashi *et al.* in 2002. To examine SAR, ten derivatives of lamellarin D were synthesized and evaluated the cytotoxicity against a human cervical carcinoma HeLa cell line in this work. It revealed that 20-OH and 8-OH of lamellarin D were important structural requirements for its cytotoxic activity, whereas the 14-OH and the two methoxy groups at C13 and C21 abolished the cytotoxicity of lamellarin (Ishibashi *et al.*, 2002). Second SAR report concerned about relationships of lamellarin D and lamellarin 501 (its unsaturated counterpart lamellarins D which was isolated from Indian ascidian *Didemnum obscurum* and named later as lamellarin χ (Reddy *et al.*, 2005)) series with cytotoxic activity against prostate (DU 145 and LN-CaP), ovarian (IGROV and IGROV-ET resistant to ecteinascidin-743) and colon (LoVo and LoVo-Dox cells resistant to doxorubicin) cancer cells. These series consisted of two groups of compounds incorporating various substituents on the three phenolic OH at positions C8, C14 and C20 of lamellarins such as non-amino compounds as group A and incorporating amino acid residues (alanine, leucine, valine, proline, phenylalanine, and tryptophan) as group B. In conclusion, this SAR study indicates that lamellarins D derivatives represent a novel family of topoisomerase I targeted antitumor agents. The C5–C6 double bond in the quinoline D-ring of lamellarins D is absolutely required to maintain an activity against topoisomerase I and a potent cytotoxic action. The three phenolic OH at positions C8, C14 and C20 are also important structural elements of the lamellarins D structure but they can be substituted with amino acid derivatives without loss of activity. The cationic proline and valine derivatives have been selected for a preclinial development to evaluate their antitumor

activity in *vivo* (Tardy *et al.*, 2004). In third report, a library of open lactone analogues of lamellarins D were prepared and tested in a panel of three human tumor cell lines (breast cancer MDA-MB-231, lung cancer A549, and colon cancer HT-29) to evaluate their cytotoxic potential. From these data, the SAR study concluded that more than 75% of the open-chain lamellarins D analogues tested showed cytotoxicity in a low micromolar GI_{50}, the drug concentration that caused 50% of cell growth inhibition after 72 hours of continuous exposure to the test molecule, range. Molecular simplification of lamellarins D by removing the lactone ring from all of the analogues and by the additional elimination of one aryl group derivatives produced a decrease in activity with respect to lamellarins D. Consequently, the oxidized derivatives showed greater activity than the corresponding reduced analogues (Pla *et al.*, 2006). Fourth SAR report was done by Chittchang *et al.* in 2009. To examine the importance of the C5-C6 double bond in the D-ring and the substitution pattern on the periphery of the lamellarin core, twenty-two naturally-occurring and three unnatural lamellarins were synthesized by Ploypradith *et al.* method (Ploypradith *et al.*, 2006) and evaluated for their cytotoxicities against eleven different cancer cells, including human non-small cell lung carcinoma A549, human multidrug-resistant, small cell lung carcinoma H69AR, human cervical adenocarcinoma HeLA, human hepatocellular carcinoma HepG2, human promyelocytic leukemia HL-60, human cholangiocarcinoma HuCCA-1, human oral epidermoid carcinoma KB, human hormone-independent breast cancer 231 MDA-MB-231, human fetal/embryonic lung fibroblast MRC-5, mouse lymphoid neoplasm P388, human hepatocellular carcinoma S102, human hormone-dependent breast cancer T47D. From cytotoxicity results, lamellarins D, X, ε, M, N, and dehydrolamellarin J are significantly more potent than the other lamellarins. Based on SAR results, there are four important structural elements that determine the cytotoxic activities of lamellarins towards cancer cells. These include the C5–C6 double bond, as well as the C7-, C8-, and C20-OH groups as shown in Figure 2. Especially, the importance of substituent at C7 of lamellarins has not been reported previously. A substitution of the hydrogen atom at this position with a hydroxyl group significantly increases the cytotoxicities of lamellarins with a C5–C6 double bond. On the other hand, methoxylation at this position may affect only slightly the cytotoxic activities of these compounds. Interestingly, the effect is

significantly more pronounced when the C7-OH is replaced by a methoxy group. This clearly decreases the cytotoxic activities of the lamellarins, especially those containing a C5–C6 double bond. These results indicate that the hydroxyl group at this position is an important structural element for cytotoxic activity of lamellarins (Chittchang *et al.*, 2009).

Table 1 Lamellarins and their marine sources.

Lamellarins	Marine sources
A-D	Prosobranch molluse *Lamellaria* sp.
C, U	Prosobranch molluse *Coriocella hibyae*
E-H	Marine ascidian *Didemnum chartaceum*
I, J-M, N triacetate, A-C, D triacetate	Marine ascidian *Didemnum* sp.
O, P	Southern Australian marine sponge *Dendrilla cactos*
Q, R	Australian marine sponge *Dendrilla cactos*
S, K	Australian tunicate *Didemnum* sp.
T-X, N, T 20-sulfate, U 20-sulfate, V 20 sulfate, Y 20-sulfate	Unidentified ascidian
Z, B 20-sulfate, C 20-sulfate, L 20-sulfate, G 8-sulfate, A-C, E, G, L, D triacetate, N triacetate	Great Barrier Reef ascidian, *Didemnum chartaceum*
α 20-sulfate	Unidentified ascidian
β, G, L	Purple unidentified encrusting marine ascidian *Didemnum* sp.
γ, α, ε, C diacetate, I, K, K diacetate, K triacetate, M, U, X triacetate	Indian ascidian *Didemnum obscurum*
ξ, η, ϕ, χ, F, I, J, K, K triacetate, L triacetate, T diacetate	Red colonial ascidian *Didemnum obscurum*

Source: Fan *et al.* (2008)

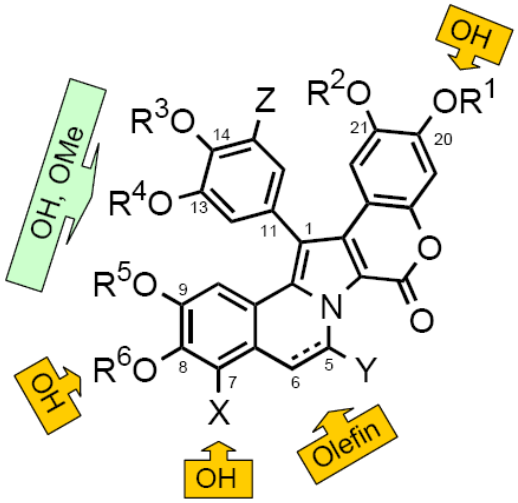


Figure 2 Important structure elements in the lamellarin skeleton, orange color refer to a group which is essential for cytotoxic activity.

Source: Chittchang *et al.* (2009)

The last SAR report was done in 2008. Twenty lamellarins were tested in a kinase assay panel comprising CDK1/cyclin B, CDK5/p25, GSK-3α/ß, PIM-1, DYRK1A, CK1. Some lamellarins are potent inhibitors of various protein kinases observed following minor changes at the lamellarin chemical structure. For example, the O-methylation at R2 results in massive loss of inhibitory activity as can be seen with lamellarin D and lamellarin α. Transposition of the substitutions at R4 and R5 of lamellarin D, resulting in lamellarin N, led to 10-fold enhanced kinase inhibitory activity and reduction in activity due to saturation of D-ring double bond (C5=C6) has previously been reported to be due to loss of planarity and therefore steric hindrance in ATP pocket of targets. Lamellarins which were totally inactive on kinases were devoid of effects on cell death. Altogether these results suggest that kinase inhibition may contribute to the effects of lamellarins on cell proliferation and cell death. However, most active compounds were active on all six kinases. A few lamellarins displayed apparent selectivity towards some kinases, suggesting that some degree of selectivity might be gained following the synthesis of more analogs (Baunbaek *et al.*, 2008).

Nowadays, QSAR plays an important role in lead structure optimization and drug design. It derives correlations between the similarities of individual compounds in term of their physicochemical properties with their biological activities (Winkler, 2002). After X-ray structures of the proteins became available, the results of QSAR regression models have been interpreted with the additional information from the three-dimensional structures (Kubinyi, 1997). In 1988 the method of CoMFA was published by Cramer *et al.* It is suited to describe ligand-receptor interactions, because it considers the properties of the ligands in their bioactive conformations. CoMFA calculates steric and electrostatic properties according to the Lennard-Jones and Coulomb potentials. According to the results of a CoMFA analysis, regions in space are identified that are favorable or unfavorable for ligand-receptor interactions (Cramer *et al.*, 1988). CoMFA is not only a popular method in 3D-QSAR analysis but CoMSIA is also the most commonly used in drug discovery. It was developed by Klebe *et al*. in 1994. CoMSIA approach calculates similarity indices in the space surrounding each of the aligned molecules in the data set by using Gaussian functions rather than Lennard-Jones and Coulombic functions. In CoMSIA, five similarity indices such as steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields were introduced in the calculations while steric and electrostatic fields were used in CoMFA calculations (Klebe *et al.*, 1994).

Among the 3D-QSAR methods, the 4D-QSAR method which was proposed by Hopfinger and co-workers in 1997, has more advantages because it can incorporate molecular flexibility, multiple alignments, while allowing for the identification of the conformation that maximizes the predicted activity for the best 4D-QSAR model. This conformation is defined as the active (or bioactive) conformation (Hopfinger, *et al.*, 1997; Romeiro, *et al.*, 2005). The fourth dimension of 4D-QSAR analysis is the dimension of ensemble sampling. It consists of two concepts which are (1) a receptor-independent (RI) 4D-QSAR whose geometry of the receptor is not available and (2) a receptor-dependent (RD) 4D-QSAR whose geometry of the receptor is known. In the QSAR model building process, 4D-QSAR analysis incorporates ligand conformational flexibility, multiple alignment exploration, and exhaustive evaluation of ligand-embedded pharmacophore groupings. Moreover, 4D-QSAR model can be

used as a virtual screen to evaluate a virtual library of compounds. It was found to yield more robust and more predictive models than the popular comparative molecular field analysis (CoMFA) in 3D-QSAR method (Krasowski *et al.*, 2002; Pasqualoto *et al.*, 2004).

In 2004, a new universal QSAR descriptor was generated by Senese and co-workers (Senese *et al.*, 2004). This method has been developed from Molecular Similarity Analysis (4D-MS). The methodology permits the generation of sets of molecular fingerprints that embed the conformational information of the molecule as well as capture its size and chemical structure. Moreover, each molecular "finger" of the molecular fingerprint is specific to a particular atom or pharmacophore type present in a molecule. Molecular fingerprints can be constructed for each specific alignment assigned to the members of a training set or library. Alignment dependent molecular fingerprints permit molecular similarity measures to be developed as a function such as the binding mode to a receptor site. However, another unique set of molecular fingerprints in this class can be developed for any molecule which are independent of alignment but do encompass the ensemble of conformational states available to the molecule. A method that can assemble a unique set of molecular fingerprints is attractive for both chemoinformatics and molecular modeling applications. Such molecular fingerprints represent descriptors that contain all information about a molecule called "4D-fingerprints". Moreover, since these descriptors are independent of alignment, they can be used in any type of molecular modeling and/or chemometrics application. That is, these 4D-fingerprints are universal descriptors (Duca and Hopfinger, 2001).

METHODS OF CALCULATIONS

1. Molecular Docking

In this thesis work, Automated Docking (AutoDock) and Genetic Optimisation for Ligand Docking (Gold) were employed to investigate the orientation and binding energies of lamellarins in the binding pocket of topoisomerase I-DNA. All starting geometries of lamellarin derivatives were constructed by using HyperChem 7.5 for windows program (Hypercube Inc., Gainesville, FL, USA, 2002). The chemical structures of lamellarin derivatives were shown in Table 2-3 and Figure 3. Each lowest energy structure was obtained from a conformational search analysis of the side chain substitutions. Then, all structure lamellarins were fully optimized at the HF/3-21G level using the GAUSSIAN 03 program (Gaussian, Inc., Wallingford, CT, USA, 2003). These optimized structure of lamellarins were used as starting geometry for molecular docking study and 3D-QSAR analysis. Additionally, the 2.10 Å resolution X-ray crystal structure of human topoisomerase I (70 Kda) in complex with the poison topotecan and covalent complex with a 22 base pair DNA duplex (Protein Data Bank 1K4T) was used as reference structure to model the lamellarin stabilization of topoisomerase I–DNA complex in Molecular docking processes.

1.1 AutoDock

Briefly, AutoDock explores the conformational space of the ligand using the Lamarkian genetic algorithm, LGA (Morris *et al*., 1998). In AutoDock process, it was performed by the automated molecular docking which is a module provided in the advanced docking program AutoDock 3.0 (The Scripps Research Institute, Molecular Graphics Laboratory, Department of Molecular Biology, La Jolla, CA, USA). The docked structures of the inhibitors were generated after a reasonable number of evaluations. The whole docking operations consist of five steps as followed.

First step, topotecan, water, and DNA duplex were removed out form the whole enzyme molecules. Topotecan was prepared as one of the inhibitors for redocking method which is used to evaluate this system in molecular docking process. DNA duplex was added CHARMM (Chemistry at HARvard Molecular Mechanics) charges and solvent parameters before connecting its back to topoisomerase I enzyme in second step.

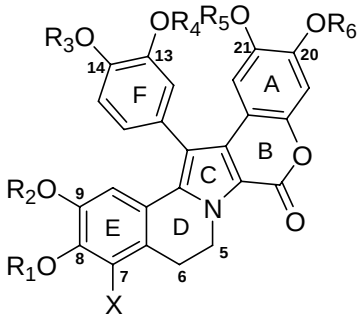
Second step, the enzyme molecules were checked for polar hydrogens and assigned for partial atomic charges. The PDBQs file was created, and the atomic solvation parameters of the macromolecules were also generated. All torsion angles of the inhibitors were defined. This allowed the conformational search of inhibitors during the process of docking.

Third step, the 3D grid was generated by the AutoGrid algorithm to evaluate the binding energies between inhibitor and enzyme. In this step, the topoisomerase I–DNA complex was embedded in the 3D grid and a probe atom was placed at each grid point. The affinity and electrostatic potential grid were calculated for each type of atom in the inhibitor. The energies of a particular inhibitor configuration were also calculated.

Fourth step, a series of the docking parameters were set on. The energy evaluations, grid spacing, and docking runs were set to 500000, 0.375, and 100, respectively. Other parameters were set as default. Atomic charges were assigned as Kollman-allatom (Weiner *et al*., 1984) for topoisomerase I–DNA complex and Gasteiger-Huckel (Gasteiger *et al*., 1980) for inhibitors.

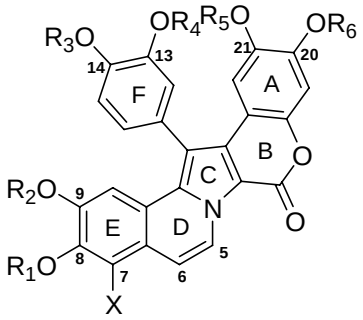
Last step, each docked complexes of inhibitor-enzyme were selected according to the criteria of interacting energy combined with geometrical matching quality. These complexes were used as the starting conformation for further energetic minimization and geometrical optimization before the final model was achieved.

Table 2 Structure and cytotoxic activity of lamellarins with a saturated D-ring.



Lamellarin	Substituent group							-Log IC_{50}	
	X	**OR_1**	**OR_2**	**OR_3**	**OR_4**	**OR_5**	**OR_6**	**T47D**	**MDA-MB-231**
A	OMe	OMe	OMe	OH	OMe	OMe	OH	-	-
C	OMe	OMe	OMe	OH	OMe	OMe	OH	5.11	5.08
E	OH	OMe	OMe	OMe	OH	OMe	OH	5.28	5.47
F	OH	OMe	OMe	OMe	OMe	OMe	OH	5.34	5.44
G	H	OH	OMe	OMe	OH	OH	OMe	5.07	4.83
I	OMe	OMe	OMe	OMe	OMe	OMe	OH	5.02	5.07
J	H	OH	OMe	OMe	OMe	OMe	OH	4.89	5.13
K	OH	OMe	OMe	OH	OMe	OMe	OH	7.04	6.40
L	H	OH	OMe	OMe	OH	OMe	OH	5.36	5.75
S	H	OH	OMe	OH	OH	OH	OH	-	-
T	OMe	OMe	OMe	OMe	OH	OMe	OH	4.88	5.06
U	H	OMe	OMe	OMe	OH	OMe	OH	4.99	5.35
V	OMe	OMe	OMe	OMe	OH	OMe	OH	-	-
Y	H	OMe	OH	OMe	OH	OMe	OH	5.14	4.10
Z	H	OH	OMe	OH	OH	OH	OMe	-	-
β	H	OH	OH	OMe	OH	OH	OH	-	-
χ	H	OH	OMe	OH	OMe	OMe	OH	5.42	5.32
K triacetate	OAc	OMe	OMe	OAc	OMe	OMe	OAc	5.18	5.33
U diacetate	H	OMe	OMe	OMe	OAc	OMe	OAc	5.10	5.46
χ triacetate	H	OAc	OMe	OAc	OMe	OMe	OAc	5.54	5.18

Table 3 Structure and cytotoxic activity of lamellarins with an unsaturated D-ring



Lamellarin	Substituent group							-Log IC_{50}	
	X	**OR_1**	**OR_2**	**OR_3**	**OR_4**	**OR_5**	**OR_6**	**T47D**	**MDA-MB-231**
B	OMe	OMe	OMe	OH	OMe	OMe	OH	6.74	5.35
D	H	OH	OMe	OH	OMe	OMe	OH	10.10	6.40
H	H	OH	OH	OH	OH	OH	OH	-	-
M	OH	OMe	OMe	OH	OMe	OMe	OH	8.02	6.95
N	H	OH	OMe	OMe	OH	OMe	OH	9.22	6.22
W	OMe	OMe	OMe	OMe	OH	OMe	OH	5.37	5.29
X	OH	OMe	OMe	OMe	OH	OMe	OH	8.25	7.12
α	H	OMe	OMe	OMe	OH	OMe	OH	6.23	5.41
ε	OH	OMe	OMe	OMe	OMe	OMe	OH	8.26	6.59
ζ	OMe	OMe	OMe	OMe	OMe	OMe	OH	7.05	5.33
η	H	OMe	OMe	OMe	OMe	OMe	OH	-	-
ϕ	OMe	OMe	OAc	OAc	OMe	OMe	OAc	-	-
Dehydro. J	H	OH	OMe	OMe	OMe	OMe	OH	10.01	6.41
Dehydro. Y	H	OMe	OH	OMe	OH	OMe	OH	7.10	6.19

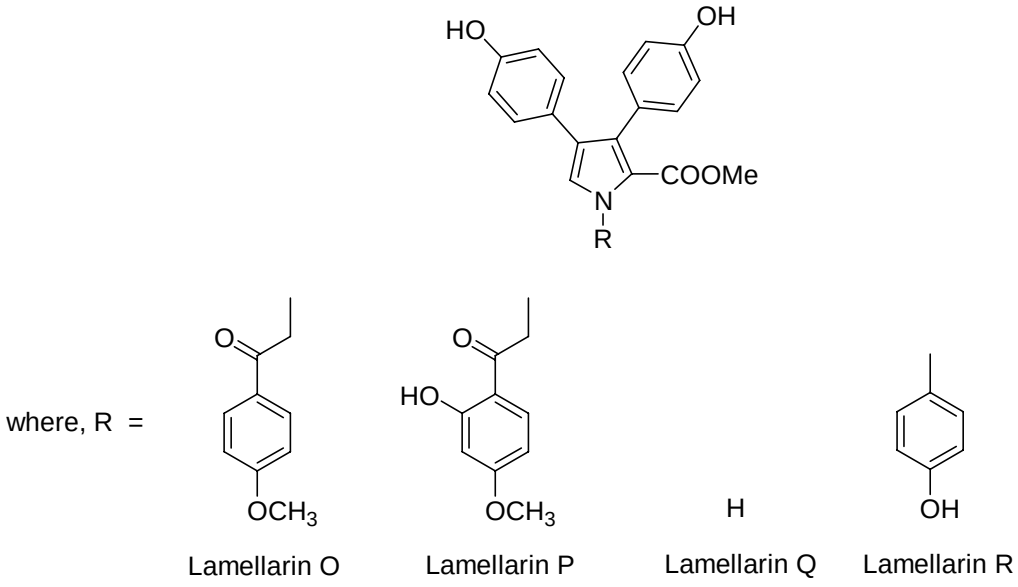


Figure 3 Chemical structure of unfused central pyrrole ring lamellarins.

1.2 Gold

Gold is an automated ligand docking program that uses a genetic algorithm (GA) docking flexible ligands into protein binding sites as well as AutoDock. In Gold process, it was run automatically by using Gold 2.2 (CCDC, Cambridge, UK, 2001) program. The docked structure of inhibitors in binding pocket were constructed and ranked by using a fitness scores (GoldScore or ChemScore). The whole Gold operations include five steps as followed.

First step, topotecan and water were removed out form the whole enzyme molecules. Topotecan was prepared as one of the inhibitors for redocking method as well as in the AutoDock process.

Second step, the enzyme molecules were added hydrogens by Sybyl 7.0 program. Then, the mol2 file was saved while deleting all lone paired electrons out of topoisomerase I enzyme.

Since each hydrogen bond energy was calculated by considering the type and the geometry of the donor-acceptor pair, inhibitors need to define atom type correctly by Sybyl 7.0 program. Then, mol2 file was saved while deleting all lone paired electrons out of inhibitors in third step.

Fourth step, a series of the docking parameters were set on. The number of generation, active site radius, population size, and number of operation were set to 15, 10, 100, and 30000, respectively. Fitness function was also chosen in this step. Other parameters were set as default.

Last step, docked complexes of inhibitor-enzymes were calculated by using GA. The fitness scores were also evaluated in the step. Then, final geometry of this complex was obtained when the system reached a good in fitness score and criteria of GA operation.

2. 3D-QSAR (CoMFA and CoMSIA) Analysis

The 26 lamellarins used in this study (Tables 1 and 2) were synthesized and purified by Ploypradith *et al.* (Ploypradith *et al*., 2006). All the compounds were solubilized in dimethyl sulfoxide (DMSO) and tested for their cytotoxic activities against T47D and MDA-MB-231 cell lines, as previously reported of Chittchang *et al.* (Chittchang *et al.,* 2009*)*. Briefly, the cells were incubated in 96-well microplates at 37°C for 48 h with the serial dilutions of the test compounds, positive control (etoposide), or negative control (DMSO). The number of surviving cells in each well was determined using crystal violet staining in order to attain the IC_{50}, which is defined as the concentration that inhibits cell growth by 50% after 48 h of continuous exposure to each test compound. The values thus obtained were then transformed by calculating their negative logarithm (i.e., -log IC_{50}), which is a standard notation to make very small numbers fit into a more comprehensible range, and larger numbers indicate more potent cytotoxicities.

Then, all lamellarin molecules were separated into two groups. The first group, consisting of 20 compounds, served as the training set. On the other hand, six compounds comprising the test set were sampled randomly to include structurally diverse molecules possessing a wide range of cytotoxic activities against both cell lines, namely lamellarins α, K, L, M, T, and U. The starting geometries of all 26 lamellarin structures were fully optimized at the HF/3-21G level using the GAUSSIAN 03 program. The partial atomic charges required for calculations of electrostatic interactions were subsequently computed using the Gasteiger-Hückel method in the SYBYL 7.0 program. Molecular alignments were then carried out using the matching method also available in the SYBYL software. The common structure used for matching alignments involves the atoms constituting the A-, B-, and C-rings, as denoted with asterisks in Figure 4. Additionally, the most active molecule towards each cell line, i.e., lamellarin D for T47D and lamellarin X for MDA-MB-231 cells, was used as the template for alignments.

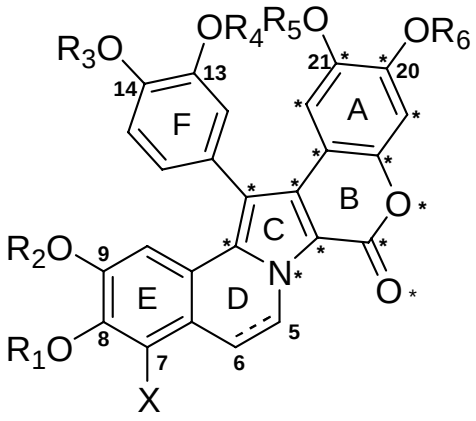


Figure 4 Core structure of lamellarins, in which the common atoms used for 3D-QSAR matching alignments are denoted by asterisks (*).

To calculate the CoMFA and CoMSIA descriptor fields, a cubic lattice was first generated around each lamellarin molecule based on its molecular volume, and a grid spacing of 2 Å was used to ensure that the grid extended by 4.0 Å beyond the molecular dimensions in all directions. In addition to a sp^3 carbon atom with a +1 charge, which is the default probe atom in SYBYL, a sp^3 oxygen atom with a -1

charge and a hydrogen atom with a +1 charge were also used as the additional probe atoms for CoMFA calculations in this study, and these atoms were placed at each lattice point. On the other hand, the steric and electrostatic fields of each aligned lamellarin were generated based on their Lennard-Jones and Coulomb potentials, respectively. The interactions between these three-dimensional fields with each probe atom were then calculated using the CoMFA standard scaling technique, in which the minimum sigma value was set at 2.0 kcal/mol, and an energy cutoff value of 30 kcal/mol was selected to not only speed up the analysis, but also reduce the amount of noises. All the calculated data were then put into a CoMFA table.

In CoMSIA, five similarity indices (including steric, electrostatic, hydrophobic, H-bond donor, and H-bond acceptor descriptors) were computed for each lamellarin molecule using the same cubic lattice employed for the CoMFA calculations, as well as the default carbon probe atom with a 1 Å radius and a +1.0 charge. The three-dimensional properties of lamellarins determined using either CoMFA or CoMSIA were then correlated with their cytotoxic activities against each cell line using partial least squares (PLS) regression analysis, and various 3D-QSAR models were subsequently derived.

The predictive ability of the derived 3D-QSAR models was evaluated by leave-one-out (LOO) cross-validation, and is expressed in terms of r^2_{cv} (also called q^2), which is defined as shown in Equation 1.

$$r^2_{cv} = (SSY\text{-}PRESS)/SSY \tag{2}$$

where, SSY represents the variance of the cytotoxic activities of molecules around the mean value, and PRESS is the prediction error sum of squares derived from the leave-one-out method. In contrast, the uncertainty of the prediction is defined as shown in Equation 2.

$$PRESS = \sum_y (y_{pred} - y_{actual})^2 \tag{3}$$

$$SSY = \sum_{y} (y_{actual} - y_{mean})^2 \quad (4)$$

The uncertainty of the prediction is defined as

$$S_{PRESS} = \sqrt{\frac{PRESS}{n-k-1}} \quad (5)$$

where n is the number of compounds used in the study, and k is the number of variables in the model (Golbraikh, *et al*., 2002, Nilsson, *et al*., 1997). For all models, a maximum number of components (i.e., noc = 6) was first used and subsequently decreased until an optimal number was obtained when the resulting cross-validated r^2_{cv} differed from the previous value by less than 0.05. The optimal number of components was then used to perform non cross-validated analyses. Briefly, the conventional correlation coefficient, r^2, was calculated based on the 20 compounds in the training set. The CoMFA and CoMSIA models with an r^2 value higher than 0.6 were subsequently validated by evaluating the correlations between the observed and predicted cytotoxic activities of the compounds in the test set, as indicated by the $r^2_{test\ set}$ values.

3. Receptor-Independent (RI) 4D-QSAR Analysis for Cytotoxicity of Lamellarins Against T47D breast cancer cells

The ten operational steps in RI 4D-QSAR have been generated in the 4D-QSAR software product, version 3.0 (The Chem21 Group, Inc., Lake Forest, IL, USA, 1999). The current of methodology are summarized in this part only in terms of modeling the lamellarin compounds, and the 10 steps of 4D-QSAR analysis.

Step 1, an initial 3D structure of each lamellarin was constructed in the neutral form using the HyperChem 7.5 software. Partial atomic charges were computed using the AM1 semiempirical method as implemented in the HyperChem

program. Each structure was minimized by using AM1 with no geometric constraint. These energy-minimized structures were used as the initial structures in the conformational ensemble sampling of step 3.

Step 2, the atoms of each molecule were classified into seven types of interaction pharmacophore elements (IPEs). Each type is represented by different number code from 0 to 6 as defined in Table 4.

Step 3, molecular dynamic simulation (MDS) was used to sample the conformational states available to each analogue and to generate its corresponding conformational ensemble profile (CEP). The MDSs are done using the MOLSIM package and MM2 force field. The temperature for the MDS is set at 300 K with a simulation sampling time of 40 ps with intervals of 0.001 ps for a total sampling of 40000 conformations of each lamellarin compound. The atomic coordinates of each conformation and its intramolecular energy sampled during the MDS are recorded every 0.02 ps for a total of 2000 “frames”, or steps, in constructing the CEP of each compound.

Step 4, three-ordered atoms in trial alignments are defined in Table 5. In this study eight alignments were explored across the overall lamellarin core structure.

Step 5, each conformation of a compound from its CEP is aligned in the grid cell lattice using the invariant coordinates of the three-ordered alignment atoms. In this study, the size of the cubic grid cells of the lattice are 1 $Å^3$, and the overall grid cell lattice size was chosen to fully enclose each compound of the training set. The normalized occupancy of each grid cell by each IPE atom type over the CEP for a given alignment forms a unique set of QSAR descriptors referred to as grid cell occupancy descriptors, GCODs. The GCODs are computed and used as the basis set of trial 4D-QSAR descriptors in 4D-QSAR analysis. Non-GCOD descriptors of the training set compounds can also be included in the trial basis set [descriptor pool]. In this particular study the logarithm of the 1-octanol/water partition coefficient (log *P*)

and the compound's molecular weight (MW) were selectively added to the trial basis set descriptors in some of the model building studies.

Step 6, a 4D-QSAR analysis generates an enormous number of trial QSAR descriptors, GCODs, because of the large number of grid cells and the seven IPEs. Partial least squares (PLS) regression analysis is used to perform a data reduction analysis between the observed dependent variable measures and the corresponding set of GCOD values.

Step 7, the most highly weighted PLS GCOD descriptors (currently the top 200), generated in step 6, are used to form the trial descriptor pool for genetic algorithm (GA) model optimization. The specific GA currently used in the 4D-QSAR software is modification of the genetic function approximation (GFA) by Roger and co-worker (Roger and Hopfinger, 1994). The GFA optimization is initiated using N (currently 300) randomly generated 4D-QSAR models. Mutation probability over the crossover optimization cycle is set at 10%. The smoothing factor, a GFA operations variable, controls the number of independent variables in the QSAR models, is varied in order to determine the optimal number of descriptors for the 4D-QSAR models. The diagnostic measures used to analyze the resultant 4D-QSAR models generated by the GFA include (i) descriptor usage as a function of crossover operation, (ii) linear cross correlation among descriptors and/or dependent variables (biological activity measures), (iii) number of significant and independent 4D-QSAR models, and (iv) indices of model significance including the correlation coefficient, r^2, leave one-out, LOO, cross-validation correlation coefficient, r^2_{cv}, and Friedman's lack of fit (LOF). In this particular 4D-QSAR application, the alignment similarity comparisons were limited to models having same number GCODs.

Step 8, steps 4-7 are repeated until all trial alignments are included in the 4D-QSAR analyses.

Step 9, the inspection and evaluation of the population of models are obtained from the set of trial alignments in this step. The goal of this step is to identify the best

and distinct set of 4D-QSAR models which is referred to as the manifold model of the analysis.

Step 10, the "active" conformation of each compound is hypothesized at this step. This conformer is achieved by identifying all conformer states sampled for each compound that are within ΔE of the global minimum energy conformation of the CEP. Currently, ΔE is set at 2 kcal/mol. Each member of the resultant set of energy-filtered conformations is then individually evaluated in the best 4D-QSAR model. The conformation within 2 kcal/mol of the apparent global minimum that predicts the highest activity in the best 4D-QSAR model is defined as the active conformation.

4. 4D-fingerprints QSAR Analysis for Cytotoxicity of Lamellarins Against T47D breast cancer cells

In 4D-fingerprints analysis, its needs to construct the main distance-dependent matrix (MDDM) and computing corresponding eigenvalues for each matrix by using 4D molecular similarity (MS), have been presented in detail in the previous work (Sendese, et al., 2004; Duca and Hopfinger, 2001). The types of atoms composing a molecule are currently defined as the IPEs shown in Table 4. A unique MDDM is constructed for each of the eight distinct and identical IPE pairs. The elements of the MDDM are defined below:

$$E_{(\nu,d\mathrm{ij})} = e^{(-\nu\langle dij\rangle)} \tag{6}$$

The "universal constant *(ν)"* in eq. 1, which is equal to 0.25, has been selected such that the difference in the sum of eigenvalues for any two arbitrary compounds with the same number, *n*, of a particular IPE type, *m*, is maximized. The term $\langle dij\rangle$ is the average distance between the atom pair ij of IPE type u and v.

$$\langle dij\rangle = \sum_k d ij(k)\, p(k) \tag{7}$$

where p(k) refers to the thermodynamic probability of the kth conformer state sampled in the assessment of conformational flexibility, and d*ij*(k) is the

corresponding distance between atom pair *i* and *j* of IPE types *u* and *v* for the kth conformer state. Then, similarity eigenvalues are derived by the diagonalization of the MDDM. For same-term IPE pairs, such as $u = v$, the MDDM are square upper/lower triangular. These matrices can be directly diagonalized. The resulting eigenvalues determined from the MDDM are normalized and ranked in numerically descending order in their eigenvector representation. The *n*th normalized eigenvalue for IPE type *m* of a compound $\alpha, \in_{mn}(\alpha)$, can be obtained by scaling the non-normalized eigenvalue $\in_{mn}'(\alpha)$ relative to the rank of its MDDM.

$$\in_{mn}(\alpha) = \in_{mn}'(\alpha)/\text{rank}(\alpha)_m \tag{8}$$

Thus, $\in_{0,3}(2)$ would correspond to the third eigenvalue of the MDDM for IPE type 0 of compound 2. Determination of eigenvalues of the MDDM for $u \neq v$, the so-called cross-terms for IPE pairs that are not the same, requires a different strategy since these matrices may, or may not, be square. In the case of rectangular MDDM ($u \neq v$), the following square MDDM are constructed

$$\text{MDDM}(u,u) = \text{MDDM}(n_u, n_v) \text{ x } \text{MDDM}(n_u, n_v)^T \tag{9}$$

$$\text{MDDM}(v,v) = \text{MDDM}(n_v, n_u) \text{ x } \text{MDDM}(n_v, n_u)^T \tag{10}$$

MDDM(*u,u*) and MDDM(*v,v*) have the same rank and trace, both have the same set of eigenvalues. Hence, for each pair of IPE ($u \neq v$)

$$\in(\alpha)_{u,v} = \{[\in(\alpha)]_{\text{MDDM}(u,u)}\}^{1/2} \tag{11}$$

According to all possible combinations of the eight IPE types, there are 36 possible molecular similarity eigenvectors from the MDDM for each compound α. The similarity eigenvectors have been calculated for the set of compounds, the estimation of molecular similarity for a pair of compounds α and β begins with a definition for molecular dissimilarity, given by

$$D_{\alpha\beta} = \sum_i |\epsilon(\alpha)_i - \epsilon(\beta)_i| \quad (12)$$

where i = ith eigenvalue in the corresponding eigenvector of a specific IPE pair. Molecular similarity is then defined as

$$S_{\alpha\beta} = (1 - D_{\alpha\beta})(1 - \varphi) \quad (13)$$

where φ = |rank(α) - rank(β)| / (rank(α) + rank(β). The rank of the matrices is essentially the number of atoms of a specific IPE type present. The φ term in eq. 8 serves to reincorporate molecular size information. Similar to the measure for dissimilarity, the similarity measure is a value between 1 and 0, where a value closer to 1 refers to compounds that are more similar, and closer to 0 refers to compounds that are more dissimilar.

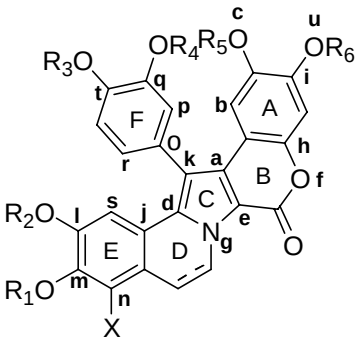
The descriptor set for α consists of all of the eigenvalues of all of the eigenvectors derived from all of the MDDM for compound α. In this work, a threshold cutoff value equal to 0.002 is applied, and those normalized eigenvalues below the threshold value are disregarded. The maximum number of significant eigenvalues specific to that data set for a particular compound and a particular IPE type, m, is determined, $\epsilon_{m,max}$. All the eigenvectors for IPE type, m, for each molecule across lamellarin data set are then assigned $\epsilon_{m,\max}$ eigenvalues for IPE type m. Eigenvectors that otherwise contain less than $\epsilon_{m,\max}$ elements have the "missing" eigenvalues set to zero. The total set of descriptors, ϵ_{total}, for a compound in the data set will be the sum of the 36 eigenvalues of $\epsilon_{m,\max}$ length which can be a large number for the data set in this work. Then, introduces some degree of estimation when creates the universal descriptor data matrix. Finally, the sets of 4D-fingerprints across each of the molecules of the training set form the trial descriptor pool to build the 4D fingerprint virtual screens. The building procedure of these virtual screens is identical to that employed in constructing the RI-4D-QSAR models. That is, steps 6 through 9 given above for the RI-4D-QSAR methodology are used.

Table 4 Set of Interaction Pharmacophore Elements (IPEs) used in the RI 4D-QSAR and 4D-fingerprint QSAR Analysis.

IPE description	Symbol	Number code
all atoms in the molecule	any	0
nonpolar atoms	np	1
polar atoms of positive partial charge	p^+	2
polar atoms of negative partial charge	p^-	3
hydrogen bond acceptor atoms	hba	4
hydrogen bond donor atoms	hbd	5
aromatic atoms	aro	6
non-hydrogen atoms[a]	hs	7

[a] hydrogen-suppressed use only in 4D-fingerprint QSAR analysis

Table 5 Set of trial alignment used in constructing the 4D-QSAR model.



Alignment	First atom	Second atom	Third atom
1	a	b	c
2	d	e	f
3	g	h	i
4	j	d	k
5	l	m	n
6	o	p	q
7	k	o	r
8	s	t	u

RESULTS AND DISCUSSION

1. Molecular Docking

1.1 AutoDock

These results are generated by using AutoDock 3.0 program. Topotecan inhibitor is docked back into the binding pocket of topoisomerase I-DNA to reproduce the crystal structures of topotecan bound to topoisomerase I-DNA enzyme which is shown in Figure 5. It is found that the root mean square deviation (RMSD) of topotecan inhibitor is 0.96 Å and the estimated free energy of binding ($G_{binding}$) is -19.91 kcal/mol. In Figure 6, it compares X-ray structure with docked structure of topotecan in the binding pocket of topoisomerase I-DNA by using superimposition method.

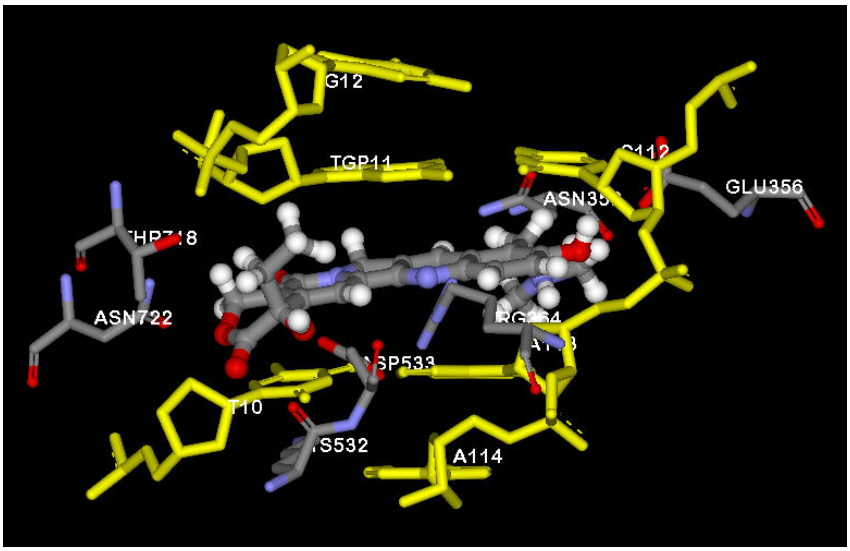


Figure 5 Topotecan surrounding with amino acids and DNA within 7.0 Å in the binding pocket of topoisomerase I-DNA. Topotecan represented by ball and stick style and colored by element (carbon: dark grey, nitrogen: blue, oxygen: red, hydrogen: white). Amino acids and DNA represented by normal element color stick and yellow stick style, respectively.

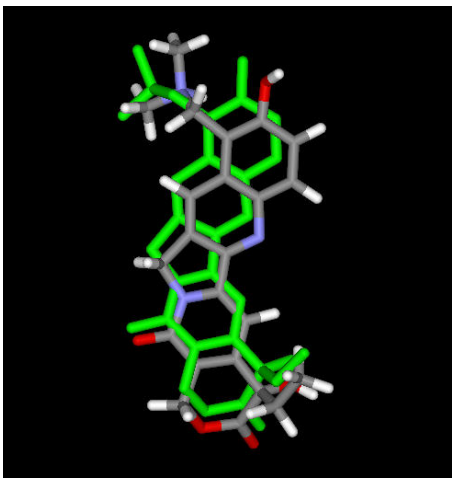


Figure 6 Superposition of topotecan obtained from X-ray crystallographic structure (green) and docking calculation structure (element color).

From docking results, there are seven amino acid residues within 7.0 Å in the binding pocket of the enzyme consisting of Asn352, Glu356, Arg364, Lys532, Asp533, Thr718, and Asn722. However, there is only one direct H-bond between the enzyme and topotecan with distance equal to 2.67 Å (between Asp533 and hydroxyl in carboxylate group of topotecan) as represented in Figure 7. Moreover, this model also shows that topotecan intercalates at the site of DNA cleavage, forming base-stacking interactions with both the upstream and downstream base pairs (Figure 7) similar to those discussed in X-ray crystal structure in previous work (Staker *et al.*, 2002). Based on the RMSD value less than 1.0 Å, it may concluded that Autodock method can be used to investigate the orientation and binding energies of lamellarins in further steps.

Due to the easily compared orientation and binding energies of all lamellarins in the binding pocket of topoisomerase I-DNA, AutoDock results are emphasized into three different lamellarin structures such as unsaturated D-ring fused (group 1), saturated D-ring fused (group 2), and unfused central pyrrole ring group (group 3). All the estimated free energies of binding, π-π interaction, and heteroatom H-bond distance of group 1, 2, and 3 lamellarins are shown in Table 6, 7, and 8, respectively. From Table 6, AutoDock calculations demonstrate that three lamellarins

F, L, and T have the estimated free energies of binding more than -19.0 kcal/mol. Four lamllarins (C, G, V, and Y) in group 1 have the estimated free energies of binding lower than -17.0 kcal/mol. Moreover, AutoDock results show that lamellarins (C, V, and Y, except G) can not intercarate into center of DNA duplex plane as shown in Figure 8. Additionally, lamellarins (O, Q, R), except lamellarin P in group 3 were also can not intercarate into center of DNA duplex plane as represented in Figure 9. In Table 7, almost all lamellarins in group 2 have the estimated free energies of binding more than -17.0 kcal/mol. Especially, lamellarin M is predicted to be the most potent compound that can bind into the binding site of topoisomerase I-DNA enzyme with the value of -19.83 kcal/mol. The orientation and interaction of lamellarin M in binding site is displayed in Figure 10. Other lamellarins in group 1 and 2 can intercalate into the pocket similar to lamellarin M with interactions as shown in Table 6-7.

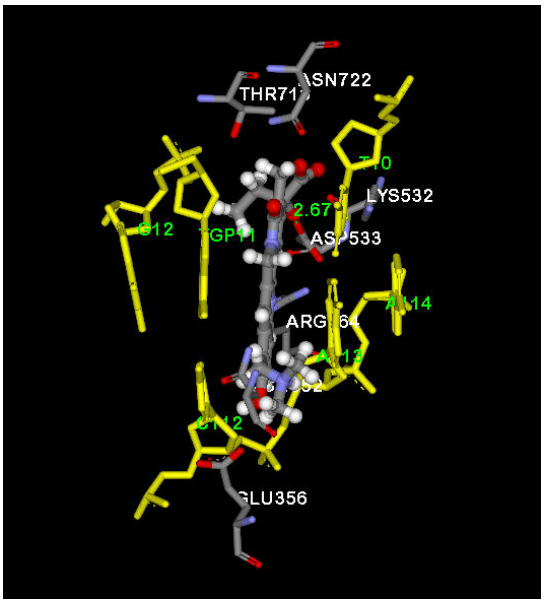


Figure 7 Base-stacking interactions and H-bond of topotecan surrounding with amino acids and DNA within 7.0 Å in the binding pocket of topoisomerase I-DNA from AutoDock calculations. Topotecan represented by ball and stick style and colored by element. Amino acids and DNA represent by normal element color stick and yellow stick style, respectively.

Table 6 Estimated free energies of binding (kcal/mol), π-π interactions, and heteroatom H-bond distances between amino acids and group 1 lamellarins (Å) in the binding pocket of topoisomerase I-DNA.

Lamellarin Group 1	$G_{binding}$ (kcal/mol)	H-bond (amino acid: Å)	π-π interactions
A*	-18.17	Asn722: 2.81	ND
C*	-16.48	Asn722: 2.61	ND
E	-18.62	Asn352: 2.83	D
		Asn722: 2.58	
F	-19.11	Asn722: 2.49	D
G*	-16.94	-	D
I	-17.68	Asn722: 2.52	D
J	-18.95	Glu356: 2.75	D
		Asn722: 2.54	
K	-18.75	Asn722: 2.74	D
L	-19.08	Glu356: 2.81	D
		Asn722: 2.57	
S	-18.72	Glu356: 2.79	D
		Asn722: 2.55	
T	-19.27	Asn722: 2.57	D
U	-18.56	Asn722: 2.60	D
V*	-16.70	-	ND
Y*	-15.66	Glu356: 2.72	ND
Z	-17.75	Asn352: 2.96	D
		Glu356: 2.75	
χ	-18.84	Glu356: 2.62	D
		Asn722: 2.51	

The carbonyl group of all lamellarins can interact with the Arg364 within distance around 3.0 Å, * do not interact with Arg364, D = detected, ND = not detected

Table 7 Estimated free energies of binding (kcal/mol), π-π interactions, and heteroatom H-bond distances between amino acids and group 2 lamellarins (Å) in the binding pocket of topoisomerase I-DNA.

Lamellarin Group 2	$G_{binding}$ (kcal/mol)	H-bond (amino acid: Å)	π-π interactions
B	-17.58	Asn352: 2.58	D
		Asn722: 2.63	
D	-18.56	Glu356: 3.10	D
		Asn722: 2.46	
H	-17.75	Asn352: 2.68	D
		Glu356: 2.90	
		Asn722: 2.96	
M	-19.83	Asn722: 2.39	D
N	-17.88	Glu356: 2.94	
		Asn722: 2.51	
W	-16.97	Asn722: 2.43	D
X	-18.41	Asn722: 2.49	D

The carbonyl group of all lamellarins can interact with the Arg364 within distance around 3.0 Å, * do not interact with Arg364, D = detected, ND = not detected

Table 8 Estimated free energies of binding (kcal/mol), π-π interactions, and heteroatom H-bond distances between amino acids and group 3 lamellarins (Å) in the binding pocket of topoisomerase I-DNA.

Lamellarin Group 3	$G_{binding}$ (kcal/mol)	H-bond (amino acid: Å)	π-π interactions
O*	-16.58	-	ND
P*	-17.16	-	D
Q*	-14.85	-	ND
R*	-17.22	-	ND

The carbonyl group of all lamellarins can interact with the Arg364 within distance around 3.0 Å, * do not interact with Arg364, D = detected, ND = not detected

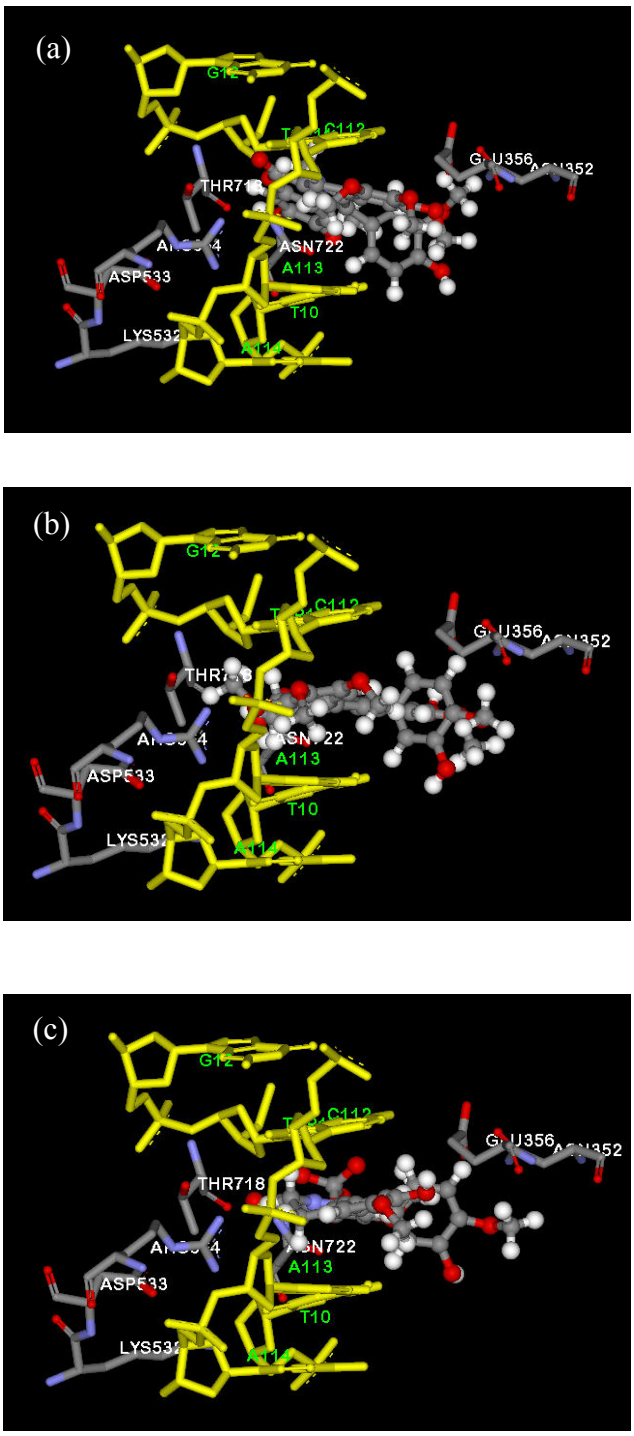


Figure 8 Orientation of lamellarin C (a), V (b), and Y (c) in 7.0 Å of the binding pocket of topoisomerase I-DNA from AutoDock calculations. Lamellaris is represented by ball and stick style and colored by element. Amino acids and DNA are represented by normal element color stick and yellow stick style, respectively.

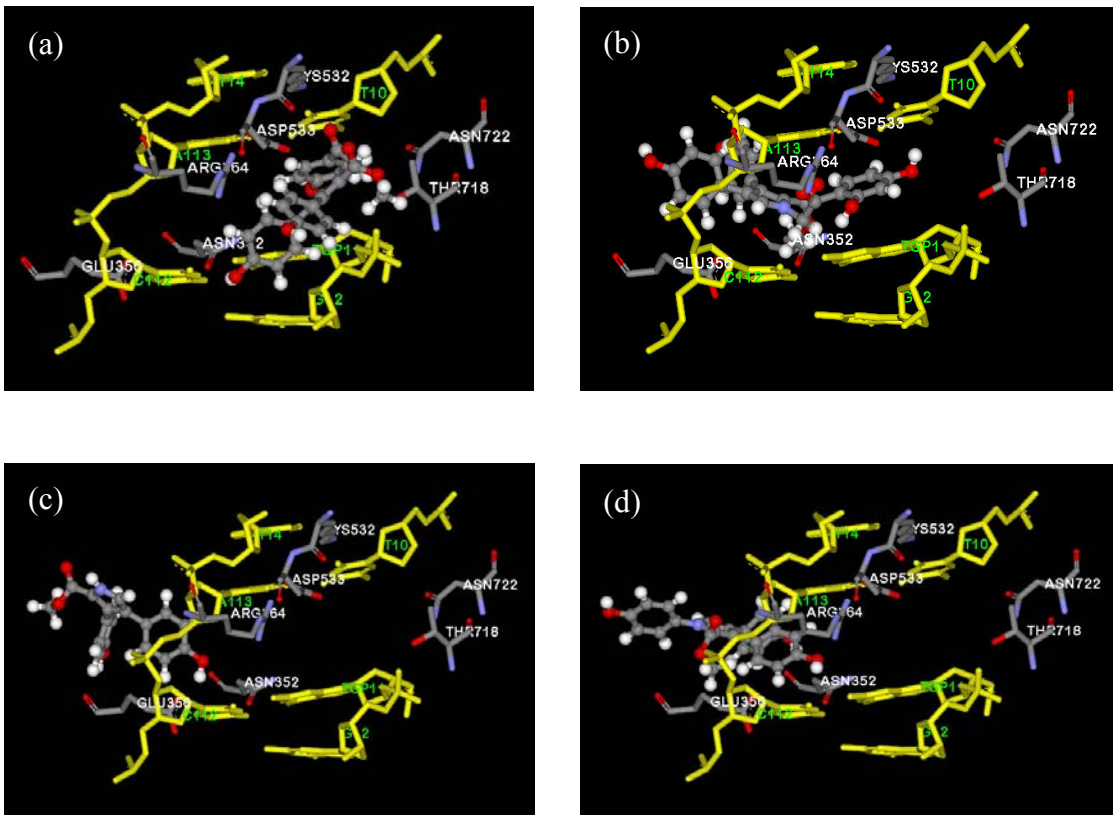


Figure 9 Orientation of lamellarin O (a), P (b), Q (c), and R (d) in 7.0 Å of the binding pocket of topoisomerase I-DNA from AutoDock calculations. Lamellaris is represented by ball and stick style and colored by element. Amino acids and DNA are represented by normal element color stick and yellow stick style, respectively.

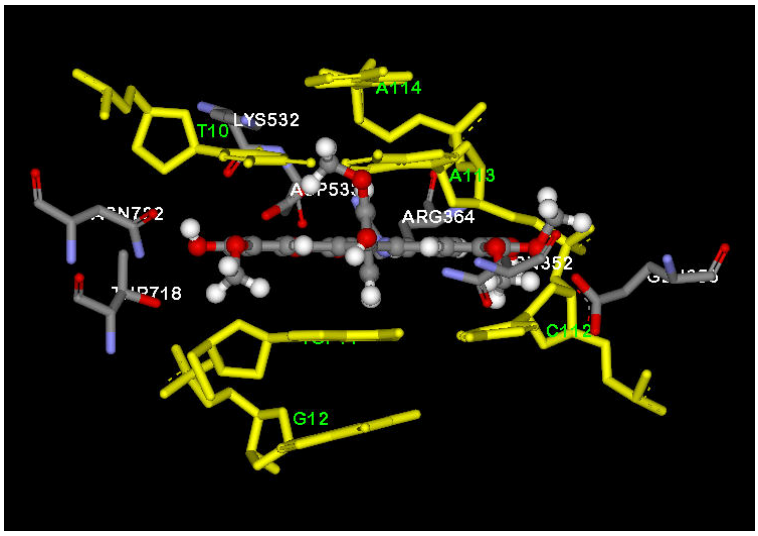


Figure 10 Orientation of lamellarin M in 7.0 Å of the binding pocket of topoisomerase I-DNA from AutoDock calculations. Lamellaris is represented by ball and stick style and colored by element. Amino acids and DNA are represented by normal element color stick and yellow stick style, respectively.

1.2 Gold

Gold 2.2 is used for docking flexible ligands (lamellarins) into topoisomerase I-DNA binding sites. Topotecan inhibitor is docked back into the binding pocket of this enzyme as well as AutoDock by using GoldScore and ChemScore fitness functions. Based on RMSD from heteroatom fit results, it is indicated that GoldScore with RMSD value of 0.79 Å is better than ChemScore with RMSD value of 1.27 Å. It means that the orientation of docking topotecan inhibitor from GoldScore calculations is closer to the orientation of topotecan inhibitor obtained from X-ray structure than the docking structure from ChemScore calculations. Figure 11 is the superposition of topotecan obtained from X-ray crystallographic structure and docking structure from GoldScore calculations. GoldScore of topotecan inhibitor is equal to 88.57.

Gold results are also discussed in term of three different lamellarin structure groups. All lamellarins are only calculated by GoldScore. Some acetate and novel natural lamellarins are added into Gold calculations to investigate orientations of these lamellarin compounds in binding pocket of topoisomerase I-DNA enzyme. All GoldScores, π-π interactions, and heteroatom H-bond distances of group 1, 2, and 3 lamellarins are shown in Table 9, 10, and 11, respectively. Since the fitness score or GoldScore is taken as the negative of the sum of the component energy terms (H-bonding term and internal energy term), larger fitness scores are better than small fitness scores.

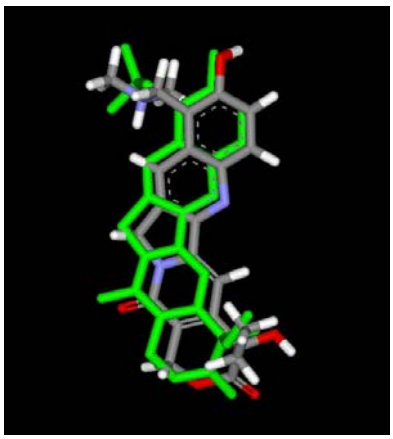


Figure 11 Superposition of topotecan obtained from X-ray crystallographic structure (green) and docking calculation structure (element color).

Based on Gold calculations, group 1 and 2 lamellarins have GoldScores in range of 70.12-87 and 71-79.08, respectively. Lamellarins in Group 3 has GoldScores in value range of 70.85-80.35. From Table 9, it demonstrates that lamellarin triacetates have GoldScores higher than lamellarins with no acetate groups. For example, lamellarin K-triacetate and K have the values of 84.10 and 73.76, respectively, while lamellarin U-diacetate 86.05 and 76.30, respectively, and lamellarin χ-triacetate and χ 87.14 and 73.59, respectively. However, it is not all the case that higher GoldScore refers to more interactions between the compound and residual. Lamellarin U-diacetate and lamellarin χ-triacetate do not have the

interactions with amino acids, while lamellarin U and lamellarin χ have interactions with Asn722, and Asn352 and Glu356, respectively. This table also shows that lamellarin E does not have H-bond with amino acids and π-π interaction with nucleotide as depicted in Figure 12. From Table 10, lamellarin N not only shows the highest in GoldScore (79.08), but also have many interactions in binding pocket as represented in Figure 13. Two lamellarins (W and ζ) have only one interaction with Arg364 at the carbonyl functional group. Almost all lamellarins in group 3 do not have H-bond with amino acids and π-π interaction with nucleotide, except lamellarin P and Q. According to Table 11, it shows that lamellarin P has two interactions with Asp533 and Asn722 as depicted in Figure 14 (a). Lamellarin Q has one interaction with Asp533 as indicated in Figure 14 (b).

Table 9 GoldScores, π-π interactions, and heteroatom H-bond distances between amino acids and group 1 lamellarins (Å) in the binding pocket of topoisomerase I-DNA.

Lamellarin Group 1	GoldScore	H-bond (amino acid: Å)	π-π interactions
A	71.80	Asn722: 2.94	D
C	73.63	Asn722: 3.04	D
E*	75.92	-	ND
F	72.87	-	D
G*	72.53	Asn352: 2.96 Glu356: 2.75	ND
I	70.12	-	ND
J	75.52	Glu356: 2.23 Asn722: 3.02	D
K	73.76	-	D
K-triacetate	84.10	-	ND
L	73.99	Glu356: 2.04 Asn722: 2.90	D

Table 9 (Continued)

Lamellarin Group 1	GoldScore	H-bond (amino acid: Å)	π-π interactions
S	70.70	Glu356: 2.32	ND
		Asn722: 2.97	
T	73.34	Asn352: 2.63	ND
U	76.30	Asn722: 2.97	ND
U-diacetate	86.05	-	ND
V	75.11	Asn722: 2.62	ND
Y	74.17	Glu356: 2.50	ND
Z	70.88	Asn352: 3.03	ND
		Glu356: 2.57	
χ	73.59	Asn352: 2.32	D
		Glu356: 2.85	
χ-triacetate	87.14	-	ND
β	70.93	Glu356: 2.05	D
		Asn722: 3.01	
γ	73.79	Asn722: 2.93	ND

The carbonyl group of all lamellarins can interact with the Arg364 within distance around 3.0 Å, * do not interact with Arg364, D = detected, ND = not detected

Table 10 GoldScores, π-π interactions, and heteroatom H-bond distances between amino acids and group 2 lamellarins (Å) in the binding pocket of topoisomerase I-DNA.

Lamellarin Group 2	GoldScore	H-bond (amino acid: Å)	π-π interactions
B	76.06	Asn352: 2.35	D
		Glu356: 2.97	
D	78.40	Glu356: 2.50	D
		Asn722: 2.47	
H	71.54	Glu356: 2.59	D
		Asn722: 2.46	
M	75.22	Asn352: 2.20	D
N	79.08	Asn352: 3.03	D
		Glu356: 3.03	
		Asn722: 2.49	
W	75.89	-	ND
X	75.08	Asn722: 2.71	D
α	74.41	Asn352: 3.00	D
		Asn722: 3.08	
ϕ	74.69	Asn352: 2.44	D
η	77.28	Asn722: 2.73	D
ε	75.84	Asn722: 2.62	D
ζ	76.96	-	ND
Dehydrolamellarin J	76.55	Glu356: 2.55	D
		Asn722: 2.40	
Dehydrolamellarin Y	73.19	Asn722: 2.93	ND

The carbonyl group of all lamellarins can interact with the Arg364 within distance around 3.0 Å, * do not interact with Arg364, D = detected, ND = not detected

Table 11 GoldScores, π-π interactions, and heteroatom H-bond distances between amino acids and group 3 lamellarins (Å) in the binding pocket of topoisomerase I-DNA.

Lamellarin Group 3	GoldScore	H-bond (amino acid: Å)	π-π interactions
O*	80.35	-	ND
P*	75.37	Asp533: 2.65 Asn722: 2.88	ND
Q	70.85	Asp533: 3.06	ND
R*	75.77	-	ND

The carbonyl group of all lamellarins can interact with the Arg364 within distance around 3.0 Å, * do not interact with Arg364, D = detected, ND = not detected

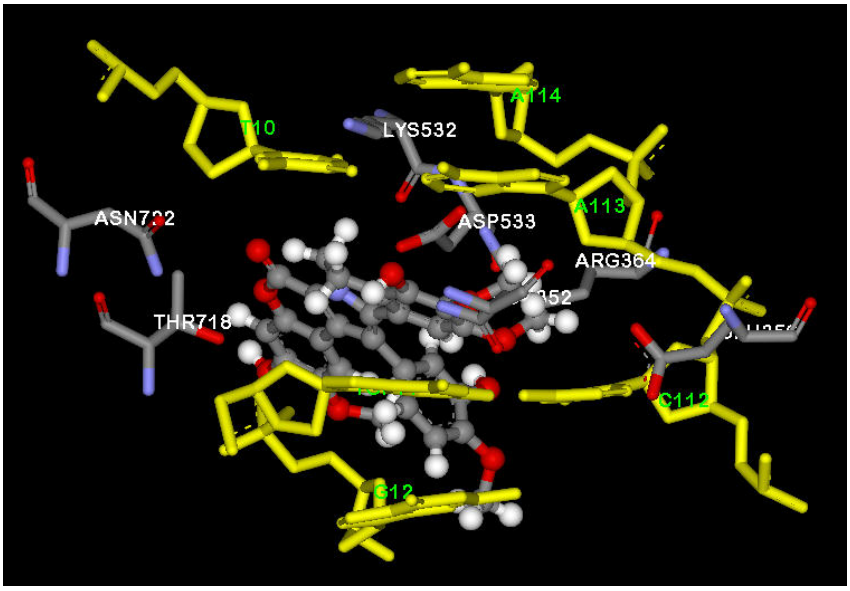


Figure 12 Orientation of lamellarin E in 7.0 Å of the binding pocket of topoisomerase I-DNA from Gold calculations. Lamellaris is represented by ball and stick style and colored by element. Amino acids and DNA are represented by normal element color stick and yellow stick style, respectively.

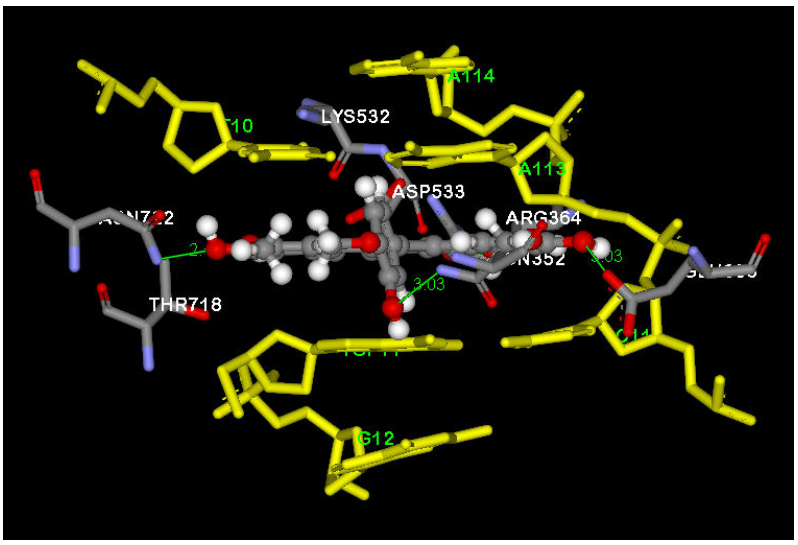


Figure 13 Orientation of lamellarin N in 7.0 Å of the binding pocket of topoisomerase I-DNA from Gold calculations. Lamellaris is represented by ball and stick style and colored by element. Amino acids and DNA are represented by normal element color stick and yellow stick style, respectively.

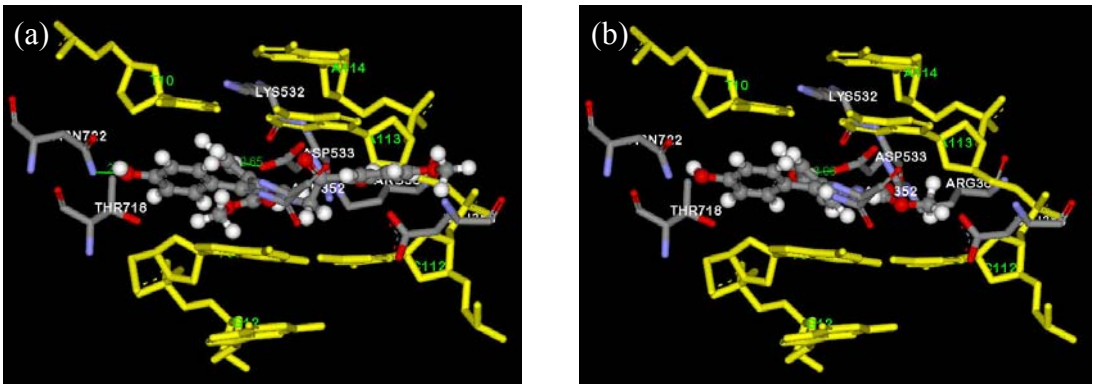


Figure 14 Orientation of lamellarin P (a) and Q (b) in 7.0 Å of the binding pocket of topoisomerase I-DNA from Gold calculations. Lamellaris is represented by ball and stick style and colored by element. Amino acids and DNA are represented by normal element color stick and yellow stick style, respectively.

2. 3D-QSAR (CoMFA and CoMSIA) Analysis

Twenty-six lamellarins used in this study can be classified into two groups that differ mainly in the nature of the C5-C6 bond in the D-ring. Eleven compounds with a saturated D-ring in Table 2 (excluding lamellarin G and acetate-containing derivatives) contained exactly the same substituents as their corresponding analogues with a C5-C6 double bond in Table 3, e.g., lamellarins C and B, lamellarins E and X, etc. Most of the compounds used in this study were naturally-occurring lamellarins, except two unnatural dehydrolamellarins J and Y, as well as three acetate-containing derivatives, that were also included in order to investigate the effect of acetylation on the cytotoxic activities of lamellarins. After alignment fo all lamellarins by using matching alignment, their orientations were represented in Figure 15. It was shown that the unsaturated lamellarins did not have a planarity of the molecule at the E and D-rings, whereas these rings of the saturated lamellarins could be aligned in the same plane of the molecule at the A, B, and C-rings. All structures in Figure 15 were used for calculations of CoMFA and CoMSIA analysis in the further step.

The cytotoxicities of lamellarins containing either a saturated or an unsaturated D-ring against the T47D and MDA-MB-231 human breast cancer cell lines are presented in Tables 1 and 2 as a negative logarithm of the IC_{50} values previously reported by Chittchang *et al.* (Chittchang *et al.,* 2009), except those of the three acetate-containing compounds. Interestingly, lamellarins with a C5-C6 single bond exhibited comparable cytotoxic activities towards both cell lines with the IC_{50} values in the same order of magnitude (Table 1). On the other hand, it is clearly demonstrated in Table 2 that all of the compounds with an unsaturated D-ring were significantly more cytotoxic to the hormone-dependent T47D cell line. Overall, lamellarins D and X were the most potent cytotoxic compounds against the T47D and MDA-MB-231 cell lines, respectively.

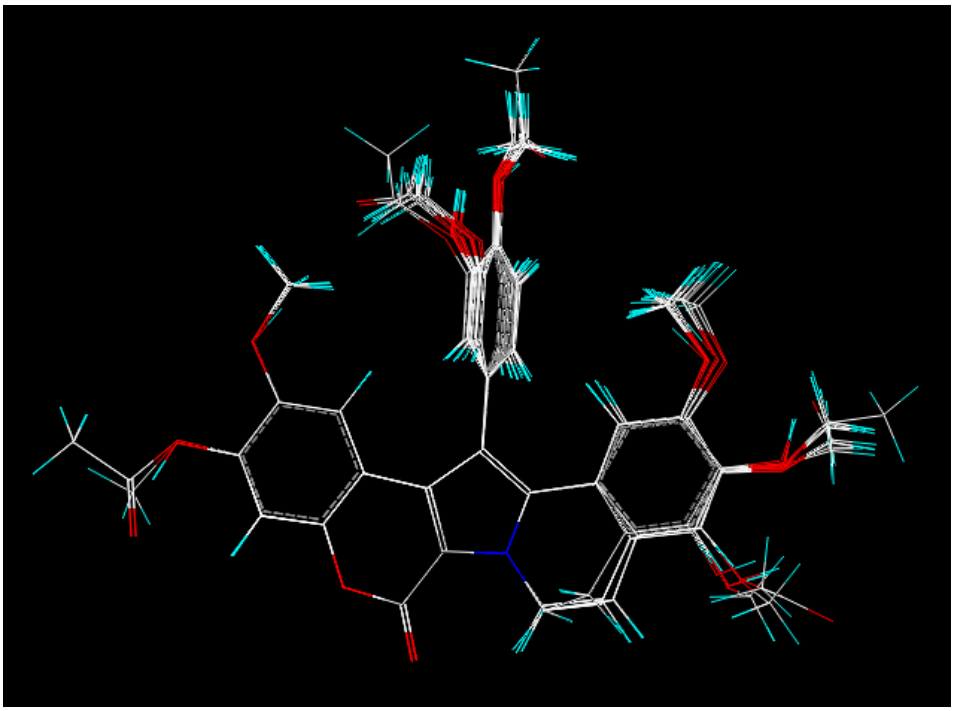


Figure 15 Structure of lamellarins, obtained from the 3D-QSAR matching alignments.

2.1 CoMFA and CoMSIA models for the cytotoxicities of lamellarins against T47D cells

Table 12 summarizes various parameters associated with the CoMFA models obtained by calculating the steric and electrostatic interactions between the aligned lamellarin molecules with each probe atom, which were subsequently correlated with their cytotoxic activities in order to identify the important interactions determining the cytotoxicities of lamellarins. All the models indicated that the changes in the steric interactions accounted for around two-thirds of the changes in the cytotoxic activities of lamellarins towards T47D cells, and the remaining 33-34% was contributed by the electrostatic interactions. Among the three CoMFA models for the T47D cell line, namely models 1 to 3 in Table 3, model 1 calculated using an sp^3 carbon as the probe atom yielded the highest predictive ability, as indicated by the r^2_{cv} (or q^2) value. However, based on the greater F-value and the $r^2_{test\ set}$, model 2 appeared

to be the best CoMFA model of the T47D cell line. Other statistical parameters were comparable among the three models.

In order to further explore whether other types of interactions also play an important role in determining the cytotoxic activities of lamellarins in both cell lines, CoMSIA was also performed, and five models were generated using different combinations of steric (St), electrostatic (El), hydrophobic (Hyd), H-bond donor (Hd), and H-bond acceptor (Ha) field types, as shown in Table 13. However, only models 7 and 8 yielded acceptable predictive ability, as indicated by the r^2_{cv} values of greater than 0.6. Additionally, whenever the hydrogen-bond donor and/or acceptor fields were included, i.e., models 9-13, the r^2_{cv} values were significantly decreased. These findings suggested that only the steric, electrostatic, and hydrophobic fields were important for predictive ability of the model. However, this did not imply that H-bond interactions were not involved in determining the cytotoxic activities of lamellarins.

Model 8 was derived from model 7 upon the exclusion of lamellarin J with the lowest activity in the training set, resulting in a model with improved r^2_{cv}, r^2, and F values, as well as a lower standard error of estimation (s). Hence, model 8 was selected as the best CoMSIA model for T47D cells. All the CoMSIA models indicated that the steric interactions accounted for approximately less than 12% of the changes in the cytotoxic activities of lamellarins, whereas the major contributions actually came from the electrostatic and hydrophobic fields.

Table 12 Summary of CoMFA results for T47D and MDA-MB-231 cell lines.

Cell line	Model	Probe atom	noc	r^2_{cv}	s-press	r^2	s	F	Steric contribution	$r^2_{test\ set}$
T47D	1	sp^3 C(+1)	6	0.717	1.143	0.963	0.414	56.144	66.0	0.466
	2	sp^3 O(-1)	6	0.659	1.255	0.965	0.405	58.959	66.9	0.628
	3	H(+1)	6	0.672	1.231	0.965	0.403	59.581	66.7	0.570
MDA-MB-231	4	sp^3 C(+1)	6	0.661	0.498	0.977	0.129	92.967	67.6	0.408
	5	sp^3 O(-1)	6	0.728	0.447	0.981	0.117	114.66	68.1	0.364
	6	H(+1)	6	0.685	0.481	0.974	0.138	81.624	66.3	0.405

noc = number of components; r^2_{cv} = cross-validated correlation coefficient; s-press = uncertainty of the prediction; r^2 = conventional correlation coefficient; s = standard error of estimation; F = F-value

Table 13 Summary of CoMSIA results for T47D and MDA-MB-231 cell lines.

Cell line	Model	Field type	noc	r^2_{cv}	s-press	r^2	s	F	Contributions	$r^2_{test\ set}$
T47D	7	St+El+Hyd	5	0.623	1.271	0.923	0.575	33.541	St = 11.8 El = 45.7 Hyd = 42.5	0.918
	8[a]	St+El+Hyd	6	0.662	1.271	0.960	0.436	48.321	St = 11.6 El = 45.8 Hyd = 42.6	0.894
	9	St+El+Hd	1	-0.090	1.907	0.325	1.500	8.688	St = 0.059 El = 0.302 Hd = 0.639	-
	10	St+El+Ha	1	-0.080	1.898	0.229	1.604	5.344	St = 0.082 El = 0.420 Ha = 0.498	-
	11	St+El+Hyd+Hd	5	0.489	1.480	0.900	0.655	25.153	St = 5.0 El = 25.6 Hyd = 27.8 Hd = 41.6	-

Table 13 (Continued)

Cell line	Model	Field type	Noc	r^2_{cv}	s-press	r^2	s	F	Contributions	$r^2_{test\ set}$
	12	St+El+Hyd+Ha	5	0.548	1.393	0.913	0.611	29.324	St = 8.0 El = 33.9 Hyd = 34.3 Ha = 23.8	-
	13	All	5	0.429	1.565	0.885	0.703	21.497	St = 4.2 El = 21.8 Hyd = 25.5 Hd = 37.2 Ha = 11.3	-
MDA-MB231	14	St+El+Hyd	6	0.608	0.536	0.954	0.184	44.956	St = 10.5 El = 48.3 Hyd = 41.1	0.595
	15[b]	St+El+Hyd	6	0.674	0.445	0.952	0.171	39.651	St = 10.2 El = 49.9 Hyd = 39.8	0.582

Table 13 (Continued)

Cell line	Model	Field type	Noc	r^2_{cv}	s-press	r^2	s	F	Contributions	$r^2_{test\ set}$
	16	St+El+Hd	6	0.131	0.798	0.954	0.184	44.670	St = 0.058 El = 0.302 Hd = 0.633	-
	17	St+El+Ha	1	-0.167	0.786	0.228	0.639	5.323	St = 0.077 El = 0.402 Ha = 0.521	-
	18	St+El+Hyd+Hd	6	0.547	0.576	0.964	0.164	57.245	St = 5.0 El = 24.8 Hyd = 20.6 Hd = 49.6	-
	19	St+El+Hyd+Ha	6	0.424	0.650	0.939	0.211	33.540	St = 7.6 El = 32.9 Hyd = 29.0 Ha = 30.6	-

Table 13 (Continued)

Cell line	Model	Field type	Noc	r^2_{cv}	s-press	r^2	s	F	Contributions	$r^2_{test\ set}$
	20	All	6	0.458	0.630	0.957	0.177	48.572	St = 4.1	-
									El = 19.0	
									Hyd = 18.7	
									Hd = 44.4	
									Ha = 13.8	

[a] Elimination of lamellarin J and [b] elimination of lamellarin Y

The predictive ability of the selected CoMFA and CoMSIA models was determined using six lamellarin compounds as test set. The CoMFA model was first considered, and the -log IC_{50} values predicted using model 2 are plotted against the experimentally-determined numbers (Figure 16 (a)). The prediction results were satisfactory for most of the compounds in the test set with an exception of lamellarin α. Exclusion of this compound from the test set significantly improved the $r^2_{test\ set}$ value for model 2 to 0.985. Similarly, the selected CoMSIA model 8 is also validated, and the results are presented in Figure 16 (b). The model appeared to overestimate the -log IC_{50} values of lamellarin α and lamellarin M, while the predicted value of lamellarin K was lower than the experimental -log IC_{50}. Nevertheless, it can be seen that the CoMFA model 2 could be used to predict the cytotoxic activity of most lamellarins in the test set towards T47D cells.

2.2 CoMFA and CoMSIA contour maps for the cytotoxicities of lamellarins against T47D cells

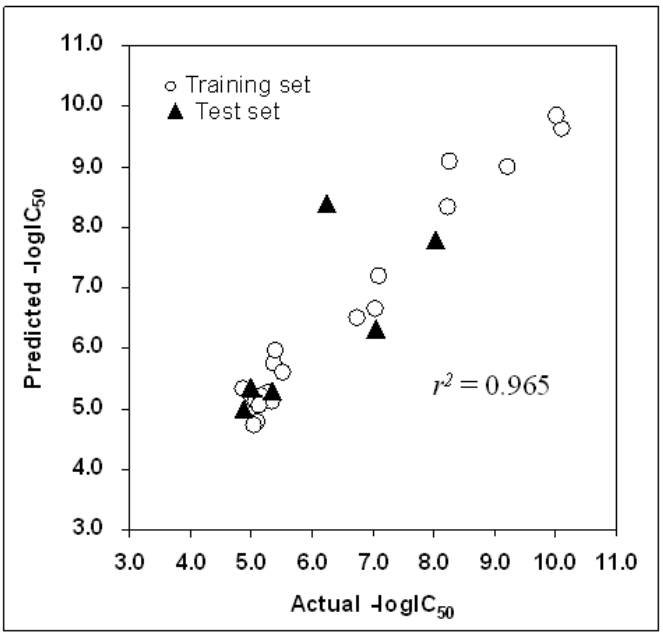
The CoMFA and CoMSIA contour maps were created using the best CoMFA and CoMSIA models for each cell line, as shown in Figures 17 and 18. The resulting contours on these maps not only provided detailed understanding of the binding pockets, but also highlighted the key structural features required for the biological activity of the molecule being considered. The green and yellow regions represented the areas where steric bulks would enhance and diminish the biological activity, respectively. On the other hand, the blue and red contours represent the regions where electropositive and electronegative group were favored for the cytotoxic activity. In order to avoid any ambiguities due to contour overlapping, all the contours were displayed in a transparent style. Additionally, lamellarin D and lamellarin X, which were the most active molecules against T47D and MDA-MB-231 cell lines, respectively, were used as the reference molecules in the corresponding contour maps.

Interestingly, Figure 17 (a) clearly shows several steric and electrostatic contours generated using CoMFA model 2 for the T47D cell line. For the pentacyclic

core, all the contours were concentrated around the two rightmost D- and E-rings (Figure 17 (a)). There were green and blue regions near the C9 group. Red contours also occupy the space between the oxygen atom of the methoxy group at C9 and the hydroxyl group at C8. These suggested that oxygen atoms at C8 and C9 were important for cytotoxic activity. The steric and electropositive group at C9 might play the important role for the activity. Considering two couples of lamellarins (lamellarin Y compared with lamellarin U and dehydrolamellarin Y compared with lamellarin α), it was shown that the cytotoxic activities of the lamellarins were decreased when C9-hydroxy group was replaced by methoxyl group. The reasons for this is due to the fact that lamellarin U and lamellarin α were classified in the test set. This might be the reason why the model 2 and 8 showed a poor predictability in -log IC_{50} value of lamellarin α. Interesting contours in the E-ring were blue contours at the hydrogen atom and methyl group at C7 and red contour located around oxygen atom at C7. These might indicate that occupancy of hydroxyl and methoxyl group was more cytotoxic than hydrogen atom at this position.

In terms of the orthogonal ring or F-ring, green contour were found around this ring. This finding indicated that the orthogonal ring of lamellarin was essential for the activity. Additionally, the red contour was found near oxygen atom and acetate group at C14 of the lamellarin. This contour showed that oxygen atom and acetate group at C14 might be important for the activity. The blue contour regions between C13 and C21 groups suggested that some electropositive groups in these areas might be advantageous for the activity.

(a)



(b)

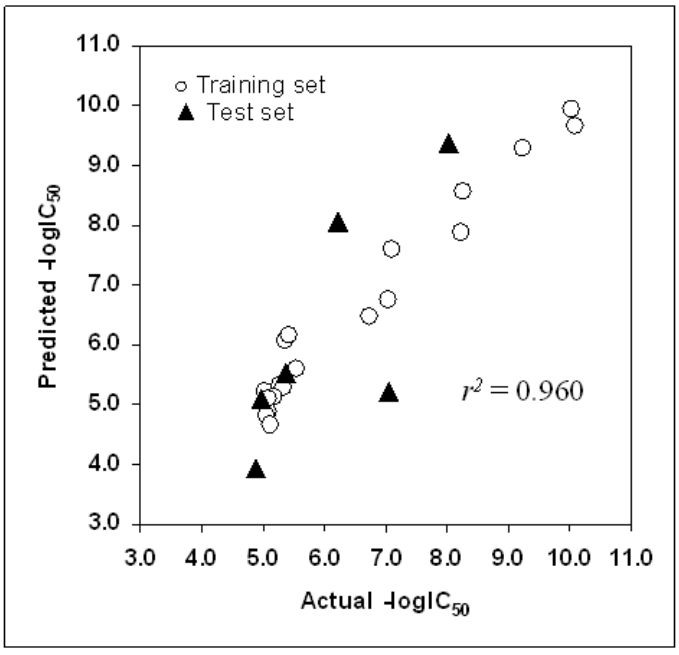


Figure 16 Plot of the predicted versus actual cytotoxic activity of lamellarins towards T47D cells. The predicted values were obtained from non-cross-validated CoMFA model 2 (a) and CoMSIA model 8 (b) for all the compounds in both the training (○) and test (▲) sets.

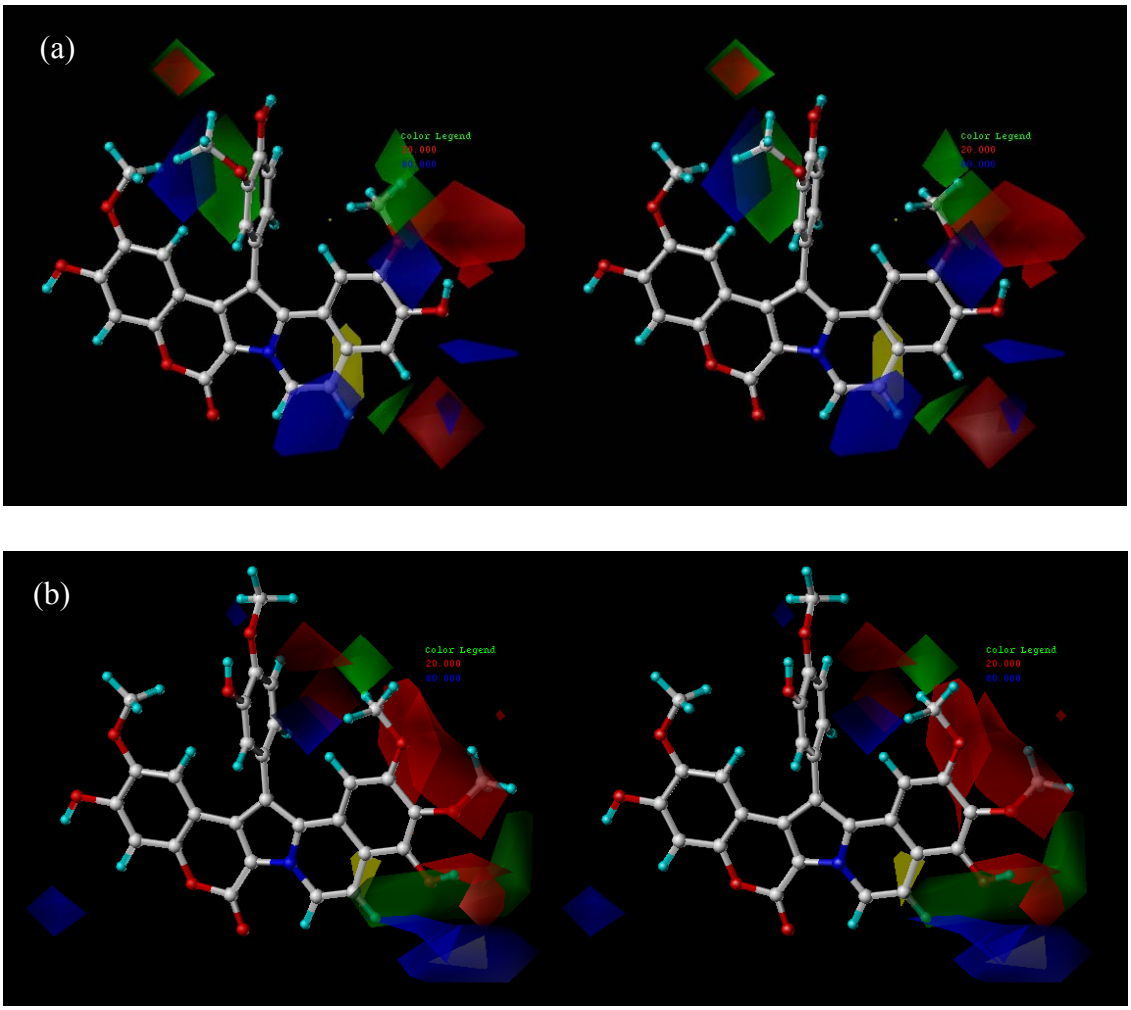


Figure 17 Stereoviews of steric and electrostatic stdev*coeff contour maps, obtained from CoMFA model 2 for T47D cytotoxicity (a) and CoMFA model 5 for MDA-MB-231 cytotoxicity (b). Lamellarin D and X are presented inside the fields as ball and stick display style of T47D CoMFA contour map and MDA-MB-231 CoMFA contour map, respectively. Sterically favored and unfavored areas are shown by green and yellow regions, respectively. Electropositive and electronegative areas are shown by blue and red regions, respectively.

Finally, yellow contour appeared behind the plane of the molecule near the C5-C6 double bond while blue and green contours occupy the space around the D-ring. The yellow contour showed that the binding pocket for this region has a limited volume. This prediction, along with the superior cytotoxicities of the lamellarins

containing a C5-C6 double bond (Tables 3), suggested that these compounds assume a structure that could fit into the binding pocket better than their counterparts with a single bond. It was observed in our preliminary molecular modeling studies that the lamellarins with a saturated D-ring were twisted, whereas the presence of the C5-C6 double bond leads to a more planar molecule that is likely to intercalate into the topoisomerase I-DNA complex more easily (Chittchang *et al.,* 2009). The blue and green contours obtained from the twisted bond of unsaturated lamellarins which can not align at the same position as shown in Figure 15.

In addition to the steric and electrostatic fields considered in CoMFA, CoMSIA helps define the contribution of the explicit hydrophobic interactions to the binding affinities of lamellarins, as shown in Figure 18. Orange and white contours referred to areas where hydrophobic groups would increase and decrease the biological activity, respectively. Since CoMSIA steric and electrostatic contours were shown in more detail than those of the CoMFA models in some positions, all three field contributions were discussed. The steric and electrostatic contours from the best CoMSIA models for the T47D cell line are shown in Figure 18 (a). There were green region at C9, and white region near oxygen atom at C9. Both contours indicated that oxygen atom at this position was essential for potent cytotoxic activity. For green region, the same reason has been dicussed in CoMFA part. The presence of steric (yellow contour near the hydrogen atom), electrostatic (blue contour near the hydrogen atom and methyl group, and red contour near the oxygen atom), and hydrophobic (white contour) fields at C8 revealed that a non-steric and electropositive group around this area were required for better activity. This suggestion has been reported by Ishibashi *et al.* and Chittchang *et al.* (Ishibashi *et al*, 2002; Chittchang *et al.,* 2009) that hydroxyl group at C8 of lamellarin D was important structure requirements for the activity.

Furthermore, yellow and green contours near hydroxyl and methoxyl group at C7 led to the idea that bulky group at this position would increase the activity, but the size of this group should not be too large. Additionally, there were blue regions near hydroxyl and methoxyl group at C7, a red region located around

oxygen atom at C7, and a white region at C7. It might indicate that a steric and electropositive oxygen-containing group such as hydroxyl and methoxyl group would be beneficial for the cytotoxicity. More hydrophobic fields (white and orange contours) were found around the C5-C6 bond. The white contour is shown at CH_2 around C5-C6 bond. The orange contour located around C5-C6 bond. Both fields were shown around this area because of the planarity difference between single and double C5-C6 bond which could not align into the same plane. These results also confirme that structural requirement is based on the double bond rather than on the single bond. The red contours located between the oxygen atom of C13 and C14 positions, and near the carbonyl of the acetate group around C20 revealed the importance of the oxygen atom of these areas for cytotoxicity.

The two green regions, which were located at C13 and C21, suggested that the substituents at these positions should be bulky groups. However, it was interesting to note that the green contour was shown at C21 only obtained due to orientation of the methoxyl group in space (see Figure 15).

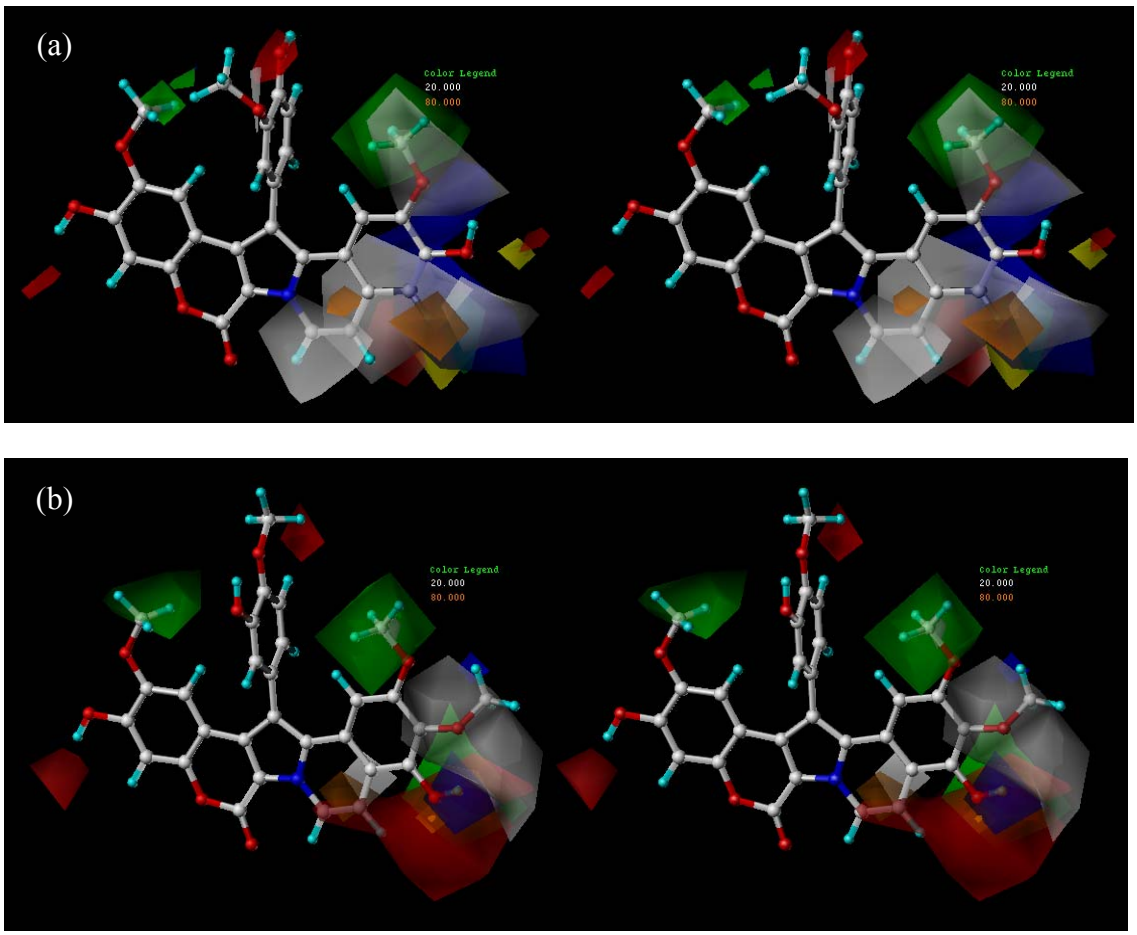


Figure 18 Stereoviews of steric, electrostatic, and hydrophobic stdev*coeff contour maps, obtained from CoMSIA model 8 for T47D cytotoxicity (a) and CoMSIA model 15 for MDA-MB-231 cytotoxicity (b). Lamellarin D and X are presented inside the fields as ball and stick display style of T47D CoMFA contour map and MDA-MB-231 CoMFA contour map, respectively. Sterically favored and unfavored areas are shown by green and yellow regions, respectively. Electropositive and electronegative areas are shown by blue and red regions, respectively. Hydrophobically favored and unfavored areas are shown by orange and white regions, respectively.

2.3 CoMFA and CoMSIA models for the cytotoxicities of lamellarins against MDA-MB-231 cells

The CoMFA and CoMSIA models for the MDA-MB-231 cell line were generated using the same compounds in the training and test sets as those used for the T47D cell line. The statistical parameters associated with the CoMFA models 4-6 for the MDA-MB-231 cell line are shown in the Table 11. Apparently, model 5 obtained using an sp^3 oxygen as the probe atom yielded not only the highest predictive ability with an r^2_{cv} value of 0.728, but also the highest F-value. Even though the other statistical parameters were not significantly different among the three models, model 5 still appears to be the best CoMFA model for the MDA-MB-231 cell line, which indicates that steric interactions were the major factor contributing approximately 68% of the changes in the cytotoxic activities of lamellarins towards the MDA-MB-231 cells.

With the sequential additions of the hydrophobic and hydrogen bonding interactions, seven CoMSIA models (models 14-20) were subsequently constructed for the MDA-MB-231 cell line, as shown in Table 12. Based on the r^2_{cv} values, the models involving only the steric, electrostatic, and hydrophobic fields, i.e., models 14 and 15, showed good predictive ability. Model 15 resulted from the exclusion of lamellarin Y with the lowest cytotoxic activity in the training set for model 14, giving rise to a significantly increased r^2_{cv} value, even though the conventional r^2 and $r^2_{test\ set}$ values for both models were not significantly different. Thus, model 15 was chosen as the best CoMSIA model for the MDA-MB-231 cell line. The contributions from steric, electrostatic, and hydrophobic interactions were determined to be 10.2%, 49.9%, and 39.8%, respectively, similar to the values obtained from the CoMSIA model 8 for the T47D cell line.

The correlations between the actual -log IC_{50} data of lamellarins against the MDA-MB-231 cell line and the values predicted based on the CoMFA model 5 and CoMSIA model 15 are shown in Figures 19 (a) and 19 (b), respectively. In the case of the CoMFA model 5 (Figure 19 (a)), lamellarin L was the only compound in

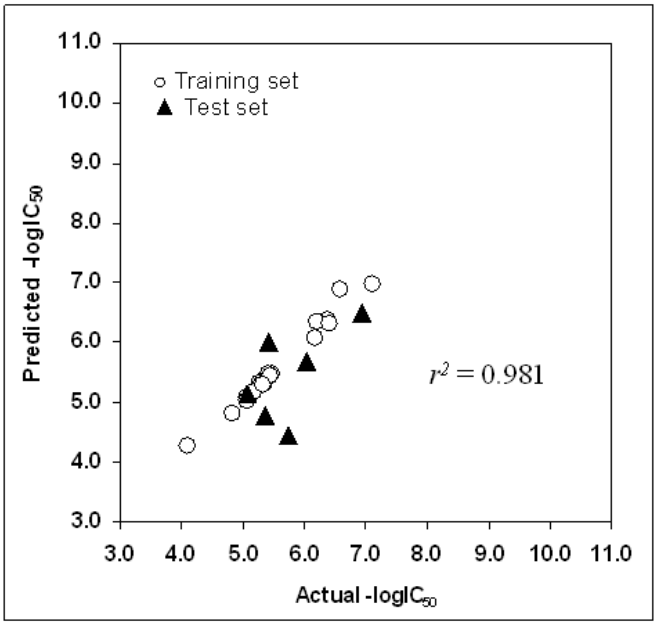
the test set, of which the predicted cytotoxicity deviated from the actual value by greater than one log unit, and the exclusion of this compound significantly increased the $r^2_{test\ set}$ from 0.364 to 0.612. A similar degree of deviation was also observed with the -log IC_{50} value of lamellarin α predicted using the CoMSIA model 15 (Figure (5 (b)). Nevertheless, these results indicated that both the CoMFA model 5 and the CoMSIA model 15 could be used to predict the cytotoxicities of lamellarins towards the MDA-MB-231 cell line.

2.4 CoMFA and CoMSIA contour maps for the cytotoxicities of lamellarins against MDA-MB-231 cells

The steric and electrostatic maps of the CoMFA analysis for the MDA-MB-231 cell line are represented in Figure 17 (b). Blue and green regions were found around hydroxy and methoxy groups at C9, and red regions located near the oxygen atom at C8 and C9. These might indicate that a steric and electropositive group at C9 would increase the cytotoxicity. However, there were only two lamellarins (Y and dehydrolamellarin Y) which have hydroxyl group at C9. Hence, it might suggest only that the oxygen atom at C8 and C9, and electropositive group at C9 would play an important role in increasing cytotoxic activity. The green, yellow, and blue contour near hydroxyl at C7 showed that C7 area preferred a steric and electropositive group but it should not be too large. These results supported by the experimental data reported in Table 2 and 3. Lamellarin K, E, F, M, X, and ε with hydroxyl group at C7 were more cytotoxic against MDA-MB-231 cell line than lamellarin C, T, I, B, W, and ζ, respectively. There was a yellow contour located behind the plane of the D-ring near the C5-C6 bond. It revealed the importance of a double bond at C5-C6 to enhance the MDA-MB-231 cytotoxic activity similar to the T47D activity. The red contours near the oxygen atom at C14 suggested that the oxygen atom at C14 would play an important role in increasing cytotoxic activity. The last contour, blue contour, near C20 indicated that occupancy of an electropositive group at this position would help the cytotoxic activity.

The steric, electrostatic, and hydrophobic contour maps of CoMSIA model for the MDA-MB-231 cell line is displayed in Figure 18 (b). Green and blue contours were found near C9. Additionally, white region was found between C8 and the oxygen atom at C9. These results can be concluded that an electropositive group at C9 enhances the cytotoxic activity. There was a green contour between C7 and C8, yellow and blue contours at C7, and red contour around C7 and the C5-C6 bond. These regions suggested that it preferred a steric group disfavoring hydrophobic interactions at C8. However, there was only two couple of lamellarins which compare substitution of hydroxyl group at C8 with methoxy group, (lamellarin L compared with lamellarin U and lamellarin N compared with lamellarin α). Three of them (L, U, and α) were the test set. At C7, a steric and electropositive oxygen-containing group which was not too large might require for increasing the activity. Both hydrophobic fields (white and orange contours) were found near C5-C6 bond. The importance of the oxygen atom also were shown near the acetate group at C14 and C20 by representing of the red regions. The last contour was green region at C21 represented the methoxyl group at this position.

(a)



(b)

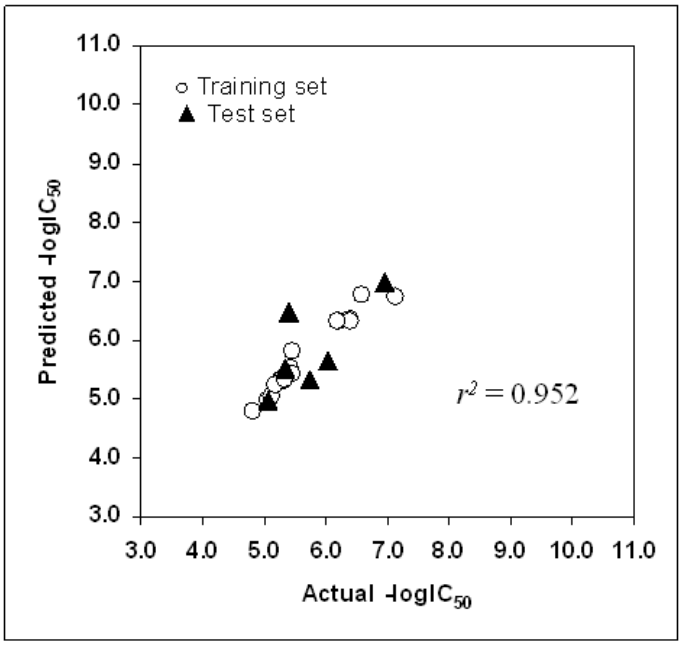


Figure 19 Plot of the predicted versus actual cytotoxic activity of lamellarins against MDA-MB-231 cells. The predicted values were obtained from non-cross-validated CoMFA model 5 (a) and CoMSIA model 13 (b) for the compounds in both the training (○) and test (▲) sets.

2.5 The common structural requirement of lamellarins as a binding pocket of the T47D compared with the MDA-MB-231 cell lines

In summary, the T47D and MDA-MB-231 receptor binding sites model are proposed as shown in Figures 20 (a) and 20 (b), respectively. By using the combination of CoMFA and CoMSIA results, the structural requirement of lamellarins at the binding pockets of the both cell lines were shown by common carbon atom of the lamellarin skeleton. Figure 20 (a) reveals that the double bond at C5-C6 bond, the electropositive groups at C9, C13, and C21, a non-steric and electropositive group at C8, and C7 should be a steric and electropositive oxygen-containing group are required for cytotoxic acitivty. Based on Figure 20, the 3D-QSAR analysis reveals that the common structure of lamellarins around the binding pocket for both cell lines was similar in many positions e.g. at C7, C9, and C5-C6 bond. More importantly, this method showed some different structural requirements of the lamellarin skeleton for both cell lines such as a non-steric and electropositive group at C8 in case of T47D cells. For the MDA-MB-231 cell line, on the other hand, C8 area seemed to require a steric group disfavoring hydrophobic interactions. Additionally, the electropositive groups at C13, and C21 in case of T47D cell line were not important for MDA-MB-231 cell line. An electropositive group at C20 in case of MDA-MB-231 cell line was not shown any significant for T47D cell line. A steric group disfavoring hydrophobic interactions requirement at C8 may account for potent cytotoxic activity of lamellarin M, X, and ε against MDA-MB-231 cell line as shown in Table 3.

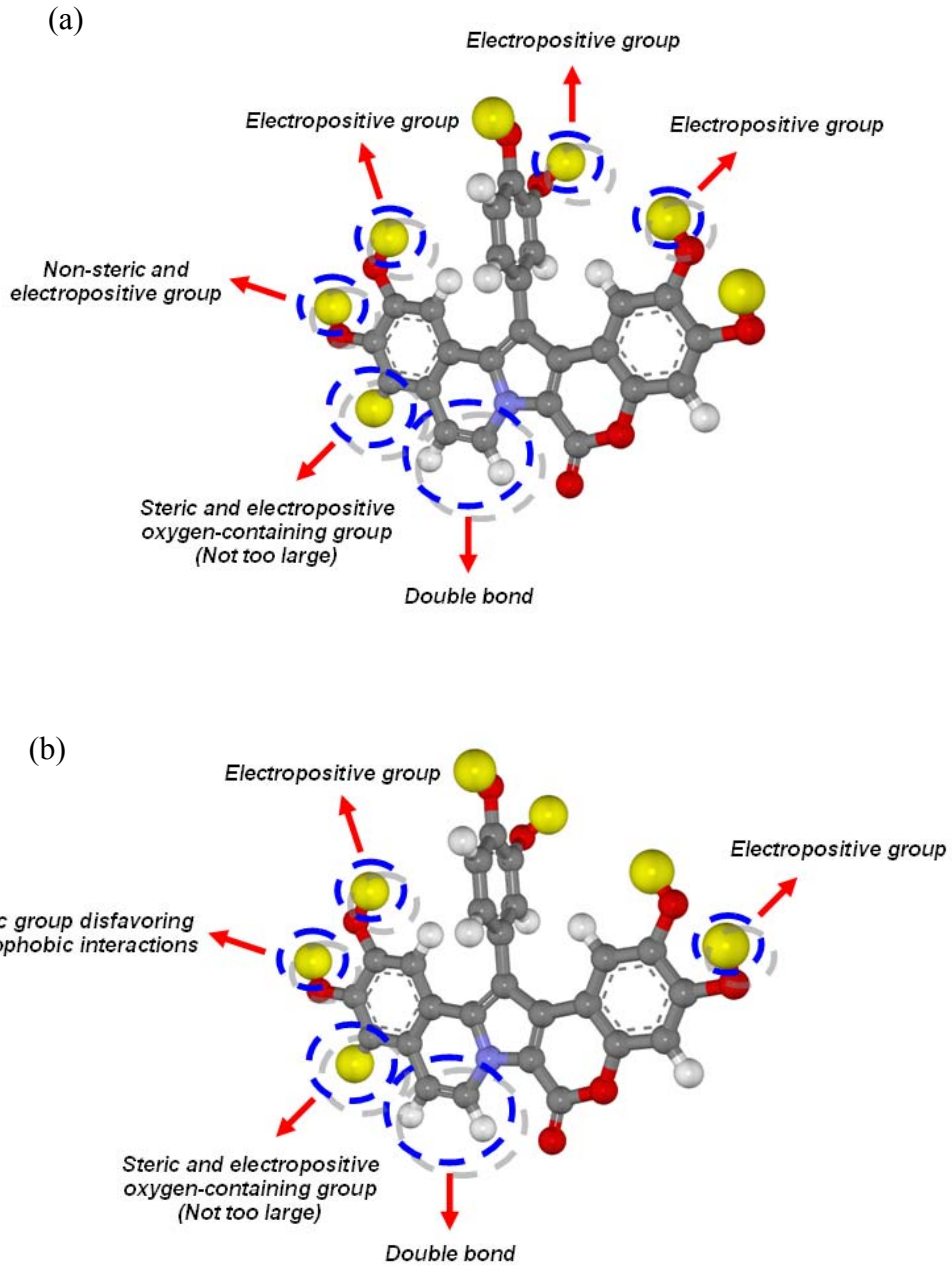


Figure 20 Structural requirements of lamellarins for T47D receptor binding site obtained from the combination of the CoMFA and CoMSIA contour maps (a) and for MDA-MB-231 receptor binding site obtained from the combination of the CoMFA and CoMSIA contour maps (b).

3. Receptor-Independent (RI) 4D-QSAR Analysis for Cytotoxicity of Lamellarins Against T47D breast cancer cells

Optimized RI-4D-QSAR models were constructed for each of the eight trial alignments listed in Table 4. Alignments 1, 2, 4, and 7 contain atoms from two rings (A and B), (B and C), (C and D), and (C and F), respectively. Alignment 5 and 6 only contain atoms from ring E and ring F, respectively. Only two alignments, 3 and 8, distribute the three-ordered atoms across three rings namely rings A, B, and C for alignment 3 and rings A, E, and F for alignment 8. The r^2 and r^2_{cv} values from the best corresponding RI-4D-QSAR models of each alignment are given in Table 14. Alignment 1 yields the poorest fits with $r^2 = 0.964$ and $r^2_{cv} = 0.929$. The differences among r^2 and r^2_{cv} of the remaining alignments are quite small, or the alignment of lamellarin is not significant to the 4D-QSAR model. However, based on the greater r^2 (0.999) and r^2_{cv} (0.998), alignment 3 appears to be the best alignment for 4D-QSAR analysis of lamellarin data set.

The optimum number of descriptors in a model using alignment 3 is determined by monitoring when r^2_{cv} becomes effectively constant, or decreases, with increasing model size. Figure 21 is a plot of the number of descriptor terms in an optimized alignment 3 model versus the corresponding r^2 and r^2_{cv}. An inspection of Figure 21 reveals that the maximum number of descriptor terms in the RI-4D-QSAR model providing additional fit to the training set data is three. There is no meaningfully enhanced model fitting by including more than three descriptor terms. Thus, the optimized RI-4D-QSAR model for the 25 lamellarins generated from alignment 3 is given by equation 14.

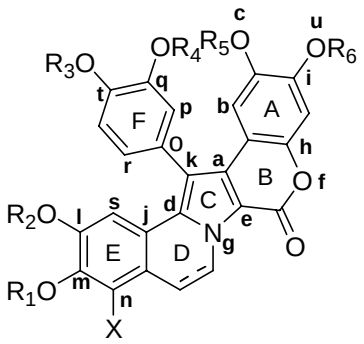
$$-\log IC_{50} = 5.14 + 16.90GC1\ (-5,6,2,np) - 56.33GC2\ (-3,4,-5,any) + 64.62GC3\ (-1,5,0,np)$$

$$r^2 = 0.971,\ r^2_{cv} = 0.947,\ n = 25 \qquad \textbf{(14)}$$

GCi(x, y, z, X) is the i^{th} GCOD descriptor term located at (x, y, z) in the reference grid cell and alignment space, and having the X type IPE as defined in Table 4. Figure 22

is a plot of the predicted, using equation 14, versus actual –log IC_{50} values. All of the predicted –log IC_{50} values are within ± 1 log unit of the corresponding observed values, and there are no outliers.

Table 14 r^2 and r^2_{cv} of trial alignment used in constructing the 4D-QSAR model.



Alignment	First atom	Second atom	Third atom	r^2	r^2_{cv}
1	a	b	c	0.964	0.929
2	d	e	f	0.996	0.992
3	g	h	i	0.999	0.998
4	j	d	k	0.997	0.995
5	l	m	n	0.995	0.984
6	o	p	q	0.997	0.993
7	k	o	r	0.999	0.997
8	s	t	u	0.999	0.995

Two GCODS (GC1 and GC3) of equation 14 correspond to nonpolar atom occupancy that increases potency because they both have positive regression coefficients with values of 16.90 and 64.62, respectively. GCOD GC2, having an 'any' IPE type, has a negative regression coefficient with value of -56.33. Consequently occupancy of the GC2 site by any type of atom will lead to a decrease in the potency of anti-breast cancer activity of the corresponding lamellarin. From an analysis of equation 14 it is found that the any IPE type at (-3,4,-5) has about three

times more of a negative effect upon -log IC_{50} than the positive effect of the nonpolar IPE type at (-5,6,2), and about the same, but opposite effect on –log IC_{50} as the nonpolar IPE type at (-1,5,0). None of the best models from GFA model optimization contain GCOD descriptors which deal with specific interactions like hydrogen bonding.

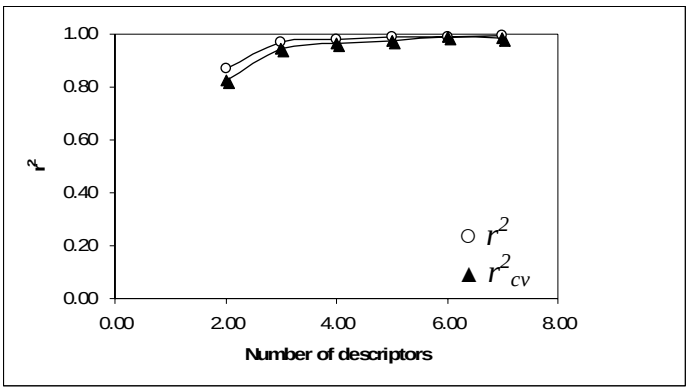


Figure 21 Plot of the number of 4D-QSAR model descriptors versus r^2 *and* r^2_{cv}.

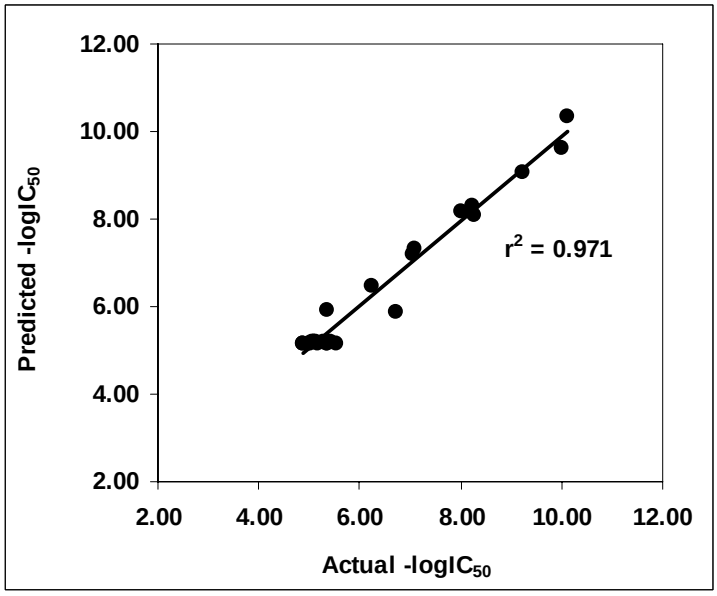


Figure 22 Predicted and actual -log IC_{50} derived by 4D-QSAR model of 25 lamellarins data set.

In order to further search for pharmacophore sites which are specifically associated with lamellarins exhibiting high cytotoxic activity, an additional RI-4D-QSAR analysis was carried out. The training set of this study was limited to the six lamellarins (D, M, N, X, ε, and Dehydrolamellarin J of Table 3) that have the highest –log IC_{50} values, and are not redundant in their structural features. The RI-4D-QSAR models were constructed and optimized by using the same methodology and alignment used to build equation 14. The best RI-4D-QSAR model from this small high activity data set of lamellarins is given by equation 15.

$$-\log IC_{50} = 10.31 - 50.52GC1\ (-1,1,-6,any) + 1.58GC2\ (-1,4,-6,np)$$

$$r^2 = 0.997,\ r^2_{cv} = 0.984,\ n = 6 \qquad \textbf{(15)}$$

The regression coefficients of the descriptors of equation 15 suggest placing any type of atom at (-1,1,-6) has about 30 times more negative effect on –log IC_{50} than the positive gain by locating a nonpolar atom or group at (-1,4,-6). Certainly equation 15 is, or borders upon, being an over-fit model. However, equation 15 and its 3D-pharmacophore are only used as adjuncts to equation 14 and its 3D-pharmacophore. That is, equation 15 is being used to provide a higher-resolution view of the SAR features most characteristic of the high activity lamellarins of the training set. Equation 14 and its 3D-pharmacophore are used outside that context.

The 3D-pharmacophores defined by equation 14 and equation 15 are shown in Figure 23 (a) and 23 (b), respectively. The reference structure superimposed on each of the 3D-pharmacophores in these two figures is the predicted active conformation of the most active compound [lamellarin D] based upon equation 14. The red spheres in Figure 23 represent those GCOD descriptor terms which have negative regression coefficients. Correspondingly, the blue spheres delineate GCOD descriptors having positive regression coefficients in the corresponding best RI-4D-QSAR equation. From an inspection of Figure 23 (a), a red sphere near C8 and C9 groups specifies a pharmacophore site where occupancy by any type of atom, or group, decreases potency since the corresponding regression coefficient -56.33. Two blue spheres are

found near C14 and C13 suggesting that substitution of nonpolar groups to occupy one or both sites is conducive to increase the cytotoxic activity of the lamellarins.

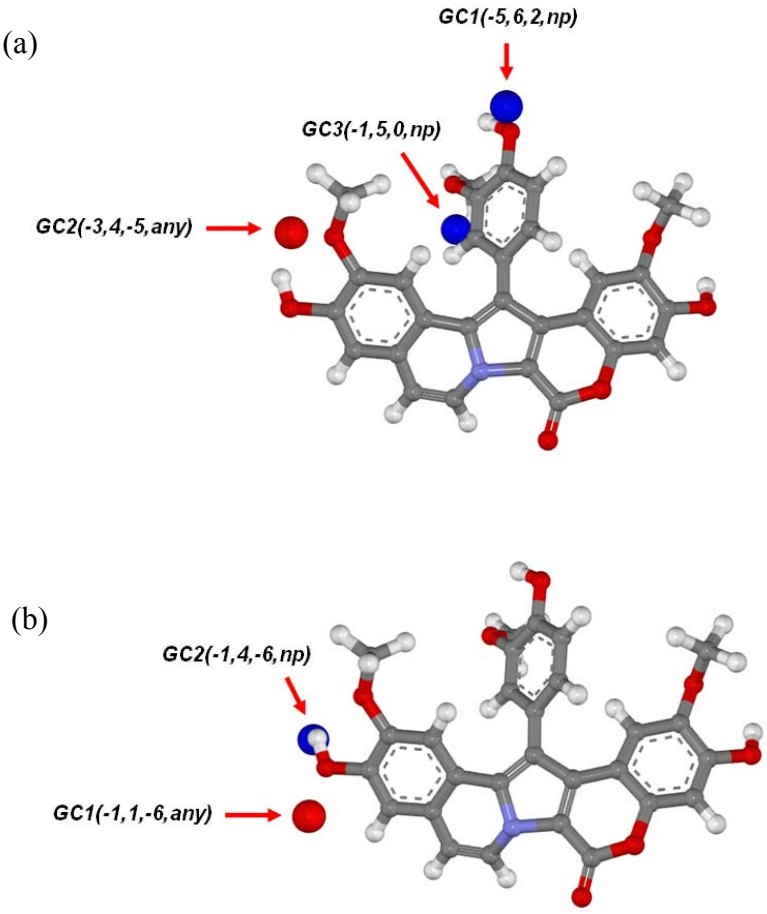


Figure 23 3D-pharmacophores of 25 lamellarins training set (a) and high activity lamellarin training set (b) 4D-QSAR models relative to predicted active conformation of the most active compound (lamellarin D). Red spheres refer to negative regression coefficient and blue spheres refer to positive regression coefficient.

The 3D-pharmacophore of the high activity model, equation 15, is represented by one red sphere GCOD located around C7 and C8, and a blue sphere (GCOD) positioned near C8 and C9. The most active compounds of the potent lamellarins seemingly achieve most of their additional -log IC_{50} potency, as compared to the less

potent lamellarins, by not having any atoms or groups at (-1,1,-6) in contrast to increasing occupancy by nonpolar atoms or groups at the GCOD located at (-1,4,-6). The 30:1 ratio of not occupying the GCOD at (-1,1,-6) as compared to having a nonpolar atom or group at (-1,4,-6) is consistent with the relative binding energy contributions of an intermolecular hydrogen bond involving the OH near (-1,1,-6) as compared to a hydrophobic binding effect due to the methyl of the methoxy group near (-1,4,-6) as is shown in Figure 23 (b).

Overall, the high activity compounds are seemingly distinguished from one another in equation 15 by their ability to form an intermolecular hydrogen bond where the hydrogen bond acceptor atom in the receptor is expected to be near (-1,1,-6). Some additional increase in $-\log IC_{50}$ can also be realized by having a hydrophobic substituent group of the ligand occupying the (-1,4,-6) site. The two GCODs of equation 15 may be a higher resolution representation of the single GC2(-3,4,-5,any) GCOD found in equation 14.

In order to evaluate the possible roles of ligand molecular weight (MW) and partition coefficient (log *P*) on cytotoxic potency, $-\log IC_{50}$, both of these properties were included as part of the trial basis set of descriptors in a GFA model optimization study. Unfortunately, no GFA model optimization could be realized. An inspection of the MW and log *P* values of the training set compounds revealed that three lamellarins (K-triacetate, χ-triacetate, and U-diacetate) have very high MWs relative to the other training set compounds, and one lamellarin (F) has a very low log *P* value relative to the other lamellarins. These four lamellarins were removed to form a revised training set, and GFA model building and optimization repeated for this 21 compound training set in the same manner as employed in developing equation 14 and equation 15. Ten best models were determined from the GFA optimization, and the residuals of fit cross-correlations between each pair of these models are given in Table 15. All pairs of the top-ten models have residuals of fit highly correlated to one another, with value of at least 0.70, indicating these 10 models are all very nearly the same model. Therefore, the best of the ten models was selected as the preferred RI-4D-QSAR model for this training set and is given by equation 16.

$$-\log (IC_{50}) = 10.31 - 4.77GC1(-2,1,-6,np) - 33.91GC2(-3,4,-5,any) - 8.12GC3(3,3,2,np)$$

$$r^2 = 0.935,\ r^2_{cv} = 0.890,\ n = 21 \qquad \textbf{(16)}$$

Table 15 Correlation between 10 models of 21 lamellarins training set.

Model no.	1	2	3	4	5	6	7	8	9	10
1	1.00									
2	0.91	1								
3	0.83	0.97	1							
4	0.93	0.93	0.90	1						
5	0.92	0.99	0.97	0.93	1					
6	0.83	0.97	0.99	0.90	0.97	1				
7	0.93	0.93	0.90	0.99	0.93	0.90	1			
8	0.94	0.79	0.72	0.93	0.80	0.73	0.93	1		
9	0.91	0.98	0.95	0.91	0.98	0.95	0.91	0.80	1	
10	0.84	0.97	0.99	0.90	0.97	0.99	0.90	0.73	0.96	1

Figure 24 is a plot of the observed versus the predicted –log IC_{50} values determined from using equation 16. The 3D-pharmacophore embedded in the RI-4D-QSAR model given by equation 16 is shown in Figure 25 with lamellarin D again the reference compound. All three GCOD descriptors of equation 16 correspond to pharmacophore sites where an increasing occupancy decreases activity. One pharmacophore site, (-3, 4, 5, any) from equation 16, is identical to a site from equation 14, while the pharmacophore site at (-2, 1,-6, np) from equation 16 is very close to the pharmacophore site of equation 15 located at (-1,1,-6, any) as can be seen by comparing Figure 25 to Figure 23. The third pharmacophore site of equation 16 located at (3,3,2), which predicts the occupany of nonpolar groups to decrease –log IC_{50}, is unique to this model as compared to equation 14 and equation 15. This new GCOD descriptor term of equation 16 and the decrease in r^2 and r^2_{cv} may be an indication of a significant pharmacophore-site dependence on one, or more, of the four lamellarins eliminated from the training set used to build equation 16 and its corresponding 3D-pharmacophore.

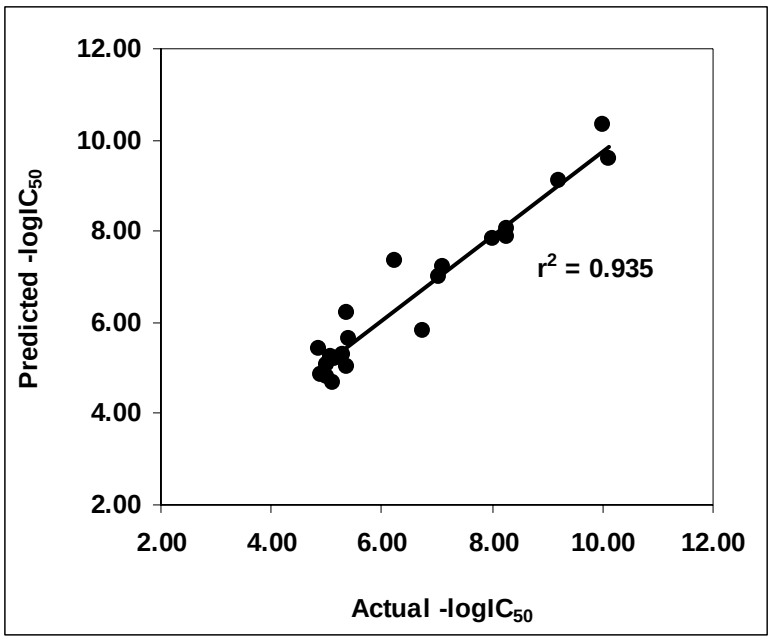


Figure 24 Predicted and actual –log IC_{50} derived by 4D-QSAR model of 21 lamellarins data set.

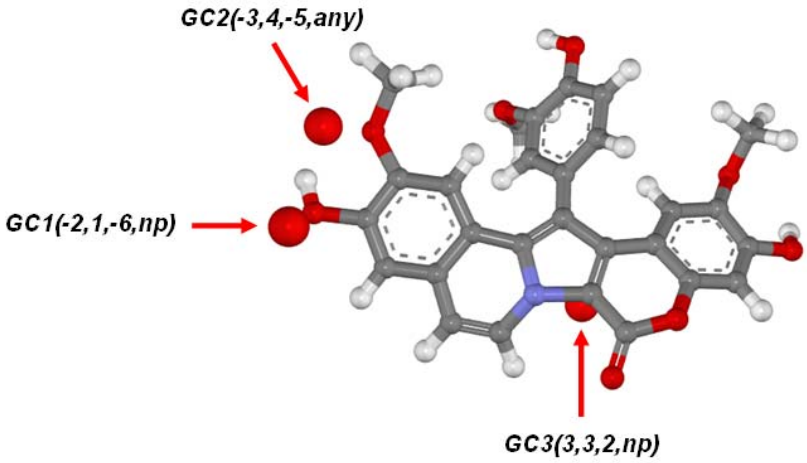


Figure 25 3D-pharmacophores of 21 lamellarins training set relative to predicted active conformation of the most active compound (lamellarin D). Red spheres refer to negative regression coefficient and blue spheres refer to positive regression coefficient.

An attempt was made to further explore if log *P* plays a role in the structure-activity relationship of the lamellarin training set by forcing overfitting in the GFA model building and optimization process. The log *P* descriptor was the only non-GCOD descriptor added to the trial basis set (descriptor pool) at step 5 of 4D-QSAR methodology. Overfit RI-4D-QSAR model were permitted under the same methodology, same alignment, and for all lamellarins in training set as used to develop equation 14. None of the 10 most significant overfit 4-term or 5-term RI-4D-QSAR models contained a log *P* descriptor term. Therefore, it was concluded that molecular lipophilicity is not a major contributing factor in the specification of the cytotoxic activity for the lamellarins studied in this analysis.

4. 4D-fingerprints Analysis for Cytotoxicity of Lamellarins Against T47D breast cancer cells

4D-fingerprint virtual high throughput screens permit a larger range of chemistry diversity to be assayed more quickly than RI-4D-QSAR models. In this study 4D-fingerprints models were generated using all 25 of the lamellarins in the training set. Two types of 4D-fingerprints can be constructed: those 4D-fingerprints explicitly dependent upon a particular alignment, and absolute 4D-fingerprints which are alignment independent. Absolute 4D-fingerprints were used in this analysis to maximize the range of lamellarin chemical diversity that could be reasonably screened. That is, a 4D-fingerprint screening model built independent of alignment is more general than its corresponding alignment-dependent screen, but at the cost of being somewhat less significant in its fit to the training set data.

The absolute 4D-fingerprints were derived for each of the 25 training set lamellarins using the modeling methodology given above in the *Methods* section. These 4D-fingerprints formed the trial basis set for model building. No non-4D-fingerprints were added to this trial descriptor pool. Model building and optimization in deriving the 4D-fingerprint QSAR equations, which are the high-throughput virtual screens, was carried in the identical fashion used to build the RI-4D-QSAR models.

Figure 26 is a plot of the number of descriptor terms in a 4D-fingerprint model versus r^2 and r^2_{cv}. The r^2_{cv} of the 4D-fingerprints of the 4- and 5-term models are very nearly the same, and r^2_{cv} behaves in something of an erratic fashion for models having 5, or more, terms. The optimized 4-descriptor term virtual screening model appears, on the basis of r^2_{cv}, to capture maximum fitting to the training set data without overfitting. Thus, the 4-term QSAR model given by equation 17 was selected as the preferred absolute 4D-fingerprint virtual screen. Equation 17 is the best 4-term model from the top-ten 4-term models derived in the GFA optimization. Table 16 shows the linear cross-correlation matrix of the residual of fit for the top-ten 4-term models. This table reveals that all pairs of models have highly correlated residuals of fit greater than 0.85 to one another. Thus, r^2_{cv} represents the best and only distinct fit to the training set data using absolute 4D-fingerprints.

$$-\log IC_{50} = -7.39 - 452.65 \in_7 (\text{any,np}) + 1357.10 \in_{11} (\text{any,hs}) + 9.58 \in_3 (p^+,\text{aro}) - 94.31 \in_2 (\text{np,hs})$$

$$r^2 = 0.831,\ r^2_{cv} = 0.719\ \text{n} = 25 \qquad \textbf{(17)}$$

For reference in defining the 4D-fingerprints, $\in_7$(any,np) represents the seventh largest eigenvalue from the MDDM of the IPEs u = (any) and v = (np) molecular similarity vector capturing all pairs of atoms in each lamellarin assigned IPEs of any and nonpolar, respectively.

The relative significance and weight of each 4D-fingerprint descriptor term in equation 17 was measured in terms of its frequency of use in the GFA model optimization process. The idea is that the more significant is a descriptor to establishing a fit to the training set data, the more often it will be used in the repetitive GFA optimization process. The frequencies of descriptor usage during GFA optimization are shown in Table 17. An inspection of Table 17 indicates that $\in_{11}$(any,hs) and $\in_7$(any,np) are the first and second important features governing the SAR of lamellarin cytotoxicity potency, respectively. Increased potency of the lamellarins arises from increasing the values of $\in_{11}$(any,hs) and/or $\in_3(p^+,\text{aro})$, while a

decrease in lamellarin cytotoxicity accompanies an increase in the values of the $\in_7$(any,np) and $\in_2$(np,hs) 4D-fingerprints. Figure 27 is a plot of –log IC_{50} values predicted using equation 12 versus the corresponding observed –log IC_{50} values.

4.1 Comparison of the 4D-fingerprints QSAR virtual screening model to RI 4D-QSAR models

The RI-4D-QSAR model given by equation 14 with three descriptor terms is a more significant fit to the training set data (r^2_{cv} = 0.947 and r^2 = 0.971) than the four descriptor 4D-fingerprint model given by equation 17 (r^2_{cv} = 0.719 and r^2 = 0.831). Presumably the inclusion of alignment information in equation 14 provides this boost in the overall fitting quality of this model as compared to equation 17. But equation 17 is not dependent on the alignment which correspondingly permits a wider range of variations lammellarin chemistry to be considered. Table 18 is the linear correlation matrix of the residuals of fit of equation 14, the RI-4D-QSAR, to equation 17, the absolute 4D-fingerprint virtual screen, as well as correlations of both models to the observed –log IC_{50} cytotoxicity values. The correlation coefficient of 0.797 between the residuals of fit from equation 14 and 17 indicates that these two models are basically the same, but equation 14, owing to inclusion of alignment, fits the training set better than equation 17.

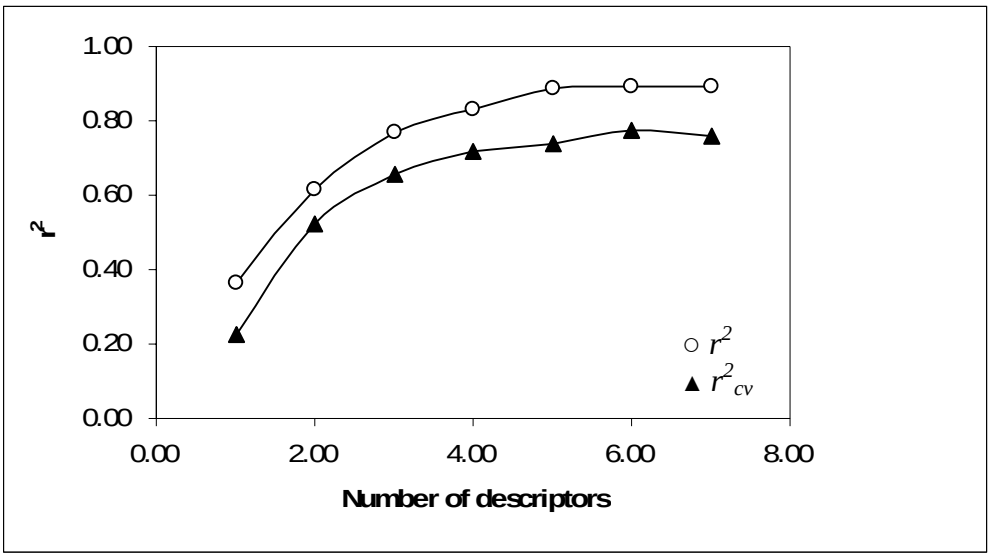


Figure 26 Plot of the number of 4D-fingerprint model descriptors versus r^2, and r^2_{cv}.

Table 16 Linear cross-correlation matrix of the top-ten model from 4 descriptors terms 4D-fingerprint model.

Model no.	1	2	3	4	5	6	7	8	9	10
1	1									
2	0.929	1.000								
3	0.955	0.889	1.000							
4	0.927	0.943	0.929	1.000						
5	0.893	0.945	0.903	0.956	1.000					
6	0.908	0.861	0.935	0.884	0.850	1.000				
7	0.949	0.897	0.988	0.931	0.910	0.939	1.000			
8	0.924	0.912	0.920	0.914	0.885	0.918	0.931	1.000		
9	0.903	0.922	0.897	0.903	0.966	0.874	0.914	0.885	1.000	
10	0.910	0.920	0.918	0.895	0.956	0.885	0.935	0.904	0.976	1.000

Table 17 Frequency and ranking of each descriptor term in 4D-fingerprint model.

	$\in_7$(any,np)	$\in_{11}$(any,hs)	$\in_3$(p^+,aro)	$\in_2$(np,hs)
Frequency	124	128	51	17
Ranking	2	1	5	11

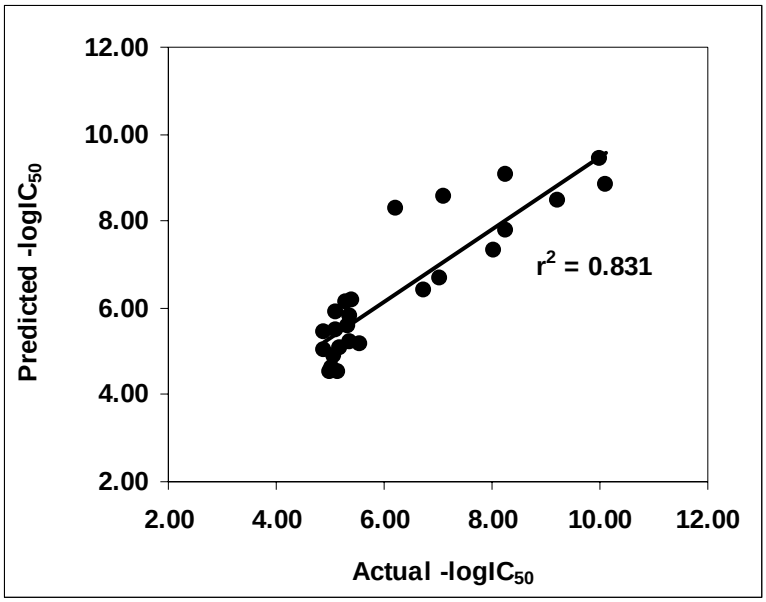


Figure 27 Predicted and actual –log IC_{50} derived by 4D-fingerprints model of 25 lamellarins data set.

Table 18 Linear correlation matrix of the residuals of fit of the 4D-QSAR (1), 4D-fingerprint (2) model, and actual cytotoxicity (3).

	1	2	3
1	1		
2	0.797	1	
3	0.972	0.823	1

CONCLUSIONS

The orientation and binding energies of all lamellarins in the binding pocket of topoisomerase I-DNA were investigated by using molecular docking methods (AutoDock and Gold). The results revealed that many lamellarins could intercalate at the site of DNA cleavage, forming base-stacking interactions with both the upstream and downstream base pairs. Several H-bond interactions also found within 7.0 Å in the binding pocket of topoisomerase I-DNA.

Two 3D-QSAR methods, CoMFA and CoMSIA, were applied to lamellarins active against T47D and MDA-MB-231 cytotoxicity. The saturated and unsaturated D-ring lamellarins, caused to different planarity at D-, E-ring structures, were focused. The satisfied CoMFA and CoMSIA models were derived and the obtained results reveal these powerful 3D-QSAR methods can be used to handle small data sets consisted of structurally different planarity in the structures. As there is no information about the target structure for both cell lines, CoMFA and CoMSIA analyses provided more details about the steric, electrostatic, and hydrophobic field requirements of lamellarins for breast cancer cytotoxicity. In addition, the CoMSIA contour maps showed a good correlation with those obtained by CoMFA contour s. Based on combination of CoMFA and CoMSIA contour maps, the results can discriminate the structural requirements between T47D and MDA-MB-231 cytotoxicity by common carbon atom of the lamellarin skeleton such as at C8, C13, C20, and C21. The interesting different structural requirement of lamellarins at C8 may play an important role with responding for cytotoxicity activity different against both cell lines. In T47D cell line, 3D-QSAR contours highlight the important of F-ring at C13 for the cytotoxicity, on the other hand, this is not occurred on the presence of contour maps around this region for MDA-MB231 cell line. The contour maps of both cell lines suggest the necessity of A-ring at C21 for the cytotoxcity against T47D cell line, but the importance of A-ring for cytotoxcity against MDA-MB-231 cell line is clearly shown at C20.

Moreover, the 3D-QSAR results revealed specific structural requirements of the lamellarins for their cytotoxic activity towards two breast cancer cell lines, including a steric and electropositive oxygen-containing group at C7, and an electropositive group at C9. Especially, the significance of the C5-C6 double bond was also found. Hence, 3D-QSAR is a useful method to explore the specific structural requirements between both types of human breast cancer and also as a guideline to design more effective inhibitors from lamellarins. These are not only help to the more detailed understanding interaction of lamellarin derivatives upon the binding of unknown receptor, but also applicable to a small lamellarins data set.

This work puts forth a 'quality for quantity' strategy to handle small data sets composed of structurally complex, difficult to synthesize compounds that can exhibit a wide-range in endpoint activity. A high-level modeling approach providing detailed structural, thermodynamic and electronic information about each complex compound of the data set is used to negate the lack-of-data drawbacks to the small size of the data set. In this study the flexibility, yet high-level of modeling sophistication of the 4D-QSAR paradigm is used to explore different subpopulations of the data set in extracting the maximum SAR information from the data set in terms of a pseudo consensus RI-4D-QSAR model and its corresponding 3D-pharmacophore. The consensus aspect to the RI-4D-QSAR modeling arises from the fact that the same methodology and parameters, including alignment, can be used in any manner across any subpopulations of the data set. As such, all resulting models are not only directly comparable, but to an appreciable extent can be combined to elucidate a high-resolution 3D-phamracophore. In addition, the 4D-fingerprint formulation of the 4D-QSAR paradigm permits alternate model generation, particularly useful in virtual screening. Still, the 4D-fingerprint models are once again directly comparable to the RI-4D-QSAR models so as to exact additional information from the data set, as well as to evaluate the self-consistency across all the models constructed. The consensus set of 4D-QSAR models expressed by equation 14-17, suggests that the ability to form a ligand-receptor intermolecular hydrogen bond and hydrophobic interactions for substituents on the E ring most modulate the cytotoxicity against T47D breast cancer cells. The optimization of this intermolecular hydrogen bond, and, to a lesser

extent, the hydrophobic interactions, are coupled to the alignment freedom of a lamellarin owing, in turn, to other possible substitutions across the molecule and their possible interactions with sites on the receptor.

Hydrophobic substitutions on the F-ring can also enhance cytotoxic potency, but given that the 3D-pharmacophore sites for these interactions arise for the entire data set, and not the restricted high activity data subset, would indicate these are likely minor binding pharmacophore sites. Attempts to force the lipophilicity of the entire lamellarin into a 4D-QSAR model were unsuccessful. Thus, the findings of 3D-pharmacophore sites, where occupancy by nonpolar atoms and/or groups can modulate activity, likely reflect specific interactions at these sites, and not global lipophilic features of the lamellarins.

The 4D-fingerprint virtual screening model, equation 17, is highly consistent with the general RI-4D-QSAR model given by equation 14. Consequently, equation 17 can be used to rapidly virtually screen prospective compounds without concern for alignment, but with the expectation that the 3D-phamracophore of equation 14 will be relevant helping to understand findings from virtual screenings.

LITERATURE CITED

Andersen, R.J., D.J. Faulkner, C.-H. He, G.D. Van Duyne and J. Clardy. 1985. Metabolites of the marine prosobranch mollusc *Lamellaria* sp. **J. Am. Chem. Soc.** 107: 5492-5495.

Anonymous. 2008. **WHO: Cancer to be world's top killer by 2010.** Availabe Source: http://www.usatoday.com/news/health/2008-12-09-cancer_N.htm, December 9, 2008.

American Cancer Society. 2007. **Cancer Facts & Figure 2007**. American Cancer Society, Atlanta.

Bailly, C. 2004. Lamellarins, from A to Z: a family of anticancer marine pyrrole alkaloids. **Curr. Med. Chem.: Anti-Cancer Agents.** 4: 363-378.

Baunbaek, D., N. Trinkler, Y. Ferandin, O. Lozach, P. Ploypradith, S. Rucirawat, F. Ishibashi, M. Iwao and L. Meijer. 2008. Anticancer Alkaloid Lamellarins Inhibit Protein Kinases. **Mar. Drugs**. 6: 514-527.

Bohm, M., J. Sturzebecher and G. Klebe. 1999. Three-dimensional quantitative structure-activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. **J. Med. Chem.** 42: 458-477.

Boik, J. 2001. **Natural Compounds in Cancer Therapy**. Quality Book, Inc., Princeton, Minnesota.

Boys, S. F. and F. Bernardi. 1970. The calculations of small molecular interaction by the difference of separate total energies some procedures with reduced error. **Mol. Phys**. 19: 553-566.

Carroll, A.R., B.F. Bowden and J.C. Coll. 1993. Studies of Australian ascidians: six new lamellarin-class alkaloids from a colonial ascidian, *Didemnum* sp. **Aust. J. Chem.** 46: 489-501.

Chittchang, M., P. Batsomboon, S. Ruchirawat and P. Ploypradith. 2009. Cytotoxicities and structure-activity relationships of natural and unnatural lamellarins towards cancer cell lines. **ChemMedChem**. In press.

Costantini, P., E. Jacotot, D. Decaudin and G. Kroemer. 2000. Mitochondrion as a novel target of anticancer chemotherapy. **J. Natl. Cancer Inst.** 92: 1042-1053.

Cragg, G.M. and D.J. Newman. 2006. Plants as source of anti-cancer agents. **Ethnopharmacology**. Eolss Publishers, Oxford.

Cramer, R.D. 3rd, D.E. Patterson and J.D. Bunce. 1988. Comparative molecular field analysis (CoMFA): 1. Effect of shape on binding of steroids to carrier proteins. **J. Am. Chem. Soc.** 110: 5959-5967.

Davis, R.A., A.R. Carroll, G.K. Pierens and R.J. Quinn. 1999. New lamellarin alkaloids from the Australian ascidian *Didemnum chartaceum*. **J. Nat. Prod.** 62: 419-424.

Debatin, K.M., D. Poncet and G. Kroemer. 2002. Chemotherapy: targeting the mitochondrial cell death pathway. **Oncogene.** 21: 8786-8803.

Dias, N. and C. Bailly. 2005. Drugs targeting mitochondrial functions to control tumor cell growth. **Biochem. Pharmacol.** 70: 1-12.

_____, H. Vezin, A. Lansiaux and C. Bailly. 2005. Topoisomerase inhibitors of marine origin and their potential use as anticancer agents. **Top Curr Chem.** 253: 89-108.

Duca, J.S. and A.J. Hopfinger. 2001. Estimation of Molecular Similarity Based on 4D-QSAR Analysis: Formalism and Validation. **J. Chem. Inf. Comput. Sci.** 41: 1367-1387.

Facompre′, M., C. Tardy, C. Bal-Mahieu, P. Colson, C. Perez, I. Manzanares, C. Cuevas, and C. Bailly. 2003. Lamellarin D: A Novel Potent Inhibitor of Topoisomerase I. **Cancer Res.** 63: 7392-7399.

Fan, H., J. Peng, M.T. Hamann and J.F. Hu. 2008. Lamellarins and related pyrrole-derived alkaloids from marine organisms. **Chem. Rev.** 108: 264-287.

Gallego, M-A., C. Ballot, J. Kluza, N. Hajji, A. Martoriati, L. Castéra, C. Cuevas, P. Formstecher, B. Joseph, G. Kroemer, C. Bailly and P. Marchetti. 2008. Overcoming chemoresistance of non-small cell lung carcinoma through restoration of an AIF-dependent apoptotic pathway. **Oncogene.** 27: 1981-1992.

Golbraikh, A. and A. Tropsha. 2002. Beware of q2!. **J. Mol. Graph. Mod.** 20: 269-276.

Gottesman, M.M. and I. Pastan. 1993. Biochemistry of Multidrug Resistance Mediated by the Multidrug Transporter. **Annu. Rev. Biochem.** 62: 385-427.

Ham, J. and H. Kang. 2002. A novel cytotoxic alkaloid of lamellarin class from a marine ascidian *Didemnum* sp. **Bull Korean Chem. Soc**. 23: 163-166.

Hopfinger, A.J., A. Reaka, P. Venkatarangan, J.S. Duca and S. Wang. 1999. Construction of a virtual high throughput screen by 4D-QSAR analysis: application to a combinatorial library of glucose inhibitors of glycogen phosphorylase b. **J. Chem. Inf. Comput. Sci.** *39*: 1151-1160.

_____, S. Wang, J.S. Tokarski, B. Jin., M. Albuquerque, P.J. Madhav and C. Duraiswami. 1997. Construction of 3D-QSAR models using the 4D-QSAR analysis formalism. **J. Am. Chem. Soc.** *119*: 10509-10524.

Ishibashi, F., S. Tanabe, T. Oda and M. Iwao. 2002. Synthesis and structure-activity relationship study of lamellarin derivatives. **J. Nat. Prod.** 65: 500-504.

Iyer, M. and A.J. Hopfinger. 2007. Treating chemical diversity in QSAR analysis: modeling diverse HIV-I integrase inhibitors using 4D-fingerprints. **J. Chem. Inf. Model.** 47: 1945-1960.

_____, T. Zheng, A.J. Hopfinger and Y.J. Tseng. 2007. QSAR analyses of skin penetration enhancers. **J. Chem. Inf. Model.** 47: 1130-1149.

Klebe, G., U. Abraham and T. Mietzner. 1994. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. **J. Med. Chem.** 37: 4130-4146.

Kluza, J., M.A. Gallego, A. Loyens, J.C. Beauvillain, J.M.F. Sousa-Faro, C. Cuevas, P. Marchetti and C. Bailly. 2006. Cancer cell mitochondria are direct proapoptotic targets for the marine antitumor drug lamellarin D. **Cancer Res**. 66: 3177-3187.

Krasowski, M. D, X. Hong, A.J. Hopfinger and N.L. Harrison. 2002. 4D-QSAR analysis of a set of propofol analogues: mapping binding sites for an anesthetic phenol on the GABAA receptor. **J. Med. Chem.** *45*: 3210-3221.

Krishnaiah, P., V.L.N. Reddy, G. Venkataramana, K. Ravinder, M. Srinivasulu, T.V. Raju, K. Ravikumar, D. Chandrasekar, S. Ramakrishna and Y. Venkateswarlu. 2004. New lamellarin alkaloids from the Indian ascidian *Didemnum obscurum* and their antioxidant properties. **J. Nat. Prod.** 67: 1168-1171.

Kubinyi, H. 1997. QSAR and 3DQSAR in drug design Part1: methodology. **Drug Discovery Today**. 2: 457-467.

Leach, A.R. 1996. **Molecular Modelling: Principles and Applications.** Addison Wesley Longman Limited, Edinburgh Gate, Harlow.

Li, Y., D. Pan, J. Liu, P.S. Kern, G.F. Gerberick, A.J. Hopfinger and Y.J. Tseng. 2007. Categorical QSAR models for skin sensitization based upon local lymph node assay classification measures Part 2: 4D-Fingerprint three-state and two-2-state logistic regression models. **Toxicological Sciences.** 99: 532-544.

_____, Y.J. Tseng, D. Pan, J. Liu, P.S. Kern, G.F. Gerberick and A.J. Hopfinger. 2007. 4D-fingerprint categorical QSAR models for skin sensitization based on the classification of local lymph node assay measures. **Chemical Research in Toxicology**. 20: 114-128.

Licht, T., I. Pastan, M.M. Gottesman and F. Herrmann. 1996. The multidrug-resistance gene in gene therapy of cancer and hematopoietic disorders. **Ann. Hematol.** 72: 184-193.

Lindquist, N., W. Fenical, G.D. Van Duyne and J. Clardy. 1988. New alkaloids of the lamellarin class from the marine ascidian *Didemnum chartaceum* (Sluiter, 1909). **J. Org. Chem.** 53: 4570-4574.

Liu, J., L. Yang, Y. Li, D. Pan and A.J. Hopfinger. 2005. Predicting permeability coefficient in ADMET evaluation by using different membranes-interaction QSAR. **J. Comput Aided Mol Des.** 19: 567-583.

_____, _____, _____, _____ and _____. 2006. Constructing plasma protein binding model based on a combination of cluster analysis and 4D-fingerprint molecular similarity analyses. **Bioorg. Med. Chem.** 14: 611-621.

Marco, E., W. Laine, C. Tardy, A. Lansiaux, M. Iwao, F. Ishibashi, C. Bailly and F. Gago. 2005. Molecular determinants of topoisomerase I poisoning by lamellarins: Comparison with camptothecin and structure-activity relationships. **J. Med. Chem.** 48: 3796-3807.

Morris, G. M., Goodsell, D. S., Halliday, R.S., Huey, R., Hart, W. E., Belew, R. K. and Olson,A. J. 1998. Automated Docking Using a LamarckianGenetic Algorithm and and Empirical Binding Free Energy Function. **J. Computational Chemistry**. 19: 1639-1662.

Nilsson, J., S. De Jong and A. K. Smilde. 1997. Multilinear PLS analysis application to 3D QSAR data set. **J. Chemometrics**. 11: 511-524.

Pasqualoto, K.F.M., E.I. Ferreira, O.A. Santos-Filho and A.J. Hopfinger. 2004. Rational design of new antituberculosis agents: Receptor-independent four-dimensional quantitative structure-activity relationship analysis of a set of isoniazid derivatives. **J. Med. Chem.** 47: 3755-3764.

Pla, D., A. Marchal, C.A. Olsen, A. Francesch, C. Cuevas, F. Albericio and M. Alvarez. 2006. Synthesis and structure-activity relationship study of potent cytotoxic analogues of the marine alkaloid lamellarin D. **J. Med. Chem.** 49: 3257-3268.

Ploypradith, P., T. Petchmanee, P. Sahakitpichan, N.D. Litvinas and S. Ruchirawat. 2006. Total synthesis of natural and unnatural lamellarins with saturated and unsaturated D-rings. **J. Org. Chem.** 71: 9440-9448.

Quesada, A.R., M.D. Garcia Gravalos and J.L. Fernandez Puentes. 1996. Polyaromatic alkaloids from marine invertebrates as cytotoxic compounds and inhibitors of multidrug resistance caused by P-glycoprotein. **Br. J. Cancer.** 74: 677-682.

Ridley, C.P., M.V.R. Reddy, G. Rocha, F.D. Bushman and D.J. Faulkner. 2002. Total synthesis and evaluation of lamellarin α 20-Sulfate analogues. **Bioorg. Med. Chem.** 10: 3285-3290.

Reddy, M.V.R., D.J. Faulkner, Y. Venkateswarlu and M.R. Rao. 1997. New lamellarin alkaloids from an unidentified ascidian from the Arabian Sea. **Tetrahedron**. 53: 3457-3466.

_____, M.R. Rao, D. Rhodes, M.S.T. Hansen, K. Rubins, F.D. Bushman, Y. Venkateswarlu and D.J. Faulkner. 1999. Lamellarin α 20-sulfate, an inhibitor of HIV-1 integrase active against HIV-1 virus in cell culture. **J. Med. Chem.** 42: 1901-1907.

_____, M. Srinivasulu, N. Satyanarayana, A.K. Kondapi and Y. Venkateswarlu. 2005. New potent cytotoxic lamellarin alkaloids from Indian ascidian *Didemnum obscurum*. **Tetrahedron.** 61: 9242-9247.

Roger, D. and A.J. Hopfinger. 1994. Application of genetic function approximation to quantitative structure-activity relationships and quantitative structure-property relationships. **J. Chem. Inf. Comput. Sci.** 34: 854-866.

Rustin, P. and G. Kroemer. 2007. Mitochondria and cancer. **Ernst Schering Foundation symposium proceedings.** 4: 1-21.

Tardy, C., M. Facompré, W. Laine, B. Baldeyrou, D. García-Gravalos, A. Francesch, C. Mateo, A. Pastor, J.A. Jiménez, I. Manzanares, C. Cuevas and C. Bailly. 2004. Topoisomerase I-mediated DNA cleavage as a guide to the development of antitumor agents derived from the marine alkaloid lamellarin D: Triester derivatives incorporating amino acid residues. **Bioorg. Med. Chem.** 12: 1697-1712.

Thomas, C.J., N.J. Rahier and S.M. Hecht. 2004. Camptothecin: current perspectives. **Bioorg. Med. Chem.** 12: 1585-1604.

Toogood, P.L. 2008. Mitochondrial drugs. **Current Opinion in Chemical Biology.** 12: 457-463.

Urban, S., M.S. Butler and R.J. Capon. 1994. Lamellarins O and P: new aromatic metabolites from the Australian marine sponge *Dendrilla cactos.* **Aust. J. Chem.** 47: 1919-1924.

_____ and R.J. Capon. 1996. Lamellarin-S: a new aromatic metabolite from an Australian tunicate, Didemnum sp. **Aust. J. Chem.** 49: 711-713.

_____, L. Hobbs, J.N.A. Hooper and R.J. Capon. 1995. Lamellarins Q and R: new aromatic metabolites from and Australian marine sponge, *Dendrilla cactos.* **Aust. J. Chem.** 48: 1491-1494.

Vanhuyse, M., J. Kluza, C. Tardy, G. Otero, C. Cuevas, C. Bailly and A. Lansiaux. 2005. Lamellarin D: A novel pro-apoptotic agent from marine origin insensitive to P-glycoprotein-mediated drug efflux. **Cancer Lett.** 221: 165-175.

Senese, C. L., J. Duca, D. Pan, A.J. Hopfinger and Y.J. Tseng. 2004. 4D-fingerprints, universal QSAR and QSPR descriptors. **J. Chem. Inf. Comput. Sci.** 44: 1526-1539.

Staker, B.L., K. Hjerrild, M.D. Feese, C.A. Behnke, Jr.A.B. Burgin and L. Stewart. 2002. The mechanism of topoisomerase I poisoning by a camptothecin analog. **PNAS.** 99:15387–15392.

Winkler, D.A. 2002. The role of quantitative structure-activity relationships (QSAR) in biomolecular discovery. **Briefings in Bioinformatics.** 3: 73-86.

APPENDIX

1. Publications

Poonsiri Thipnate, Montakarn Chittchang, Nopporn Thasana, Patchreenart Saparpakorn, Poonsakdi Ploypradith, and Supa Hannongbua*. 3**D-QSAR analysis for cytotoxicity of lamellarins against human hormone-dependent T47D and hormone-independent MDA-MB-231 breast cancer cells**. *Submitted to Journal of Computer-Aided Molecular Design.*

Poonsiri Thipnate, Jianzhong Liu, Supa Hannongbua* and A. J. Hopfinger. **3D-Pharmacophore Mapping Using 4D-QSAR Analysis for the Cytotoxicity of Lamellarins Against Human Hormone-Dependent T47D Breast Cancer Cells**. *Submitted to Journal of Chemical Information and Modeling*

2. Oral Presentations

Thipnate, P., M. Chittchang, N. Thasana, P. Ploypradith, P. Saparpakorn and S. Hannongbua. **3D-QSAR analysis of lamellarins against human hormone dependent breast cancer (T47D).** 12th Annual National Symposium, on Computational Science and Engineering (ANSCSE 2008), Ubon Ratchathani University, Thailand, 27-29 March 2008.

3D-QSAR analysis of lamellarins against human hormone dependent breast cancer (T47D)

Poonsiri Thipnate[1,2] Montakarn Chittchang[3] Nopporn Thasana[3] Poonsakdi Ploypradith[3] Patchreenart Saparpakorn[1,2] Supa Hannongbua[1,2,*]

[1]*Department of Chemistry, Faculty of Science, Kasetsart University, Jatuchak, Bangkok, Thailand 10900*

[2] *Center of Nanotechnology KU, and NANOTEC Center of Excellence at Kasetsart University, Kasetsart University, Jatuchak, Bangkok, Thailand 10900*

[3] *Chulabhorn Research Institute, Vibhavadee-Rangsit Highway, Bangkok 10210, Thailand*

**E-mail: fscisph@ku.ac.th, Tel: +66-2-562-5555 ext 2140*

Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) were applied to 25 lamellarins active against human hormone dependent breast cancer (T47D). Both methods were used to examine the structural requirement for T47D cytotoxic activity of lamellarins. All lamellarins were separated into 2 groups, 20 compounds served as training set and 5 compounds served as test set. The structures of 25 lamellarins were constructed and fully optimized by ab initio molecular orbital calculations at HF/3-21G level of theory. The best CoMFA and CoMSIA models yield satisfactory predictive ability with r^2_{cv} = 0.659 and 0.662, respectively. The derived CoMFA model indicates the importance of steric contribution (66.9%) and electrostatic interactions (33.1%) for T47D cytotoxicity. CoMSIA model reveals that steric, electrostatic, and hydrophobic interactions corresponding to T47D cytotoxic activity amount to 11.6%, 45.8%, and 42.6%, respectively. 3D-QSAR models were validated using a test set of 5 lamellarins. Almost all predicted activities fall within one log unit of the actual activity. The contour maps obtained from the CoMFA and CoMSIA models agree with experimental results and can be used in the design of more potent cytotoxic compounds.

REFERENCES

1. Facompre, M.; Tardy, C.; Bal-Mahieu, C.; Colson, P.; Perez, C.; Manzanares, I.; Cuevas, C.; Bailly, C. *Cancer Research* **2003**, 63, (21), 7392-7399.
2. Bailly, C. *Current Medicinal Chemistry - Anti-Cancer Agents* **2004**, 4, (4), 363-378.
3. Marco, E.; Laine, W.; Tardy, C.; Lansiaux, A.; Iwao, M.; Ishibashi, F.; Bailly, C.; Gago, F. *Journal of Medicinal Chemistry* **2005**, 48, (11), 3796-3807.

Thipnate, P., M. Chittchang, N. Thasana, P. Ploypradith, P. Saparpakorn and S. Hannongbua. **Cytotoxicity and 3D-QSAR analysis of lamellarins against human hormone dependent breast cancer (T47D) and human hormone independent breast cancer (MDA-MB231).** 17th European Symposium on Quantitative Structure-Activity Relationships (EuroQSAR2008), Uppsala, Sweden, 21-26 September 2008.

Cytotoxicity and 3D-QSAR Analysis of Lamellarins Against Human Hormone Dependent Breast Cancer (T47D) and Human Hormone Independent Breast cancer (MDA-MB231)

Poonsiri Thipnate [1, 2, 3], Montakarn Chittchang, [4] Nopporn Thasana, [4] Poonsakdi Ploypradith, [4, *] Patchreenart Saparpakorn, [1, 2] and Supa Hannongbua [1, 2, *]

[1]*Department of Chemistry, Faculty of Science, Kasetsart University, Jatuchak, Bangkok, Thailand, 10900*

[2] *Center of Nanotechnology KU, and NANOTEC Center of Excellence at Kasetsart University, Kasetsart University, Jatuchak, Bangkok, Thailand 10900*

[3]*Chemistry Department, Faculty of Science and Technology, Rajabhat Petchaburi University, Petchaburi, Thailand, 76000*

[3] *Chulabhorn Research Institute, Vibhavadee-Rangsit Highway, Bangkok, Thailand, 10210*

E-mail: thipnate@gmail.com
**E-mail: poonsakdi@cri.or.th and fscisph@ku.ac.th*

ABSTRACT

Twenty five of marine alkaloid lamellarins were synthesized and evaluated for cytotoxicity against T47D and MDA-MB231 cell line. All lamellarins were solubilized in DMSO and tested for their cytotoxic activity against T47D and MDA-MB-231 cell lines. The number of surviving cells in each well was determined using crystal violet staining, as described below, in order to determine the IC_{50} after 48 h of continuous exposure to each test compound. Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) were applied for determine the common structural requirement of both human breast cancer. All lamellarins were separated into 2 groups, 20 compounds served as training set and 5 compounds served as test set. The structures of 25 lamellarins were constructed and fully optimized by ab initio molecular orbital calculations at HF/3-21G level of theory. The best CoMFA and CoMSIA models of T47D cell line yield satisfactory predictive ability with r^2_{cv} = 0.659 and 0.662, respectively. For MDA-MB231 cell line, the r^2_{cv} of CoMFA model is 0.728 and the r^2_{cv} of CoMSIA model is 0.674. The derived CoMFA model indicates the importance of steric contribution (66.9%) and electrostatic interactions (33.1%) for T47D cytotoxicity and steric contribution with 68.1% and electrostatic interactions with 31.9% for MDA-MB231 cytotoxicity. CoMSIA model reveals that steric, electrostatic, and hydrophobic interactions corresponding to T47D cytotoxic activity amount to 11.6%, 45.8%, and

42.6%, respectively and MDA-MB231 amount to 10.2%, 49.9%, and 39.8%. 3D-QSAR models were validated using a test set of 5 lamellarins. Almost all predicted activities of test set fall within one log unit of the actual activity except lamellarin α and M for T47D cell line and lamellarin L for MDA-MB231 cell line. The contour maps obtained from the CoMFA and CoMSIA models agree with experimental results and structural requirement of lamellarins both cell lines will be very useful for design of more potent anti-breast cancer compounds.

REFERENCES

1. Ishibashi, F.; Tanabe, S.; Oda, T.; Iwao, M., Synthesis and structure-activity relationship study of lamellarin derivatives. *Journal of Natural Products* **2002,** 65, (4), 500-504.
2. Facompre, M.; Tardy, C.; Bal-Mahieu, C.; Colson, P.; Perez, C.; Manzanares, I.; Cuevas, C.; Bailly, C., Lamellarin D: A Novel Potent Inhibitor of Topoisomerase I. *Cancer Research* **2003,** 63, (21), 7392-7399.
3. Bailly, C., Lamellarins, from A to Z: A family of anticancer marine pyrrole alkaloids. *Current Medicinal Chemistry - Anti-Cancer Agents* **2004,** 4, (4), 363-378.
4. Marco, E.; Laine, W.; Tardy, C.; Lansiaux, A.; Iwao, M.; Ishibashi, F.; Bailly, C.; Gago, F., Molecular determinants of topoisomerase I poisoning by lamellarins: Comparison with camptothecin and structure-activity relationships. *Journal of Medicinal Chemistry* **2005,** 48, (11), 3796-3807.
5. Fan, H.; Peng, J.; Hamann, M. T.; Hu, J. F., Lamellarins and related pyrrole-derived alkaloids from marine organisms. *Chemical Reviews* **2008,** 108, (1), 264-287.

3. Poster Presentations

Thipnate, P., P. Ploypradith and S. Hannongbua. **The orientation and interaction of lamellarin derivatives in the binding pocket of DNA-topoisomerase I complex.** 31st Congress on Science and Technology in Thailand. Technopolis, Suranaree University of Technology, Nakhon Ratchasima, Thailand, 8-20 October 2005.

การศึกษาการวางตัวและอันตรกิริยาของอนุพันธ์ลาเมลลารีนภายในโพรงการจับที่เป็นสารประกอบเชิงซ้อนระหว่างดีเอนเอกับทอพอไอโซเมอเรส I โดยวิธีโมเลคิวลาร์ด็อกกิ้ง

พูนศิริ ทิพย์เมตร[1], พูนศักดิ์ พลอยประดิษฐ์[2], สุดา หาญหนองบัว[1]

Poonsiri Thipmte[1], Poonsakdi Ploypradith[2], Supa Hannongbua[1]

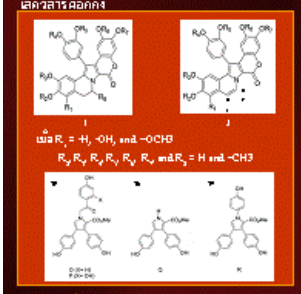
[1] ภาควิชาเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ กรุงเทพ 10900

[2] ภาควิชาเคมี คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพ 10210

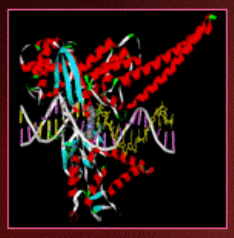
STT 2005

บทนำ

ลาเมลลารีนเป็นพอลีอะโรมาติคอัลคาลอยด์ชนิดหนึ่ง สามารถแบ่งออกเป็น 3 กลุ่มใหญ่ๆตามโครงสร้างดังรูปที่ 1 โครงสร้างที่ 1 ซึ่งเป็นวงแหวนหกเหลี่ยมที่ตำแหน่งคาร์บอนแปล-คาร์บอนค้ำเป็นพันธะเดี่ยว ตัวอย่างของกลุ่มนี้ได้แก่ ลาเมลลารีน A C E-G I-L S T-V Y Z และ β โครงสร้างที่ 2 ที่ตำแหน่งคาร์บอนแปล-คาร์บอนค้ำเป็นพันธะคู่ ดังเช่น ลาเมลลารีน B D H M N W X และ α และโครงสร้างที่ 3 โมเลกุลของวงไพรอลจะไม่ติดกับวงอะโรมาติกที่อยู่ถัดไปดังโครงสร้าง 3ก-3ค ในรูปที่ 1 ลาเมลลารีนได้ถูกนำมาศึกษาคุณสมบัติในการยับยั้งทอพอไอโซเมอเรส I บางชนิดมีฤทธิ์เป็นพิษทั้งต่อเซลล์มะเร็งที่เคีดการ ต่อต้านยาหลายๆชนิด และเซลล์มะเร็งเริ่มดื้อต่อต้านยา โดยเฉพาะอย่างยิ่งในเซลล์มะเร็งต่อมลูกหมาก (prostate cancer cells) และเซลล์มะเร็งเม็ดเลือดขาว (leukemia cells) สาเหตุที่ทำให้ลาเมลลารีนมีฤทธิ์ในการต่อต้านมะเร็งแตกต่างกันเนื่องจากโครงสร้างพื้นฐานและหมู่แทนที่แตกต่างกัน ดังนั้นในงานวิจัยนี้จึงทำการศึกษาความสัมพันธ์ของโครงสร้างและพลังงานกับฤทธิ์ในการ ยับยั้งทอพอไอโซเมอเรส I และทำการศึกษาการ วางตัวและอันตรกิริยาของอนุพันธ์ลาเมลลารีนภายในโพรงการจับที่เป็นสารประกอบเชิงซ้อนระหว่างดีเอนเอกับทอพอไอโซเมอเรส I โดยใช้วิธีโมเลคิวลาร์ด็อกกิ้ง



รูปที่ 1 โครงสร้างของอนุพันธ์ลาเมลลารีน



รูปที่ 2 โครงสร้างผลึกของเอนไซม์ทอพอไอโซเมอเรส I (1K4T)

ระเบียบวิธีการคำนวณ

คำนวณหาโครงสร้างสามมิติที่สมบูรณ์ของลาเมลลารีนด้วยวิธีแอบอินิชิโอ (HF/3-21 G)

นำโครงสร้างผลึกที่ได้จากอกเรย์ดิฟแฟรกชันของทอพอไอโซเมอเรสของมนุษย์เรสโซลูชัน 2.10 Å ในรูปสารประกอบเชิงซ้อนกับทอพอที่แคน (topotecan) และสารประกอบโควาเลนซ์กับสายดีเอนเอที่มีคู่เบส22 คู่ (Protein data bank, 1K4T) ดังรูปที่ 2

ทำการเพิ่มประจุโดยใช้โปรแกรม Sybyl เวอร์ชัน 7.0

ทำการเพิ่มอะตอมของไฮโดรเจน และประจุโดยใช้โปรแกรม Sybyl เวอร์ชัน 7.0

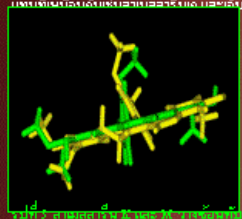
ทำการล็อคกิ้งโดยเลือกใช้ระเบียบวิธี genetic algorithm ซึ่งใช้โปรแกรม Autodock เวอร์ชัน 3.0 ในการศึกษาอันตรกิริยาของลาเมลลารีนกับสารประกอบเชิงซ้อนระหว่างดีเอนเอกับทอพอไอโซเมอเรส I และเปรียบเทียบโครงสร้างโดยนำมาวางซ้อนทับกัน

กิตติกรรมประกาศ

ขอขอบคุณมหาวิทยาลัยราชภัฏเทพสตรีสำหรับทุนสนับสนุนการวิจัย และขอขอบคุณโครงการพัฒนาบัณฑิตศึกษาและงานวิจัยด้านปิโตรเคมี (ADB-MUA) I CAC ศูนย์บริการคอมพิวเตอร์มหาวิทยาลัยเกษตรศาสตร์ และสถาบันวิจัยและพัฒนาแห่งมหาวิทยาลัยเกษตรศาสตร์สำหรับอุปกรณ์ทางคอมพิวเตอร์และสิ่งอำนวยความสะดวกต่างๆ

ผลการทดลองและวิจารณ์ผลการทดลอง

จากโครงสร้างของลาเมลลารีนทั้ง 3 กลุ่ม พบว่าเมื่อนำโครงสร้างในกลุ่มที่ 1 และ 2 มาวางซ้อนทับกันสามารถสังเกตความแตกต่างของโครงสร้างทั้งสองได้โดยในกลุ่มที่ 2 สามารถแบ่งเป็นสองระนาบทำมุมกัน 90 – 94 องศา ในทางตรงกันข้ามกลุ่มที่ 1 ไม่เป็นระนาบเดียวกัน ถึงแม้จะมีหมู่แทนที่เหมือนกันแต่ตำแหน่งการแทนที่และรีงสร้างที่ 3



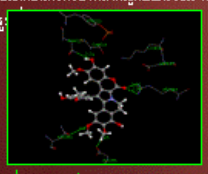
รูปที่ 3 ลาเมลลารีน K และ M วางซ้อนทับกัน

(เมื่อ ลาเมลลารีน K (กลุ่มที่ 1) = สีเขียว ลาเมลลารีน M (กลุ่มที่ 2) = สีชมพู)

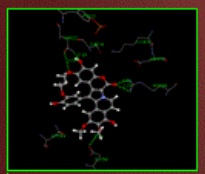
พบว่าลาเมลลารีนในกลุ่มที่ 1 และ 2 สามารถแทรกเข้าประจำอยู่ระหว่างสายดีเอนเอเช่นเดียวกับทอพอที่แคน โดยเกิดพันธะไฮโดรเจนระหว่างไฮดรอกซิลกับคาร์บอนิลของเอนไซม์ได้หลายตำแหน่ง เช่นที่ตำแหน่งแอสพาราจีน722 (ลาเมลลารีน K ระยะทาง =2.74 Å ลาเมลลารีน M = 2.39 และ3.56 Å) ที่ตำแหน่งแอสพาราจีน352 (ลาเมลลารีน K =3.41 Å) ที่ตำแหน่งอาร์จีนีน364 (ลาเมลลารีน K = 2.39 Å และ3.02 Å ลาเมลลารีน M=2.73 และ2.83 Å) และที่ตำแหน่งกลูตามีน 356 (ลาเมลลารีน K = 2.85 Å ลาเมลลารีน M=2.94 Å) ดังรูปที่ 4 และ 5 แต่ในกลุ่มที่ 3 ไม่สามารถแทรกเข้าไปอยู่ระหว่างสายดีเอนเอและไม่เกิดอันตรกิริยาที่สำคัญกับเอนไซม์ ดังรูป

ตารางที่ 1 พลังงานสุดท้ายที่ได้จากการล็อคกิ้ง (final docked energy) ของอนุพันธ์ลาเมลลารีน

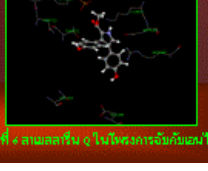
ลาเมลลารีน	พลังงานสุดท้ายที่ได้จากการล็อคกิ้ง (กิโลแคลอรี่ ต่อ โมล)
A	-18.17
E	-17.58
D	-18.67
E	-18.62
G	-16.94
J	-18.95
K	-18.75
L	-19.08
M	-19.25
N	-17.88
S	-18.72
T	-19.27
U	-19.44
W	-17.91
X	-18.41
Z	-17.75
β	-18.42



รูปที่ 4 ลาเมลลารีน K ในโพรงการจับกับเอนไซม์ทอพอไอโซเมอเรส I



รูปที่ 5 ลาเมลลารีน M ในโพรงการจับกับเอนไซม์ทอพอไอโซเมอเรส I



รูปที่ 6 ลาเมลลารีน Q ในโพรงการจับกับเอนไซม์ทอพอไอโซเมอเรส I

สรุปผลการทดลอง

ในโพรงการจับที่เป็นสารประกอบเชิงซ้อนระหว่างดีเอนเอกับทอพอไอโซเมอเรส I โดยวิธีโมเลคิวลาร์ด็อกกิ้ง พบว่าลาเมลลารีนทั้งกลุ่มที่ 1 และ 2 ส่วนใหญ่สามารถแทรกเข้าไปอยู่ระหว่างสายโซ่ดีเอนเอที่คำสังแคดออกได้ และเกิดพันธะไฮโดรเจนกับทอพอไอโซเมอเรสได้หลายตำแหน่ง แต่ลาเมลลารีนกลุ่มที่ 3 ไม่สามารถแทรกเข้าไปอยู่ระหว่างสายโซ่ดีเอนเอและไม่เกิดพันธะไฮโดรเจนกับเอนไซม์ ดังนั้นสามารถสรุปได้ว่าลาเมลลารีนกลุ่มที่ 1 และ 2 เกิดอันตรกิริยากับทอพอไอโซเมอเรส I ได้ดีกว่าลาเมลลารีนกลุ่มที่ 3 จากผลการศึกษานี้สามารถนำไปใช้พัฒนายาที่จะนำไปใช้ในการรักษาโรคมะเร็งต่อไปใน

เอกสารอ้างอิง

1. Marco, E.; Laine, W.; Tardy, C.; Lansiaux, A.; Iwao, M.; Ishibashi, F.; Bailly, C.; Gago, F. J. Med. Chem. 2005, 48, 3796-3807.
2. Facompre, M.; Tardy, C.; Bal-Mahieu, C.; Colson, P.; Perez, C.; Manfait, M.; Lansiaux, J. Cancer Res. 2003, 63, 7392-7399.

Thipnate, P., P. Ploypradith, N. Thasana, M. Chittchang and S. Hannongbua. **The orientation and interaction of new lamellarin alkaloids in the binding pocket of DNA-topoisomerase I complex.** International Conference on Modeling in Chemical and Biological Engineering Sciences, Rama Garden Hotel, Bangkok, Thailand, 25-27 October 2006.

THE ORIENTATION AND INTERACTION OF NEW LAMELLARIN ALKALOIDS IN THE BINDING POCKET OF DNA-TOPOISOMERASE I COMPLEX

Poonsiri Thipnate[1], Poonsakdi Ploypradith[2], Nopoon Thasana[2], Montakarn Chittchang[2], Supa Hannongbua*
[1,*] Laboratory for Computational & Applied Chemistry, Physical Chemistry Division, Department of Chemistry, Faculty of Science, Kasetsart University, Bangkok, 10900, Thailand
[2] Chulabhorn Research Institute, Vibhavadee-Rangsit Highway, Bankok 10210, Thailand

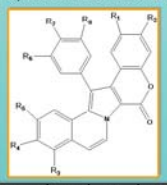
Introduction

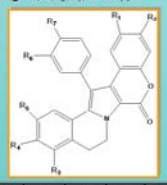
In recent years, lamellarin derivatives form a group of more than 30 polyaromatic pyrrole alkaloids, which were first isolated from the prosobranch mollusc *Lamellaria sp.* and later obtained from the genus *Didemnum*. Some of them have been displayed potent cytotoxic activites against both multidrug-resistant tumor cell lines and their corresponding parental cell lines. The mechanisms of action of inhibition of human topoisomerase I by lamellarin D was reported in 2003. In this years, new lamellarin alkaloids (as in lamellarin γ, α, and ε) were isolated from the Indian ascidian *Didemnum obscurum.* Moreover, four new lamellarin alkaloids (as in lamellarin ξ, η, ϕ, and χ) isolated from ascidian *Didemnum obscurum* in 2005 have been shown cytotoxic activity against colorectal cancer cells. Therefore, in this work, we focus on investigation of the orientation and interaction of seven new lamellarin alkaloids in binding pocket of DNA-topoisomerase I complex.

Methodology

Structures and conformations of lamellarin such as lamellarin γ, α, ε, ξ, η, ϕ, and χ were fully optimized at HF/3-21G level. The Substitute groups of this alkaloids were shown in **Table 1**. The 2.10 Å resolution X-ray crystal structure of human topoisomerase I (70 Kda) in complex with the poison topotecan and covalent complex with a 22 base pair DNA duplex (Protein Data Bank 1K4T) was used as reference structure to model the lamellarins stabilization of topoisomerase I –DNA complex. Then, hydrogen atoms and charges were added by Sybyl 7.0. The genetic algorithm implemented in the docking program Autodock 3.0 was used to explore lamellarins interaction modes with topoisomerase I-DNA complex.

Table 1 Substituted groups of new lamellarin alkaloids
a) saturated lamellarins (e.g. χ and γ)
b) unsaturated lamellarins (e.g. α, ε, ξ, η, and ϕ)

a) 

b) 

Lamellarins	R^1	R^2	R^3	R^4	R^5	R^6	R^7	R^8
χ	OMe	OH	H	OH	OMe	OMe	OH	H
γ	OMe	OH	OH	OMe	OMe	OMe	H	OMe
α	OMe	OH	H	OMe	OMe	OH	OMe	-
ϕ	OMe	OH	OMe	OMe	OH	OMe	OH	-
η	OMe	OH	H	OMe	OMe	OMe	OMe	-
ε	OMe	OH	OH	OMe	OMe	OMe	OMe	-
ξ	OMe	OH	OMe	OMe	OMe	OMe	OMe	-

Acknowledgment

Rajabhat Petchaburi University is acknowledged for financial support. The Postgraduate Education and Research on Petroleum and Petrochemical Technology (MUA-ADB), Kasetsart University Research and Development Institute (KURDI), LCAC and computing center of KU are also grateful for research facilities.

References

[1] Bailly, C. *Curr. Med Chem.: Anti-Cancer Agents* **2004**, 4, 363-378.
[2] Facompre, M.; Tardy, C.; Bal-Mahieu, C.; Colson, P.; Perez, C.; Manzanares, I.; Cuevas, C.; and Bailly, C. *CANCER RESEARCH* **2003**, 63, 7392-7399.
[3] Krishnaiah, P.; Reddy, V. L. N.; Venkataramana, G.; Ravinder, K.; Srinivasulu, M.; Raju, T. V.; Ravikumar, K.; Chandrasekar, D.; Ramakrishna, S.; and Venkateswarlu, Y. *J. Nat. Prod.* **2004**, 67, 1168-1171.
[4] Reddy, S. M.; Srinivasulu, M.; Satyanarayana, N.; Kondapi, A. K.; and Venkateswarlu, Y.; *Tetrahedron* **2005**, 61, 9242-9247.

Results & Discussion

For redocking, the root mean square deviations of topotecan was 0.96 Å and the binding energy was -19.87 (kcal/mol). We observed some interaction such as H-bonding of amino acid and lamellarins within particular distance 7 Å and the energy of binding for docking (shown in Figure 1). It was found that all new lamellarin alkaloids intercalated into the duplex DNA as same as topotecan and lamellarin D. Based on binding energies, lamellarin χ (-19.04 kcal/mol) binds more stronger than lamellarin γ (-18.88 kcal/mol), α (-18.04 kcal/mol), ε (-18.15 kcal/mol), ξ (-17.27 kcal/mol), η (-17.74 kcal/mol), and ϕ (-17.60 kcal/mol). Hydrogen bonds between the lamellarins and specific amino acid residues of protein such as Asn722, Arg364, and Glu356 further stabilize the ternary structure like lamellarin D, and some of them may have weak H-bonding and stacking interaction to DNA at cleavage site.

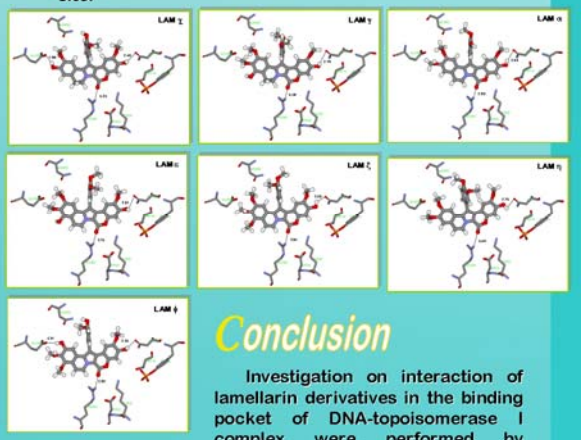


Figure 1 New lamellarins in the binding pocket of DNA-topo I complex within particular distance 7 Å

Conclusion

Investigation on interaction of lamellarin derivatives in the binding pocket of DNA-topoisomerase I complex were performed by molecular docking (using AutoDock 3.0). It was found that seven new lamellarin alkaloids isolated from the Indian ascidian *Didemnum obscurum* displayed intercalation into the duplex DNA as same as lamellarin D which isolated from the prosobranch mollusc *Lamellaria sp.*

Thipnate, P., J. Liu, M. Chittchang, N. Thasana, P. Ploypradith, A.J. Hopfinger and S. Hannongbua. **4D-QSAR of lamellarin derivatives active against human hormone-dependent breast cancer (T47D).** 6th Princess Chulabhorn International Science congress, Shangri-La Hotel, Thailand, 25-29 November 2007.

4D-QSAR of Lamellarin Derivatives Active Against Human Hormone-Dependent Breast Cancer (T47D)

Poonsiri Thipnate[1,2], Jianzhong Liu[3], Montakarn Chittchang[4], Nopporn Thasana[4], Poonsakdi Ploypradith[4], Anton J. Hopfinger[3], and Supa Hannongbua[1,2,*]

[1] Department of Chemistry, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand, [2] Center of Nanotechnology KU, Kasetsart University and NANOTEC Center of Excellence at Kasetsart University, [3] College of Pharmacy MSC09 5360, 1 University of New Mexico, Albuquerque, New Mexico 87131-0001 [4] Chulabhorn Research Institute, Vibhavadee-Rangsit Highway, Bankok 10210, Thailand.

The University of New Mexico

Introduction

Lamellarin, marine polyaromatic pyrrole alkaloids, have been isolated from different sources such as ascidians, mollusks, and sponges. Among more than 30 lamellarins, many lamellarins exhibit inhibition of HIV-I integrase, cytotoxic activities against tumor cells *in vitro*, and insensitive to Pgp-mediated drug efflux.[1] Especially, some lamellarins have been shown cytotoxic activities against human breast cancer (T47D) which is the one of serious problem for woman in the would. However, there is only one target enzyme topoisomerase IB that has been known mechanism. Nowadays, it still lacks of knowledge about structure-activity relationships of lamellarin derivatives active against T47D. Therefore, receptor-independent four dimensional quantitative structure-activity relationship (RI 4D-QSAR) analysis is applied for exploring more details about QSAR of lamellarin anti-breast cancer due to the geometry of receptor no available. This analysis incorporates conformation and alignment freedom into the development of 3D QSAR models by performing ensemble averaging (the fourth dimension). The descriptors in 4D-QSAR analysis are the grid cell (spatial) occupancy measures of the atoms composing each molecule in the training set realized from the sampling of conformation and alignment spaces.[2-3]

In this study, 25 lamellarins, central pyrrole ring fused to adjacent aromatic rings and with the quinoline moiety containing a C5-C6 single bond and double bond as shown in Fig. 1(a) and 1(b), respectively, are used to develop RI 4D-QSAR models for provides new knowledge to design the potent anti-breast cancer candidate and for understanding enzyme active site topography.

Lamellarin	R_1	R_2	R_3	R_4	R_5	R_6	R_7
C	OMe	OMe	OMe	OH	OMe	OMe	OH
E	OH	OMe	OMe	OMe	OH	OMe	OH
F	OH	OMe	OMe	OMe	OMe	OMe	OH
G	H	OH	OMe	OMe	OH	OH	OMe
I	OMe	OMe	OMe	OMe	OMe	OMe	OH
J	H	OH	OMe	OMe	OMe	OMe	OH
L	H	OH	OMe	OMe	OH	OMe	OH
T	OMe	OMe	OMe	OMe	OH	OMe	OH
U	H	OMe	OMe	OMe	OH	OMe	OH
Y	H	OMe	OH	OMe	OH	OMe	OH
χ	H	OH	OMe	OH	OMe	OMe	OH
K 6'-bromide	OAc	OMe	OMe	OAc	OMe	OMe	OAc
U 6'-bromide	H	OMe	OMe	OMe	OAc	OMe	OAc
χ 6'-bromide	H	OAc	OMe	OAc	OMe	OMe	OAc

All bioactivity are not shown

Lamellarin	R_1	R_2	R_3	R_4	R_5	R_6	R_7
B	OMe	OMe	OMe	OH	OMe	OMe	OH
D	H	OH	OMe	OH	OMe	OMe	OH
M	OH	OMe	OMe	OH	OMe	OMe	OH
N	H	H	OH	OMe	OMe	OH	OMe
W	OMe	OMe	OMe	OMe	OH	OMe	OH
α	H	OMe	OMe	OMe	OH	OMe	OH
X	OH	OMe	OMe	OMe	OH	OMe	OH
ε	OH	OMe	OMe	OMe	OMe	OMe	OH
ζ	OMe	OMe	OMe	OMe	OMe	OMe	OH
Dehydrolm. J	H	OH	OMe	OMe	OMe	OMe	OH
Dehydrolm. Y	H	OMe	OH	OMe	OH	OMe	OH

Figure 1. Structure of lamellarins: (a) unsaturated lamellarins (b) saturated lamellarin

Methodology

Ten operation steps of RI 4D-QSAR [4]

- *Generate the initial 3D models for all compounds in the training set.*
- *Perform a conformational ensemble sampling of each compound to generate its conformational ensemble profile, CEP.*
- *Select the trial set of interaction pharmacophores elements, IPEs.*
- *Select a trial alignment.*
- *Generate the reference grid cell space. Place each conformation of each compound in the reference grid cell space according to the alignment and record the grid cell occupancy profile, GCOP, for each IPE and choice in occupancy measure. The resulting composite sets of grid cell properties constitute the set of grid cell occupancy descriptors, GCODs.*
- *Perform a PLS data reduction of the entire set of GCODs against the biological activity measures.*
- *Use the most highly weighted PLS GCODs, and any other user selected descriptors, for the initial basis set in a GA analysis.*
- *Return to STEP 4 and repeat STEP 4-7 unless all trial alignments have been included in the analysis.*
- *Select the optimum set of 3D-QSAR models with respect to alignment and any of the methodology parameters.*
- *Adopt the lowest energy conformer state, from the set sampled for each compound, which predicts the maximum activity using the optimum 3D-QSAR model, as the "active" conformation (shape).*

Results & Discussion

4D-QSAR models were constructed for each 8 trial alignments listed in Table 1. Both r^2 and xv-r^2 from the best corresponding 4D-QSAR models were shown in Table 2. This table , the alignment of lamellarin derivatives is not significant to the 4D-QSAR model. Based on the highest r^2 and xv-r^2, alignment 3 is the best alignment.

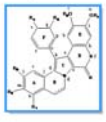


Table 1. Set of trial alignment used in constructing the 4D-QSAR model

Alignment	First atom	Second atom	Third atom
1	a	b	c
2	d	e	f
3	g	h	i
4	j	k	l
5	l	m	n
6	o	p	q
7	k	o	r
8	s	t	u

Table 2. Cross-validation correlation coefficient, xv-r^2, and correlation coefficient, r^2, of the best 4D-QSAR model for each trial alignment

Alignment	1	2	3	4	5	6	7	8
r^2	0.944	0.966	0.999	0.997	0.966	0.997	0.999	0.999
xv-r^2	0.929	0.962	0.998	0.995	0.964	0.963	0.997	0.998

Fig. 2, there is not much improvement in r^2 and xv-r^2 by addition of more descriptor than 3 descriptors. Therefore, all consequent procedures for set up the 4D-QSAR model were performed by using alignment 3 and 3 descriptor.

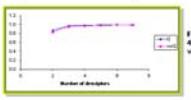


Figure 2. Plot of the number of 4D-QSAR model descriptors versus r^2, and xv-r^2

The best 4D-QSAR model from 25 lamellarins in training set is

$$-\log(IC_{50}) = 5.14 + 16.90GC1\ (-5,6,2,np) - 56.33GC2\ (-3,4,-5,any) + 64.62GC3\ (-1,5,0,np)$$
$$r^2 = 0.971,\ xv\text{-}r^2 = 0.947,\ n = 25 \qquad (1)$$

When GC_i (x, y, z, X) is the i th significant GCOD descriptor located at (x, y, z) and having the X type IPE as defined in table 3, n is number of compound used for set up

All predicted activities from eq. (1) are within ± 1 log unit of activity. Hence, 4D-QSAR model shows a good predictive ability for lamellarin derivative.

To distinguish the 3D-pharmacophore of high activity compound from all 25 lamellarins training set, 4D-QSAR analysis of six lamellarins (D, M, N, X, ε, and Dehydrolam J) were constructed by same method.

Table 3. Set of Interaction Pharmacophore Elements (IPEs) used in the RI 4D-QSAR Analysis

IPE description	Symbol	Number code
all atoms in the molecule	any	0
nonpolar atoms	np	1
polar atoms of positive partial charge	p+	2
polar atoms of negative partial charge	p-	3
hydrogen bond acceptor atoms	hba	4
hydrogen bond donor atoms	hbd	5
aromatic atoms	aro	6

The best 4D-QSAR model from high activity lamellarins training set is

$$-\log(IC_{50}) = 10.31 - 50.52GC1\ (-1,1,-6,any) + 1.58GC2\ (-1,4,-6,np)$$
$$r^2 = 0.997,\ xv\text{-}r^2 = 0.984,\ n = 6\ (2)$$

3D-pharmacophores embedded in the 4D-QSAR model given by eq. 1 and 2, are shown in Fig. 3. The reference 3D structures used in Fig. 3 are predicted active conformation of the most active compound (lamellarin D). Two red spheres near hydroxy (R_2) group specify pharmacophore sited that decrease potency when occupied all atom types (any) since they have regression coefficient equal to -56.33 for 25 lamellarins training set and -50.52 for high activity lamellarins training set. However, blue sphere, positive regression coefficient, was found near hydroxy (R_2) group for high activity lamellarins training set. Nonpolar atoms occupy of this GCOD will increase compound potency. The other two GCODs for the complete range in activity have positive regression coefficient with nonpolar IPE type. These mean that nonpolar atoms near hydroxyl (R_4) group and methoxy (R_5) group will improve compounds potency. From Fig. 3, there is no particular GCOD site that separates low and high activity. Rather, the space around the hydroxy (R_2) group is free and nonpolar occupancy above this hydroxy group near the methoxy (R_3) group is possible to be the key factors for high activity. Moreover, 4D-QSAR model of high activity compounds provided self-consistent pharmacophore of 25 compounds training.

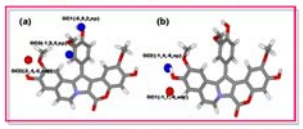


Figure 3. The 3D-pharmacophores of 25 lamellarins training set (a) and high activity lamellarin training set (b) 4D-QSAR models relative to predicted active conformation of the most active compound (lamellarin D). Red spheres refer to negative regression coefficient and blue spheres refer to positive regression coefficient

Reference

1. Bailly, C. *Current Medicinal Chemistry - Anti-Cancer Agents* **2004**, 4, (4), 363-378.
2. Hopfinger, A. J.; Wang, S.; Tokarski, J. S.; Jin, B.; Albuquerque, M.; Madhav, P. J.; Duraiswami, C. *J. Am. Chem. Soc.* **1997**, 119, (43), 10509-10524.
3. Liu, J.; Pan, D.; Tseng, Y.; Hopfinger, A. J. *J. Chem. Inf. Comput. Sci.* **2003**, 43, 2170-2179.
4. 4D-QSAR User's Manual, Version 3.0, The ChemBats21 Group, Inc., 1780 Wilson Dr., Lake Forest, IL 60045, 2003.

Acknowledgement

This work is supported by Commission on Higher Education, Faculty of Science, Kasetsart Univeristy (KU), and Rajabhat Petchaburi University. We also kindly acknowledge support from Professor A. J. Hopfinger, College of Pharmacy, University of New Mexico. Moreover, The Postgraduate Education and Research Program in Petroleum, Petrochemical Technology and Advanced Materials, KU Research and Development Institute (KURDI), LCAC and computing center of KU, Center of Nanotechnology KU, and NANOTEC Center of Excellence at KU are also grateful for research facilities.

Kasetsart University

CURRICULUM VITAE

Name and Family Name:	Miss Poonsiri Thipnate
Nationality:	Thai
Religion:	Buddhism
Birth Date:	16/10/1975
Birth Place:	Pitsanulok province
Age:	33 Years
Sex:	Female
Position:	Chemistry lecturer, Department of Chemistry, Faculty of Science, Phetchaburi Rajabhat University
Mailing address:	81/2, Moo 2, Singhawat road, Bankhlong, Pitsanulok, 65000
E-mail Address:	nui_tipanet@yahoo.com, thipnate@gmail.com

Academic qualifications:

1. **Bachelor degree in Science**
   Chang Mai University
2. **Master degree in Chemistry**
   Chang Mai University
3. **Doctor of Philosophy in Chemistry**
   Kasetsart University

Scholarships:

1997-1999	National Metal and Materials Technology Center, MTEC, Thailand supports master degree program
1999-2001	Phetchaburi Rajabhat Institute, Thailand supports master degree program
2004-2007	Phetchaburi Rajabhat Institute, Thailand supports doctor of philosophy program
2007	Commission on Higher Education, Ministry of Education, Thailand supports for visiting at University of New Mexico, USA (4 months)

Scholarships:

2007	Faculty of Science, Kasetsart University partially supports for visiting at University of New Mexico
2007-2008	Graduate School Kasetsart University Scholarship supports for thesis work in doctor of philosophy program
2008	Phetchaburi Rajabhat Institute, Thailand supports oral presentation in EuroQSAR 2008 symposium at Uppsala, Sweden
2008	Faculty of Science, Kasetsart University supports oral presentation in EuroQSAR 2008 symposium at Uppsala, Sweden
2008	Graduate School Kasetsart University Scholarship supports oral presentation in EuroQSAR 2008 symposium at Uppsala, Sweden
2008	EuroQSAR 2008 symposium supports for registration free

Research Interests:

Computational Chemistry, Computer-Aided Molecular Modeling and Drug Design, 3D-QSAR, 4D-QSAR, Polymer Science, Biomedical Polymer

Referee at Kasetsart University	Assoc.Prof. Dr. Supa Hannongbua Address: Department of Chemistry, Faculty of Science, Kasetsart University, Bangkok, 10900, Thailand Telephone: +662-5625555 ext.2140 Fax: +662-5625555 ext. 2176 E-mail Address: fscisph@ku.ac.th