In-silico comparative study of inhibitory mechanism of Plant Serine Proteinase Inhibitors

Chekkara Venkata Sathya Siva Prasad[1]*, Saurabh Gupta[1], Alex Gaponenko[2] & Murli Dhar Tiwari[1]

[1]Division of Applied Sciences & IRCB, Indian Institute of Information Technology, Deoghat, Jhalwa, Allahabad 211012, India; [2]Koltzov Institute of Development Biology, Russian Academy of Sciences, Moscow, Russia; Chekkara Venkata Siva Prasad – Email: shiva@iiita.ac.in; Phone: +91532 2922201, Fax: +91532 2430006; *Corresponding author

Abstract:
The nematodes like root-knot and cyst are plant-parasitic pest found in horticultural and agricultural crops. They do damages in the roots of plants as a result losses million tons of production. High cost of nematicides and environment safety concern has necessitated finding of some alternative methods. Under Integrated Pest Management (IPM) such problems are solving significantly by means of target gene inhibition, agrobacterium mediated transformation etc. One of this strategy use Plant Proteinase Inhibitors (PIs) gene which are used to control the proteolysis mechanism of Pest by inhibiting gut Serine Proteinase (SP). Present work investigates the utility of computer aided methods to study the mechanism of Protein-Protein interactions and thereby inhibition of Serine Proteinase by PIs. Hence 3D models of Serine Proteinase as well as Serine Proteinase Inhibitors (SPIs) generated using homology modeling. Validations of constructed models have been done by PROCHECK, VERIFY3D, ERRAT and PROSA. Prediction of Protein interacting surface patches and site specific protein docking was performed by using ZDOCK Server. Backbone refinement of output protein complexes was executed in Fiber Dock server. Interaction study between SP and SPIs complexes shows their comparative inhibition efficacy, measured in terms of number of hydrogen bonds, Van dar wall attraction and docking energy. This work reported that *Vigna marina* and *Phaseolus oligospermus* are having better inhibition efficiency in comparison to other inhibitors.

Keywords: Serine Proteinase, Serine Proteinase Inhibitors, *Vigna marina, Phaseolus oligospermus,* Integrated Pest Management, Modeling, Docking

Background:

Presently sedentary plant parasitic cyst nematodes (*Heterodera spp.* and *Globodera spp.)* and root-knot nematodes (*Meloidogyne spp.)* are the world's most damaging agricultural pests. They reduce crop production and quality of seeds causing billion dollar losses per annum [1]. Crop safety relies predominantly on the use of environmentally toxic synthetic nematicides. Rather than using chemical pesticides on a large scale, other alternative need to dig into that are based on system-oriented science and technology. Such confined systems will decrease inputs of chemicals and will not generate harmful output such as pesticides residues [2]. To carry out this aim it is essential to increase the resistance of plant to pest and nematodes through

promising control of Integrated Pest Management (IPM). IPM includes various strategies to control the pest like Mechanical control, Biological controls etc. The crop protection under biological controls principle includes the inhibition of target gene by suitable protein inhibitors [3]. Plant Serine Proteinase Inhibitors (SPIs) are natural defense-related proteins frequently present in seeds and certain plant tissue which prevents plant by herbivores and wounding. SPIs gens are inhibiting the target protease enzymes which are responsible for proteolytic activity in nematodes [4]. Numerous structural biology approaches are facilitating to check insight inhibition mechanism of protease by SPIs. Protein-Protein docking is one of the suited methods to find out the interaction between two proteins. The current work

demonstrated the comparative insight of structural inhibition pattern of SPIs with respect to Serine proteinase of nematodes.
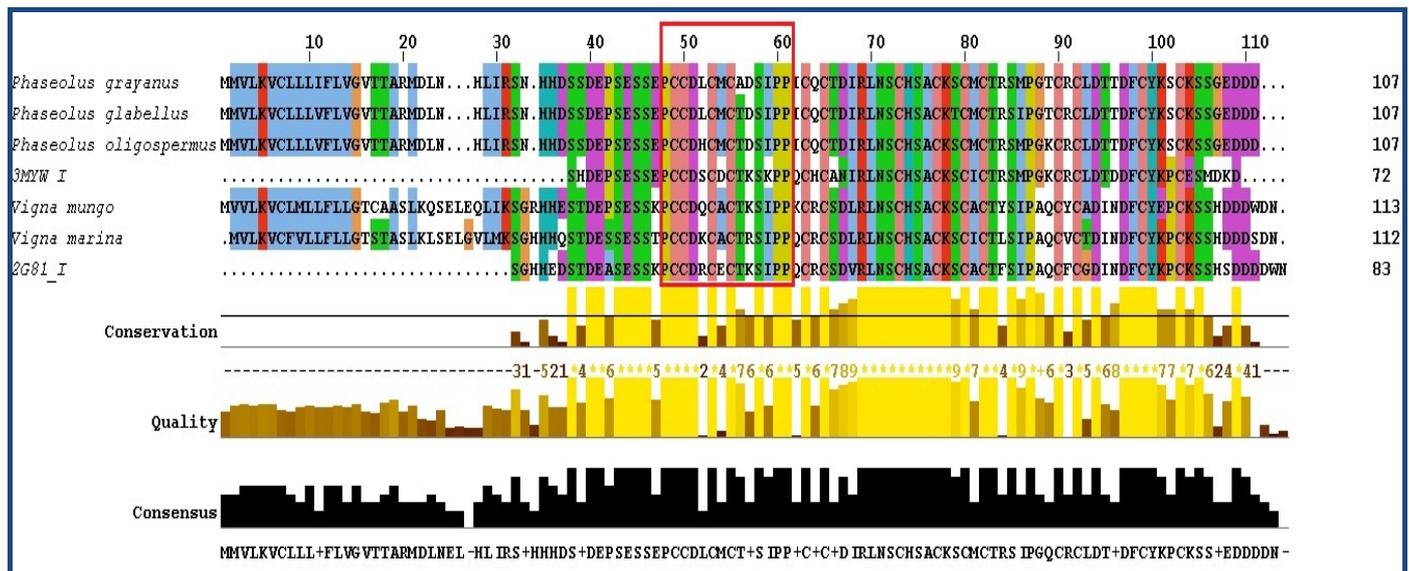

Figure 1: Multiple sequence alignment of all SPIs with their respective template structures for modeling. Rectangular red box covered residues are having high probability to participate during protein-protein interaction.
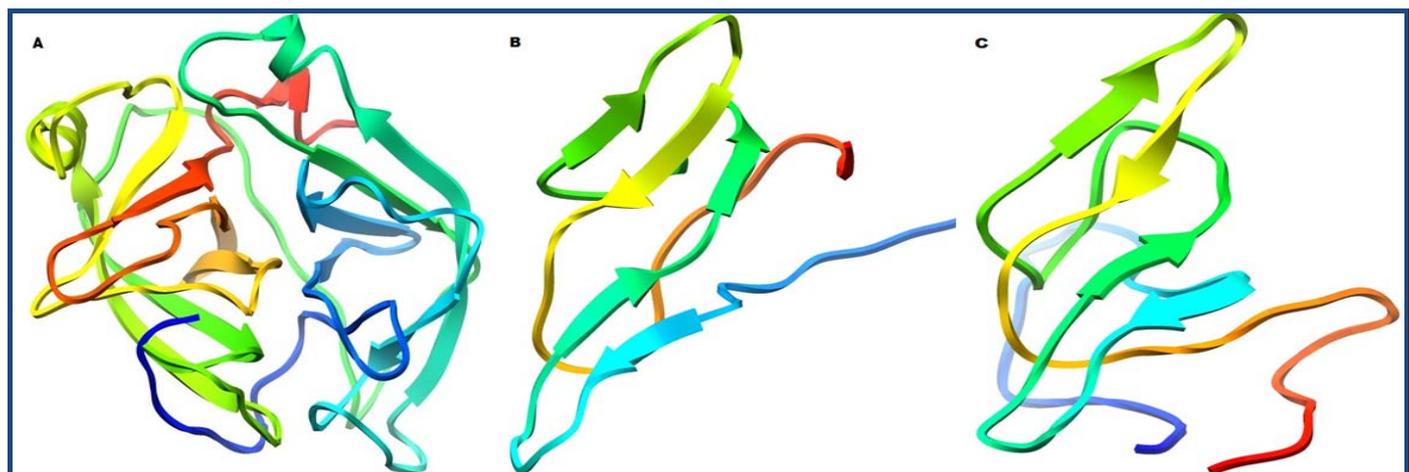

Figure 2: Cartoon representation of catalytic domain of Serine Proteinase of Heterodera glycines **(A)**, Serine Proteinase Inhibitors of *Phaseolus* Family **(B)** and *Vigna* Family **(C)**.

Methodology:

Molecular model of Serine Proteinase of Heterodera glycines

The protein sequence Serine Proteinase of *Heterodera glycines* (SP_HG) was downloaded from National Centre for Biotechnology Information (NCBI) **[5].**The FASTA sequence of the SP_HG was subjected to PSI-BLAST against PDB database for the selection of best homologous templates **[6, 7].** The best template was selected based on similarity, percentage of identity, expectation value, bit scores and query coverage area. A 3D model of SP_GH was generated by Modeller 9v10 **[8].** The chain-A of Earthworm Fibrinolytic Enzyme Component A from Eisenia Fetida, PDB: 1M9U was identified as the best template for comparative modeling **[9]**. This 3D structure was generated by X-ray diffraction studies and resolution was 2.30Å. The R-factor of the structure was 0.191 and R-free value was 0.236. A bundle of 100 models are formed through random generation and subsequently the good model was obtained.

Molecular models of Plant Serine Proteinase inhibitors

Following Plant Serine Proteinase Inhibitors (SPIs) of *Phaseolus glabellus* (SPI_PG), *Phaseolus grayanus* (SPI_PGn), *Phaseolus oligospermus* (SPI_PO), *Vigna mungo* (SPI_VM), *Vigna marina* (SPI_VMn) sequences having Accession No. CAQ17032, CAM88858, CAO82009, ABD97865, ABD97867 respectively were obtained from NCBI **[5].** Each Sequence of proteinase inhibitors was subjected to PSI-BLAST against PDB database for the selection of best homologous templates **[6, 7].** The chain-I of The Bowman-Birk type inhibitor from Mung bean, PDB ID: 3MYW and Bowman-Birk Inhibitor from *Vigna unguiculata* Seeds having, PDB ID: 2G81 were identified as the best model for comparative modeling of *Phaselous* family and *Vigna* family inhibitors respectively **[10, 11].** Both 3D structures were generated by X-ray diffraction studies and resolution were 2.50Å for 3MYW and 1.55 Å for 2G81. Out of Twenty generated models one good model was obtained for each plant proteinase inhibitors using Modeller 9v10 **[8].**

BIOINFORMATION

Minimization and molecular evaluation

Energy minimization 3-D model of SP_HG and all SPIs were performed using Amber force field for 10000 steps with ten update interval without fixing any atom in chimera **[12, 13]**. Stereo- chemistry of all models was evaluated by PROCHEK and WHAT-IF **[14, 15]**. The overall quality factor of non-bonded interactions between different atoms was calculated by ERRAT **[16]**.Compatibility of amino acids in each 3D model was checked in VERIFY-3D **[17, 18]**. X-Ray and NMR spectroscopy structural validation of structures was predicted by PROSA server **[19]**.

Designation of Protein-Protein interacting sites and Docking

A combine study of various PSI-BLAST results of SPIs showing the specific region of protease binding site of their inhibitor with respect to crystal structures was obtained. Multiple sequence alignment of SPIs sequences with respect to crystal structure helps us to find out protease binding sites shown in **(Figure 1)**. In addition to that we also find out detailed protein interacting sites for all SPIs and SP_HG using online server PPi-pred **[20]**. Protein Interface site specific docking of all SPIs with SP_HG was performed using ZDOCK server **[21]**. ZDOCK output files were subjected in FiberDock server for flexible induced-fit backbone and side chain refinement of the protein complexes **[22, 23]**.

Molecular Interaction Study

The molecular interaction plots between SP_HG and SPIs were generated using Dimplot in LIGPLOT software **[24]**. The 3D structure and detailed interaction of all proteins complexes were visualized using Accelrys DS Visualizer software **[25]**.
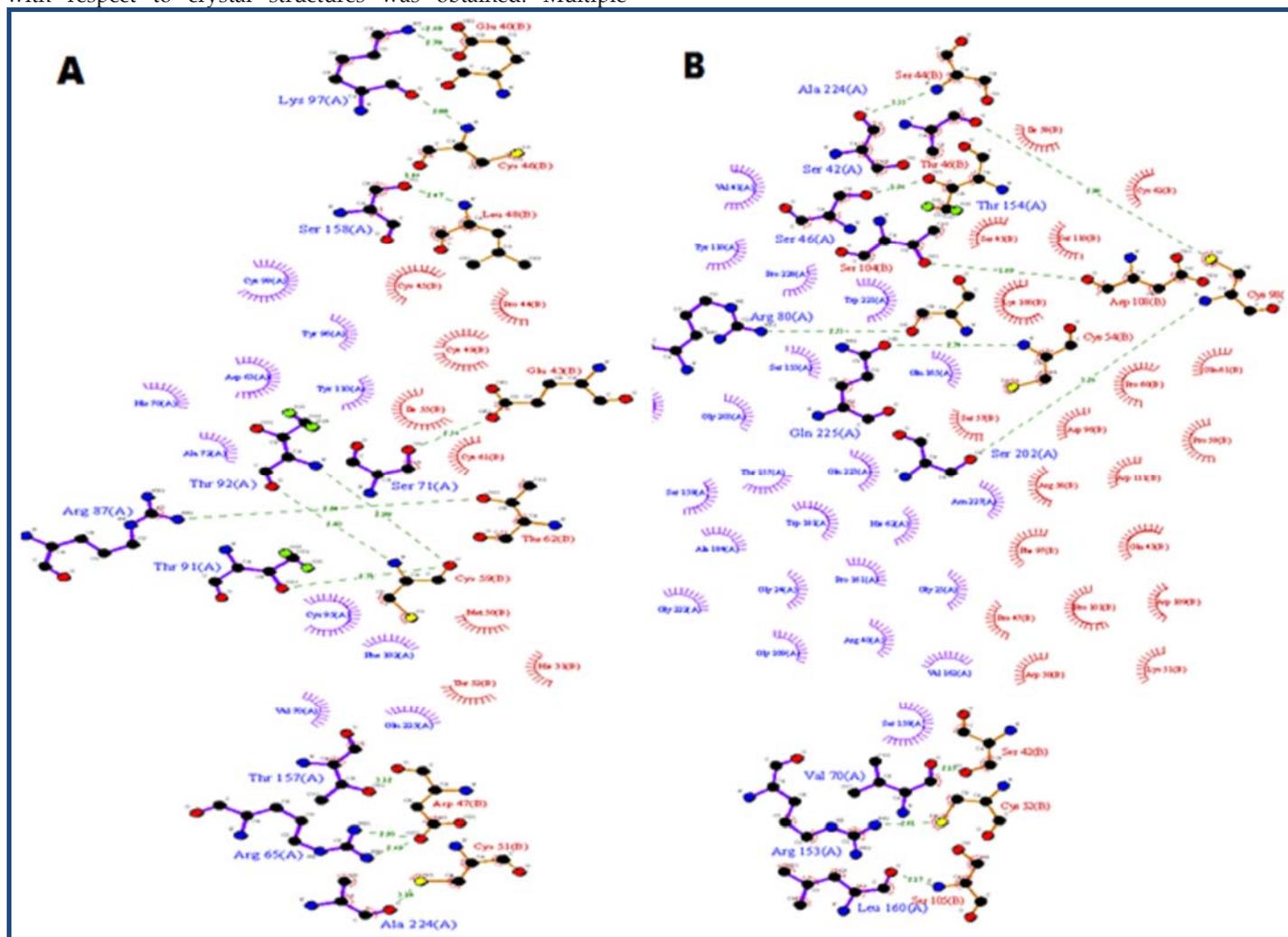

Figure 3: Molecular Interaction plots of docked complexes of *Phaseolus oligospermus* **(A)** *Vigna marina* **(B)** with Serine proteinase inhibitor of *Heterodera glycines.* Hydrogen bonds with their bond length between Protein interface residues shown in green dotted line.

Discussion:

Finding of good proteinase inhibitor using structural biology approaches would be helpful for experiment biologist to explore their research in pest management. In this connection we have generated the 3D- models of SPIs and SP_HG. **Figure 2** shows the catalytic domains of SP_HG **(A)** and Serine proteinase inhibitors of Phaseolus Family **(B)** and Vigna Family **(C).** The dependability and quality of these models were ensured by evaluating the backbone conformation, angles and bond lengths based on Psi/Phi Ramachandran plot using PROCHECK **[14]**. Analysis of each model with ProSA web interface brings out that Z-Score and over all residues energies are all well within the limits **[15].** ERRAT score shows overall quality factor for non-bonded atomic interactions and higher

the score means better the quality of models (accepted range for a high quality model is >50) **[16].** VERIFY-3D score indicates good sequence-to-structure agreement because most of the amino acids score for all models have > 0.2 values shown in **Table 1 (see supplementary material) [17, 18].** The above said validating parameters assuring the quality and reliability of models. To get insight of structural inhibition mechanism of proteinase inhibitors need to have the protein interacting sites. The protein interacting sites were identified using a web based server PPi-pred **[20]**. Most probable interacting residues are common with the crystal structure substrate binding sites which are given in **Table 2 (see supplementary material)**. The Docking of each SPIs with SP_HG deducing a reliable results, in terms of minimum Global, Attractive and Repulsive Van-der wall, Atomic contact and Hydrogen bond energies shown in **Table 3 (see supplementary material).** In addition to that molecular interaction plots are also certifying the protein-protein interaction between protein complexes. The interaction study inferring that the Serine Proteinase inhibitor of *Vigna marina* and *Phaseolus oligospermus* are the potent inhibitors in comparison to others within the same family of SPIs. Molecular interacting plot of SPI_PO with SPI_HG **(Figure 3A)** shows total 13 hydrogen bonds formation among eight amino acids of SPI_PO and nine residues of SPI_HG. **Figure 3B** shows the molecular interaction plots between SPI_VMn and SP_HG which having eight hydrogen bonds.The nine residues of SPI_VMn and eight residues of SP_HG are participating in bond formation. The comparative studies between these two inhibitors conclude that SPI_PO is the most capable inhibitor among them. This finding will help to biologist in nematodes control.

Conclusion:

Now a day's protection of crop against Nematodes is being big challenge for the researchers with taking concern of environmental safety and economical balance. Our present work will help to researchers in biological pest control. We have developed high quality 3D models of Plant Serine Proteinase Inhibitors. The comparative *In-sillico* inhibitory mechanisms of these inhibitors were identified against Serine Proteinase of *Heterodera glycines.* The inhibitory efficiency of each serine proteinase was calculated on the ground of Number of hydrogen bonds, Van der Waal attraction, and docking energy. *Vigna marina* and *Phaseolus oligospermus* proteinase inhibitor forming major hydrogen bonds with minimum docking energy was concluded as potent inhibitors in their family.

References:
[1] Nicol JM *et.al* Genomics *and Mol Gen of Plant-Nematode Interactions.* DOI c 10.1007/978-94-007-0434-3_2.
[2] Boulter D *et al. Biochemistry.* 1993 **34:** 1453 [PMID: 7764318]
[3] Haq SK *et al. Arch Biochemistry Biophysics.* 2004 **431:**145 [PMID: 15464737]
[4] Huma H *et al. Biot & Mole Biology Rev.* 2007 **2:**68
[5] http://www.ncbi.nlm.nih.gov/
[6] http://www.pdb.org/pdb/home/home.do
[7] http://www.ncbi.nlm.nih.gov/blast/Blast.cgi/(PSI)
[8] http://salilab.org/modeller/
[9] Tang Y *et al. J Mol Biol.* 2002 **321:** 57 [PMID: 12139933]
[10] Lin G *et al. Eur J Biochem.* 1993 **212:** 549 [PMID: 8444191]
[11] Barbosa JA *et al. Biophys J.* 2007 **92:** 1638 [PMID: 17142290]
[12] Junmei Wang *et al. J Comput Chem.* 2004 **25:** 1157
[13] Pettersen EF *et al. J Comput Chem.* 2005 **25:** 1605 [PMID: 15264254]
[14] Laskowski RA *et.al. J Appl Cryst.* 1993 **26:**283
[15] Vriend G J *Mol Graph.* 1990 **8:**52[PMID:2268628]
[16] Colovos C *et al. Protein Science.* 1993 **2:**1511 [PIMD: 8401235]
[17] Bowie JU *et al. Science.*1991 **253:** 164 [PMID:1853201]
[18] Luthy R *et al. Nature* 1992 **356:** 83[PMID: 1538787]
[19] http://prosa.services.came.sbg.ac.at/prosa.php/
[20] http://bmbpcu36.leeds.ac.uk/ppi_pred/overview.html
[21] Chen R *et al. Proteins* 2003 **52:** 80 [PMID:12784371]
[22] Mashiach E *et al. Nucleic Acids Res.* 2008 **36:** w229 [PMID: 18424796]
[23] Mashiach E *et al. Nucleic Acids Res.* 2008 **36:** w2296 [PMID: 20077569]
[24] Wallace AC *et al. Prot Eng.* 1995, **8:**127[PMID:7630882]
[25] http://accelrys.com/products/discovery-studio/visualization-download.php

BIOINFORMATION

Supplementary material:

Table 1: Ramachandran Plot scores and other protein validations score for 3D model of serine proteinase of *Heterodera glycines* and all Plant serine proteinase inhibitors

Ramachandran Plot Scores	SP_HG	SPI_PG	SPI_PGn	SPI_PO	SPI_VM	SPI_VMn
% Amino acid in most favored regions	73.6%	82.5%	80.6%	88.8%	90.1%	73.7%
% Amino acid in additional allowed regions	21.3%	14.4%	16.1%	9.0%	9.9%	22.4%
% Amino acids in generously allowed regions	3.6%	2.1%	3.2%	1.1%	0.0%	3.9%
% Amino acids in disallowed regions	1.5%	1.0%	0.0%	1.1%	0.0%	0.0%
ProSA Z-Score	-5.82	-3.49	-3.44	-3.49	-3.13	-3.76
Errat Score	86.97	78.00	95.08	75.82	81.96	77.41
Overall G-factor	-0.37	-0.28	-0.36	-0.14	-0.19	-0.38
Verfy_3D score	0.73	0.35	0.51	0.38	0.33	0.51

Table 2: Most favorable protein interacting residues for SP_HG and SPIs predicted by the PPI-pried server.

Proteins	Favorable protein binding site Amino acid
SP_HG	40A, 41V, 46S, 62H, 66S, 68I,101S,103V,104N, 105N, 107N, 109G, 110Y, 113D, 153R, 158S, 202S, 205S, 221, 222G, 224A, & 225Q
SPI_PG	44P, 45C, 46C, 47D, 49C, 51C, 53D, 54S, 55I, 56P, 57P, 92T, 94F, 93D 95C, 96Y, &97K
SPI_PGn	44P, 46C, 47D, 48K, 49C, 50M, 51C, 52A, 53D, 54S, 57P, 58I, 59C, 60Q, 61C, 62T, 94C, 95Y & 107D
SPI_PO	44P, 46C, 47D, 48H, 49C, 50M, 51C, 53D, 54S, 57P, 58I, 59C, 60Q, 61C, 62T, 94C, 95Y & 96K
SPI_VM	47K, 49C, 51D, 52Q, 53C, 55C, 56T, 57K, 58S, 59I, 61P, 62K, 63C, 64R, 69R, 73C, &102P
SP_VMn	42S, 45S, 46T, 47P, 48C, 50D, 51L, 52C, 54C, 56P, 57S, 58I, 59P, 60P, 61Q & 62C

Table 3: Various energy values for each best complex of SPIs with SP_HG generated by Fiber Dock Server. Attractive and Repulsive Van-der wall (VDW) forces for each complex are positive and negative respectively. ACE stand for Atomic Contact Energy whereas HB shows minimum Hydrogen bond energy.

Sr. No	Proteinase Inhibitors	Global Energy (Kj/mole)	Attractive VDW (Kj/mole)	Repulsive VDW (Kj/mole)	ACE (Kj/mole)	HB (Kj/mole)
1	SPI_PG	-24.34	-21.27	11.38	-8.91	-2.19
2	SPI_PGn	-24.08	-17.85	15.49	-12.92	-1.33
3	SPI_PO	-25.81	-17.82	8.79	-7.39	-2.16
4	SPI_VM	-23.43	-16.24	14.97	-10.96	-1.09
5	SPI_VMn	-24.14	-14.79	13.62	-11.79	-2.91