Optimization of Neural Networks Architecture for Impact Sensitivity of Energetic Molecules

Soo Gyeong Cho,[*] Kyoung Tai No,[†,‡,*] Eun Mee Goh, Jeong Kook Kim, Jae Hong Shin,[†] Young Dae Joo,[†] and Seeyearl Seong[†]

Agency for Defense Development, P.O. Box 35-5, Yuseong, Daejeon 305-600, Korea. [*]E-mail: soocho@add.re.kr
[†]Research Institute of Bioinformatics & Molecular Design, Seoul 120-749, Korea
[‡]Department of Biotechnology, Yonsei University, Seoul 120-749, Korea. [*]E-mail: ktno@bmdrc.org
Received September 13, 2004

We have utilized neural network (NN) studies to predict impact sensitivities of various types of explosive molecules. Two hundreds and thirty four explosive molecules have been taken from a single database, and thirty nine molecular descriptors were computed for each explosive molecule. Optimization of NN architecture has been carried out by examining seven different sets of molecular descriptors and varying the number of hidden neurons. For the optimized NN architecture, we have utilized 17 molecular descriptors which were composed of compositional and topological descriptors in an input layer, and 2 hidden neurons in a hidden layer.

Key Words : Energetic molecule, Explosives, Impact sensitivity, Neural networks, Molecular descriptor

Introduction

Predicting sensitivity is of significant importance in deriving novel energetic molecules because safe handling is one of the most important issues to the scientists and engineers who handle energetic molecules.[1] Among various aspects of sensitivity, impact sensitivity is closely related to many accidents in working places. Experimentally, the impact sensitivity is measured by drop weight impact test, where a height of 50% probability in causing an explosion ($H_{50\%}$) was measured when hit by a hammer with a standard weight.[2] Obtaining reliable experimental results is known to be relatively difficult because of the nature of impact sensitivity test. Apparently, prediction of the impact sensitivity is also a difficult task.[3] This difficulty is probably associated with puzzles in initiation mechanisms of explosion caused by mechanical impact, which is still considered as a black box. In addition, large errors present in some of experimental data also hamper for explosives scientists to find a good correlation between some of specific molecular descriptors and the impact sensitivity of energetic molecules.[4-7] Although some scientists report successes in finding a reasonable correlation between them, this appears to be only possible in a certain molecular type with a small number of energetic molecules. Thus, if one wants to derive a predictive methodology for impact sensitivity to cover various types of energetic molecules, neural networks (NN) may be a good choice in current circumstances.[8]

Database and Selection of Descriptors

We used a database archived by Storm, Stine, and Kramer (SSK), where experimental values of impact sensitivity for a variety of explosive molecules had been collected.[7] Since experimental data of impact sensitivity varied widely depending upon instrumental types, experimental conditions and others, the database of SSK was not augmented by others. The database of SKK contained 258 explosive molecules over a variety of molecular types such as nitroaromatics, nitropyridines, nitroimidazoles, nitropyrazoles, nitrofurazans, nitrotriazoles, nitropyrimidines, nitramines, and nitrate esters. They contained 2,4,6-trinitrotolune (TNT), 1,3,5-trinitrohexahydro-1,3,5-triazine (RDX), and octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (HMX). We removed 12 molecules of which the experimental values were ambiguous. Initially, 246 explosive molecules were employed for our NN study. Later, 234 explosive molecules were used by removing 12 insensitive molecules whose experimental values were given as 'greater than 320 cm'. The minimum energy conformations of the molecules were obtained with CVFF and energy minimization procedure in Cerius2 commercial software.[9] Thirty-nine molecular descriptors were calculated for each molecule, and summarized in Table 1. All the compositional including oxygen balance, topological descriptors, and the number of rotatable bonds, HB donor and acceptors were computed manually. The electronic descriptors, LUMO_MOPAC, HOMO_MOPAC, Dipole_MOPAC, and HF_MOPAC were calculated with the AM1 hamiltonian[10] of the MOPAC package implemented in the Cerius2 program. Sum of atomic polarizabilities (A_{pol}) and electrical moments were calculated with atomic polarizabilities and net atomic charges, respectively. The descriptors related with the molecular shape, 26-34 in the Table 1, were calculated from three dimensional structures of the minimum energy conformers optimized by the procedure mentioned above. The water/octanol partition coefficient and molecular reflective index were calculated using the method proposed by Ghose[11] in the Cerius2 program.

We used back-propagation of error artificial neural net-

Table 1. Thirty-nine molecular descriptors calculated for this study

Number	Parameters	Number	Parameters
1	Oxygen balance[a]	21	Apol[e]
2	Number of N=N bonds	22	Magnitude of dipole moments
3	Number of C=O bonds	23	X component of dipole moment
4	Number of CO_2 groups	24	Y component of dipole moment
5	Number of $C(sp^2)$-NO_2 bonds	25	Z component of dipole moment
6	Number of $C(sp^3)$-NO_2 bonds	26	Radius of gyration
7	Number of N-NO_2 bonds	27	Area
8	Number of O-NO_2 bonds	28	Molecular weight
9	Number of cyclic rings	29	Molecular volume
10	Number of NH_2 groups	30	Density
11	Number of OH groups	31	Magnitude of PMI[f]
12	Number of $C(NO_2)_3$ groups	32	X component of PMI
13	Number of C atoms	33	Y component of PMI
14	Number of H atoms	34	Z component of PMI
15	Number of O atoms	35	Number of rotatable bond
16	Number of N atoms	36	Numbers of HB[g] acceptors
17	LUMO_MOPAC[b]	37	Numbers of HB donors
18	Dipole_MOPAC	38	Water/octanol partition coefficient
19	HF_MOPAC[c]	39	Molar reflective index
20	HOMO_MOPAC[d]		

[a]Oxygen balance = $100(2n_O - n_H - 2n_C - 2n_{COO})$ / molecular weight (where n_O, n_H, and n_C represent the number of O, H, and C atoms, respectively, and n_{COO} represents the number of carboxyl groups) [b]Lowest unoccupied molecular orbital energy calculated in MOPAC (AM1 hamiltonian). [c]Heat of formation calculated in MOPAC (AM1 hamiltonian). [d]Highest occupied molecular orbital calculated in MOPAC (AM1 hamiltonian). [e]A_{pol}: sum of atomic polarizabilities. [f]PMI: principal moment of inertia. [g]HB: hydrogen bond.

works (BPE-ANN) for the prediction of $H_{50\%}$.[12] The validation set consisted of 22 molecules, and the test set utilized 11 molecules. Molecules in both the validation and test sets were selected evenly over different molecular types. The rest of the molecules were employed for training NN architecture. Only 1 hidden layer was considered. In the training of NN, total 10,000 iterations were carried out with validating the training process per 100 iterations. Seven different sets of molecular descriptors were constructed by considering the types of molecular descriptors. These sets were summarized in Table 2. Initially, we employed three subsets of molecular descriptors, *i.e.* **S1**, **S2**, and **S3** sets. The set, **S1**, was composed of compositional and topological descriptors, and the **S2** set was composed of electronic descriptors solely. In **S3** set, a subset of molecular descriptors was constructed by choosing some of descriptors from various types. We also constructed the **S4** set by choosing 10 descriptors among those used by Nefati *et al.*[8] We were unable to employ all the descriptors used by Nafati *et al*. because the molecular descriptors pool is not the same as the one computed by Nefati *et al*. In these sets, we strove to make the number of molecular descriptors to be close to 10. We also constructed three more sets, *i.e*. **S5**, **S6**, and **S7**, by extending the numbers of molecular descriptors from **S1**, **S2**, and **S3** set, respectively.

Results and Discussion

Logarithmic impact height of 50 % probability, $\log(H_{50\%})$, was used in a regression because we believed that $\log(H_{50\%})$

Table 2. Sets of molecular descriptors used this study

Identification	Description of the set
S1	Collection of compositional and topological descriptors (11 descriptors: 2, 3, 5, 6, 7, 8, 9, 13, 14, 15, 16)[a]
S2	Collection of electronic descriptors (13 descriptors: 17, 18, 19, 20, 21, 22, 27, 28, 30, 36, 37, 38, 39)
S3	Collection of various types of descriptors (12 descriptors: 1, 5, 6, 7, 8, 14, 15, 16, 18, 22, 30, 39)
S4	Sets of descriptors similar to the one used by Nefati *et al.* (10 descriptors: 1, 2, 5, 6, 7, 8, 14, 15, 16, 28).
S5	More compositional and topological descriptors were added from **S1** (17 descriptors: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 35).
S6	More electronic descriptors were added from **S2** (20 descriptors: 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 31, 32, 33, 34, 36, 37, 38, 39).
S7	More descriptors were added from **S3** (19 descriptors: 1, 5, 6, 7, 8, 9, 12, 13, 14, 15, 16, 17, 28, 20, 27, 28, 31, 36, 37).

[a]Numbers in parenthses are molecular descriptors which construct the set. See Table 1 for the number and explanation for the descriptors.

represented the nature of impact sensitivity better than the raw impact height did. We calculated the correlation coefficient (r^2) and standard error of prediction (SEP) for each set.[13] These criteria were utilized in deciding which NN architecture was more appropriate for our purpose. We varied the number of neurons in the hidden layer from 1 to 5 for each set. In the NN architectures employing all 39 molecular descriptors, only 1 or 2 neurons were put in the hidden layer. By restricting the numbers of neurons in the hidden layers as described previously, we were able to keep the ratio between the number of training molecules and the number of adjustable weights to be larger than 2.1. This forced to eliminate the possibility of chance correlation.

We illustrated r^2 and SEP values of various NN architectures employed in this study in Figure 1and 2, respectively. As shown in these figures, the best result was obtained from **39(all)-2-1**,[14] which employed whole 39 molecular descriptors and 2 hidden neurons. The correlation coefficient from **39(all)-2-1** was 0.848, and SEP was calculated to be 0.175. Although the NN architecture with 39 molecular descriptors provided better results than those with any other subsets of molecular descriptors, we were reluctant to use this architecture for our future predictive model because utilizing all the 39 molecular descriptors appeared to be unrealistic. The NN architectures with **S1** and **S4** sets furnished better results than those with **S2** and **S3** sets. In these NN architectures, the results were slightly improved by increasing the number of neurons in the hidden layer, with some exceptions in **S1** and **S4** sets. The SEP value calculated by **11(S1)-5-1** architecture was 0.207, and r^2 became 0.782, while the SEP value calculated by **10(S4)-3-1** architecture was 0.211, and r^2 was 0.777. In the results with **S2** set, r^2 values were in the range from 0.610 to 0.699, and SEP values were from 0.281 to 0.245. The results with **S3** set were slightly better than those with **S2** set. We also performed NN studies by using 3 more subsets, *i.e.* **S5**, **S6**, and **S7** sets, with an increased number of molecular descriptors. As shown in Figure 1 and 2, the increase of numbers of molecular descriptors didn't always ameliorate the results. Specifically, the NN results with **S6** and **S7** sets

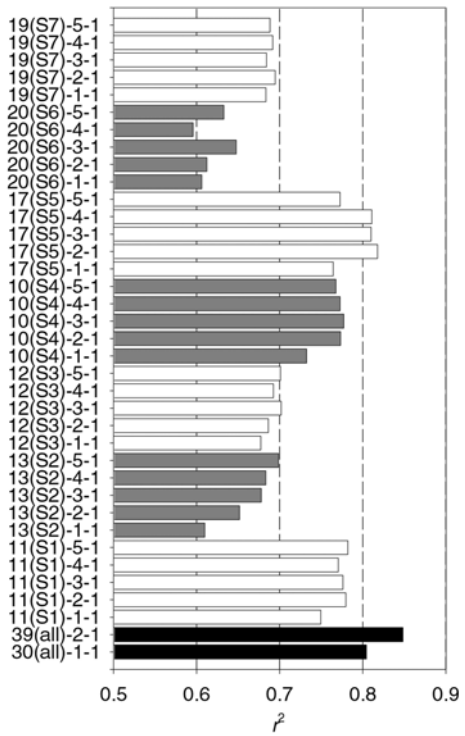

Figure 1. Comparison of r^2 calculated by various NN architectures.

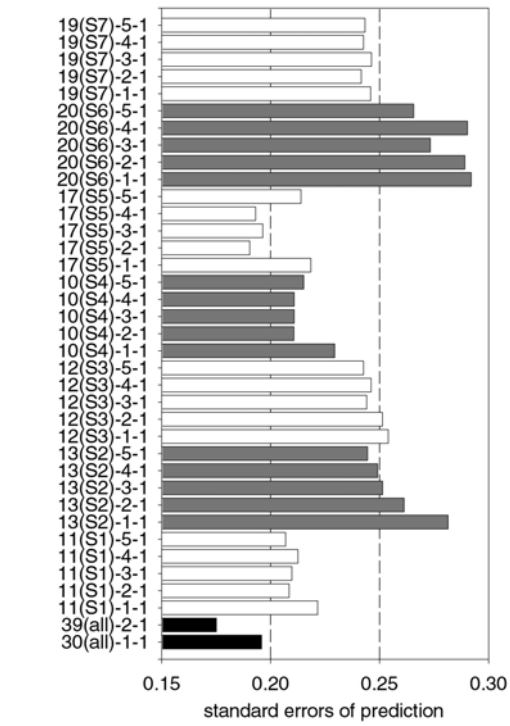

Figure 2. Comparison of SEP calculated by various NN architectures.

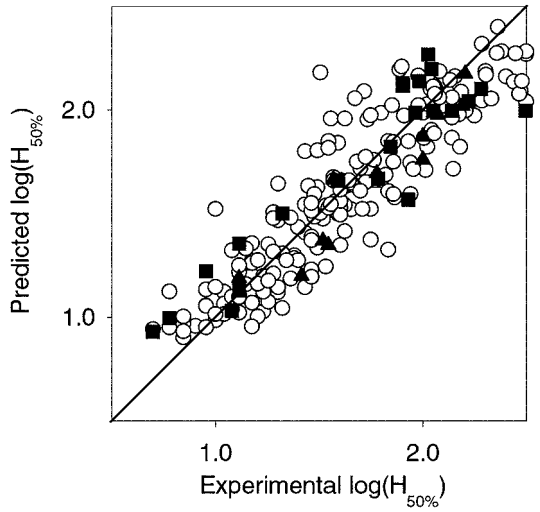

Figure 3. Plot of predicted impact sensitivity values against experimental ones. Molecules in the training, and validation, and test sets are marked as ○ , ■ , and ▲ , respectively. The solid line represents perfect agreement between experiment and prediction.

Table 3. Impact Sensitivity Values ($H_{50\%}$)[a] Predicted with NN Architecture by Us (**17(S5)-2-1**) and by Nefati *et al*. Along with Experimental Values

Serial No.	Ident. Code Given by Us[b]	Ident. Code by Nefati *et al.*	Compound Names	Our Result	Exp. Value[c]	Result by Nefati *et al.*
1	01_01	-[d]	Hexanitrobenzene	11	12	-
2	01_02	ar57	Benzotrifuroxan	54	50	27
3	01_03	-	1,2,4,5-Tetranitrobenzene	28	27	-
4	01_04	ar1	2,3,4,5,6-Pentanitroaniline	20	15	17
5	01_05	ar12	1,3,5-Trinitrobenzene	72	100	81
6	01_06	ar7	Picric acid	53	87	58
7	01_07	ar3	2,4,6-Trinitroresorcinol	41	43	44
8	01_08	ar22	2,4,6-Trinitrophloroglucinol	33	27	37
9	01_09	-	2,3,4,6-Tetranitroaniline	38	41	-
10	01_10	ar16	2,4-Dinitroresorcinol	109	296	203
11	01_12	ar14	2,4,6-Trinitroaniline	96	177	111
12	01_13	ar8	2,4,6-Trinitro-3-aminophenol	74	138	81
13	01_14	ar4	2,3,4,6-Tetranitroaniline	38	41	31
14	01_15	ar20	1,3-Diamino-2,4,6-trinitrobenzene	119	320	147
15	01_16	ar43	1-Hydroxy-3,5-diamino-2,4,6-trinitrobenzene	97	120	121
16	01_19	ar13	2,4,6-Trinitrobenzonitrile	72	140	85
17	01_20	ar10	2,4,6-Trinitrobenzoic acid	106	109	93
18	01_21	ar17	2,4,6-Trinitroanisole	102	192	147
19	01_23	ar18	1,3-Dimethoxy-2,4,6-trinitrobenzene	131	251	223
20	01_24	ar2	2',2',2'-Trinitroehtyl-2,4,6-trinitrobenzoate	23	24	23
21	01_25	ar6	2',2',2'-Trinitroethyl-3,5-dinitrobenzoate	43	73	41
22	01_26	ar5	2',2',2'-Trinitroethyl-3,5-dinitrosalicylate	37	45	35
23	01_27	ar49	1,4,5,8-Tetranitronapthalene	111	100	128
24	01_28	-	2',2'-Dinitropropyl-2,4,6-trinitrobenzoate	107	214	-
25	01_29	ar9	2,2',4,4',6,6'-Hexanitrobiphenyl	54	85	49
26	01_30	ar23	3-Hydroxy-2,2',4,4',6,6'-hexanitrobiphenyl	46	42	47
27	01_31	ar24	3,3'-Dihydroxy-2,2',4,4',6,6'-hexanitrobiphenyl	41	40	45
28	01_32	ar21	2,2',4,4',6,6'-Hexanitrodiphenylamine	65	48	66
29	01_33	ar25	3,3'-Diamino-2,2',4,4',6,6'-hexanitrobiphenyl	88	132	80
30	01_34	-	2,2',4,4',6-Pentanitrobenzophenone	129	54	-
31	01_35	ar51	2,2',2",4,4',4",6,6',6"-Nonanitro-m-terphenyl	44	39	55
32	01_36	ar52	2,2",4,4',4",6,6',6"-Octanitro-m-terphenyl	75	63	83
33	01_37	-	2,2'2",4,4",5',6,6"-Octanitro-p-terphenyl	75	40	-
34	01_38	ar54	2,2',2",4,4",6,6',6"-octanitro-p-terphenyl	75	59	82
35	01_39	-	Dodecanitroquaterphenyl	37	40	-
36	01_40	ar56	Azo-bis-2,2',4,4',6,6'-hexanitrobiphenyl	16	40	29
37	02_01	ar29	2,4,6-Trinitrobenzaldehyde	96	36	48
38	02_02	ar33	2,4,6-Trinitrobenzaldoxime	87	42	63
39	02_03	ar31	2,4,6-Trinitrotoluene	134	160	80
40	02_04	ar58	1-Dinitromethyl-3-nitrobenzene	121	105	62
41	02_05	ar32	2,4,6-Trinitrobenzylalcohol	107	52	74
42	02_06	-	2,4,6-Trinitro-m-cresol	108	191	-
43	02_07	ar26	1-(2,2,2-Trinitroethyl)-2,4,6-trinitrobenzene	16	13	17
44	02_08	-	2,4,6-Trinitrostyrene	151	32	-
45	02_09	ar28	1-(2,2,2-Trinitroethyl)-2,4-dinitrobenzene	28	31	21
46	02_10	ar38	3,5-Dimethyl-2,4,6-trinitrophenol	113	77	169
47	02_11	ar27	1-(3,3,3-Trinitropropyl)-2,4,6-trinitrobenzene	25	21	24
48	02_12	ar60	1-(3,3,3-Trinitropropyl)-2,4-dinitrobenzene	51	31	37
49	02_13	ar35	3-Methyl-2,2'4,4',6,6'-hexanitrobiphenyl	47	53	57
50	02_14	ar37	3-Methyl-2,2'4,4',6-pentanitrobiphenyl	146	143	93
51	02_15	ar61	Hexanitrostilbene	99	39	55

Table 3. Continued

Serial No.	Ident. Code Given by Us	Ident. Code by Nefati *et al.*	Compound Names	Our Result	Exp. Value	Result by Nefati *et al.*
52	02_16	-	2,2',4,4',6,6'-Hexanitrobibenzyl	128	114	-
53	02_17	ar36	3,3'-Dimethyl-2,2',4,4',6,6'-hexanitrobiphenyl	129	135	60
54	03_01	py1	2,4,6-Trinitropyridine-1-oxide	29	20	33
55	03_02	py2	3,3',5,5'-Tetranitro-2,2'-azopyridine	38	56	35
56	03_03	py3	2,6-Bis-(picrylazo)-3,5-dinitropyridine	12	33	35
57	03_04	-	2,6-Bis-(picrylamino)-3,5-dinitropyridine	79	63	-
58	03_05	py5	3,5-Bis-(picrylamino)-2,6-dinitropyridine	79	92	92
59	03_06	py6	2,6-Bis(picrylamino)pyridine	169	192	236
60	04_01	-	2,4,5-Trinitroimidazole	18	68	-
61	04_02	im2	2,7-Dinitroimidazole	46	105	69
62	04_05	im5	4,4',5,5'-Tetranitrobiimidazole	33	37	40
63	04_07	im7	2,4-Dinitro-1-picrylimidazole	44	46	46
64	04_08	-	2-Nitro-1-picrylimidazole	92	312	-
65	04_09	-	4-Nitro-1-picrylimidazole	92	161	-
66	04_10	im10	1-Picrylimidazole	164	314	264
67	05_02	pr2	4-Nitro-1-picrylpyrazole	92	112	102
68	05_04	pr4	3,5-Dinitro-1-methyl-4-picrylpyrazole	78	118	108
69	05_05	-	3,5-Dinitro-1-methyl-4-picrylaminopyrazole	92	274	-
70	05_06	-	1,4-Dipicrylpyrazole	115	314	-
71	05_07	-	4-Nitro-1-picryl-3-picrylaminopyrazole	80	149	-
72	06_01	-	3-Amino-4-nitrofurazan	52	27	-
73	06_02	-	4,4'-Dinitro-3,3'-bifurazan	22	13	-
74	06_03	fu3	3-Nitro-4-picrylaminofurazan	48	60	60
75	06_04	fu4	3-Amino-4-picrylaminofurazan	119	120	192
76	06_05	-	2,5-Dipicryl-1,3,4-oxadiazole	64	20	-
77	06_06	-	3,4-Bis-(picrylamino)-furazan	89	71	-
78	06_07	fu7	3,5-Bis-(picrylamino)-1,2,4-oxadiazole	89	95	113
79	07_02	tr2	3-Nitro-1,2,4-triazole-5-one	77	291	175
80	07_05	tr5	4-Methyl-3,5-dinitro-1,2,4-triazole	64	155	121
81	07_06	tr6	5,5'-Dinitro-3,3'-bi-1,2,4-triazole	73	153	115
82	07_07	tr7	4-(2-Nitroethyl)-3,5-dinitro-1,2,4-triazole	48	35	60
83	07_08	tr22	4-Nitro-2-picryl-1,2,3-triazole	63	68	69
84	07_11	tr11	4-Picrylamino-1,2,4-triazole	144	314	218
85	07_12	tr12	3-Amino-5-picrylamino-1,2,4-triazole	162	230	256
86	07_13	tr13	4-(2,4-Dinitrobenzyl)-3,5-dinitro-1,2,4-triazole	107	96	138
87	07_16	tr16	3,5-Bispicrylamino-1,2,4-triazole	120	240	127
88	07_17	tr17	N,N'-Dipicryl-5,5'-dinitro-3,3'-bi-1,2,4-triazole	50	138	62
89	08_01	tr19	4-Nitro-1,2,3-triazole	34	25	25
90	08_03	tr21	4-Nitro-1-picryl-1,2,3-triazole	23	9	27
91	08_05	-	1-(3',5'-Dinitrophenyl)-4-nitro-1,2,3-triazole	54	56	-
92	08_06	-	1-(3',4'-Dinitrophenyl)-4-nitro-1,2,3-triazole	54	51	-
93	08_07	-	1-Picryl-1,2,3,-triazole	54	10	-
94	08_08	tr26	2-Picryl-1,2,3-triazole	130	200	184
95	08_09	-	1-(3'-Amino-2',4',6'-trinitrophenyl)-1,2,3-triazole	70	31	-
96	08_10	-	4-Picrylamino-1,2,3-triazole	65	103	-
97	08_11	-	4,6-Dinitro-1-picrylbenzotriazole	34	40	-
98	08_12	tr29	5,6-Dinitro-1-picrylbenzotriazole	34	35	27
99	08_13	tr30	1-Picryl-4-picrylamino-1,2,3-triazole	39	35	36
100	08_14	-	2,6-Dipicrylbenzo-[1,2-d: 4,5-d']-bistriazole-4,8-dione	162	95	-
101	08_15	-	1,7-dipicrylbenzo-[1,2-d: 4,5-d']-bistriazole	26	38	-
102	08_16	-	1,5-Dipicrylbenzo-[1,2-d: 4,5-d']-bistriazole	26	40	-
103	09_01	-	2,4,6,2',2",4",6"-Heptanitro-4',6'-diaza-m-terphenyl	76	58	-

Table 3. Continued

Serial No.	Ident. Code Given by Us	Ident. Code by Nefati *et al.*	Compound Names	Our Result	Exp. Value	Result by Nefati *et al.*
104	09_02	-	2,4,6,4',2",4",6"-Heptanitro-2',6'-diaza-m-terphenyl	76	58	-
105	09_03	-	1-Picryl-2-picrylamino-1,2-dihydropyrimidine	150	106	-
106	09_04	-	5-Nitro-2,4,6-tris-(picrylamino)-pyrimidine	92	201	-
107	10_01	al47	1,1,1,3-Tetranitrobutane	21	33	23
108	10_02	al49	1,1,1,3,5,5,5-Heptanitropentane	10	8	9
109	10_03	al79	1,1,1,6,6,6-Hexanitro-3-hexyne	11	7	9
110	10_04	al80	1,1,1,6,6,6-Hexanitro-3-hexene	12	17	14
111	10_05	al73	3,3,4,4-Tetranitrohexane	112	80	139
112	10_06	al75	2,2,4,4,6,6-Hexanitroheptane	38	29	31
113	10_07	al74	2,2,4,6,6-Pentanitroheptane	86	56	79
114	11_01	al45	2,2,2-Trinitroethylcarbamate	25	18	22
115	11_02	al81	2,2-Dinitro-1,3-propanediol	57	110	90
116	11_03	al46	Methyl-2,2,2-trinitroehtyl carbonate	28	28	33
117	11_04	al82	4,4,4-Trinitrobutyramide	50	40	60
118	11_05	al50	Bis-(2,2,2-trinitroethyl)-carbonate	11	16	9
119	11_06	-	Methylene-bis-N,N'-(2,2,2-trinitroacetamide)	12	9	-
120	11_07	al83	Bis-(trinitroethoxy)-methane	13	17	10
121	11_08	al51	N,N'-Bis-(2,2,2-trinitroehtyl)-urea	12	17	13
122	11_09	al52	5,5,5-Trinitropentanone-2	81	125	117
123	11_10	al54	Ethyl-2,2,2-trinitro-ethylcarbonate	62	81	97
124	11_11	al84	N-(2-propyl)-trinitroacetamide	100	112	169
125	11_12	al85	Bis-(trinitroethyl)-oxalate	15	15	10
126	11_13	al53	2,2,2-Trinitroethyl-4,4,4-trinitrobutyrate	14	18	14
127	11_14	al86	Bis-(trinitroethyl)oxamide	15	13	15
128	11_15	al87	Trinitroethyl-2,2-dinitropropylcarbonate	21	15	23
129	11_16	al88	N-Trinitroethyl-4,4,4-trinitrobutyramide	15	18	16
130	11_17	al89	1,5-Bis-(Trinitroethyl)biuret	17	24	20
131	11_18	-	N-(t-Butyl)-trinitroacetamide	127	110	-
132	11_19	al55	Tris-(2,2,2-trinitroethyl)orthoformate	9	7	9
133	11_20	al91	1,1,1,7,7,7-Hexanitro-4-heptanone	18	34	23
134	11_21	-	Methylene-bis-(trinitroethyl)carbamate	15	27	-
135	11_22	-	2,2-Dinitropropyltrinitrobutyrate	52	151	-
136	11_23	-	2,2,2-Trinitroethyl-4,4-dinitrovalerate	52	70	-
137	11_24	-	Bis-(2,2-dinitropropyl)carbonate	117	300	-
138	11_25	al66	2,2-Dinitropropyl-4,4,4-trinitrobutyramide	59	72	61
139	11_26	-	Bis-(trinitropropyl)urea	25	23	-
140	11_27	al94	Bis-(1,1,1-trinitro-2-propyl)urea	25	19	29
141	11_28	al95	Bis-(trinitroethyl)fumarate	21	14	18
142	11_29	al96	Trinitroethyl-bis-(trinitroethoxy)-acetate	10	6	10
143	11_30	al58	4,4,4-Trinitrobutyricanhydride	26	30	30
144	11_31	al60	Bis-(2,2,2-trinitroethyl)-succinate	29	30	35
145	11_32	-	Bis-(2,2-dinitropropyl)oxalate	166	227	-
146	11_33	al59	N,N'-Bis-(3,3,3-trinitropropyl)-oxamide	35	45	46
147	11_34	al68	2,2,2-Trinitroethyl-4,4-dinitrohexanoate	90	138	102
148	11_35	-	2,2-Dinitrobutyl-4,4,4-trinitrobutyramide	68	101	-
149	11_37	al57	Nitroisobutyl-4,4,4-trinitrobutyrate	160	279	235
150	11_38	al61	Tetrakis-(2,2,2-trinitroethyl)-orthocarbonate	9	7	8
151	11_39	al62	Methylene-bis-(4,4,4-trinitrobutyramide)	60	113	75
152	11_40	al63	Ethylene-bis-(4,4,4-trinitrobutyrate)	96	120	76
153	11_41	al69	N,N-Bis-(2,2-dinitropropyl)-4,4,4-trinitrobutyramide	48	72	72
154	11_42	-	Bis-(2,2,2-trinitroethyl)-4,4-dinitroheptanedioate	36	68	-
155	11_43	al70	2,2-Dinitropropane-1,3,1,3-diol-bis-(4,4,4-trinitrobutyrate)	36	50	44

Table 3. Continued

Serial No.	Ident. Code Given by Us	Ident. Code by Nefati *et al.*	Compound Names	Our Result	Exp. Value	Result by Nefati *et al.*
156	11_44	al72	Bis-(2,2,2-Trinitroethyl)-4,4,6,6,8,8-hexanitro-undecanedioate	42	32	44
157	12_01	al1	N,N'-Dinitromethanediamine	15	13	15
158	12_02	al97	N-Nitro-N-methylformamide	95	316	241
159	12_03	al2	N,N'-Dinitro-1,2-ethanediamine	37	34	40
160	12_04	al16	Methyl-2,2,2-trinitroethylnitramine	12	9	10
161	12_05	al98	Trinitroethylnitroguanidine	13	15	13
162	12_06	al3	Cyclotrimethylenetrinitramine	17	26	19
163	12_07	al4	N-Methyl-N,N'-dinitro-1,2-ethanediamine	98	114	113
164	12_08	al99	Trinitroethylcyanomethylnitramine	13	11	9
165	12_09	al17	Bis-(2,2,2-trinitroethyl)-nitramine	9	5	7
166	12_10	al100	N-Methyl-N-nitro(trinitroethyl)-carbamate	16	17	12
167	12_11	-	N,N'-Dimethyl-N,N'-dinitrooxamide	155	79	-
168	12_12	al102	N-Nitro-N-(trinitroethyl)-glycinamide	20	17	15
169	12_13	al5	Cyclotetramethylenetetranitramine	14	29	20
170	12_14	al6	N,N'-Dinitro-N-[2-(nitroamino)ethyl]-1,2-ethanediamine	46	39	41
171	12_15	al10	1,3,3,5,5-Pentanitropiperidine	17	14	13
172	12_16	al18	2,2,2-Trinitroethyl-3,3,3-trinitropropylnitramine	10	6	8
173	12_17	al19	N,N'-Bis-2,2,2-trinitroethyl-N,N'-dinitromethanediamine	9	5	7
174	12_18	al103	Trinitroehtyl-N-ethyl-N-nitro-carbamate	27	19	23
175	12_19	al104	Trinitroethyl-2-methoxyethylnitrate	26	42	28
176	12_20	al20	N-methyl-N'-trinitroethyl-N,N'-dinitro-1,2-ethanediamine	15	11	18
177	12_21	al9	N,N'-3,3-Tetranitro-1,5-pentanediamine	39	35	33
178	12_22	al35	N-Nitro-N-(3,3,3-trinitropropyl)-2,2,2-trinitroethyl carbamate	15	9	9
179	12_23	al21	2,2,2-Trinitroethyll-N-(2,2,2-trinitroethyl)nitramino acetate	15	9	9
180	12_24	al22	2,2,2-Trinitroethyl-4-nitrazavalerate	53	35	42
181	12_25	al105	Trinitropropyl-(2,2,2-dinitropropyl)-nitramine	15	17	15
182	12_26	al36	2',2',2'-Trinitroethyl-2,5-dinitrazahexanoate	20	15	21
183	12_27	-	2,2,2-Trinitroethyl-3,3-dinitrobutyl nitramine	15	20	-
184	12_28	al11	N-(2,2-Dinitropropyl)-N,2,2-trinitro-1-propanamine	38	29	26
185	12_29	al106	1,7-Dimethoxy-2,4,6-trinitrazaheptane	87	166	90
186	12_30	al7	N,N'-Dinitro-N,N'-bis-[2-(nitroamino)ethyl]-1,2-ethanediamine	47	53	61
187	12_31	al107	Bis-(trinitroethyl)-2,4-dinitrazapentanedioate	10	10	10
188	12_32	-	2,2-Dinitropropyl-5,5,5-trinitro-2-nitrazapanetanoate	21	16	-
189	12_33	al109	Trinitroethyl-5,5-dinitro-3-nitrazahexanoate	21	25	18
190	12_34	-	2,2,2-Trinitroethyl-2,5,5-trinitro-2-azahexanoate	21	22	-
191	12_35	al110	N-Nitro-N,N'-bis-(trinitropropyl)-urea	13	21	12
192	12_36	al38	2,2,2-Trinitroethyl-2,4,6,6-tetranitro-2,4-diazaheptanoate	13	18	12
193	12_37	al24	Bis-(2,2,2-trinitroethyl)-3-nitrazaglutarate	15	14	15
194	12_38	al39	N,N'-Dinitro-N,N'-bis-(3,3,3-trinitropropyl)-oxamide	11	9	11
195	12_39	al111	Bis-(trinitroethyl)-2,4,6-trinitrazaheptanedioate	10	13	10
196	12_40	al40	2,2,6,9,9-Pentanitro-4,oxa-5-oxo-6-azadecane	98	47	55
197	12_41	al25	1,1,1,3,6,9,11,11,11-nonanitro-3,6,9-triazaundecane	10	12	13
198	12_42	-	N-(2,2-Dinitrobutyl)-N-2,2-trinitro-1-butanamine	114	80	-
199	12_43	al44	N,N'-Dinitro-N,N'-bis-(3-nitrazabutyl)-oxamide	85	90	67
200	12_44	al13	2,2,4,7,9,9-hexanitro-4,7-diazadecane	42	72	34
201	12_45	al113	N,N'-Dinitromethylene-bis-(4,4,4-trinitro)butyramide	19	13	14
202	12_46	al41	1,1,1,5,7,10,14,14,14-Nonanitro-3,12-dioxa-4,11-dioxo-5,7,10-triazatetradecane	10	11	11
203	12_47	al112	Bis-(5,5,5-trinitro-3-nitrazapentanoyl)methylenedinitramine	9	15	10
204	12_48	al27	1,1,1,4,6,6,8,11,11,11-Decanitro-4,8-diazaundecane	10	11	11
205	12_49	al26	1,1,1,3,6,6,9,11,11,11-Decanitro-3,9-diazaundecane	10	10	11
206	12_50	al28	Bis-(2,2,2-trinitroethyl)-4-nitraza-1,7-heptanedioate	32	29	32

Table 3. Continued

Serial No.	Ident. Code Given by Us	Ident. Code by Nefati *et al.*	Compound Names	Our Result	Exp. Value	Result by Nefati *et al.*
207	12_51	al30	Bis-(2,2,2-trinitroethyl)-3,6-dinitraza-1,8-octanedioate	17	29	20
208	12_52	al114	Bis-(trinitroethyl)-2,5,8-trinitrazanonanedioate	12	17	16
209	12_53	al42	N,N'-Dinitro-N,N'-bis-(3,3-dinitrobutyl)oxamide-triazatetradecane	82	37	44
210	12_54	al29	1,1,1,3,6,9,12,14,14,14-Decanitro-3,6,9,12-tetrazatetradecane	10	19	18
211	12_55	al115	Bis-(trinitroethyl)-5,5-dinitro-2,8-dinitrazanonanedioate	12	12	14
212	12_56	-	2,2,4,7,7,10,12,12-Octanitro-4,10-diazatridecane	44	44	-
213	12_57	al15	2,2,5,7,7,9,12,12-Ocatanitro-5,9-diazatridecane	44	37	37
214	12_58	al116	1,4-Bis-(5,5,5-trinitro-2-nitrazapentanoate)-2-butyne	19	16	17
215	12_59	al43	1,1,1,18,18,18-hexanitro-3,16-dioxa-4,15-dioxo-5,8,11,14-tetranitraoctadecane	12	19	23
216	12_60	-	1,1,1,3,6,6,8,10,10,13,15,15,15-tridecanitro-3,8,13-triazapentadecane	9	23	-
217	12_61	-	2,2-Dinitropropanediol-bis-(5,5-dinitro-2-nitraza-hexanoate)	92	138	-
218	13_01	al119	1,2,3-Propanetrioltrinitrate	13	20	10
219	13_02	al32	N-(2,2,2-Trinitroethyl)-nitraminoethyl nitrate	11	7	8
220	13_03	-	2,2-Bis(nitroxymethyl)-1,3-propanediol dinitrate	15	13	-
221	13_04	al33	3-[N-(2,2,2-Trinitroethyl) nitramino]-propylnitrate	19	12	13
222	13_05	al34	3,5,5-Trinitro-3-azahexyl nitrate	32	21	26
223	13_06	al120	1,9-Dinitrato-2,4,6,8-tetranitrazanonane	10	10	11
224	13_07	-	4,4,8,8-Tetranitro-1,11-dinitrato-6-nitrazaundecane	36	87	-
225	14_03	-	3,5-Dinitroisoxazole	22	29	-
226	14_04	mi4	1,4-Dinitroglycoluril	51	100	79
227	14_05	mi5	1-Nitro-2,5-bis(trinitromethyl)-pyrrolidine	15	6	9
228	14_06	mi6	5-Picrylaminotetrazole	41	36	33
229	14_07	-	3,3,9,9-Tetranitro-1,5,7,11-Tetraoxaospiro-(5.5)-undecane	39	66	
230	14_08	al64	N-(2,2,2-Trinitroethyl)-3,3,5,5-tetranitropiperidine	16	18	17
231	14_09	mi7	N-(Trinitropropyl)-3,3,5,5-tetranitropiperidine	23	29	24
232	14_11	-	3,6-Bis-(picry1amino)-s-tetrazine	38	61	-
233	14_12	-	1,3,5,-Tris(1-oxo-5,5,5)trinitro-3-nitrazapentyl)-s-triazacyclohexane	13	13	-
234	14_13	-	Tripicryl-s-triazene	77	85	-
235	-[e]	al2[f]	Bis(2,2-dinitrobutyl)nitramine	-	79[g]	78
236	-	al56	1,1,1,7,7,7-Hexanitro-2-heptanone	-	34[g]	22
237	-	al76	Bis(2,2-dinitropropyl)oxalate	-	229[g]	151
238	-	al8	Tetryl	-	32[g]	36
239	-	ar102	5-(2-Nitroguanidino)trinitro-1,3-benzenediamine	-	47[g]	67
240	-	ar103	1,2-Bis(3,5-diaminotrinitroanilino)ethane	-	186[g]	92
241	-	ar104	5-Aminostyphnic acid	-	33[g]	69
242	-	ar105	5-Formamidostyphnic acid	-	74[g]	80
243	-	ar107	Diamino-2,4,6-trinitrobenzonitrile	-	182[g]	157
244	-	ar108	Dicyanopicramide	-	93[g]	141
245	-	ar109	3,5-Diaminotrinitrobenzamide	-	219[g]	193
246	-	ar110	5-Aminotrinitroisophalamide	-	170[g]	207
247	-	ar112	Pentanitrobenzene	-	17[g]	13
248	-	ar42	Methyl-2,4,6-trinitrobenzoate	-	89[g]	212
249	-	ar44	2,4,6-Trinitrophenetole	-	191[g]	274
250	-	ar48	Ammonium picrate	-	135	146
251	-	ar62	1,3,5-Triazido-2,4,6-trinitrobenzene	-	37[g]	54
252	-	ar64	Bis(2,4,6-trinitrophenyl)diazine	-	37[g]	24
253	-	ar65	7-Amino-4,6-dinitrobenzofurazan	-	60[g]	64
254	-	ar66	4,6-Dinitrobenzofuroxan	-	30[g]	50
255	-	im3	Ammonium-2,4,5-trinitroimidazole	-	50	42
256	-	im6	Diammonium-4,4',5,5'-tetranitrobiimidazolate	-	105	142
257	-	mi1	1-Ammonium-5-nitrotetrazole	-	30[g]	33

Table 3. Continued

Serial No.	Ident. Code Given by Us	Ident. Code by Nefati *et al.*	Compound Names	Our Result	Exp. Value	Result by Nefati *et al.*
258	-	mi2	Ethylenediammonium-di-5-nitrotetrazole	-	42[g]	15
259	-	mi3	3,5-Dinitroglycoluril	-	29[g]	24
260	-	mi8	5-(3,5-Diaminotrinitroanilino)tetrazole	-	31[g]	48
261	-	pr1	Ammonium 3,5-dinitropyrazole	-	158	224
262	-	tr4	Ammonium 3,5-dinitro-1,2,4-triazolate	-	110	106
263	-	tr8	3-Nitro-1-picryl-1,2,4-triazole	-	67	69

[a]Unit in cm. [b]Our identification of explosive molecules followed the classification of SSK database. For details, see the tables of ref. 7(a). [c]Except when noted, experimental values were taken from ref. 7. [d]The compounds denoted by hyphen in this column were excluded in the NN studies by Nefati *et al.* [e]The compounds denoted by hyphen in this column were excluded in our NN studies. [f]Both this molecule and N,N'-dinitro-1,2-ethanediamine (serial no.: 159) were assigned as 'al2' by Nefati. [g]These explosive molecules were not found in SSK database.

went worse than the results with **S2** and **S3** sets. On the other hand, **S5** set which was expanded from **S1** set improved the result significantly, when compared with the one with **S1** set. The best result was obtained from **17(S5)-2-1** architecture, and was presented in Table 3 along with experimental values. This result was also depicted in Figure 3. The SEP and r^2 values were 0.190 and 0.818, respectively. We also checked whether this architecture depended upon the selection of different molecules in validation and test sets. Nine more predictions were performed by switching molecules in validation and test sets. These predictions provided SEP and r^2 values which were quite similar to the original prediction.

Since our study was initiated by following the footsteps of Nefati *et al.*,[8] it would be worthwhile to mention the NN architecture and results obtained by Nefati *et al.* They employed total 204 explosive molecules among which 181 molecules were taken from the database of SSK, which is the same database we utilized. Nefati *et al.* appeared to take a significant number of explosive molecules from other databases. As it was mentioned previously, the impact sensitivity test results were sometimes substantially different due to testing environment as well as the instrument. Thus, mingling impact sensitivity values from various sources without careful adjustments may hamper the outcome of total NN analyses.

Nefati *et al.* computed 39 molecular descriptors for each molecule. Among these descriptors, two-thirds were compositional and topological descriptors. They employed 18 topological, 6 compositional, 2 indicator, and 13 electronic descriptors. Not only Nefati et. al. employed a large number of topological descriptors, but also electronic parameters were greatly biased toward electrostatic potential distribution of the specific bonds. Eight out of 13 electronic descriptors were derived from X-NO_2 bonds. Although we also employed the same number of molecular descriptors as Nefati *et al.* did, we believed that our selection of molecular descriptors was more balanced in nature. Our molecular descriptor set was composed of 6 compositional, 14 topological, and 19 electronic descriptors. We reduced the number of topological parameters by eliminating some types of bonds which we felt were less influential in determining the impact sensitivity. In addition, we attempted to utilize a variety of different types of electronic parameters most of which were quite often utilized in many QSPR scientists.[15]

Due to a considerable difference in the explosive molecules utilized in optimizing the NN architecture, it may be difficult to compare our results with those of Nefati *et al.* However, one set of molecular descriptors we utilized, namely **S3**, was a subset of molecular descriptors Nefati *et al.* used. Our best architecture with **S3**, *i.e.* **10(S3)-3-1** architecture, provided 0.211 of SEP, and 0.777 of r^2, respectively. This result was poorer than that obtained by Nefati *et al*. Again, as mentioned previously, one should not compare this result directly with that of Nefati *et al*. because the sets of molecules and descriptors were different. We knew that the absolute values of correlation coefficients and SEP might change slightly by either adding or removing some of explosive molecules. One of the important findings through our studies was the trend of predictive nature with certain types of molecular descriptors. When we applied 7 different subsets of molecular descriptors, our results showed that the subsets composed of compositional and topological descriptors provided better results than those composed of electronic descriptors. This trend was clearly deduced from our study, *i.e.* that NN results by using **S1**, **S4**, or **S5** subsets were much better than those by using **S2** or **S6** subsets. Similar trend was also observed by Nefati *et al*., although their investigation in electronic descriptors was quite limited. We believe that this trend is an important factor in applying NN study in predicting impact sensitivity of explosive molecules because calculating compositional and topological descriptors are much easier than calculating electronic descriptors. This **17(S5)-2-1** NN architecture is currently applied to predict the impact sensitivity of various novel energetic molecules in the Agency for Defense Development (ADD).

Concluding Remarks

We performed NN studies to predict impact sensitivity of energetic molecules, and attempted to derive an optimum NN architecture for our future prediction. From the database of SSK, 234 molecules were selected, and 39 molecular

descriptors of these molecules were computed. We constructed 7 subsets of molecular descriptors by combining about 10 to 20 descriptors based on the types of descriptors. Our results showed that the subsets composed of compositional and topological descriptors provide better results than those composed of electronic descriptors. The best result was obtained by employing **17(S5)-2-1** architecture. Since the molecular descriptors employed in **17(S5)-2-1** architecture were quite easy to compute, this architecture will be of great use in predicting the impact sensitivity of novel energetic molecules. As good predictive tools in estimating the impact sensitivity of novel energetic molecules were hardly available in the scientific community of energetic molecules, our optimized NN architecture may serve as a practical solution in guiding new energetic molecules in terms of safety.

Acknowledgement. K.T.N thanks for research fund supported by the Basic Research Program of ADD, Korea.

References

1. Perssen, P.-A.; Holmberg, R.; Lee, J. *Rock Blasting and Explosives Engineering*; CRC Press: Baca Raton, FL, 1993.
2. (a) Koehler, J.; Meyer, R. *Explosives*, 4th ed.; VCH: Weinheim, Germany, 1993; (b) Coffey, C. S.; DeVost, V. F. *Propel. Explos. Pyrotech.* **1995**, *20*, 105.
3. Brill, T. B.; James, K. J. *Chem. Rev.* **1993**, 2667.
4. (a) Zeman, S.; Krupka, M. *Propel. Explos. Pyrotech.* **2003**, *28*, 301; (b) Zeman, S. *Propel. Explos. Pyrotech.* **2003**, *28*, 308; (c) Zeman, S.; Krupka, M. *Propel. Explos. Pyrotech.* **2003**, *28*, 249; (d) Zeman, S. *Propel. Explos. Pyrotech.* **2000**, *25*, 66.
5. (a) Politzer, P.; Murray, J. S.; Lane, P.; Sjoberg, P.; Adolph, H. G. *Chem. Phys. Lett.* **1991**, *181*, 78; (b) Politzer, P.; Murray, J. S. *Mol. Phys.* **1995**, *86*, 251.
6. Rice, B.; Hare, J. J. *J. Phys. Chem. A* **2002**, *106*, 1770.
7. (a) Storm, C. B.; Stine, J. R.; Kramer, J. F. *Sensitivity Relationship in Energetic Materials*; Los Alamos National Laboratory, LA-UR-89-2936, 1989; (b) Storm, C. B.; Stine, J. R.; Kramer, J. F. In *Chemistry and Physics of Energetic Materials*; Bulusu, S. N., Ed.; Kluwer Academic Press: Dordrecht, Netherlands, 1990; p 605.
8. Nefati, H.; Cense, J.-M.; Legendre, J.-J. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 804.
9. *Cerius Modeling Environment*, Relaese 4.0; Accelrys Inc.: San Diego, 1999.
10. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
11. Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. *J. Phys. Chem. A* **1998**, *102*, 3762.
12. (a) Zupan, J.; Gasteiger, J. *Neural Networks in Chemistry and Drug Design*, 2nd ed.; Wiley-VCH: Weinheim, Germany, 1999; (b) *Neural Networks in QSAR and Drug Design*; Devillers, J., Ed.; Academic Press: London, 1996.
13. SEP is also called as SDEP (standard deviation error of prediction). For a detailed explanation, see (a) Wold, S.; Johansson, E.; Cocchi, M. In *3D-QSAR in Drug Design. Theory Methods and Applications*; Kubynyi, H., Ed.; ESCOM: Leiden, Netherlands, 1993; pp 523-550; (b) Myers, R. H. *Classical and Modern Regression with Applications*, 2nd ed.; Duxbury Press: Belmont, CA, 1990.
14. NN architecture was denoted as the number of molecular descriptors in the input layer (the set of molecular descriptors) – the number of hidden neurons – the number of output.
15. (a) Cho, D. H.; Lee S. K.; Kim, B. T.; No, K. T. *Bull. Korean Chem. Soc.* **2001**, *22*, 388; (b) Karelson, M.; Lobonov, V. S. *Chem. Rev.* **1996**, *96*, 1027; (c) Lee, K. W.; Kwon, S. Y.; Hwang, S.; Lee, J.-U.; Kim, H. *Bull. Korean Chem. Soc.* **1996**, *17*, 147.