BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 10(4)

Molecular designing and *in silico* evaluation of darunavir derivatives as anticancer agents

Manoj kumar Mahto[1, 2], Nanda Kumar Yellapu[3], Ravendra Babu Kilaru[4], Naga Raju Chamarthi[4] & Matcha Bhaskar[2]*

[1]Department of Biotechnology, Acharya Nagarjuna University, Guntur, AP, India, 522510; [2]Division of Animal Biotechnology, Department of Zoology, Sri Venkateswara University, Tirupati, AP, India, 517502; [3]Biomedical Informatics Center, Vector Control Research Center, Indian Council of Medical Research, Pondicherry, India, 605006; [4]Department of Chemistry, Sri Venkateswara University, Tirupati, AP, 517502, India; Matcha Bhaskar – Email: matchabhaskar2010@gmail.com; Phone: +91-877-2289304 Ext: 111; Fax: +91-877-2289555; *Corresponding author

Received April 11, 2014; Accepted April 17, 2014; Published April 23, 2014

Abstract:
Darunavir is a synthetic nonpeptidic protease inhibitor which has been tested for anticancer properties. To deduce and enhance the anticancer activity of the Darunavir, we have modified its reactive moiety in an effective way. We designed 9 analogues in ChemBioOffice 2010 and minimized using the LigPrep tool of Schrödinger 2011. These analogues can obstruct the activity of other signalling pathways which are implicated in many tumors. Results of the QikProp showed that all the analogues lied in the specified range of all the pharmacokinetic (ADMET) properties required to become the successful drug. Docking study was performed to test its anticancer activity against the biomarkers of the five main types of cancers i.e. bone, brain, breast, colon and skin cancer. Grid was generated for each oncoproteins by specifying the active site amino acids. The binding model of best scoring analogue with each protein was assessed from their G-scores and disclosed by docking analysis using the XP visualizer tool. An analysis of the receptor-ligand interaction studies revealed that these nine Darunavir analogues are active against all cancer biomarkers and have the features to prove themselves as anticancer drugs, further to be synthesized and tested against the cell lines.

Key Words: Darunavir, Cancer, EGFR, VEGFR2, Docking, HIV, ADMET

Background:
Cancer is a disease caused by normal cells changing so that they grow in an uncontrolled way. The uncontrolled growth causes a lump called a tumour to form. If not treated, the tumor can cause problems in one or more of the following ways: Spreading into normal tissues nearby, causing pressure on other body structures, spreading to other parts of the body through the lymphatic system or bloodstream. Cancer is one of the most dreadful diseases in the modern era, which is taking millions of life every year worldwide **[1].** Day by day it becomes more dangerous as it can affect any age group and any ethnicity. Cancer is not one disease but a conglomeration of more than 100 diseases, with lots of heterogeneity even within a single tumor type **[2]**. Nowadays various types of cancers are reported every month that spread with various mechanisms. There are over 200 different types of cancers for the reason that there are over 200 different types of body cells. The most frequently diagnosed cancers are bone, brain, breast, colon and skin cancers. There are many known causes of cancers like exposure to chemicals, drinking excess alcohol, excessive exposure to sunlight, and genetic differences, to name a few **[3]**.

Bone cancer can be either primary or secondary. Primary bone cancer starts in the bone which is uncommon, while the cancer

that spreads to the bone from the other parts of the body is secondary. With the spread of the disease, various proteins like TGF beta-I, VEGFR, MMP2, MMP9, collagen-X, bone sialoprotein(BSP), decorin, osteocalcin(OC), osteonectin(ON), osteopontin(OPN), and versican gets expressed. Alkaline phosphatase **[4]** (AP), tartrate-resistant acid phosphatase (TRAP) is a known biomarker for bone cancer. A tumor biomarker is a measurable substance in cancer patients, which reflects the presence of a tumor **[5, 6].**

Solid tumors in brain are the leading cause of death in children. Overexpressed telomerase, MMP-2, MMP-9, aberrant mRNA transcripts of EGFR, cytokerin 14 and 15, TGFα, TGF-β1, STAT-1, HLA-C, GST-2, and nuclear factor- kappa beta(NF-kB) are some of the most important biomarkers along with VEGFR-2 for prognostic or predictive valve in brain tumors **[7]**. The Combination of MMP-2 and VEGFR provides superior accuracy compared to any other combination or individual biomarker for the presence of disease, decrease with treatment, and can be tracked from source tissue to urine **[8].**

Ductal carcinoma is the most common type of breast cancer other than lobular carcinoma. ER positive, HER2 positive, and Triple negative/triple positive are the three main types of breast cancers. About 75% of all breast cancers are ER positive, they grow in response to estrogen hormone, about 65% of these are also PR positive, and they grow in response to another hormone, progesterone. In about 20-25% of breast cancers, the cancer cells make high quantity of HER2 (human epidermal growth factor receptor 2). It is one of the most aggressive and fast growing cancer. Triple negative breast cancers lack the estrogen and progesterone receptors and do not overexpress the HER2 proteins. Majority of the breast cancers caused by *BRCA1* gene are triple negative.

The American society of clinical oncology(ASCO) has recommended eight different protein related tumor markers for breast cancers: CA 15–13, CA 27–29, carcinoembryonic antigen, estrogen receptor (ER), progesterone receptor, human epidermal growth factor receptor 2 (HER2), urokinase plasminogen activator (uPA), and plasminogen activator inhibitor (PAI)-1. Starting 3 are biomarkers for monitoring, next 3 markers are for treatment planning and last 2 are biomarkers for recurrence risk prediction respectively. Other potential markers include p53, TGF beta-1, cathepsin D, kallikrein 14, cyclin B1, D1 and E1 **[9].**

Colon cancer and rectal cancer, collectively called as colorectal cancer, have many similar characteristics, is the third most common cancer in both men and women. About 95% of colorectal cancers develop in the glandular cells of the lining. Some of the genes that are important for the development of colon cancer are: APC gene (adenomatous polyposis coli), TP53 gene, k-Ras gene, MSH2 and MLH1 genes. TGF beta-1, COX-2, B-raf, TRAP and EGFR are the potential biomarker in colon cancer **[10].**

Skin cancer is by far the most common type of cancer. It can be divided into nonmelanoma and melanoma. Basal cell and squamous cell carcinoma are non-melanomus and treatment is very effective. Melanoma is a cancerous growth of melanocytes, develops in skin and usually brown or black, but can appear pink, tan, or even white. Melanoma may also develop in other parts of the body that contains melanocytes including the meninges, the digestive tract, the eyes and lymph nodes. MC1R, CDKN2A, RAS, RAF and BCL2 are some of the causative genes for melanoma skin cancer. Mutations in BRAF genes were observed in 66% of melanoma cell lines examined **[11, 12]**. B-raf is the best biomarkers for skin cancers.

Darunavir (brand name Prezista) is an antiretroviral drug, used for human immunodeficiency virus (HIV) infection. Darunavir is always used in combination with the ritonavir and other anti-HIV medicines. It functions as protease inhibitor (PIs) which slows down the progress of HIV infection by reducing the amount of virus in the body. Many anti-HIV drugs have been reported as anti-cancer drugs and lots of new research works are still going on. Earlier, Darunavir (DRV), Lopinavir (LPV) and Ritonavir (RTV) were examined *in vitro* and *in vivo* for its anticancer activity on Primary effusion lymphoma (PEL) cell lines by *Kariya et al* **[13]**. PEL is a non-Hodgkin lymphoma which occurs mainly in patients with advanced AIDS.

In our study we have designed the analogues of the Darunavir drug, its ADMET prediction and testing its anticancer activity against the biomarkers of the five main types of cancers i.e. bone, brain, breast, colon and skin cancers. Therefore, the most important aim of the work is to find out the anti-proliferative activity of Darunavir analogues by docking studies.

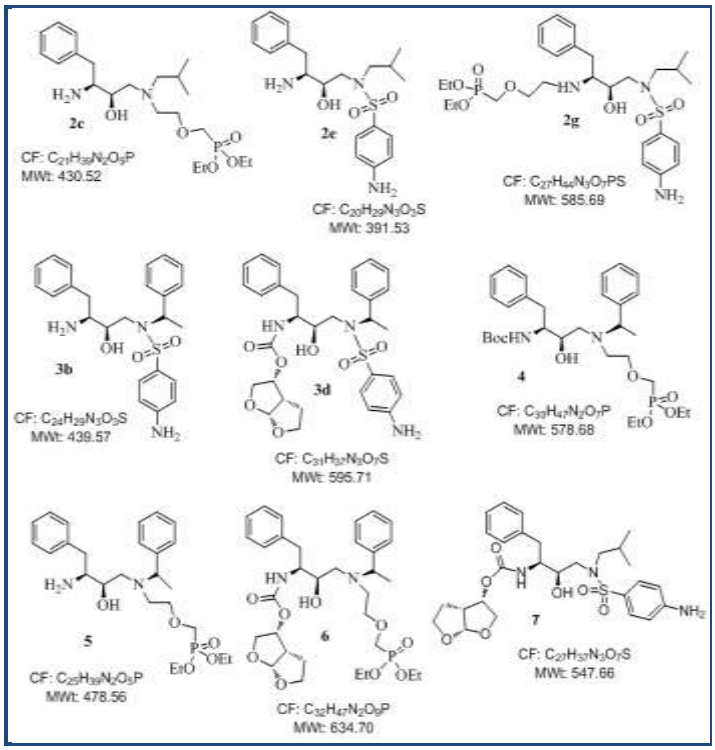

Figure 1: Structures of all the Darunavir analogues

Methodology:

Ligand preparation

We have taken the stereoselective and chemically active group for the designing of Darunavir analogues. The chemical structures of 9 analogues were drawn using ChemBioOffice 2010 and saved in mol file format shown in **(Figure 1).** LigPrep 2.5 is used from Schrödinger software to take 2D structures and

produce the corresponding low-energy 3D structures with correct chiralities **[14]**. Minimization is performed using OPLS_2005 force field and Epik ionizer at the standard pH of 5 to 9 and with maximum number of conformers per structure as 1000 with RMSD 1.0 Å.

ADMET Studies

QikProp 3.4 is performed on minimized structures to calculate the ADME properties for assessing the disposition and potential toxicity of ligand with in an organism and the overall pharmacological properties of these molecules **Table 1 A & B (see supplementary material)** justify that the molecules are biologically active without any toxic functional groups. QikProp predicts physically significant descriptors and pharmaceutically relevant properties like absorption, distribution, metabolism, excretion and toxicity (ADMET) **[15]**.

Protein preparation

We have selected specific cancer protein biomarkers, which are expressible in maximum number of patients, for each type of cancer to perform the interaction studies with designed anticancer compounds. Human is the source organism of all the selected Protein biomarkers, their PDB IDs and active sites information are given in **Table 2 (see supplementary material)**. Prior to docking, it is important to identify the binding site in the target protein. Protein preparation wizard from Schrödinger is used to locate and fix structural defects in raw state, (missing hydrogen atoms) then deleting unwanted chains, waters and het groups and finally we have performed optimization of the fixed structure. As active site is in the chain A of all the proteins, so we have deleted all chains except chain A. Energy minimization of all proteins was performed using OPLS_2005 force field and active site amino acid in all proteins are specified for the receptor grid generation.

Grid generation

Receptor grid generation was performed on all the minimized proteins using OPLS_2005 force field to define the active site, in the range of < 20Å radius by specifying amino acids name given in (**Table 2).**

Glide

Glide (Grid-based Ligand Docking with Energetics) uses a hierarchical series of filters to search for possible locations of the ligand in the active-site region of the receptor. On a grid, the shape and properties of the receptor are characterized by different sets of fields that provide gradually more accurate scoring of the ligand poses. Glide is used here to find the favourable interactions between ligand molecule and receptor molecule. We have used the XP (extra-precision) mode of Glide because it uses only good ligand poses to perform docking and it gives better result than the SP (standard-precision) mode **[16]**. Finally, we got the result in GScore (GlideScore) scoring function given in **Table 3 (see supplementary material).**, which uses the various components (vdW-Van der Waals energy, Coul- Coulomb energy, Lipo -Lipophilic term, HBond-Hydrogen-bonding term, Metal -Metal-binding term, BuryP -Penalty for buried polar group, RotB-Penalty for freezing rotatable bonds, Site -Polar interactions in the active site.) and the formula is given below:

GScore = 0.065*vdW + 0.130*Coul + Lipo + Hbond + Metal + BuryP + RotB + Site

By this docking study we came to know that most of our designed ligands are interacting to various Oncoproteins with sufficient selectively and specificity. The docking analysis is done and the results are presented in the form of image given in **(Figure 2)**, and hydrogen bond interactions between receptor and ligands, given in the **Table 4 (see supplementary material).**

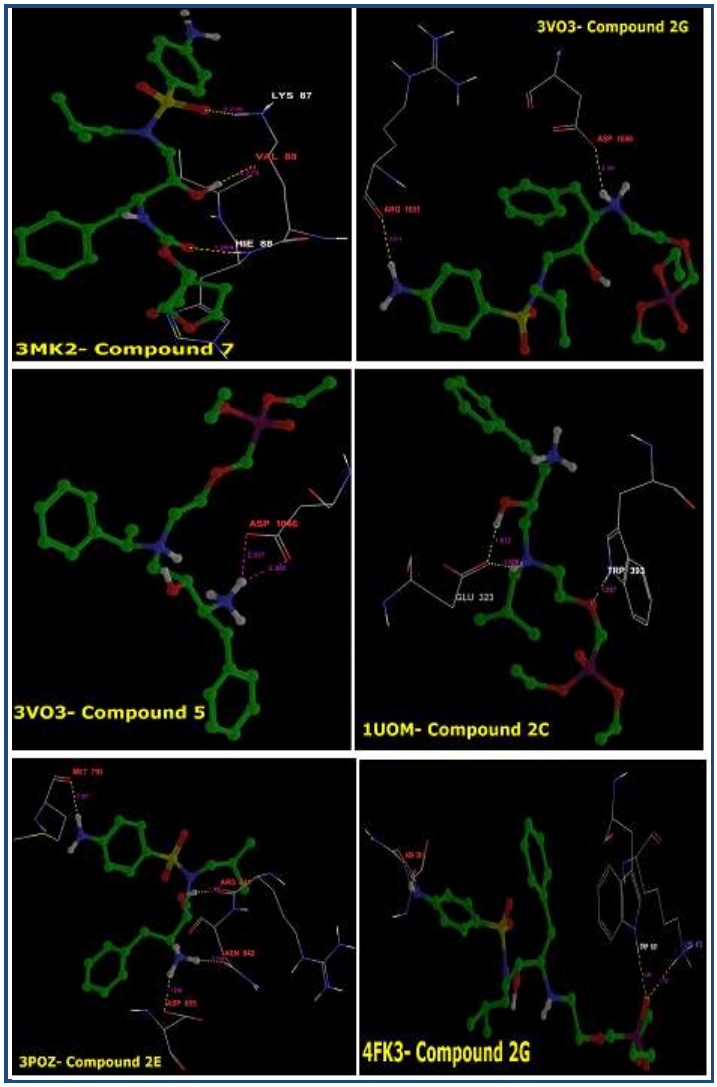

Figure 2: Docking of each protein with best ligands

Results & Discussion:

Since a long time anti-HIV drugs are being used in cancer treatments. After development of HIV protease in 1990, their potential anti-tumor properties have been under investigation because of their success in treating HIV-related Kaposi's sarcoma **[17]**. Recent studies have shown that protease inhibitors might be used to treat cancer patients, who are not infected with HIV as they have direct anti-angiogenic and antitumor effects that are unrelated to their antiviral activity **[18]**.

We have selected the ligand conformations with least potential energy for each of nine individual analogues after geometry optimization. Results of the QikProp shows that all the ligands lied in the specified range of all the pharmacokinetic properties described by Professor William L. Jorgensen to become the

successful drug. The most important pharmaceutical descriptor in QikProp is stars, which are the combination of all properties, a few numbers of stars reflects more drug likeliness in the range of 0 to 5, and we have the valve of 0 & 1 for four and five designed analogues respectively. Range of some of the other important parameters are, predicted range of central nervous system activity (CNS) is -2 (inactive) to +2 (active), molecular weight of the analogues must be within 130.0 – 725.0, total solvent accessible surface area (SASA) range is 300.0 – 1000.0, estimated number of hydrogen bonds donated (DonorHB) by the solute to water molecules in an aqueous solutions is 0.0 – 6.0, estimated number of hydrogen bonds accepted (acceptHB) by the solute to water molecules in an aqueous solutions is 2.0 – 20.0, predicted octanol/gas partition coefficient (QPlogPoct) is 8.0 – 35.0, predicted water/gas partition coefficient (QPlogPw) is 4.0 – 45.0, predicted octanol/water partition coefficient (QPlogPo/w) is -2.0 – 6.5.0, predicted aqueous solubility, log S (QPlogS) is -6.5 – 0.5, predicted brain/blood partition coefficient (QPlogBB) is -3.0 – 1.2, Van der Walls surface area of polar nitrogen and oxygen atoms (PSA) is 7.0 – 200.0, and number of violations of Jorgensen's rule of three is maximum 3 **[19]**.

Out of many protein biomarkers available for the each respective type of cancer, we have selected the one which has maximum sensitivity towards the detection of the disease and treatment response to anticancer drugs. Active sites of each receptor were taken from the crystallographic structure obtained from the PDB database. Docking studies revealed that all the analogues had good inhibitory activity and binding affinity towards the Oncoproteins. Analogues with the highest binding energy for the alkaline phosphatase, estrogen receptor, epidermal growth factor receptor and B-Raf Oncoproteins are 7, 2C, 2E and 2G respectively. Ligands 2G & 5 both has the highest and same Glide score for the brain cancer biomarker vascular endothelial growth factor receptor-2 and most importantly they have formed the hydrogen bonds to the same active site amino acid (ASP 1046) determined by X-ray crystallography. All other ligands have also interacted to the Oncoproteins with good number of hydrogen bonds. Darunavir has best activity (G score = -7.0 Kcal/mol) against skin cancer and has the same activity (G score = -5.0 Kcal/mol) compared to analogue 7 against bone cancer.

Conclusion:

All the nine designed analogues have good pharmacokinetic properties to become successful drugs and they have interacted to all the five biomarkers with sufficient selectivity, specificity and reflected the optimum binding energy to form the stable docked complex. Most of the analogues have better binding energy to oncoproteins with compared to Darunavir. All the analogues have good ADMET properties to become successful drugs. Based on these good results with oncoproteins, we are proposing the synthesis of these ligand molecules and testing on the specific cancer cell lines.

Acknowledgement:

We thank Raghu Rangaswamy, Senior Director and Vinod Devaraji IT consultant from Schrödinger for providing academic evaluations software and continuous support to undertake this work.

References:

[1] Sporn MB, *Lancet*. 1996 **347**: 1377 [PMID: 8637346]
[2] Brücher BL *et al. BMC Cancer.* 2014 **14**: 186 [PMID: 24629025]
[3] Fidler IJ, *Cancer Res.* 1978 **38**: 2651 [PMID: 354778]
[4] Bonucci E & Nanci A, *Ital J Anat Embryol.* 2001 **106**: 129 [PMID: 11729947]
[5] Sawyers CL, *Nature* 2008 **452**: 548 [PMID: 18385728]
[6] Mishra A & Verma M, *Cancers (Basel).* 2010 **22:** 190 [PMID: 24281040]
[7] Russell MD *et al. Front Pediatr.* 2013 **1:** 7 [PMID: 24400253]
[8] Smith ER *et al. Clin Cancer Res.* 2008 **14**: 2378 [PMID: 18413828]
[9] Agarwal R *et al. Clin Cancer Res*. 2009 **15**: 3654 [PMID: 19470724]
[10] How J *et al. Histochem Cell Biol*. 2014 [PMID: 24429833]
[11] Haluska FG *et al. Clin Cancer Res*. 2006 **12:** 2301s [PMID: 16609049]
[12] Davies H *et al. Nature*. 2002 **417:** 949 [PMID: 12068308]
[13] Kariya R *et al. Cancer Lett.* 2014 **342**: 52 [PMID: 24012878]
[14] LigPrep, version 2.5, Schrödinger, LLC, New York, NY, 2012
[15] QikProp, version 3.5, Schrödinger, LLC, New York, NY, 2012.
[16] Friesner RA *et al. J Med Chem.* 2006 **49:** 6177 [PMID: 17034125]
[17] Chow WA *et al. Lancet Oncol.* 2009 **10**: 61 [PMID: 19111246]
[18] Monini P *et al. Nat Rev Cancer.* 2004 **4**: 861 [PMID: 15516959]
[19] Meraj K *et al. Bioinformation.* 2012 **8**: 1139 [PMID: 23275710]

Edited by P Kangueane

Citation: **Mahto *et al.*** Bioinformation 10(4): 221-226 (2014)

BIOINFORMATION

open access

Supplementary material:

Table 1(A): QIKPROP 3.4 predictions of ADMET for the 9 compounds in the study

Analogues	#stars	CNS	Mol_MW	Dipole	SASA	Volume	DonorHB	accptHB	QPlogPoct
2c	1	0	430.523	2.719	689.217	1372.444	3	11.4	22.776
2e	0	-2	391.527	8.356	638.374	1216.387	4.5	8.2	24.213
2g	1	-2	585.694	7.914	829.966	1706.856	3.5	15.4	30.959
3b	0	-2	439.571	7.738	650.778	1284.745	4.5	8.2	25.618
3d	0	-2	595.709	3.841	756.523	1611.707	3.5	13.1	30.372
4	1	1	578.684	3.705	794.608	1707.184	2	12.9	27.067
5	1	0	478.567	2.853	716.701	1461.32	3	11.4	24.529
6	1	0	634.705	6.978	868.145	1846.836	2	16.3	31.129
7	0	-2	547.665	9.89	728.98	1527.978	3.5	13.1	29.501

Table 1(B): QIKPROP 3.4 predictions of ADMET for the 9 compounds in the study

Analogues	QPlogPw	QPlogPo/w	QPlogS	QPlogBB	QplogKp	IP(eV)	EA(eV)	HOA	PSA	Rule of three
2c	14.718	1.659	0.402	-0.433	-4.795	9.058	-0.258	1	81.631	0
2e	16.315	1.624	-1.77	-1.387	-4.853	8.836	0.457	2	105.428	0
2g	20.341	2.16	-1.311	-1.785	-3.414	8.705	0.347	1	126.407	0
3b	17.05	2.346	-1.948	-1.14	-3.994	8.686	0.366	2	104.877	1
3d	19.86	3.261	-3.41	-1.394	-1.62	8.94	0.609	2	134.964	0
4	15.122	4.134	-1.706	-0.469	-1.418	8.897	-0.375	1	89.463	0
5	15.394	2.232	-0.014	-0.5	-4.662	8.988	-0.227	1	83.204	1
6	18.756	3.34	-1.681	-0.759	-1.988	8.863	-0.14	1	121.21	0
7	19.235	2.529	-3.013	-1.544	-2.306	8.583	0.348	2	136.991	0

Table 2: Protein, biomarkers and active site information

S.No	Cancer type	Protein Biomarker	PDB ID	X-ray Resolutions [Å]	Active site Amino Acid
1.	Bone cancer	Alkaline phosphatase	3MK2	1.89	Asp86A
2.	Brain cancer	VEGFR-2	3VO3	1.52	Glu885A,Lys868A,Cys919A,Asp1046A
3.	Breast cancer	Estrogen Receptor	1UOM	2.28	Asp351,Glu353A,Arg394A
4.	Colon cancer	EGFR	3POZ	1.50	Met793A
5.	Skin cancer	B-Raf	4FK3	2.65	Leu525A-His540A

Table 3: Docking Analysis of 9 compounds with Oncoproteins

		Darunavir Analogues								
Proteins PDB IDs	**DRV**	**2C**	**2E**	**2G**	**3B**	**3D**	**4**	**5**	**6**	**7**
3MK2	-5.0	-3.7	-3.7	-2.8	-3.9	-3.3	-3.0	-3.9	-2.3	-5.0
3VO3	-4.7	-5.8	-5.8	-6.2	-5.2	-5.1	-5.3	-6.2	-5.5	-4.3
1UOM	-2.9	-6.9	-6.4	-6.4	-5.2	-1.5	-3.8	-5.8	-5.2	-2.5
3POZ	-6.0	-3.2	-7.5	-4.1	-5.6	-5.1	-5.3	-4.1	-1.4	-4.7
4FK3	-7.0	-5.0	-5.3	-6.5	-5.1	-6.4	-5.8	-2.6	-2.6	-5.5

Table 4: Hydrogen Bond Analysis

S.No	Protein & Analogues Name	No of H-Bond	Amino Acid name	H-Bond Distances (Å)
1.	3MK2-7	3	LYS 87	2.236
			HIS 88	1.994
			VAL 99	2.279
2.	3VO3-2G	2	ASP 1046	2.164
			ARG 1051	2.011
3.	3VO3-5	2	ASP 1046	2.037
				2.085

4.	1UOM-2C	3		1.612
			GLU 323	1.928
			TRP 393	1.857
5.	3POZ-2E	4	MET 793	2.097
			ARG 841	1.900
			ARG 842	2.104
			ASP 855	1.936
6.	4FK3-2G	3	LYS 473	2.122
			TRP 531	2.364
			ASN 581	2.068