

# *Internet* **Electronic** Journal of **Molecular Design**

January 2009, Volume 8, Number 3, Pages 29–41

Editor: Ovidiu Ivanciuc

## **Analog Based Studies on Cathepsin B Inhibitors to Design Novel Lead Compounds for the Treatment of Cancer**

L. Jayashankar,<sup>1</sup> B. Syama Sundar,<sup>1</sup> Shruti Awasthi,<sup>2</sup> and B. Madhulika<sup>2</sup>

<sup>1</sup> Department of Pharmacy, Acharya Nagarjuna University, Guntur, India

<sup>2</sup> GVK BioSciences Pvt. Ltd., Informatics Division, Balanagar, Hyderabad, India

Received: July 7, 2009; Revised: October 21, 2009; Accepted: November 28, 2009; Published: December 4, 2009

### **Citation of the article:**

L. Jayashankar, B. S. Sundar, S. Awasthi, and B. Madhulika, Analog Based Studies on Cathepsin B Inhibitors to Design Novel Lead Compounds for the Treatment of Cancer, *Internet Electron. J. Mol. Des.* 2009, 8, 29–41, <http://www.biochempress.com>.

## Analog Based Studies on Cathepsin B Inhibitors to Design Novel Lead Compounds for the Treatment of Cancer

L. Jayashankar,<sup>1,\*</sup> B. Syama Sundar,<sup>1</sup> Shruti Awasthi,<sup>2</sup> and B. Madhulika<sup>2</sup>

<sup>1</sup> Department of Pharmacy, Acharya Nagarjuna University, Guntur, India

<sup>2</sup> GVK BioSciences Pvt. Ltd., Informatics Division, Balanagar, Hyderabad, India

Received: July 7, 2009; Revised: October 21, 2009; Accepted: November 28, 2009; Published: December 4, 2009

*Internet Electron. J. Mol. Des.* 2009, 8 (3), 29–41

### Abstract

**Motivation.** We have performed analog based design along with pharmacophore hypothesis development on a series of cathepsin B inhibitors. We intend to utilize the pharmacophore information to undertake 3D searches on large databases to identify novel lead candidate against cathepsin B inhibitors for the treatment of cancer.

**Method.** Molecular field analysis (MFA) and receptor surface analysis (RSA) methods have been carried out to derive best QSAR models. Catalyst version 4.7 was used to generate pharmacophore models.

**Results.** The QSAR models developed by MFA and RSA methods have  $r^2$  values of 0.77 and 0.879, respectively. The best quantitative pharmacophore model showed two hydrogen bond acceptor, one hydrophobic aliphatic and one ring aromatic feature.

**Conclusions.** The 3D QSAR study shows that hydrophobic groups are responsible for an increase in activity, and indicate that RSA predicts better than MFA. Further, the knowledge of the four–feature pharmacophore hypothesis for cathepsin B inhibitors can be very useful for virtual screening to design more potent lead compounds.

**Keywords.** Cathepsin B; quantitative structure–activity relationships; QSAR; pharmacophore.

### Abbreviations and notations

CAT–B, cathepsin B

RSA, receptor surface analysis

MFA, molecular field analysis

QSAR, quantitative structure–activity relationships

## 1 INTRODUCTION

Cathepsin B (CAT–B), a lysosomal cysteine protease, has been implicated in the pathology of a number of important human diseases. It is also known to be involved in various disease stages, such as inflammation, trauma, muscular dystrophy and tumors. The wide–ranging functions of this enzyme therefore make it an attractive drug discovery target [1–4]. We have performed pharmacophore and 3D–QSAR studies for developing novel cathepsin B inhibitors [5–15] using the Catalyst 4.7 and Cerius2 program suite. QSAR equations has been generated for 50 cathepsin B inhibitors employing molecular field analysis (MFA) as well as receptor surface analysis (RSA)

\* Correspondence author; phone: E–mail: [jsl.pharma@gmail.com](mailto:jsl.pharma@gmail.com)

using genetic function approximation (GFA) as regression method. The best equations with training set consisting 35 molecules, produced  $r^2$  value of 0.770 and  $r^2_{cv}$  value of 0.632 in MFA–model and  $r^2$  value of 0.879 and  $r^2_{cv}$  of 0.826 in the RSA–model (Table 1). For the 15 test set molecules predicted activities have correlation of 0.724 and 0.816 for MFA and RSA with observed activities [16–18]. Pharmacophore models were generated using 25 molecules as training set. The best quantitative pharmacophore model consists 1) two hydrogen bond acceptor 2) one hydrophobic aliphatic and 3) one ring aromatic feature. For the training set the accuracy in predicting active and inactive compounds was 90%. The best pharmacophore hypothesis yielded a RMS deviation of 0.828 and a correlation coefficient of 0.961 with a cost difference (null cost – total cost) of 52.246. The obtained pharmacophore models were validated on 50 test molecules to give correlation value of 0.886. For the test set, the accuracy in predicting active compounds was greater than 8%, while 15% and 5% representing both false positives and negatives respectively [19–21].

## 2 MATERIALS AND METHODS

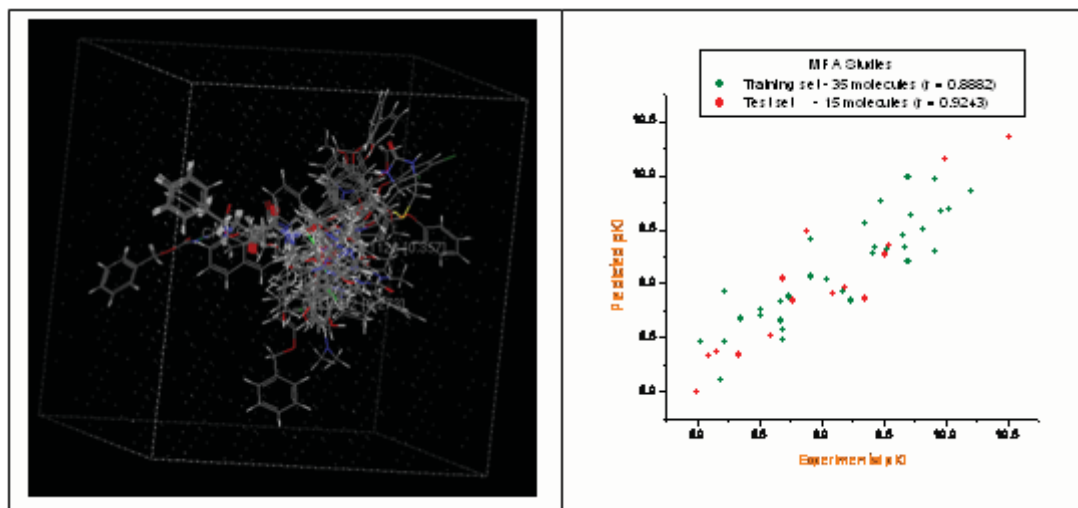
All molecular modeling works were performed on a Silicon Graphics Octane2 R12000 computer running Irix 6.5.12 (SGI, 1600 Amphitheatre Parkway, Mountain View, CA 94043) using Cerius2 4.6 and Catalyst 4.7 version [22,23]. 35 molecules forming the training set were used to generate the QSAR equation. For MFA studies molecular field was created using proton and methyl groups as probes, which represent electrostatic and steric fields respectively. For RSA studies chemical properties namely charge, electrostatic potential, hydrogen bonding propensity and hydrophobicity associated with each surface point were calculated. For generating equations, only 10% of the total descriptors whose variance was highest were considered for further analysis. Regression analysis was carried out using G/PLS method consisting of over 50,000 generations with a population size of 100 (Table 2 and 3). Catalyst version 4.7 was used to generate pharmacophore models. 25 molecules forming the training set were used to generate Hypogen hypothesis. All structures were built and minimized within the Catalyst software package, and conformational analysis of each molecule was implemented using the poling algorithm. Several hypotheses were generated from a collection of conformational models of compounds spanning activities of 4–5 orders of magnitude.

## 3 RESULTS AND DISCUSSION

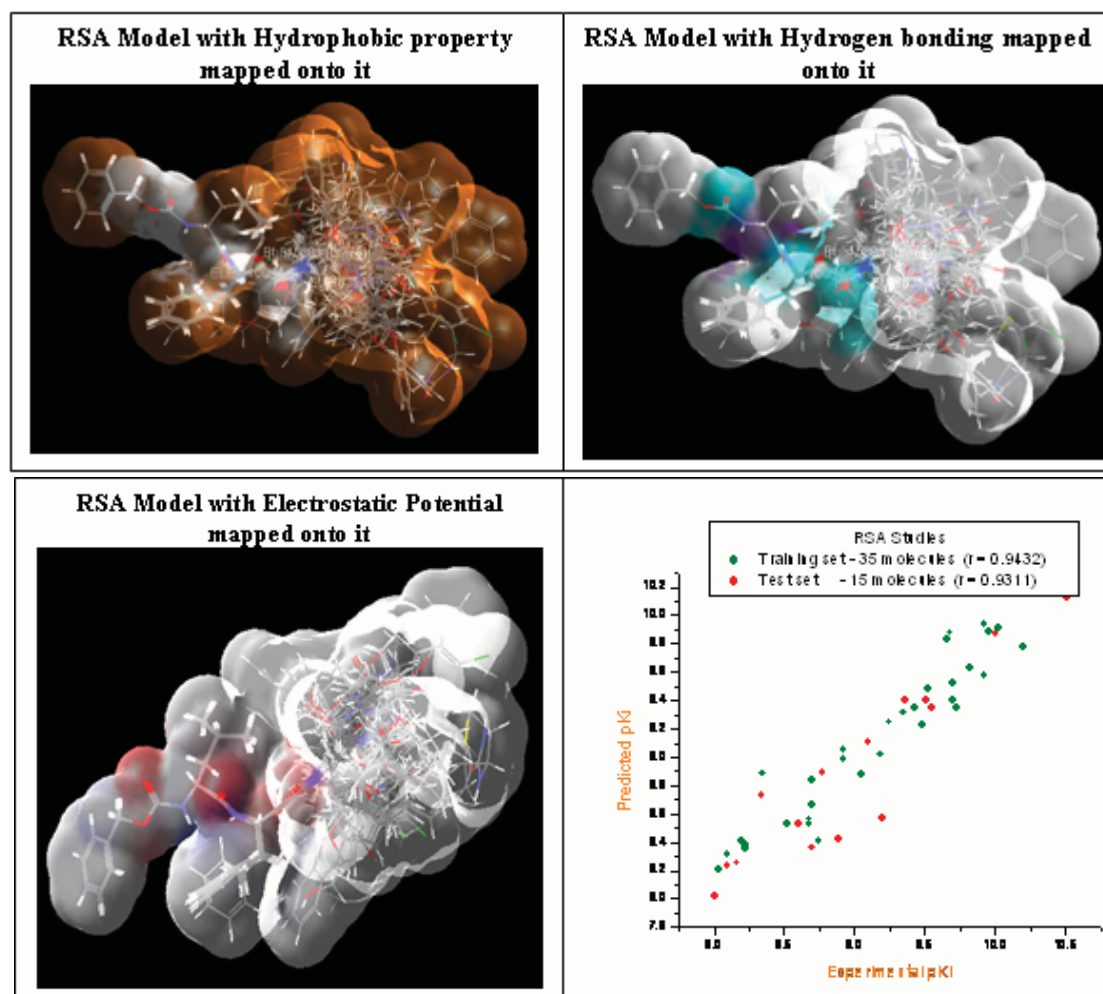
### 3.1 Molecular Field Analysis (MFA)

MFA equation that for the probe point of CH3 at position 540 in MFA grid indicates bulky groups are favored to decrease the activity. Stereo view of MFA grid is shown in Figure 1.

$$\text{MFA equation: Activity} = 9.10648 + 0.030798 * \langle \text{CH3/972} \rangle + 0.185231 \langle \text{CH3/540} \rangle + 0.182041 * 11.7526 - \langle \text{CH3/540} \rangle - 0.025185 * \langle \text{CH3/1128} \rangle + 0.310553 \langle \text{CH3/1128} \rangle - 0.44600$$



**Figure 1.** Stereo view of rectangular molecular field surrounding aligned molecules. Some of the field descriptors, which are involved in the equation, are indicated. Correlation of MFA = 0.77.

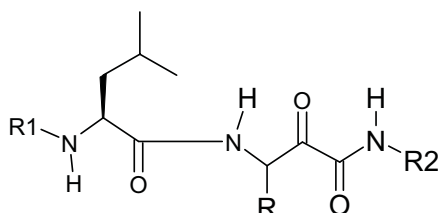


**Figure 2.** Stereo view of the receptor surface which represents the vitral active site. Some of the RSA descriptors that constitute the equation are labeled. Correlation of RSA = 0.879.

**Table 1.** Statistical Parameters for MFA and RSA

No	Statistical parameter	Value (MFA)	Value (RSA)
1	R <sup>2</sup>	0.770	0.879
2	F-test	26.548	36.524
3	R <sup>2</sup> cross validated	0.632	0.826
4	Press	4.926	4.604
5	Predicted R <sup>2</sup>	0.724	0.816

**Table 2.** Training Set with Experimental and Predicted Activity



Peptidyl R-Keto Amides R1-Leu-AA-CONH-R

No.	R1	AA	R2	Exp. K <sub>i</sub> (nM)	Exp. pK <sub>i</sub> (nM)	RSA <sub>pred</sub>	MFA <sub>pred</sub>
1	Z	Phe	(CH <sub>2</sub> ) <sub>2</sub> Ph	9.3	8.031	8.456	8.211
2	Z	Abu	n-Pr	8	8.096	8.328	8.32
3	Z	Abu	CH <sub>2</sub> -4-pyridyl	6.4	8.193	8.106	8.412
4	Z	Abu	CH <sub>2</sub> -2-furyl	6	8.221	8.464	8.355
5	Z	Phe	Et	6	8.221	8.932	8.383
6	Z	Abu	CH <sub>2</sub> -2-tetrahydrofuryl	4.5	8.346	8.673	8.886
7	Z	Abu	(CH <sub>2</sub> ) <sub>2</sub> OH	4.5	8.346	8.672	8.886
8	Z	Abu	CH <sub>2</sub> -2-pyridyl	3	8.522	8.764	8.536
9	Z	Phe	n-Pr	3	8.522	8.7	8.536
10	Z	Phe	CH <sub>2</sub> CH(OH)Ph	2.1	8.677	8.836	8.532
11	Z	Phe	CH <sub>2</sub> CH(OH)C <sub>6</sub> H <sub>4</sub> (4-OPh)	2.1	8.677	8.656	8.565
12	Z	Abu	(CH <sub>2</sub> ) <sub>3</sub> -1-pyrrolidin-2-one	2	8.698	8.486	8.665
13	Z	Abu	(CH <sub>2</sub> ) <sub>2</sub> O(CH <sub>2</sub> ) <sub>2</sub> OH	2	8.698	8.57	8.843
14	Z	Abu	CH <sub>2</sub> C <sub>6</sub> H <sub>3</sub> (3,5-(OCH <sub>3</sub> ) <sub>2</sub> )	1.8	8.744	8.879	8.406
15	Z	Abu	CH <sub>2</sub> -3-pyridyl	1.2	8.92	9.064	8.988
16	Z	Abu	(CH <sub>2</sub> ) <sub>2</sub> -2-(N-methylpyrrole)	1.2	8.92	9.403	9.05
17	Z	Abu	CH <sub>2</sub> -C <sub>6</sub> H <sub>7</sub> (1,3,3-(CH <sub>3</sub> ) <sub>3</sub> -5-OH)	0.89	9.05	9.043	8.883
18	Z	Abu	(CH <sub>2</sub> ) <sub>2</sub> NH-biotinyl	0.66	9.18	8.924	9.02
19	Z	Phe	Me	0.57	9.244	8.844	9.247
20	Z	Abu	CH <sub>2</sub> -2-quinolinyl	0.45	9.346	9.56	9.315
21	Z	Abu	CH <sub>2</sub> CH(OH)Ph	0.37	9.431	9.34	9.349
22	Z	Abu	Me	0.33	9.481	9.765	9.226
23	Z	Phe	CH <sub>2</sub> CH(OH)C <sub>6</sub> H <sub>4</sub> (4-OCH <sub>2</sub> Ph)	0.38	9.42	9.276	9.353
24	Z	Abu	CH <sub>2</sub> -1-isoquinolinyl	0.3	9.522	9.322	9.48
25	Z	Abu	CH <sub>2</sub> CH(OH)C <sub>6</sub> H <sub>2</sub> (3,4,5-(OCH <sub>3</sub> ) <sub>3</sub> )	0.22	9.657	9.45	9.833
26	Z	Abu	CH <sub>2</sub> CH(OH)C <sub>6</sub> H <sub>4</sub> (4-OPh)	0.21	9.677	9.34	9.88
27	Z	Abu	(CH <sub>2</sub> ) <sub>3</sub> C <sub>6</sub> H <sub>5</sub>	0.2	9.698	9.985	9.403
28	Z	Abu	(CH <sub>2</sub> ) <sub>2</sub> -2-pyridyl	0.2	9.698	9.202	9.523
29	Z	Abu	CH <sub>2</sub> CH(OCH <sub>3</sub> ) <sub>2</sub>	0.19	9.721	9.629	9.351
30	Z	Abu	CH <sub>2</sub> CH(OH)C <sub>6</sub> F <sub>5</sub>	0.15	9.823	9.497	9.626
31	Z	Abu	CH <sub>2</sub> CH(OH)C <sub>6</sub> H <sub>4</sub> (4-OCH <sub>3</sub> )	0.12	9.92	9.294	9.934
32	Z	Abu	CH <sub>2</sub> CH(OH)C <sub>6</sub> H <sub>4</sub> (3-OPh)	0.12	9.92	9.966	9.582
33	Z	Abu	CH <sub>2</sub> CH(OC <sub>2</sub> H <sub>5</sub> ) <sub>2</sub>	0.11	9.958	9.674	9.881
34	Z	Abu	CH <sub>2</sub> CH(OH)C <sub>6</sub> H <sub>4</sub> -3-OC <sub>6</sub> H <sub>3</sub> (3,4-C <sub>12</sub> )	0.094	10.026	9.684	9.917
35	Z	Abu	CH <sub>2</sub> CH(OH)-1-C <sub>10</sub> H <sub>7</sub>	0.063	10.2	9.859	9.778

### 3.2 Receptor Surface Analysis (RSA)

The term ELE/4106 in the RSA equation indicates electronegative groups are favored to enhance the activity. RSA Model with Hydrophobic property and Hydrogen bonding mapped onto it is shown in Figure 2.

$$\text{MFA equation: Activity}_1 = 8.55202 + 220.171 * < -0.59102 - \text{“ELE/4106”} > \\ + 52.0302 * < \text{“ELE/5876”} - 1.00686 > - 1.3303 * < - 0.645906 - \text{“ELE/6391”} > \\ + 10.3418 * < -0.017168 - \text{ELE/5891”} >$$

### 3.3 Pharmacophore Hypothesis Generation

Training set consists of 25 compounds tested against cathepsin B was used to develop Pharmacophore hypotheses. A total of 10 hypotheses were generated and its different cost values, correlation coefficients (r), RMS deviations, and pharmacophore feature definitions are listed in (Table 4).

**Table 3.** Test Set with Experimental and Predicted Activity

No.	R1	AA	R2	Exp. $K_i$ (nM)	Exp. $pK_i$ (nM)	RSA <sub>pred</sub>	MFA <sub>pred</sub>
1	Z	Abu	(CH <sub>2</sub> ) <sub>3</sub> -1-imidazolyl	9.9	8.004	7.991	8.024
2	Z	Abu	(CH <sub>2</sub> ) <sub>3</sub> -2-tetrahydroisoquinolyl	8	8.096	8.331	8.24
3	Z	Abu	(CH <sub>2</sub> ) <sub>3</sub> -4-morpholyl	6.9	8.161	8.365	8.257
4	Z	Phe	CH <sub>2</sub> Ph	4.6	8.337	8.34	8.735
5	Z	Abu	(CH <sub>2</sub> ) <sub>2</sub> -4-morpholyl	2.5	8.602	8.521	8.531
6	Z	Abu	CH <sub>2</sub> C <sub>6</sub> H <sub>5</sub>	2	8.698	9.044	8.364
7	Z	Abu	CH <sub>2</sub> -8-caffeinyl	1.7	8.769	8.842	8.89
8	Z	Abu	(CH <sub>2</sub> ) <sub>2</sub> C <sub>6</sub> H <sub>5</sub>	1.3	8.886	9.48	8.421
9	Z	Phe	CH <sub>2</sub> -2-pyridyl	0.8	9.096	8.904	9.103
10	Z	Phe	CH <sub>2</sub> CH(OH)C <sub>6</sub> H <sub>4</sub> (3-OPh)	0.64	9.193	8.96	8.575
11	Z	Abu	(CH <sub>2</sub> ) <sub>2</sub> C <sub>6</sub> H <sub>4</sub> (4-OCH <sub>3</sub> )	0.44	9.356	8.861	9.398
12	Z	Abu	(CH <sub>2</sub> ) <sub>2</sub> C <sub>6</sub> H <sub>4</sub> (3-OCH <sub>3</sub> )	0.31	9.508	9.272	9.406
13	Z	Abu	(CH <sub>2</sub> ) <sub>5</sub> OH	0.28	9.552	9.352	9.351
14	Z	Abu	CH <sub>2</sub> CH(OH)C <sub>6</sub> H <sub>4</sub> -3-OC <sub>6</sub> H <sub>3</sub> (3,4-C <sub>12</sub> )	0.1	10	10.156	9.87
15	Z	Abu	(CH <sub>2</sub> ) <sub>2</sub> -3-indolyl	0.031	10.508	10.36	10.124

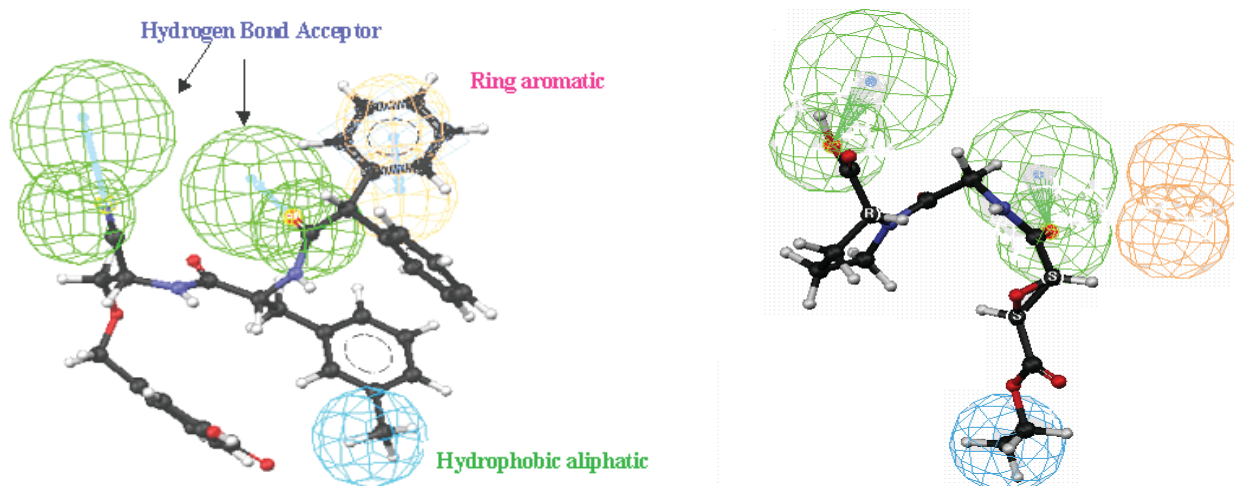
**Table 4.** 10 Pharmacophore Hypotheses Generated Using 25 Training Set Molecules

Hypothesis	Total cost	Cost Difference (Null cost–Total cost)	RMS deviation	Correlation	Features Definition
1	108.884	52.124	0.828	0.961	AAHR
2	109.794	51.214	1.004	0.936	AAHR
3	110.772	50.236	1.166	0.914	AAHR
4	112.752	48.256	1.244	0.902	AAHR
5	113.146	47.862	1.244	0.886	AAHR
6	114.754	46.254	1.290	0.862	AAHR
7	115.252	45.756	1.265	0.841	AAHR
8	116.744	44.264	1.274	0.821	AAHR
9	116.792	44.216	1.292	0.802	AAHR
10	120.792	40.216	1.311	0.786	AAHR

A – Hydrogen bond acceptor, H – Hydrophobic aliphatic and R – Ring aromatic.

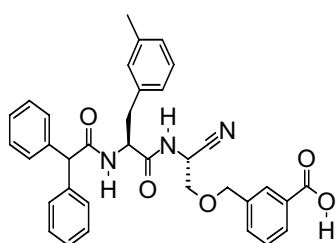
Null cost of 10 top scored hypotheses is 161.008. Fixed cost value is 101.312 Configuration is 12.352. All cost units are in bits.

For the training set the accuracy in predicting active and inactive compounds was 90%. The selected pharmacophore hypothesis yielded a RMS deviation of 0.828 and a correlation coefficient of 0.961 with a cost difference of 52.246. The best pharmacophore model was validated on 50 test molecules to give correlation value of 0.886. For the test set, the accuracy in predicting active compounds was greater than 8%, while 15% and 5% representing both false positive and negative respectively. The mapping of Hypothesis1 model onto an active training set compound 90 ( $IC_{50} = 1.8$  nM) is shown in Figure 3.

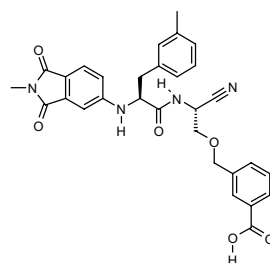


**Figure 3.** Pharmacophore. Mapping to most active compound in training set ( $IC_{50} = 1.8$  nM) and mapping to inactive compound in training set ( $IC_{50} = 15000$  nM).

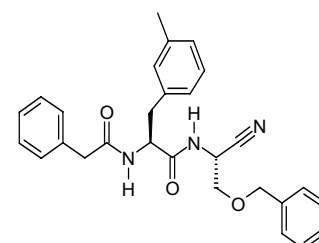
**Chemical Structures of the 26 Training Set Molecules Applied to HypoGen Pharmacophore Generation (Cathepsin B Activities Are Given as  $IC_{50}$  Values).**



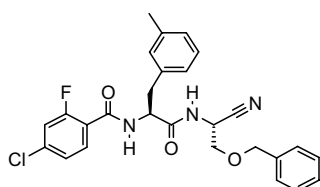
Compound-1  
 $IC_{50} = 1.8$  nM



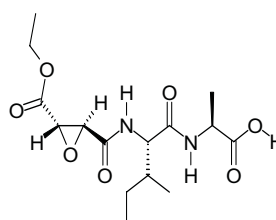
Compound-2  
 $IC_{50} = 4.9$  nM



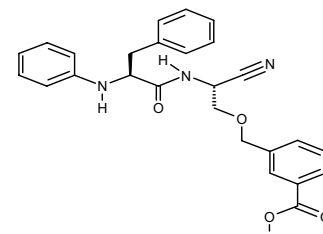
Compound-3  
 $IC_{50} = 10$  nM



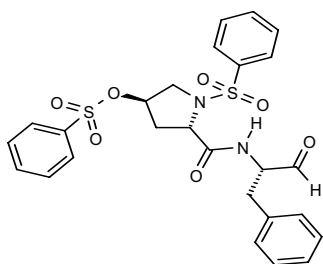
Compound-4  
 $IC_{50} = 17$  nM



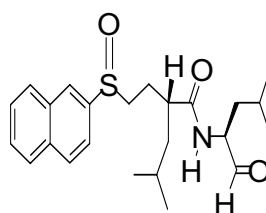
Compound-5  
 $IC_{50} = 23$  nM



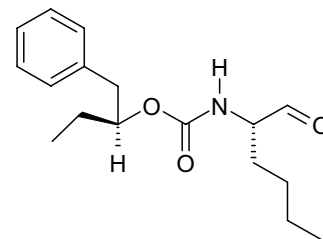
Compound-6  
 $IC_{50} = 30$  nM



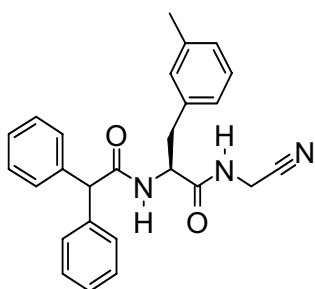
Compound-7  
IC<sub>50</sub> = 22 nM



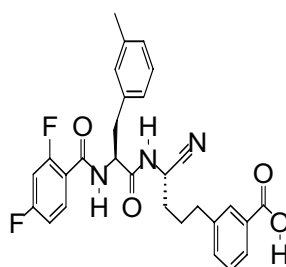
Compound-8  
IC<sub>50</sub> = 60 nM



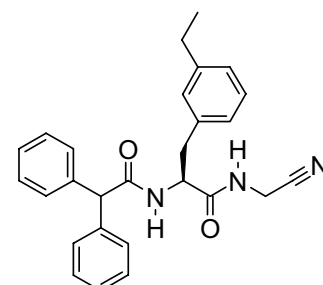
Compound-9  
IC<sub>50</sub> = 60 nM



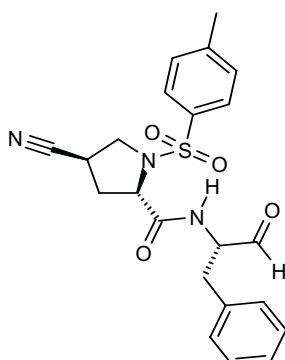
Compound-10  
IC<sub>50</sub> = 87 nM



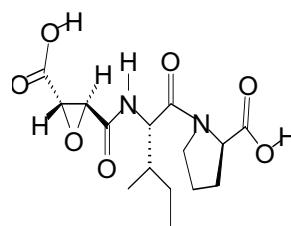
Compound-11  
IC<sub>50</sub> = 110 nM



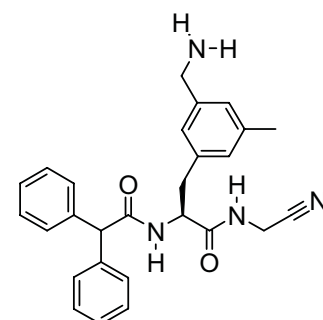
Compound-12  
IC<sub>50</sub> = 140 nM



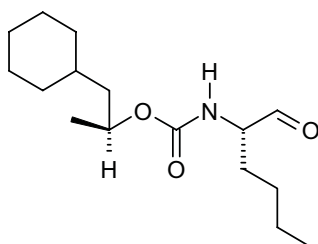
Compound-13  
IC<sub>50</sub> = 180 nM



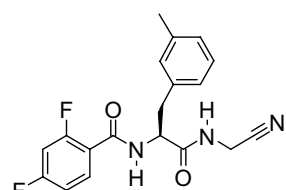
Compound-14  
IC<sub>50</sub> = 280 nM



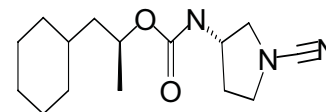
Compound-15  
IC<sub>50</sub> = 360 nM



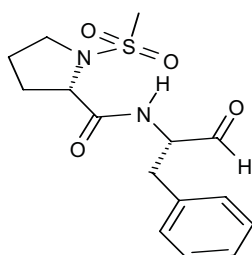
Compound-16  
IC<sub>50</sub> = 400 nM



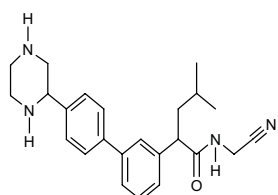
Compound-17  
IC<sub>50</sub> = 590 nM



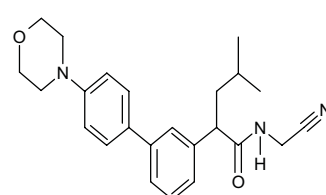
Compound-18  
IC<sub>50</sub> = 680 nM



Compound-19  
IC<sub>50</sub> = 880 nM

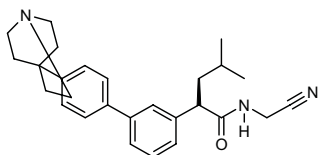


Compound-20  
IC<sub>50</sub> = 1200 nM

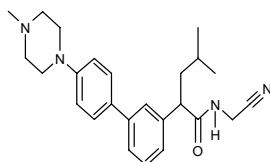


Compound-21  
IC<sub>50</sub> = 1700 nM

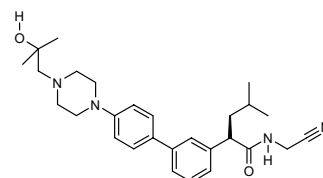




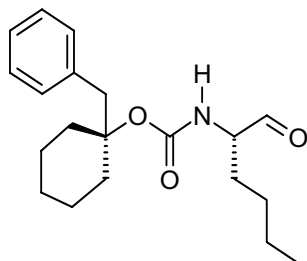
Compound-22  
IC<sub>50</sub> = 2200 nM



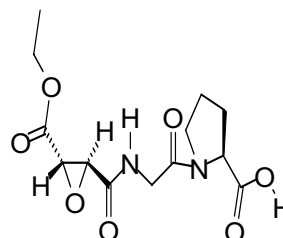
Compound-23  
IC<sub>50</sub> = 3800 nM



Compound-24  
IC<sub>50</sub> = 6400 nM

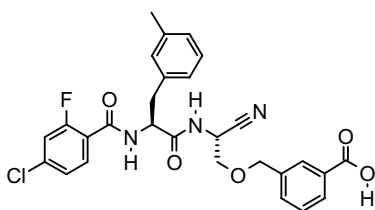


Compound-25  
IC<sub>50</sub> = 12000 nM

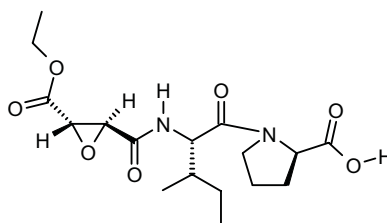


Compound-26  
IC<sub>50</sub> = 15000 nM

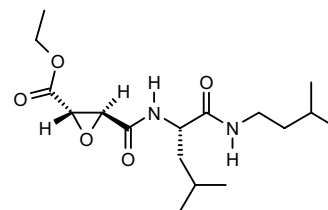
**Chemical Structures of the 50 Cathepsin B-antagonists used as Test Set for Validation of the Predictive Power of the HypoGen Pharmacophore**



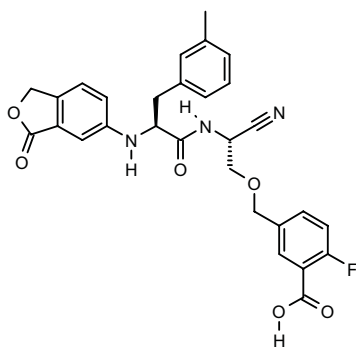
Compound-1  
IC<sub>50</sub> = 2 nM



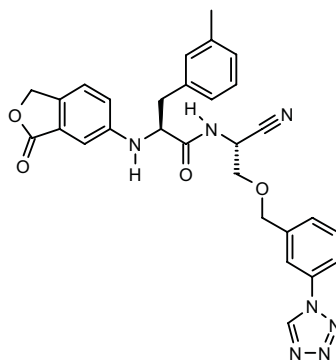
Compound-2  
IC<sub>50</sub> = 2.3 nM



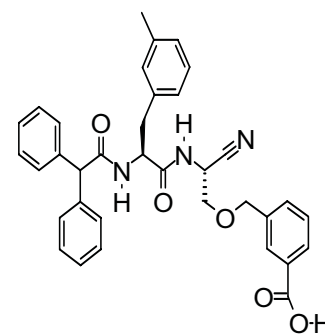
Compound-3  
IC<sub>50</sub> = 3.4 nM



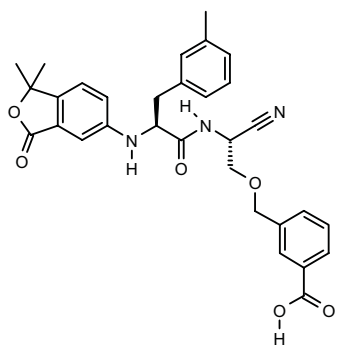
Compound-4  
IC<sub>50</sub> = 4.1 nM



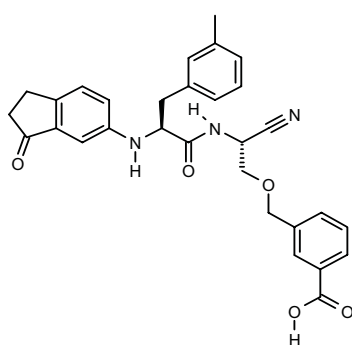
Compound-5  
IC<sub>50</sub> = 5 nM



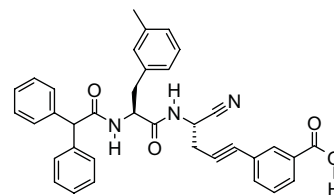
Compound-6  
IC<sub>50</sub> = 5 nM



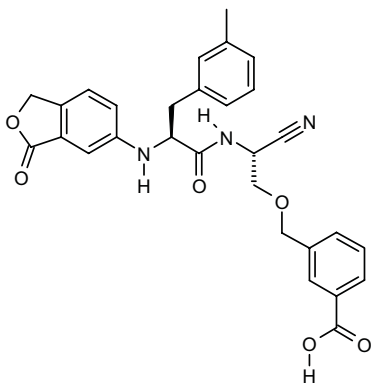
Compound-7  
IC<sub>50</sub> = 5.3 nM



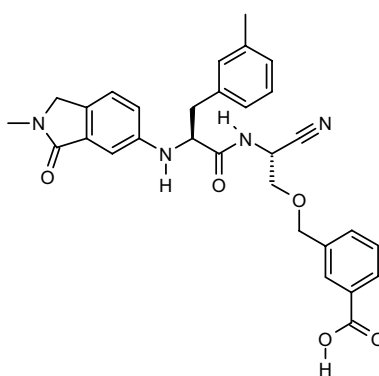
Compound-8  
IC<sub>50</sub> = 5.3 nM



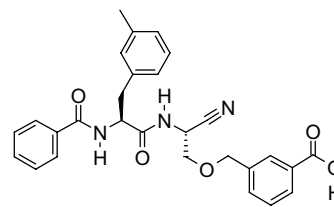
Compound-9  
IC<sub>50</sub> = 5.6 nM



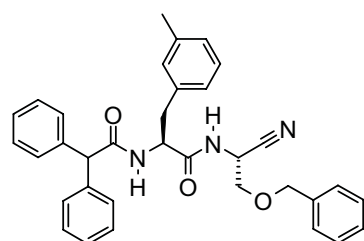
Compound-10  
IC<sub>50</sub> = 6.5 nM



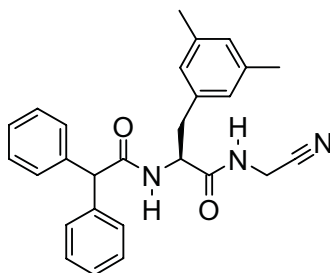
Compound-11  
IC<sub>50</sub> = 9.3 nM



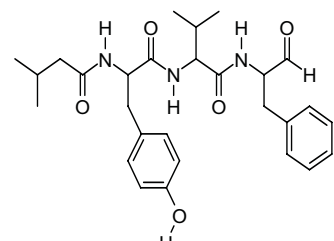
Compound-12  
IC<sub>50</sub> = 9.4 nM



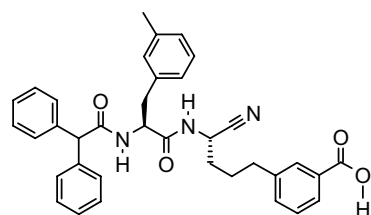
Compound-13  
IC<sub>50</sub> = 10.2 nM



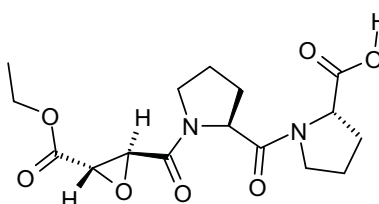
Compound-14  
IC<sub>50</sub> = 11.9 nM



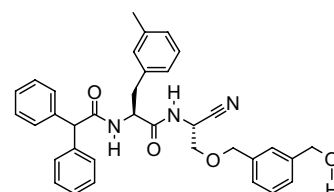
Compound-15  
IC<sub>50</sub> = 12 nM



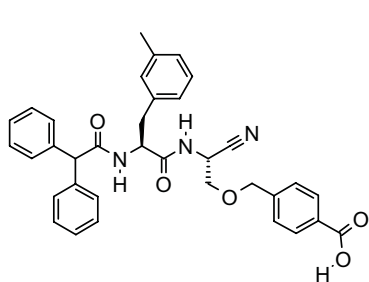
Compound-16  
IC<sub>50</sub> = 18 nM



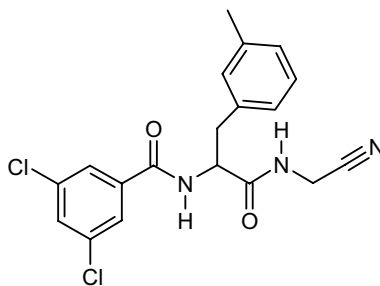
Compound-17  
IC<sub>50</sub> = 25 nM



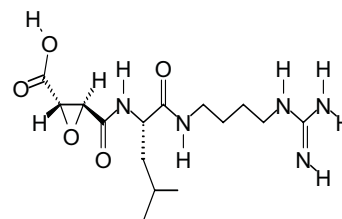
Compound-18  
IC<sub>50</sub> = 29 nM



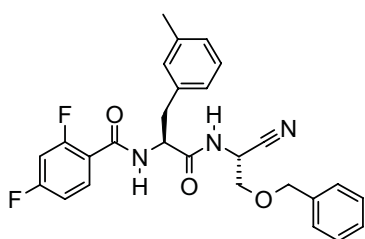
Compound-19  
IC<sub>50</sub> = 30.7 nM



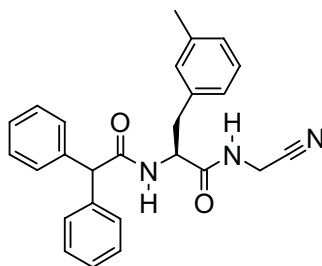
Compound-20  
IC<sub>50</sub> = 31 nM



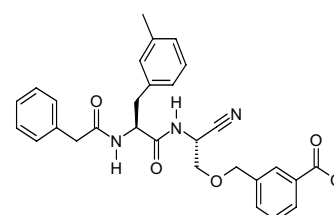
Compound-21  
IC<sub>50</sub> = 35 nM



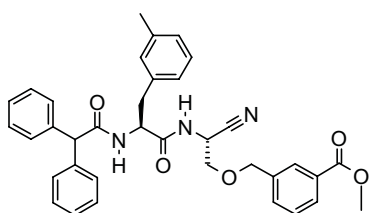
Compound-22  
IC<sub>50</sub> = 42 nM



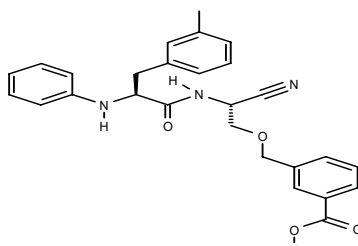
Compound-23  
IC<sub>50</sub> = 45 nM



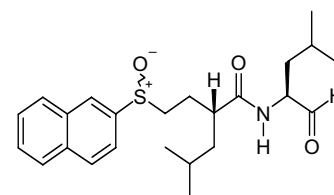
Compound-24  
IC<sub>50</sub> = 47 nM



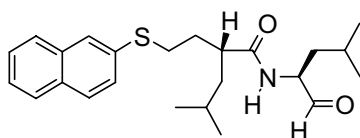
Compound-25  
IC<sub>50</sub> = 47 nM



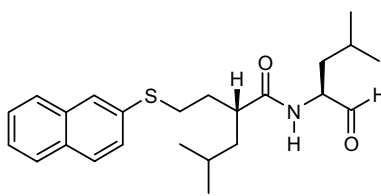
Compound-26  
IC<sub>50</sub> = 49 nM



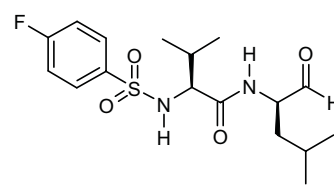
Compound-27  
IC<sub>50</sub> = 60



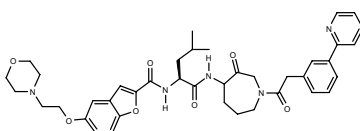
Compound-28  
IC<sub>50</sub> = 60 nM



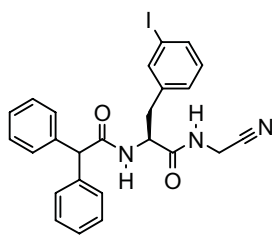
Compound-29  
IC<sub>50</sub> = 60 nM



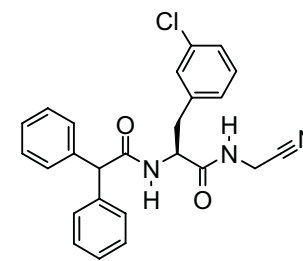
Compound-30  
IC<sub>50</sub> = 78 nM



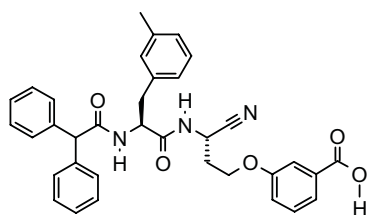
Compound-31  
IC<sub>50</sub> = 100



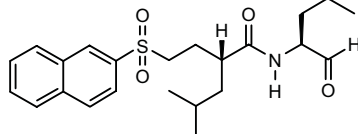
Compound-32  
IC<sub>50</sub> = 121



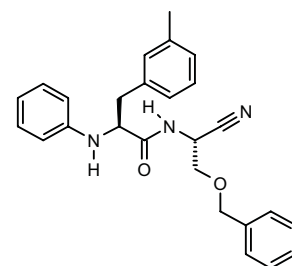
Compound-33  
IC<sub>50</sub> = 130



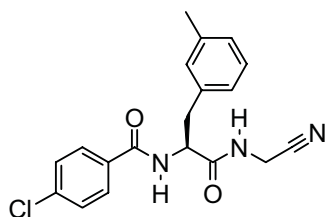
Compound-34  
IC<sub>50</sub> = 178 nM



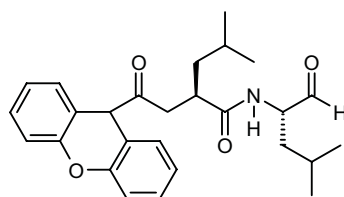
Compound-35  
IC<sub>50</sub> = 150 nM



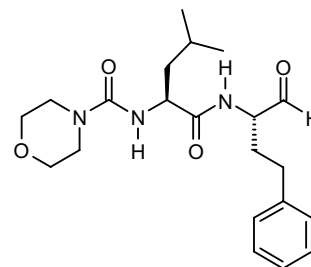
Compound-36  
IC<sub>50</sub> = 194 nM



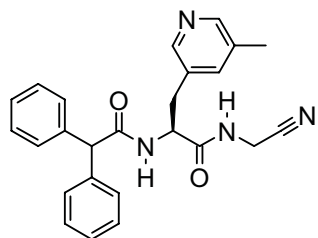
Compound-37  
IC<sub>50</sub> = 308 nM



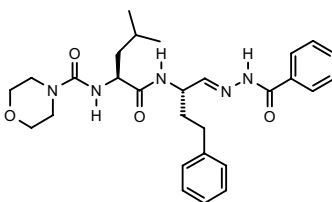
Compound-38  
IC<sub>50</sub> = 440 nM



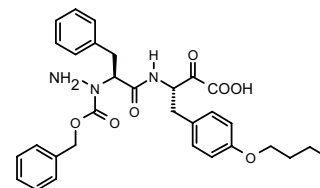
Compound-39  
IC<sub>50</sub> = 640 nM



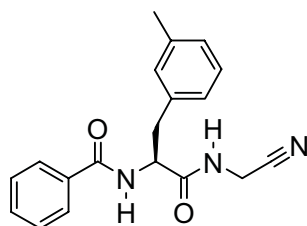
Compound-40  
IC<sub>50</sub> = 652 nM



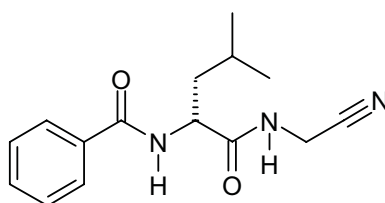
Compound-41  
IC<sub>50</sub> = 700 nM



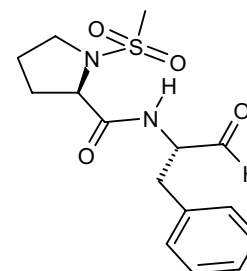
Compound-42  
IC<sub>50</sub> = 836 nM



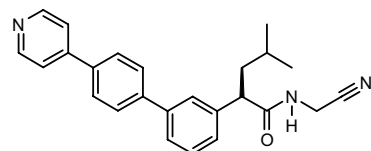
Compound-43  
IC<sub>50</sub> = 1900 nM



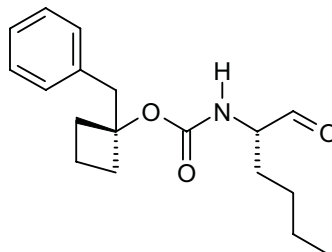
Compound-44  
IC<sub>50</sub> = 1903 nM



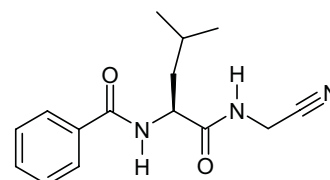
Compound-45  
IC<sub>50</sub> = 3000 nM



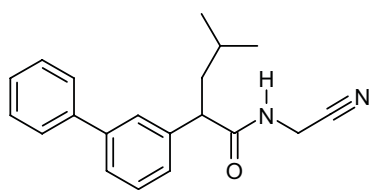
Compound-46  
IC<sub>50</sub> = 3329 nM



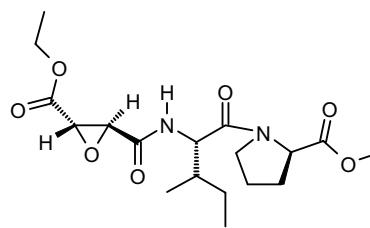
Compound-47  
IC<sub>50</sub> = 3400 nM



Compound-48  
IC<sub>50</sub> = 3560 nM



Compound-49  
IC<sub>50</sub> = 3560 nM



Compound-50  
IC<sub>50</sub> = 4100 nM

## 4 CONCLUSIONS

The 3D QSAR study shows that hydrophobic groups are responsible for increase in activity, which shows that RSA model predicts better than MFA. Further, the knowledge of this four-feature pharmacophore hypothesis for cathepsin B inhibitors can be very useful for virtual screening to design more potent lead moieties for the treatment of various types of cancer.

### Acknowledgment

We sincerely thank Dr. J. A. R. P. Sarma, GVK Biosciences for providing lab facilities and helpful discussions.

## 5 REFERENCES

- [1] Paul D. Greenspan, N-Arylamino nitriles as bioavailable peptidomimetic inhibitors of cathepsin B, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 4121–4124.
- [2] Sinha AA, Morgan JL, Buus RJ, Ewing SL, Fernandes ET, Le C, Wilson MJ, Cathepsin B expression is similar in African-American and Caucasian prostate cancer patients, *Anticancer Res.* **2007**, *27*, 3135–3141.
- [3] Giusti I, D'Ascenzo S, Millimaggi D, Taraboletti G, Carta G, Franceschini N, Pavan A, Dolo V, Cathepsin B mediates the pH-dependent proinvasive activity of tumor-shed microvesicles, *Neoplasia*. **2008**, *10*, 481–488.
- [4] Ha SD, Martins A, Khazaie K, Han J, Chan BM, Kim SO, Cathepsin B is involved in the trafficking of TNF- $\alpha$ -containing vesicles to the plasma membrane in macrophages, *J Immunol.* **2008**, *181*, 690–697.
- [5] Leung-Toung R.; Li W.; Tam T.F.; Kaarimian K, Thiol-Dependent Enzymes and Their Inhibitors: A Review, *Med Chem.* **2002**, *9*, 979–1002.
- [6] Charles L. Cywin, Raymond A. Firestone, The design of potent hydrazones and disulfides as cathepsin S inhibitors, *Bioorg. Med. Chem.* **2003**, *11*, 733–740.
- [7] Paul D. Greenspan, Kirk L. Clark, Scott D. Cowen, Leslie W. McQuire, Ruben A. N-Arylamino nitriles as bioavailable peptidomimetic inhibitors of cathepsin B, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 4121–4124.
- [8] Joel Robichaud, Renata Oballa, Peppi Prasit, Jean-Pierre Falguyret, A Novel Class of Nonpeptidic Biaryl Inhibitors of Human Cathepsin K, *J. Med. Chem.* **2003**, *46*, 3709–3727.
- [9] Hans-Hartwig Otto† and Tanja Schirmeister, Cysteine Proteases and Their Inhibitors, *Chem. Rev.* **1997**, *97*, 133–172.
- [10] Robert W. Marquis, Inhibition of the Cysteine Protease Cathepsin K, *Annu. Rep. Med. Chem.* **2004**, *39*, 79–98.
- [11] Paul D. Greenspan, Kirk L. Clark, Ruben A, Identification of Dipeptidyl Nitriles as Potent and Selective Inhibitors of Cathepsin B through Structure-Based Drug Design, *J. Med. Chem.* **2001**, *Nov14*, 4524–4534.
- [12] David N. Deaton, Anne M. Hassell, Robert B. McFadyen, Aaron B. Miller, Larry R. Miller, Lisa M. Shewchuk, Francis X. Tavares, Derril H. Willard, Lois L. Wright, Novel and potent cyclic cyanamide-based cathepsin K inhibitors, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1815–1819.
- [13] Rabindranath Tripathy, Zi-Qiang Gu, Derek Dunn, Shobha E. Senadhi, Mark A. Ator, Sankar Chatterjee, P2-proline-derived inhibitors of calpain I, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2647–2652.
- [14] Schirmeister T, Klockow A, Cysteine Protease Inhibitors Containing Small Rings, *Mini. Rev. Med. Chem.* **2003**, *3*, 585–596.
- [15] Schirmeister T, Klockow A, Cysteine Protease Inhibitors Containing Small Rings, *J. Med. Chem.* **1993**, *36*, 3472–3480.

- [16] J. A. R. P. Sarma, G. Rambabu, K. Srikanth, D. Raveendra and M. Vithal, Analogue Based Design of MMP-13 (Collagenase-3) Inhibitors, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2689–2693.
- [17] Sairam KV, Sarma JA, Desiraju GR. 3D-QSAR studies of some [[1-aryl(or benzyl)-1-(benzenesulphonamido)methyl] phenyl] alkanolic acid derivatives as thromboxane A2 receptor antagonists, *Drug Des Discov.* **2003**, *18*, 47–51.
- [18] Vema A, Panigrahi SK, Rambabu G, Gopalakrishnan B, Sarma JA, Desiraju GR, Design of EGFR kinase inhibitors: a ligand-based approach and its confirmation with structure-based studies, *Bioorg Med Chem.* **2003**, *11*, 4643–53.
- [19] Boppana K, Dubey PK, Jagarlapudi SA, Vadivelan S, Rambabu G, Knowledge based identification of MAO-B selective inhibitors using pharmacophore and structure based virtual screening models, *Eur J Med Chem.* **2009** Mar 11.
- [20] Vadivelan S, Sinha BN, Tajne S, Jagarlapudi SA, Fragment and knowledge-based design of selective GSK-3beta inhibitors using virtual screening models, *Eur J Med Chem.* **2009**, *44*, 2361–71.
- [21] Vadivelan S, Sinha BN, Irudayam SJ, Jagarlapudi SA, Virtual screening studies to design potent CDK2-cyclin A inhibitors, *J Chem Inf Model.* **2007**, *47*, 1526–35.
- [22] Catalyst, Version 4.7, Accelrys, Inc. (Previously known as Molecular Simulations Inc.): 9685 Scranton Road, San Diego, CA 92121, **2000**.
- [23] Cerius2, Version 4.6, Accelrys, Inc. (Previously known as Molecular Simulations Inc.): 9685 Scranton Road, San Diego, CA 92121, **2000**.