Analysis of Quantitative Structure-Activity Relationship, Pharmacophore, and Molecular Docking of Tetracyclic Indenoquinoline Derivatives as Anticancer Agents

Nur Syamsi Dhuha[1]*, Daryono Hadi Tjahjono[2]

[1]Pharmacy Departement/ Alauddin Islamic State University, South Celebes, Indonesia
[2]School of Pharmacy/ Bandung Institute of Technology, West Java, Indonesia

Abstract: Topoisomerases have been the focus in the treatment of some diseases, such as bacterial gyrase (topoisomerase II) and topoisomerase IV which are the target of two classes of antibiotic drug: quinolones and coumarins; likewise Topoisomerase I and topoisomerase II are in cancer treatment. QSAR analysis has been performed on *tetracyclic indenoquinoline* derivatives as anticancer agents in three cell lines (HeLa, A-549, and MCF-7 cells). Ligand interaction analysis was performed by docking the derivatives against topoisomerase I receptor. Furthermore, the design of new derivatives and their prediction activities were based on the result of QSAR analysis and their interaction using MOE software. Design of new derivatives resulted 10 compounds that have been predicted better in biological activity than that of analogous compounds (camphtetocine) and the test compounds.

Keywords: QSAR, tetracyclic indenoquinoline, HeLa, A-549, MCF-7, pharmacophore, docking molecule.

Introduction

Topoisomerases enzymes were found on all type of virus to humans. These enzymes having monomer subunit structure acted to regulate DNA supercoiling by catalyzing the winding and unwinding of DNA strands[1,2]. The topoisomerases contain a nucleophilicthyrosine in which it uses to promote strand scission. Transesterification between an enzyme tyrosyl and a DNA phosphate group leads to the breakage of DNA backbone bond and the formation of a covalent enzyme-DNA intermediate as presented in figure 1[3,4,5].

Camptothecins works by inhibiting enzyme and DNA dissociation or segregation, so the formed slit will be permanent. In normal cells, this damage can be repaired moreefficiently than cancer cells. Cracks or breaks in the DNA permanent strand can lead to cell death[6].

Two water-soluble derivatives, topotecan and irinotecan, have been accepted by FDA as drugs of ovarian cancer, cervical, lung, and colon. As others anti-neoplastic drugs, problems of toxicity and the emergence of multidrug resistance through induction of ATP-dependent P-glycoproteins and other efflux pumps confines clinic efficacy of these molecules[7,11].

Therefore, attempts to design inhibitors that can overcome this limitation become very important. It is illustrated well by indenoisoquinoline derivatives which have the similar mode of action to camptothecin and its analog but have better chemical stability, form a complex that is longer with TOPO I, and are not the substrates of efflux pumps[8-10].

Based on data of activity indenoquinoline tetracyclic compounds series, this research conducted the study of quantitative structure activity relationships (QSAR) of inhibiting cancer cell lines: (i) breast cancer cells (MCF-7); (ii) the lung cancer cells (A-549); (iii) the cell of cervical cancer (HeLa).Further research is searching pharmacophore and docking molecule of indenoquinoline tetracyclic derivative compounds existed in TOPO I receptor. Based on phsycochemical properties of the compounds, the found pharmacophore and the interactions of compounds against receptors, we then proposed a new derivatives compounds activity which are more optimal than those of previous compounds.
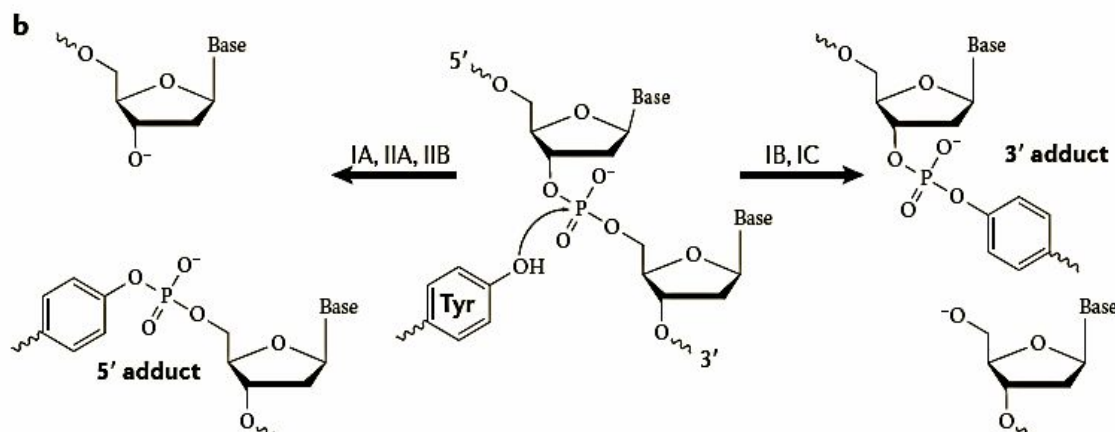
Figure 1. The DNA cleavage reaction. Topoisomerases catalyse strand scission by forming a reversible, covalent enzyme–DNA adduct through their active-site tyrosine. Type IB and IC topoisomerases become attached to 3′ DNA ends, and type IA, IIA and IIB topoisomerases attach to 5′ DNA ends[l3]

Experimental

Materials and Methods

Material

Chemical structure data of Tetracyclic indenoquinoline derivatives and the anticancer activity in HeLa cells, A-549, and MCF-7 are in *Journal of Bioorganic and Medicinal Chemistry Syntesis and evaluation as potential anticancer agents of novel tetracyclic indenoquinoline derivatives by* Chakrabarty Croft, Marko, and Moyna[11].The x-ray crystallographic 3D structure of topoisomerase I (PDB code: 1T8I) was downloaded from online Protein Data Bank (http://www.rcsb.org/pdb/explore/explore.do?structureId=1T8I). 3D structures of Tetracyclic Indenoquinoline derivatives were built using *Hyperchem*®*Release 8.02*[12].

Docking of tetracyclic indenoquinolinederivatives to DNA binding region of topoisomerase I were carried out by using MOE 2009.10[13].

Quantitative structure-activity relationship

Compound that have been drawn then they were optimized using Ab initio with the parameter set of a minimum base on Hyperchem®8.02. The optimized compounds were calculated by the phsycochemical properties usingMOE 2009.10. Analysis method used analysis of multilinier regression approach with SPSS 17.0[14] where the psycochemical properties became descriptors and named independent variable. The biological activity in the initial data were modified with the formula (Log 1/C).The values then became dependent variable. Then the equation models that had a high statistical parameters were validated using Leave One-Out Cross Validation methods (LOOCV) with a q^2 parameters.

Searching pharmacophore of tetracyclic indenoquinoline series

Searching pharmacophore of indenoquinoline tetracyclic series were the ligand annotation point quest that had an important role in generating the activity of TOPO I inhibition. Searching pharmacophore began with looking at the interacions on analog compound (camphtetocin) against TOPO I via crystal protein of NMR results downloaded from the RSCB.PDB[15] (code: 1T8I) through software MOE 2009.10.

Docking molecule

Preparation of ligand was done by displaying test ligand then it would be protonized the ligand for adding hydrogen and partial charge. The receptor preparation was done by displaying the receptor then the solvent was removed from the structure. Made a binding pocket as a target of the drug. The best docking positions were based on RMSD<2 and lowest scoring value.

Design new tetracyclic indenoquinoline derivative compounds

Design of new derivative compounds made by taking a previous compound that had a best anticancer biological activity. On designing new derivative compounds, the presence of aromatic ether functional group must be considered. After that, compound modification was made on the aromatic region by considering the phsycochemical properties from the QSAR model. Then, moleculer docking has been done to see the interaction of the new compounds to TOPO I receptor.

Results and Discussion

Quantitative structure-activity relationship

After microscopic data or phsycochemical properties as descriptors have been collected, the statistical calculation was done to connect a lot of descriptors to biological activities. Parameters were used to select the best regression equations is r, r^2, SE, Fisher criterion value, q^2, and value of Pearson correlation. The best regression equation and selectedequations of tetracyclic indenoquinoline compound series are presented on below table.

Hela cell line

Basedon QSAR analysis, the pshycochemical properties influence the biological activity of tetracycline indenoquinoline derivative against HeLa cells are globularity, total energy, heat of formation, LUMO energy, and electronic energy. Four pshycochemical properties are electronic parameters and another is steric parameter. This is in line with the theory of chemical structure of drug and its pshycochemical properties[16] the chemical reactivity and drug ability to interact with its receptor depend on its electronic-structure composition, properties, and interactions of all electrons in the molecule. The Total Energy (E_{tot}) is the total energy of electrons in the molecule or in the other words the minimum energy of molecules conformation. LUMO energy (E_{LUMO}) is the lowest energy of orbital which contains no electrons in a molecule[17]. Molecules with a small LUMO energy is more able to accept electrons than with the high LUMO energy. As a result, LUMO is so close to the affinity of electron that is a measure of the molecular electrophilisities. The Affinity of electron close to the energy of electronic. It showed that reactivity of tetracyclin indenoquinoline derivative compounds in inhibition of cell line largely was determined by the distribution of electrons in the molecule. In addition, the electronic parameters, globularity as a steric factor also affected the activity of these derivatives. Globularity was calculated from the ratio between volume and surface area of the molecule. On the equations, coefficient of globularity was signed by positive. This indicated that the highest of globularity or the perfection of molecular topology, it would be the highest inhibition activity to the HeLa cell line.

A-549 cell line

On the selected equation, the psycochemical properties were influence to the inhibition activity of tetracyclin indenoquinoline derivatives, coefficient of partition (Log P), globularity (glob), electronic energy (E_{ele}), LUMO energy (E_{LUMO}), and total energy (E_{tot}).On the equation, coefficient of globularity was signed by positive.It indicated that the highest of globularity or the perfection of molecular topology, it would be the highest inhibition activity to the A-549 cell line. Electronic and LUMO energy were signed by negative on the equation so that it can be said that the more electronic energy reduced, the more inhibition activity of the compound series increased. Like with LUMO enery, the higher opportunity of compound to accept electrons (small E_{LUMO} value), the higher inhibition activity to the A-549 cell line. On the other hand, the total energy was signed by positive so that the higher minimum energy of molecular conformation, the higher biological activity of compounds. Another descriptor was globularity related to the stereo of structure. Positive

coeffisient showed that the increase in globularity or the perfection of structure shape can improve biological activity. Parameter with correlated to hydrophobic descriptor is coeffisient of partition (Log P). Based on QSAR study, the coefficient of partition provides a contribution to inhibition activities. Coefficient of partitionaffects greatly the characteristics of drug transport or how the drug reaches the site of action (pharmacophore). Coefficient of partition determines which human chains can be reached by drug compound. On equation, coefficient of Log P shows negative. This indicates that the decrease in solubility of lipids will also increase inhibition activity of compound.

Mcf-7 cell line

MCF-7 cell is a cell culture that represents breast cancer cells. In QSAR studies, the pshycochemical properties of tetracyclic indenoquinoline derivatives whics influences the biological activities is volume, E_{LUMO}, ASA_H, E_{HOMO}, and mr.As the other two cell lines, the equation contains electronic descriptor. E_{LUMO} and E_{HOMO} is descriptor which is related to the chemical reactivity. In a chemical reaction, electron transfer occurs from the HOMO to LUMO. On the structure of hydrogen bonding, electron transfer occurs from molecules that havelone-pair electrons or weak (π) electrons to molecules that have a more polarized bonds. This study is related to the results of docking molecule that interaction of the compound series is done by hydrogen bond interaction.On the equation, coefficient of E_{LUMO} was signed by negative sign which indicates the compound has a high opportunity to accept electron, then it will increase the biological activity. Whereas the coefficient of E_{HOMO} signed by positive shows if the compound have a high opportunity to provide electrons, thus its activities will be increasing. Another descriptor is ASA_H which related to the partial charge of molecule and its conformation. ASA_H or accessible surface area (Å) of all hydrophobic is surface area that can be passed by water on all hydrophobic compounds. Coefficient on the equation is negative. It shows that the smaller accessible surface area, the higher biological activity of compound. Another descriptor is mr or *molecular refractivity.* mr belongs to the 2D descriptors and steric parameters. Based on reference, mr is often related to the measure of non-lipophilic molecular interaction. mr is also a parameter that indicates bulky structure. The last descriptor is volume which depends on structure connectivity and conformations. Volume is defined as van der waals volume calculated using grid approximation with spacing 0,75 Å.

Searching pharmacophore

Pharmacophore concept plays an important role in guiding the process of drug discovery. It assist medicinal chemists in gaining insight into the interaction between the ligand and the receptor when the receptor structure could not be proved experimentally[18].The first interaction is seen by exposure between the two aromatic groups on DA 113 amino acid residual (figure 2). The second interaction occurs when amino acid residual ARG 364 as donor to the N group then hydrogen bonding occurs. Compound synthesized by Shubhashis et al were 20 compounds. They are divided into 2 series. First, 4a-4j, N group on second cyclic bonded with ARG 364 is modified with addition to be NH. Whereas in the second series, 5a-5j, two cyclic above is retained.
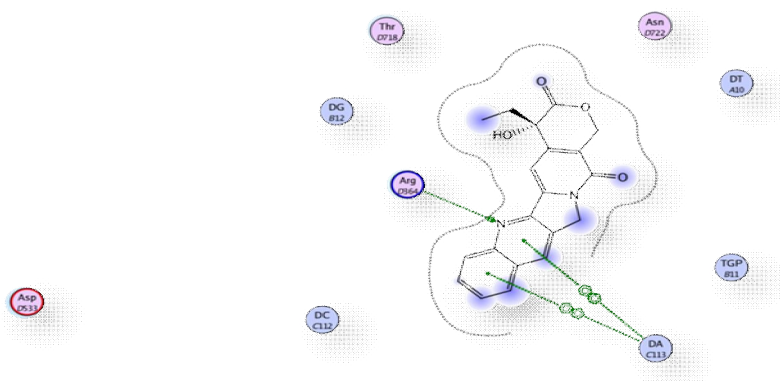
Fig. 2. Interaction of Camphtetocin ligand to TOPO I receptor
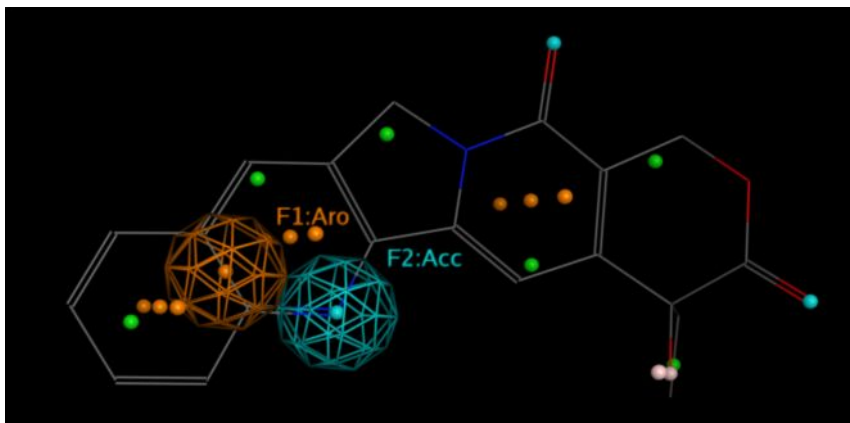
Fig. 3. Camphtetocin ligand and the analog Pharmacophore

After in vitro test with MTT colometric assay on three cancer cells, inhibition of growth cells is significantly different.The first series, compounds 100-15.00 times is less active than camphtotecin as analog compound.While the second series shows the result can be compared fairly with the analogue compound. Furthermore, there are some compounds that have larger activity. It becomes the compelling reasons that both pharmacophore are due to the camphtotecin and the analog.

Next, With query pharmacophore feature on MOE 2009.10, pharmacophore file was made and used as a base in a docking simulation as presented in figure 3.

Pharmacophore files(*.PH4)thathave been made become filter in performing docking molecular simulation. Validation of molecular docking method is done by redocking with native ligands on the binding site. RMSD value is 0,7909 and Scoring Function (Kcal/mol) is -19,9862. This indicates that the method is valid (RMSD < 2) where the position of the docking ligand is similar to the original ligand.

The table above indicates that the bonding occuring for all compounds is hydrogen bonding with near bonding distance. The Scoring function values are small. It also pointed out that this interaction takes a place spontaneously. Hydrogen bonding can probably occur due to electron transfer among molecules. This is in line with the predictors on the equations that dominates three cell lines are an electronic parameter.

The interaction occurs in a test compounds involving a number of amino acids, namely DA 113, ASN 352, ASN 722, ARG 364, LYS 425, LYS 751, DC 112, and THR 718. While the interaction happens on the analog compound (camphtotecin) involves only amino acids DA 113 and ARG 364. The docking of compounds 4a-4g do not have any bonding or exposure among the aromatic rings in its cyclic. Some of other compounds of the series have exposure to aromatic moieties with residual amino acid on DA 113 similar to analog compounds, but it is not supported with bonding on the amino residual acid ARG 364. The interaction probably make the first series of compound has far less activities than the analogue compounds. Some compounds of the second series do not show any bonding at result of docking molecules.

Design new tetracyclic indenoquinoline

New compounds that have been built in 3D structure through Hyperchem 8.02 were calculated the descriptors by using software MOE. Next, the predicted biological activity values were calculated by prediction equation models. The best activity of new derivative compounds were docked to TOPO I reseptor.

Based on results, compound 10 has a better activity against HeLa cells compared to 5j. Modification is made in the form of the addition of p- SOCH3 benzene on the alkyl compounds of serie 5. The changed function group effects in higher globularity than 5j. The total energy and electronic energy of compound 22 is also lower than 5j. Each steric and electronic parameter influenced in increasing activity. In interactional study, there is a hydrogenbonding with amino acid residual ARG 364. Besides, some aromatic groups also interacted with amino acid residual DA 113 and LYS 425. These two amino acids were related to the interactional study of camphtotecine as analog compound.

The new derivative compound which had the best activity to A549 cell line was 2. The predicted biological activity GI_{50} of compound 2 is 0,000153 whereas 5a is 0,005. It showed that compound 2 had good inhibition characteristic to A549 cell line. The made modification was to substitute alkyl group with 2-Floro, 4-Ometil Benzen. This modification gave globularity compound 2 bigger than compound 5a whereas electronic energy and LUMO energy compound 2 smaller than 5a. Similar to HeLa cell line, the change value of steric parameter and electronic increased activity to A549 cell line significantly. The docking result of compound 2 to the receptor had 1 hydrogen bonding to amino acid residual ARG 364 and 3 aromatic rings exposed with amino acid residual DA 113 and LYS 425.

The new derivative compounds which had the best predicted activity to MCF-7 cell line was compound 8. The modification was done by alkyl group substitution $pN(CH_3)_2$Benzene. The modification influenced van der walls volume from compound 8 to be higher than compound 5a. similar to E_{HOMO} compound 8 which was bigger than compound 5a, this modification impacted very well to activity. This interaction occurring to compound 8 with the reseptor of topoisomerase I was atom N bonding with hydrogen with acid amino residual ARG 364. Then, there were 3 aromatic rings exposed with aromatic side from amino acid residual DA 113 and TGP 11

Table 1. Best Regression equations for cell lines and the statistical parameters.

Cell Lines	Equation Models	r^2	q^2	Coefficient ofrrelation of GI_{50} experiments and GI_{50} Predictions	Significance of Pearson Correlation of the descriptors and biological activities
HeLa	Equation 2	0,983	0,9565	0,9844	3 SS
	Equation 5	0,982	0,9621	0,9829	3 SS
A549	Equation 29	0,972	0,9456	0,9720	3SS
	Equation 32	0,972	0,9457	0,9719	3SS
MCF7	Equation 13	0,973	0,951465	0,9706	1SS
	Equation 15	0,970	0,953861	0,9733	2 SS

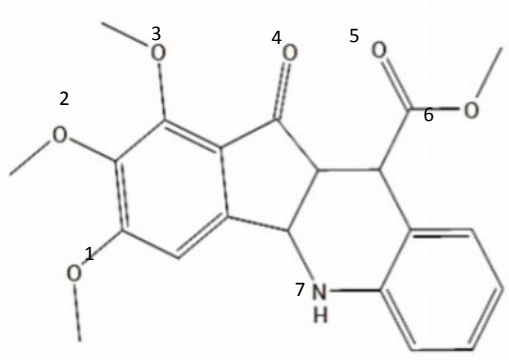**Table 2.Selected Equations of Tetracyclic Indenoquinoline Compound Series**

Biological Activity	Selected Equation
Log 1/GI_{50} Sel HeLa	$-3{,}418 + 9{,}144 \text{ glob} + 9{,}038\times10^{-5} \text{ AM1\_E} - 0{,}008 \text{ AM1\_HF} - 6{,}369 \text{ AM1\_LUMO} - 7{,}058\times10^{-6} \text{ AM1\_Eele}$
Log 1/GI_{50} Sel A-549	$-3{,}312 - 0{,}493 \text{ Log P} + 4{,}146 \text{ glob} - 6{,}282\times10^{-6} \text{ AM1\_Eele} - 6{,}467 \text{ AM1\_LUMO} + 5{,}858\times10^{-5} \text{ AM1\_E}$
Log 1/GI_{50} Sel MCF-7	$36{,}399 + 0{,}206 \text{ Vol} - 9{,}35 \text{ AM1\_LUMO} - 0{,}0494 \text{ ASA\_H} + 3{,}849 \text{ AM1\_HOMO} - 5{,}017 \text{ mr}$

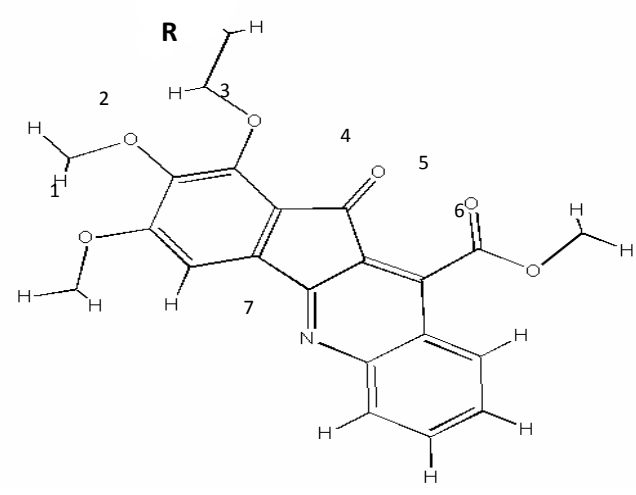**Table 3. The DockingResultof Tetracyclic indenoquinoline derivatives**

No	Compounds	RMSD	*Docking Score* (S) (kkal/mol)	Number of bonding		Bond distances (Å)	The amino acid residual	Functional group that is bonded
1	4a	1,524	-15,2519	-	-	-	-	-
2	4b	1,1418	-12,7000	Hydrogen	3	1,88	DA 113	HN no 7[don]
						2,71	ASN 352	O no 4[acc]
						2,03	ARG 364	O no 1[acc]
3	4c	1,6899	-13,2920	Hydrogen	1	2,15	LYS 751	O no 5[acc]
4	4d	1,1084	-18,9850	Hydrogen	2	2,05	LYS 425	O no 5[acc]
						2,82	DC 112	O no 4[acc]
5	4e	1,8149	-13,4401	Hydrogen	3	2,86	DA 113	O no 3[acc]

						2,03	ARG 364	O no 1[acc]
						2,61		O no 2[acc]
6	4f	1,2325	-14,2602	Hydrogen	1	2,07	ARG 364	O no 1[acc]
7	4g	1,0073	-17,4915	-	-	-	-	-
8	4h	1,0486	-13,2137	Hydrogen	1	2,09	DA 113	HN no 7[don]
9	4i	1,7385	-15,1581	-	-		-	-
10	4j	1,6010	-9,1368	Hydrogen	1	2,58	ASN 722	O no 1[acc]
11	5a	0,6479	-13,9075	Hydrogen	4	3,78	THR 718	O no 3[acc]
						2,78		O no 2[acc]
						2,49	DA 113	O no 5[acc]
						2,43	ARG 364	N no 7[acc]
12	5b	1,0625	-14,0034	Hydrogen	2	3,66	THR 718	O no 2[acc]
						2,40	ARG 364	O no 7[acc]
13	5c	0,9980	-13,5741	Hydrogen	2	2,14	ARG 364	N no 7[acc]
						3,61	THR 718	O no 2[acc]
14	5d	1,1920	-13,2745	-	-	-	-	-
15	5e	0,7365	-13,5667	-	-	-	-	-
16	5f	0,9146	-14,4403	Hydrogen	1	3,96	THR 718	O no 2[acc]
17	5g	1,3048	-21,0999	Hydrogen	1	2,04	ARG 364	O no 4[acc]
18	5h	0,7971	-14,6211	Hydrogen	1	2,30	DA 113	O no 5[acc]
19	5i	1,6393	-14,5362	-	-	-	-	-
20	5j	1,2767	-14,9280	Hydrogen	4	2,45	ARG 364	O no 2[acc]
						2,14		O no 1[acc]
						2,75	DA 113	O no 3[acc]
						2,55	ASN 722	O pada R[acc]
21	Kamptotesin	0,7909	-19,9862	Hydrogen	1	2,24	ARG 364	N no 7[acc]

References:

1. DNATopoisomerase, Protein of the Months, 2013. Available from: <https://www.ebi.ac.uk/interpro/potm/2006_1/Page1.htm>. [Accessed on: 2 February 2016].
2. Champoux, James J., DNA TOPOISOMERASES: Structure, Function, and Mechanism, Annu. Rev. Biochem, 2001, 70, 369–413.
3. M. Vos, S., Tretter, E. M., Schimdt, B. H., dan Berder, J. M., All tangled up: how cells direct, manage and exploit topoisomerase function, Nature Reviews, 2011, 12, 827-841.
4. Wang, J. C.,Cellular Roles of DNA Topoisomerases: a Molecular Perspective, Nature Publishing Group, 2002, 3, 430-440.

5. Hsiang, Y.H., Leroy F. Liu, Monroe E. Wall, et al. Cancer Research: DNA Topoisomerase I-mediated DNA Cleavage and Cytotoxicity of Camptothecin Analogues, Cancer Res, 1989, 49, 4385-4389.
6. DNA Topoisomerase, Protein of the Months, 2013. Available from: <https://www.ebi.ac.uk/interpro/potm/2006_1/Page2.htm>. [Accessed on: 3 February 2016].
7. Gottesman, M. M., Tito Fojo,Bates, S. E. Multidrug resistance in cancer: role of ATP–dependent transporters, Nature Reviews Cancer, 2002, 2, 48.
8. Strumberg, D.; Pommier, Y.; Paull, K.; Jayaraman, M.; Nagafuji, P.; Cushman, M., Synthesis of Cytotoxic Indenoisoquinoline Topoisomerase I Poisons J. Med. Chem. 1999, 42, 446.
9. Pommier, Y.; Cushman, The Indenoisoquinolines Non-Camptothecin Topoisomerase I Inhibitors: Update and Perspectives, M. Mol. Cancer Ther. 2009, 8, 1008.
10. Beretta, G. L., Zuco, V., Perego, P., Zaffaroni, N., Targaetting DNA Topoisomerase I with Non-Camptothecin Poisons, Curr. Med. Chem. 2012, 19, 1238.
11. Chakrabarty, Shubhashis, Michael S. Croft, Melissa G. Marko, Guillermo Moyna, Syntesis and evaluation as potential anticancer agents of novel tetracyclic indenoquinoline derivatives, Bioorganic and Medicinal Chemistry , 2013, 21, 1143-1149
12. Hyperchem. Hyperchem^{TM} Relaease 8.02. Molecular Modelling System, Hypercube Inc.
13. Molecular Operating Environment. MOE molecular operating Environment, Chemical Computing Group, Montreal, Quebec, Canada
14. SPSS, 2008, SPSS Statistic 17.0 Release 17.0.1, Polar Engineering and Consulting, www.winwrap.com.
15. Human DNA Topoisomerase I (70 Kda) In Complex With The Poison Camptothecin and Covalent Complex With A 22 Base Pair DNA Duplex, 2005. Available from: <http://www.rcsb.org/pdb/explore/explore.do?structureId=1T8I>. [Accessed on: 2 February 2016].
16. Nogrady, Thomas, Raslim Rasyid, and Amir Musadad. Kimia Medisinal: Pendekatan Secara Biokimia Terbitan Kedua. ITB, 1992, 8-42.
17. Todeschini, Roberto, and Viviana Consonni, Handbook of Molecular Descriptor. WILEY-VCH Verlag GmbH, Weinheim, 2000, 362.
18. Abraham, Donald J., Burgers's Medicinal Chemistry and Drug Discovery: Drug Discover, 2010, 1, 71.

*****