Original Article

3D-QSAR study with pharmacophore-based molecular alignment of hydroxamic acid-related phosphinates that are aminopeptidase N inhibitors

Huawei Zhu[1], Hao Fang[1], Luyao Wang[2], Wenxiang Hu[2], Wenfang Xu[1,*]

[1] *Institute of Medicinal Chemistry, School of Pharmaceutical Sciences, Shandong University, Ji'nan, Shandong, China;*
[2] *Department of Chemistry, Capital Normal University, Beijing, China.*

ABSTRACT: 3D-QSAR models for a series of aminopeptidase N inhibitors were developed based on comparative molecular field analysis (CoMFA) and comparative molecular similarity analysis (CoMSIA). GALAHAD, a pharmacophore generation module involving a genetic algorithm, was used to generate the pharmacophore model of the series of inhibitors. Molecules both in the training set and test set were aligned to the pharmacophore model. Values for the CoMFA model were r^2 = 0.992, q^2 = 0.586, SEE = 0.111, and F (8, 10) = 191.263. Values for the CoMSIA model were r^2 = 0.990, q^2 = 0.776, SEE = 0.123, and F (7, 11) = 156.68. This model can help not only in improving current understanding of enzyme-ligand interactions but also in predicting the activity of derivatives and designing new compounds with better potency.

Keywords: Aminopeptidase N, 3D-QSAR, CoMFA, CoMSIA

Introduction

Aminopeptidase N (APN)/CD13 (EC 3.4.11.2) is a transmembrane protease belonging to the M1 family of Zinc-dependent metalloexopeptidase (*1,2*). It is widely expressed on the surface of a variety of human tissues and cell types and especially on renal and intestinal brush border cells (*1,3,4*). It is also considered to be the receptor of TGEV and HEV229E and *Bacillus thuringensis* Cry1A toxin (*5-7*). Furthermore, many studies have indicated that APN is overexpressed on tumor cells and plays an essential role in extracellular matrix degradation and invasion of tumor cells (*8,9*). Therefore, APN inhibitors have been used to suppress tumor cell invasion (*10*). Bestatin, which is the sole APN inhibitor on the market to date, can treat leukemia with an inhibiting capacity in the micromole range (*11,12*). Thus, a number of laboratories are designing and synthesizing APN inhibitors with a new scaffold, expecting with greater APN inhibiting ability. A series of hydroxamic acid-related phosphinates compounds designed and synthesized by Marcin Drag and his co-workers has shown good APN-inhibiting activity (*13*). Here, 3D-QSAR analysis with molecular alignment based on a pharmacophore model has been performed on this series of APN inhibitors, using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) with partial least-square fit to predict the steric, electrostatic, hydrophobic and H-bond molecular field interactions for this activity (*14,15*).

Correspondence to: Dr. Wenfang Xu, Institute of Medicinal Chemistry, School of Pharmaceutical Sciences, Shandong University, Ji'nan, Shandong, China; e-mail: xuwenf@sdu.edu.cn

Materials and Methods

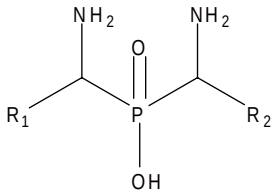
Computer modeling was conducted on a Dell Precision workstation. All calculations were done with the programs SYBYL7.0 (*16*) and SYBYL7.3 (*17*) with default values except those specially referred to. A set of 23 hydroxamic acid-related phosphinates derivatives (*13*) has been chosen for 3D-QSAR study. The affinity of compounds binding to APN was represented by the IC_{50} value (μM) in the literature (Table 1). Log ($1/IC_{50}$) values (pIC_{50}) were used to derive 3D-QSAR models (*see* Table 3).

Pharmacophoric conformation determination

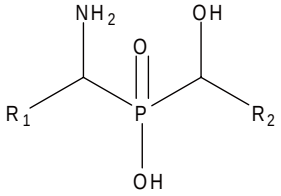
Compounds described in the literature were racemic mixtures, with two chiral carbon atoms included in each. The pharmacophoric conformation of the compounds was first determined. Compound **16** was selected as the template molecule because it had the greatest activity. The four isomers of compound **16** were built with a SYBYL/Sketch module and the

Table 1. Structure of compounds used in this study and their bioactivity

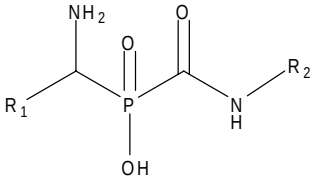
Compounds	R_1	R_2	IC_{50} (μM)
1	$-C_6H_5$ (Phenylglycine)	–H (Gly)	42
2	$-C_6H_5$ (Phenylglycine)	$-CH_3$ (Ala)	46
3	$-C_6H_5$ (Phenylglycine)	$-CH_2CH_2CH_3$ (nor-Val)	17
4	$-C_6H_5$ (Phenylglycine)	$-CH_2CH(CH_3)_2$ (Leu)	4
5	$-C_6H_5$ (Phenylglycine)	$-C_6H_5$ (Phenylglycine)	0.6
6	$-C_6H_5$ (Phenylglycine)	$-CH_2CH_2$Phe (hPhe)	0.71
7	$-CH_2CH_2$Phe (hPhe)	$-CH_2CH_2$Phe (hPhe)	0.83
8	$-CH_2CH(CH_3)_2$ (Leu)	–H (Gly)	76
9	$-CH_2CH(CH_3)_2$ (Leu)	$-CH_3$ (Ala)	155
10	$-CH_2CH(CH_3)_2$ (Leu)	$-CH_2CH_2CH_3$ (nor-Val)	50
11	$-CH_2CH(CH_3)_2$ (Leu)	$-CH_2CH(CH_3)_2$ (Leu)	50



Compounds	R_1	R_2	IC_{50} (μM)
12	$-C_6H_5$ (Phenylglycine)	$-CH_3$ (Ala)	48
13	$-C_6H_5$ (Phenylglycine)	$-(CH_2)_3CH_3$ (nor-Leu)	28
14	$-C_6H_5$ (Phenylglycine)	$-CH_2CH(CH_3)_2$ (Leu)	12
15	$-C_6H_5$ (Phenylglycine)	$-C_6H_5$ (Phenylglycine)	0.55
16	$-C_6H_5$ (Phenylglycine)	$-CH_2CH_2$Phe (hPhe)	0.24
17	$-CH(CH_3)_2$ (Val)	$-CH_2CH_2$Phe (hPhe)	0.5
18	$-CH_2CH(CH_3)_2$ (Leu)	$-(CH_2)_3CH_3$ (nor-Leu)	3
19	$-CH_2CH(CH_3)_2$ (Leu)	$-CH_2CH_2$Phe (hPhe)	1



Compounds	R_1	R_2	IC_{50} (μM)
20	$-CH_2CH(CH_3)_2$ (Leu)	$-C_6H_5$ (Phenylglycine)	35
21	$-CH_2CH(CH_3)_2$ (Leu)	$-CH_2$Phe (Phe)	0.53
22	$-CH(CH_3)_2$ (Val)	$-C_6H_5$ (Phenylglycine)	160
23	$-CH(CH_3)_2$ (Val)	$-CH_2$Phe (Phe)	0.93



Powell method was used with a Tripos force field in the energy-minimization process; convergence criterion was 4.184 J/mol and a Gasteiger-Hückel charge was given. Then, the compounds were docked into the active site of E. coli APN (PDB Code: 2DQM) using the SYBYL/FlexX module. The docking result suggested both chiral carbon atoms had an *S*-configuration when the binding model was consistent with that described in literature (Figure 1). The other 22 molecules were built and minimized based on the template.

Pharmacophore-based molecular alignment

The series of compounds reported in the literature can be divided into three subtypes according to their scaffold, but alignment by traditional molecular alignment methods, such as fit atoms or match atoms, is difficult. Under such conditions, pharmacophore-based molecular alignment is an ideal way to align molecules in different scaffolds. The SYBYL new pharmacophore alignment module GALAHAD (Genetic Algorithm with Linear Assignment for Hypermolecular Alignment of Datasets) (*18*) aligns a set of molecules that share a common mode of biological activity and develops corresponding pharmacophore models. Using a sophisticated genetic algorithm and a multi-objective scoring function, GALAHAD takes into account energetics, steric similarity, and pharmacophoric overlap, while accommodating conformational flexibility, ambiguous stereochemistry, alternative ring configurations, multiple partial match constraints, and alternative feature mappings among molecules. Here compounds **5**, **6**, **16**, **17**, and **21** were chosen as the data source to generate pharmacophore models with GALAHAD (Figures 2 and 3). Compounds both in the training set and test set were aligned to the pharmacophore model using the Align Molecules to Template Individually method. The alignment conformation of all molecules is shown in Figure 4.

3D-QSAR analysis

3D-QSAR analysis was performed on the aligned compounds using the CoMFA and CoMSIA methods. In the CoMFA model, a hybrid sp^3 carbon atom with a positive charge was used as a probe to compute the CoMFA steric and electrostatic fields. The lattice size and probe step size were adjusted automatically. Default parameters were used for other values. Partial least squares (PLS) regression was separately performed on the compounds. The Leave-One-Out method, with 2.0 kcal/mol as the column filtering value, was first used

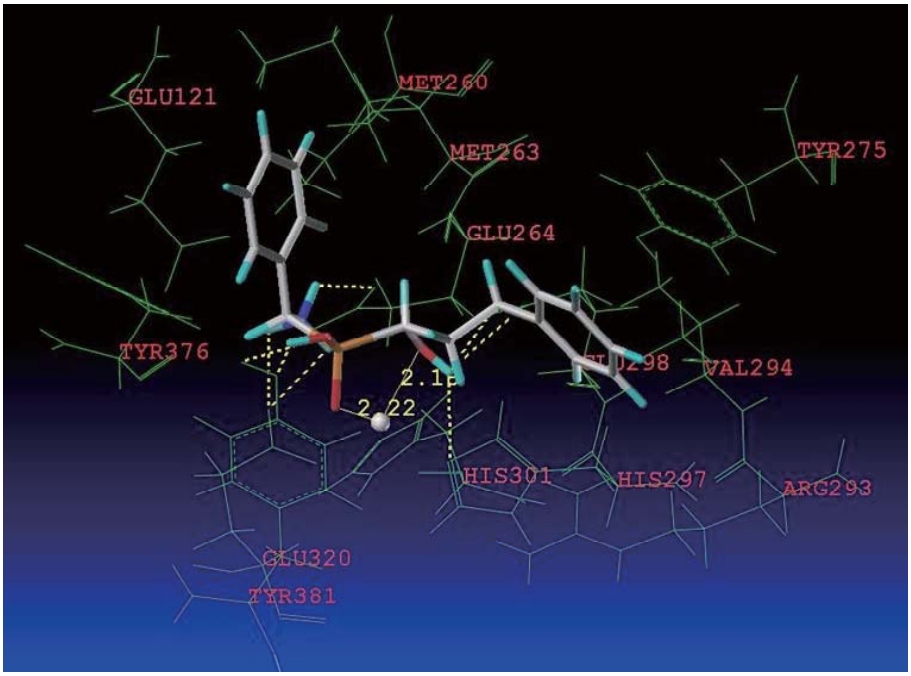


Figure 1. Compound **16** docked into the activity site of E.coli APN with the two chiral carbon atoms in an *S*-configuration.

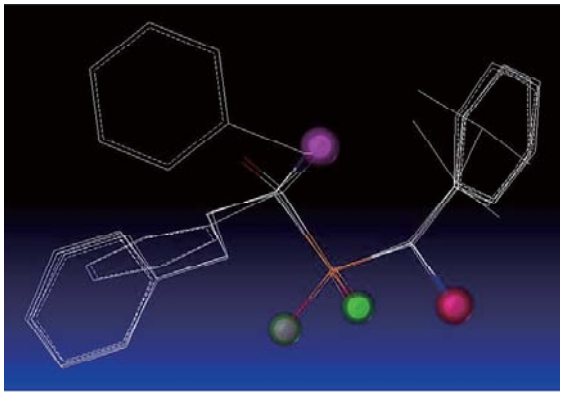


Figure 2. Pharmacophore model generated by GALAHAD with compounds **5**, **6**, **16**, **17** and **21**.

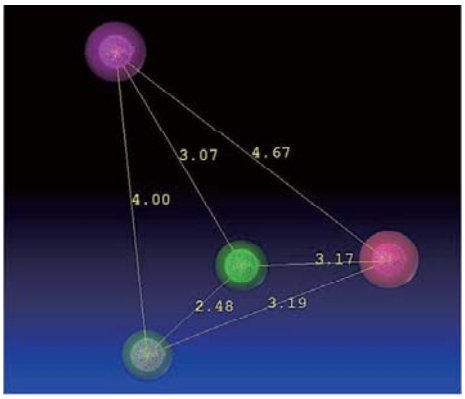


Figure 3. The relative spatial locations of the pharmacophore elements. Pink for positive nitrogen, magenta for H-bond donor, green for H-bond acceptor. The yellow line represents the relative distance of each element, with length in angstroms.

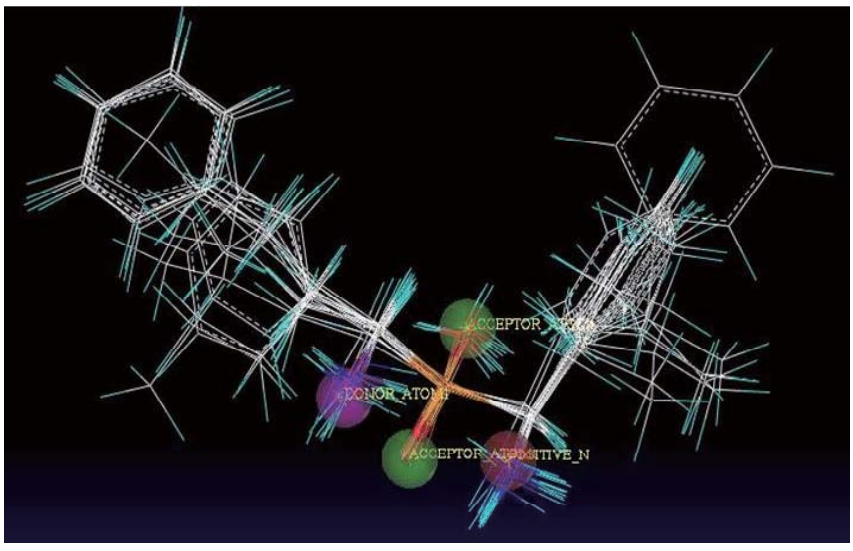


Figure 4. Pharmacophore-based alignment of the training set and the test set.

to obtain the optimum number of components (ONC), and then the CoMFA model was computed with non-cross-validation PLS at the ONC. Computation of the CoMSIA model was similar to that for the CoMFA model, and apart from steric and electrostatic fields the CoMSIA model took into account hydrophobic and H-bond fields.

Results and Discussion

Pharmacophore model

Four elements were included in the pharmacophore model which generated from compounds **5**, **6**, **16**, **17**, **21**, including two H-bond acceptors (AA1, AA2), one H-bond donor (DA), and one positive nitrogen site (PN). Their relative locations and relative distance are shown in Figure 3.

Table 2. Summary of PLS results

Statistical parameters	CoMFA	CoMSIA
r^2_{CV}	0.586	0.776
ONC	7	6
r^2	0.992	0.990
SEE	0.111	0.123
F_{ratio}	191.263	156.68
Fraction of field contributions		
Steric	59.4%	13.9%
Electrostatic	40.6%	39%
Hydrophobic	-	30%
Donor	-	11.8%
Acceptor	-	15.3%

CoMFA and CoMSIA models

PLS analysis was performed on CoMFA/CoMSIA field values and experimental bioactivity values (PIC_{50}) of the training set. The CoMFA and CoMSIA statistical analysis is summarized in Table 2. Statistical data show an r^2_{cv} of 0.586 for the CoMFA model and 0.776 for the CoMSIA model, which indicates good internal predictive ability for both models. The models developed also displayed an r^2_{NCV} of 0.992 and 0.990 for the CoMFA and CoMSIA models, respectively.

QSAR model verification

To test the predictive ability of the models, 19 molecules in a training set and a test set of 4 molecules excluded from the model derivation were used (Table 3). For the test set, a predictive correlation coefficient r^2_{pred} of 0.772 for the CoMFA model and 0.839 for the CoMSIA model indicates good external predictive ability for the models. Figure 5 shows the scatterplot of the actual pIC_{50} *vs* predicted pIC_{50}.

Contour analysis

The contour maps of CoMFA (electrostatic and steric) and CoMSIA (electrostatic, steric, hydrophobic, donor, and acceptor) are shown in Figure 6.

The steric contour maps of the CoMFA and CoMSIA models are shown in Figure 6a and 6c. The phenyl group on the R_2 substitute site was located in the green region, which suggested that bulky group substitution would help to increase potency. For example,

Table 3. Actual activity versus predicted activity and their residuals

Compounds	Actual pIC_{50}	CoMFA Predicted pIC_{50}	Residuals (CoMFA)	CoMSIA Predicted pIC_{50}	Residuals (CoMSIA)
1	4.38	4.35	0.03	4.38	0
2	4.34	4.42	-0.08	4.34	0
3*	4.77	4.68	0.09	4.51	0.26
4	5.40	5.18	0.22	5.21	0.19
5	6.22	6.20	0.02	6.19	0.03
6	6.15	6.12	0.03	6.13	0.02
7	6.08	6.12	-0.04	6.01	0.07
8	4.12	4.11	0.01	4.13	-0.01
9	3.81	3.81	0	3.87	-0.06
10*	4.30	4.98	-0.68	4.75	-0.45
11	4.30	4.35	-0.05	4.31	-0.01
12	4.32	4.34	-0.02	4.36	-0.04
13	4.55	4.46	0.09	4.46	0.09
14	4.92	5.11	-0.19	5.05	-0.13
15	6.26	6.31	-0.05	6.31	-0.05
16*	6.62	5.95	0.67	6.11	0.51
17	6.30	6.28	0.02	6.35	-0.05
18*	5.52	5.71	-0.19	4.98	0.54
19	6.00	6.00	0	5.97	0.03
20	4.46	4.32	0.14	4.19	0.27
21	6.28	6.39	-0.11	6.21	0.07
22	3.80	3.93	-0.13	4.09	-0.29
23	6.03	5.92	0.11	6.07	-0.04

* Molecules in test set.

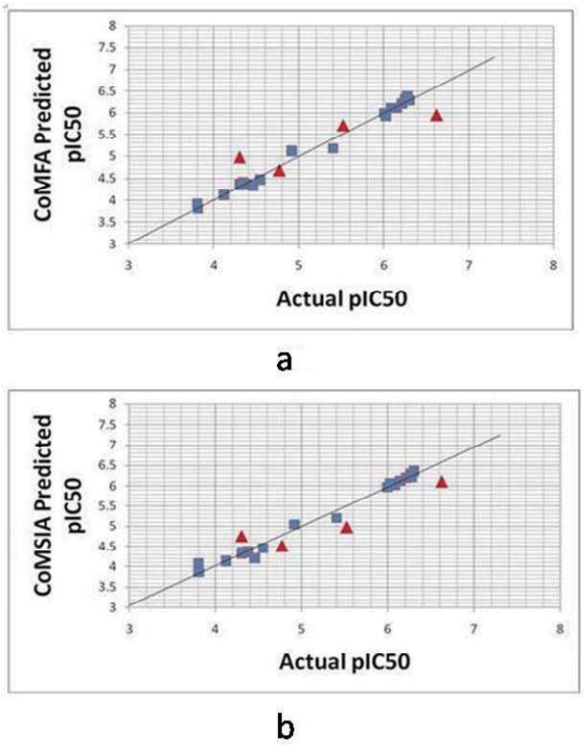


Figure 5. Scatterplot of actual pIC_{50} *vs* CoMFA (a) and CoMSIA (b) predicted pIC_{50}. ■, Training set; ▲, Test set.

compounds **1** (pIC_{50} = 4.38), **2** (pIC_{50} = 4.34), **8** (pIC_{50} = 4.12), and **9** (pIC_{50} = 3.81 M) exhibited low inhibitory activity due to the lack of a steric bulk substituent in the R_2 position. While the compounds with $pIC_{50} > 6$, such as compounds **5** (pIC_{50} = 6.22), **6** (pIC_{50} = 6.15), **7** (pIC_{50} = 6.08), **15** (pIC_{50} = 6.26), **16** (pIC_{50} = 6.62), **17** (pIC_{50} = 6.3), and **21** (pIC_{50} = 6.28), have phenyl or phenethyl group substitution in the R_2 position. The yellow contour maps existed near the R_1 substitute site, which indicate that bulky group substitution at this position will lower potencies, such as compounds **10** (pIC_{50} = 4.30) and **11** (pIC_{50} = 4.30) with a long and flexible isobutyl group in the R_1 substitute.

CoMFA and CoMSIA electrostatic contour maps are shown in Figure 6b and 6d. Both electron positive and electron negative substitution in the R_2 position significantly affect inhibitory activity because most blue and red regions are located around the R_2 position. CoMSIA H-bond donor/acceptor contour maps (Figure 6e and 6f) indicate favoring of a H-bond acceptor rather than H-bond donor groups near the hydroxyl group of compound **15**.

CoMSIA hydrophobic contour maps are shown in Figure 6g. The yellow and white contour maps highlight areas where hydrophobic and hydrophilic properties are preferred. The presence of a large yellow contour map beside the R_2 substituent suggests that its occupancy by hydrophobic groups would favor inhibitory activity. In the most active compounds R_2 was a phenyl group (*i.e.*, compounds **5** and **15**) or a phenethyl group (*i.e.*, compounds **6**, **7**, **16**, and **17**).

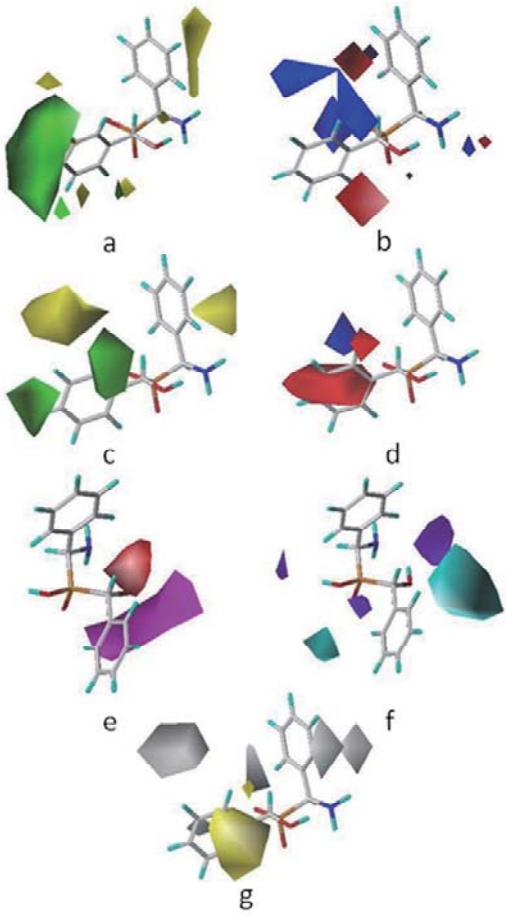


Figure 6. (a) CoMFA steric fields (color code: favored, green; disfavored, yellow); (b) CoMFA electrostatic fields (color code: increase in positive charge favored, blue; increase in negative charge favored, red); (c) CoMSIA steric fields (color code: favored, green; disfavored, yellow); (d) CoMSIA electrostatic fields (color code: increase in positive charge favored, blue; increase in negative charge favored, red); (e) CoMSIA acceptor fields (color code: favored, magenta; disfavored, red); (f) CoMSIA donor fields (color code: favored, cyan; disfavored, purple); (g) CoMSIA hydrophobic fields (color code: favored, yellow; disfavored, white). The reference molecule is **15**.

Conclusion

The present study examined a 3D-QSAR model of a set of APN inhibitors, and pharmacophore-based molecular alignment provided a QSAR model with a high level of predictability. To the extent known, this is the first time the GALAHAD module has been introduced in the process of generating CoMFA and CoMSIA models. The models derived in this study can help to elucidate the binding models of APN inhibitors and may lead to design of new acid-related phosphinates exhibiting a high level of APN inhibition.

Acknowledgements

The authors wish to thank Ms. Mei Zhou (Beijing Hongcam Software Technologies Co., Ltd.) for her helpful instruction in using the SYBYL/GALAHAD module.

References

1. Antczak C, De Meester I, Bauvois B. Ectopeptidases in pathophysiology. BioEssays 2001; 23:251-260.
2. Taylor A. Aminopeptidases: structure and function. FASEB J 1993; 7:290-298.
3. Antczak C, De Meester I, Bauvois B. Transmembrane proteases as disease markers and targets for therapy. J Biol Regul Homeost Agents 2001; 15:130-139.
4. Jardinaud F, Banisadr G, Noble F, Mélik-Parsadaniantz S, Chen H, Dugave C, Laplace H, Rostène W, Fournié-Zaluski MC, Roques BP, Popovici T. Ontogenic and adult whole body distribution of aminopeptidase N in rat investigated by *in vitro* autoradiography. Biochimie 2004; 86:105-113.
5. Delmas B, Gelfi J, L'Haridon R, Vogel LK, Sjöström H, Norén O, Laude H. Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. Nature 1992; 357:417-420.
6. Yeager CL, Ashmun RA, Williams RK, Cardellichio CB, Shapiro LH, Look AT, Holmes KV. Human aminopeptidase N is a receptor for human coronavirus 229E. Nature 1992; 357:420-422.
7. Nakanishi K, Yaoi K, Nagino Y, Hara H, Kitami M, Atsumi S, Miura N, Sato R. Aminopeptidase N isoforms from the midgut of Bombyx mori and Plutella xylostella-their classification and the factors that determine their binding specificity to Bacillus thuringiensis Cry1A toxin. FEBS Lett 2002; 19:215-220.
8. Saiki I, Fujii H, Yoneda J, Abe F, Nakajima M, Tsuruo T, Azuma I. Role of aminopeptidase N (CD13) in tumor-cell invasion and extracellular matrix degradation. Int J Cancer 1993; 54:137-143.
9. Menrad A, Speicher D, Wacker J, Herlyn M. Biochemical and functional characterization of aminopeptidase N expressed by human melanoma cells. Cancer Res 1993; 53:1450-1455.
10. Fujii H, Nakajima M, Aoyagi T, Tsuruo T. Inhibition of tumor cell invasion and matrix degradation by aminopeptidase inhibitors. Biol Pharm Bull 1996; 19:6-10.
11. Fujisaki T, Otsuka T, Gondo H, Okamura T, Niho Y, Ohhinata A, Abe F. Bestatin selectively suppresses the growth of leukemic stem/progenitor cells with BCR/ABL mRNA transcript in patients with chronic myelogeneous leukemia. Int Immunopharmacol 2003; 3:901-907.
12. Xu WF, Li QB. Progress in the development of aminopeptidase N (APN/CD13) inhibitors. Curr Med Chem Anticancer Agents 2005; 5:281-301.
13. Drag M, Grzywa R, Oleksyszyn J. Novel hydroxamic acid-related phosphinates: Inhibition of neutral aminopeptidase N (APN). Bioorg Med Chem Lett 2007; 17:1516-1519.
14. Cramer RD, Patterson DE, Bunce JD. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 1988; 110:5959-5967.
15. Klebe G, Abraham U, Mietzner T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem 1994; 37:4130-4146.
16. Sybyl, version 7.0; Tripos Inc.: St. Louis, MO, 2004.
17. Sybyl, version 7.3; Tripos Inc.: St. Louis, MO, 2007.
18. GALAHAD, Tripos, St. Louis, MO; http://www.tripos.com/data/SYBYL/GALAHAD_9-7-05.pdf.

(*Received December 13, 2007; Revised January 29, 2008; Accepted February 13, 2008*)