J Comput Aided Mol Des (2010) 24:49–56
DOI 10.1007/s10822-009-9313-0

Artificial neural network study on organ-targeting peptides

Eunkyoung Jung · Junhyoung Kim · Seung-Hoon Choi · Minkyoung Kim · Hokyoung Rhee · Jae-Min Shin · Kihang Choi · Sang-Kee Kang · Nam Kyung Lee · Yun-Jaie Choi · Dong Hyun Jung

Received: 6 September 2009 / Accepted: 1 December 2009 / Published online: 18 December 2009

Abstract We report a new approach to studying organ targeting of peptides on the basis of peptide sequence information. The positive control data sets consist of organ-targeting peptide sequences identified by the peroral phage-display technique for four organs, and the negative control data are prepared from random sequences. The capacity of our models to make appropriate predictions is validated by statistical indicators including sensitivity, specificity, enrichment curve, and the area under the receiver operating characteristic (ROC) curve (the ROC score). VHSE descriptor produces statistically significant training models and the models with simple neural network architectures show slightly greater predictive power than those with complex ones. The training and test set statistics indicate that our models could discriminate between organ-targeting and random sequences. We anticipate that our models will be applicable to the selection of organ-targeting peptides for generating peptide drugs or peptidomimetics.

Keywords Neural network · Organ-targeting peptide · ROC score · VHSE descriptor

Introduction

Successful drug development requires not only the optimization of pharmacological specificity and potency, but also a method for efficient drug delivery to the target site. Targeted drug delivery to specific disease sites would

E. Jung · J. Kim · S.-H. Choi · M. Kim · H. Rhee · D. H. Jung (✉)
Insilicotech Co. Ltd, A-1101 Kolontripolis, 210, Geumgok-Dong, Bundang-Gu, Seongnam-Shi 463-943, Korea
e-mail: dhjung@insilicotech.co.kr

E. Jung
e-mail: jungek@insilicotech.co.kr

J. Kim
e-mail: seva7411@nate.com

S.-H. Choi
e-mail: shchoi@insilicotech.co.kr

M. Kim
e-mail: mkkim@insilicotech.co.kr

H. Rhee
e-mail: hifiverhee@insilicotech.co.kr

J.-M. Shin
SBScience Co. Ltd, Sung-Ok BD, Sunae-Dong, Bundang-Gu, Seongnam-Shi 463-825, Korea
e-mail: sbscience@gmail.com

K. Choi
Department of Chemistry, Korea University, 1, Anam-Dong 5-Ga, Seongbuk-Gu, Seoul 136-701, Korea
e-mail: kchoi@korea.ac.kr

S.-K. Kang · N. K. Lee · Y.-J. Choi
School of Agriculture Biotechnology, Seoul National University, San 56-1, Shilim-Dong, Kwanak-Gu, Seoul 151-742, Korea
e-mail: haman@unitel.co.kr

N. K. Lee
e-mail: lnk025@snu.ac.kr

Y.-J. Choi
e-mail: cyjcow@snu.ac.kr

increase the efficacy and deliver the drugs right to the specific organ, thereby reducing the chances for side-effects in other healthy tissues and organs. One approach to targeted drug delivery is based on a unique set of marker molecules on endothelial surfaces of individual tissues or organs [1–3]. At least some of the markers specifically expressed on vascular endothelial cells are vascular receptors that mediate signals transduced by angiogenic or vasculogenic factors during tumor or normal cell growth [4–6]. Tissue-specific markers such as vascular receptors have revealed prospective molecular targets that may be used to direct therapies to specific tissues [6, 7]. Therefore, the targeted drug delivery system based on specific markers of vascular endothelial cells can be used to achieve more efficient delivery to specific tissues or organs [8–10].

In vivo biopanning with phage display peptide libraries has been used to identify targeting ligands against vascular receptors selectively expressed on endothelial cells [4, 5]. In vivo phage display makes it possible to isolate peptides that home selectively to the target site by means of a peptide binding to vascular receptors selectively expressed in specific tissues or organs within a living organism [11]. Isolation of peptides that home to selective vascular beds in vivo was first reported in 1996 [1]. Since then, there have been many studies for identification of tissues or organ homing peptides by in vivo phage display. For example, Rajotte et al. [2] reported that GFE-1 (CGFECVRQCP-ERC) peptide shows 35-fold selectivity over a control phage in lung homing. Arap et al. [8] screened the peptide SMSIARL which homes to the prostate 15 times more than a control phage. Recently, Arap et al. [12] screened a phage library in a patient to identify peptide ligands homing to specific tissues, such as the bone marrow, skin, fat, muscle or prostate. However, these experimental in vivo methods are rather labor-intensive and not easily applicable to high-throughput screening.

Rapid developments in biotechnology and peptide synthesis have made it possible to exploit the unique pharmacological activities of peptides; thousands of different peptides have been designed, synthesized and subjected to a range of screening procedures and biological assays. To analyze the vast amounts of biological data on peptides, quantitative structure–activity relationship (QSAR) models have been successfully employed.

In this article, we propose a new artificial neural network method to study organ-targeting peptides based on sequence information. A group of organ-targeting peptides were selected from a random phage-peptide library using the 'peroral phage display technique' in which a phage-peptide library is administered orally to rats and the organ-targeting phages are collected from the internal organs [13]. Using the sequence set of the selected phage-displayed peptides, we constructed an artificial neural network model to screen organ-targeting peptides using various descriptors of the physicochemical properties.

Materials and methods

Preparation of organ-targeting peptides

For selection of peptide ligands which could bind certain inside organ tissues specifically after absorption from the intestinal epithelia, in vivo phage display experiments were performed as described in our previous study [13]. A random phage peptide library with more than 1.2×10^9 independent clones (Ph.D.-C7C: New England BioLabs, USA) was orally administered into rats (Sprague–Dawley, 12-weeks-old, male) and peptide sequences were selected from the phages which were recovered from four representative inside organs: liver, lung, spleen, and kidney. Three rounds of phage display biopanning were conducted consecutively to enhance the specificity of the selected peptide ligands to organ tissues. After the third round of biopanning, a total of 847 peptide sequences were identified from randomly selected individual phage recombinants rescued from four representative inside organs, kidney, liver, lung, and spleen from which 207, 218, 218 and 204 peptide sequences were identified, respectively. Overall procedure of in vivo phage display in this study was depicted in Fig. 1.

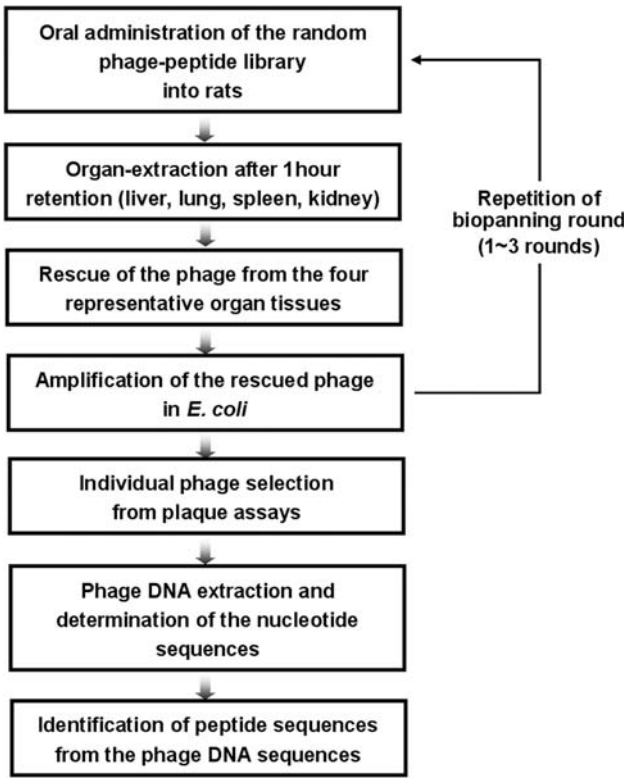

Fig. 1 Flow chart of the in vivo phage display procedure

Data sets

The positive control data set of peptides that can home to each organ was obtained from the organ-targeting hepta-peptide sequences identified by the peroral phage display experiments. The decoy negative control data set was generated from random sequences that had the same frequencies of occurrence of each amino acid residue as in the Ph.D.-C7C phage library. The decoy sequences were then compared with the positive control data and any common sequences were removed from the decoy negative control data. Also, any decoy sequences that are much similar to positive control sequences were removed because some sequences that have high sequence similarity with organ-targeting peptides will be also possible to home to the organ. About 80% of each data set was used for network training and the remaining data were used for the test set to validate the trained network.

Descriptors

Three types of amino acid descriptors, VHSE, Z3 and Z5, were used to encode important features of individual peptide sequences. The VHSE descriptor is composed of eight variables for each amino acid and is derived from the Principal Components Analysis (PCA) on independent families of 18 hydrophobic properties, 17 steric properties, and 15 electronic properties, which are included in total 50 physicochemical variables of 20 coded amino acids [14]. For a heptapeptide, $7 \times 8 = 56$ variables were used to build models based on this descriptor. The Z3 descriptor is composed of three z-scales for each amino acid [15]. The z_1 scale describes the hydrophobic properties of amino acid residues, the z_2 scale describes the steric properties, and the z_3 scale describes the electronic properties in the z-scales. The Z5 descriptor is composed of five variables as the extended z-scales and can be interpreted as lipophilicity, size/polarizability, and electronic properties of the amino acids [16].

Neural network model

We used the machine-learning method to derive structure–activity relationships. The calculations were carried out on a Pentium 2.2 GHz machine using the nnet of the VR 7.2 package [17] for feed-forward neural networks and for multinomial log-linear models. We used a three-layer neural network architecture containing a single hidden layer in which the number of neurons was increased from 0 to 2. This network consisted of a multilayer system of neurons, with each neuron in a given layer fully connected to all the neurons in the two adjacent levels. A neural network was trained to map a set of input data to a corresponding set of output data by iterative adjustment of the weights. The activation function of the hidden layer units is the logistic function and the output units are linear. The Broyden-Fletcher-Goldfarb-Shanno (BFGS) method was used as the optimization function. To help the optimization process and to avoid over-fitting, the weight decay was set at 0.001 and n-fold cross-validations were performed. The n is the number of "folds" for cross-validation. For example, setting n to 5 implies 5-fold cross-validation, in with 4/5 of the data are used to train a model and the remaining 1/5 to test the model. The process is repeated 5 times so that each sample is left out once. The maximum number of iterations for network training was 50,000 and the other parameters were given as the default values set by the nnet of the VR 7.2 package. Before the learning network was applied, the input value of the positive control was 0.9 and that of the negative control was 0.1.

Evaluation

To score the models, the receiver operating characteristics (ROC) score, which is the area under the ROC curve [18], was used for each training and test set. The score is 1 for a perfect classification and 0.5 for a random classification. All the ROC scores reported here were generated from a leave-some-out cross-validation of real and decoy set.

Validation using decoy set

We prepared a supplementary model trained with a decoy set [19] and compared that model with the model trained with a real data set for validation of the ability of our models to discriminate between organ-targeting and non-targeting peptides. The positive as well as the negative control data of the decoy set were generated from random peptides. The decoy set was prepared carefully to ensure that there was no redundant peptide in the random positive control data and no overlap between the random positive and the random negative control data.

Results

Using the peroral phage display technique, we isolated peptides independently targeting to four different mouse organs (kidney, liver, lung and spleen) from phages randomly selected from 10^5 to 10^7 clones (see the Methods section for details). These organ-targeting peptides were used as the positive control set for further analysis.

Decoy negative control data were generated as described in Data Set section and the number of peptides in decoy negative control data was same as that in the positive control set for each tissue. Table 1 shows the number of

Table 1 The number of data sets

Organs	Kidney	Liver	Lung	Spleen
Training Set	330	348	348	326
Test Set	84	88	88	82

sequences in training and test set. We utilized a feed-forward neural network for our sequence-based organ-targeting prediction. Nine models were derived for training data set by varying the type of peptide descriptor and/or the number of neurons in the (single) hidden layer. The predictive features of the resulting model are illustrated in Fig. 2, which shows that our model can distinguish effectively between kidney targeting and non-targeting peptides.

Table 2 shows the prediction accuracy of our models using VHSE, Z3 and Z5 descriptors. The ROC score was used as the primary yardstick of performance since it provides an overview of the possible cut-off levels for performance test. The table shows that VHSE descriptors tend to produce slightly better training models than the others. As for the neural network architecture, increasing the number of neurons in the hidden layer improved the ability of our models to predict the organ-targeting peptides in the training set. However, no such tendency was apparent in the statistics for the test set. This is presumably due to overtraining of the networks as mentioned in our previous study [13].

To evaluate the robustness of our models, we performed leave-5%-out cross-validation, which is analogous to leave-one-out cross-validation [20]: 5% of the sequences are kept as validation data. Table 3 shows the result of 20 rigorous tests. The standard deviation of the 20 ROC scores for test sets is relatively large for all models. Especially, the models using Z3 and Z5 descriptors show larger variations of the ROC scores. With respect of the robustness of models, the VHSE descriptor seems to be better in performance. We will discuss this result in more detail later.

To test the reliability of the peptide sequences from the phage-display experiment as the positive control set and to validate the strength of our model in predicting organ-targeting peptides, a separate decoy set was generated as

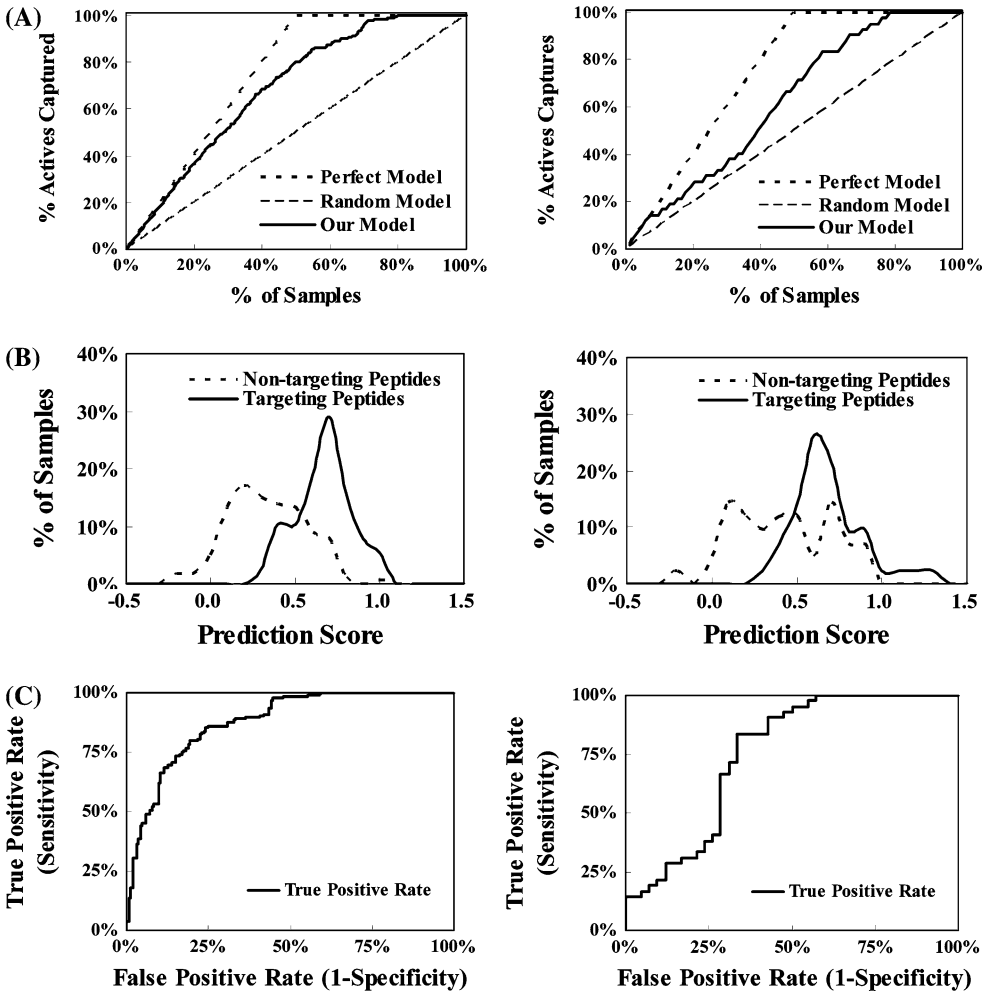

Fig. 2 Predictive features of the model. The model was constructed with zero neuron in a hidden layer and one in an output layer using VHSE descriptors. **a** Enrichment curve, **b** histogram actives vs. model values, and **c** receiver operating characteristic (ROC) curve. The features for the training and test set were plotted in the *left* and *right panels*, respectively

Table 2 Prediction accuracy for models with various network architectures

Descriptor	No. of HN[a]	Kidney		Liver		Lung		Spleen	
		Training	Test	Training	Test	Training	Test	Training	Test
VHSE	0	0.88	0.75	0.86	0.75	0.86	0.76	0.88	0.76
	1	0.96	0.70	0.97	0.67	0.98	0.73	0.99	0.55
	2	1.00	0.71	1.00	0.53	1.00	0.60	1.00	0.71
Z3	0	0.78	0.71	0.73	0.63	0.71	0.60	0.69	0.68
	1	0.86	0.71	0.83	0.61	0.83	0.62	0.82	0.56
	2	0.91	0.62	0.88	0.55	0.88	0.56	0.90	0.64
Z5	0	0.85	0.70	0.79	0.67	0.77	0.80	0.80	0.69
	1	0.93	0.70	0.91	0.59	0.90	0.70	0.91	0.67
	2	0.99	0.57	0.98	0.64	0.95	0.78	0.96	0.66

The network architecture A–B–C indicates the total number of descriptors in an input layer, where A is (7, the sequence length of a peptide) × (the number of descriptors for each amino acid), B and C are the numbers of neurons in hidden and output layers, respectively. For instance, the network architecture (7 × 8)–0–1 specifies a model constructed with zero neuron in hidden layer and one in output layer using the VHSE descriptor. All the models have one neuron in output layer

[a] The number (B) of neurons in a hidden layer

random positive control in the training and test set. We constructed a supplementary model trained with the decoy set and compared that model with the model trained with the real data set. The predictive features presented in Fig. 3 indicate that the model constructed with the decoy set do not discriminate between virtually generated organ-targeting and non-targeting peptides. Also, the validation result for the model with network architecture (7 × 8)-0-1 (see Table 4) suggest that the predictive power for the test set is considerably greater when the model is constructed with the real set than with the decoy. This result confirms that the model has found meaningful patterns in the set of organ-targeting peptides.

More detailed statistics about the predictive capacities of our models are listed in Table 5, which shows a truth table analysis of the binary outcome based on organ-targeting. The results show that our models are sensitive in predicting organ-targeting peptides rather than specific in screening out non-targeting peptides.

Discussion

We performed an artificial neural network study on organ-targeting peptides based on homing peptide sequences information obtained by in vivo screening. Although the standard deviation of ROC scores over the different test runs was slightly large, the models seem to show reasonable performance. The models trained with decoy sets did not show any prediction capability but the peptide sequences collected from the in vivo experiment served well as the positive control sets for the QSAR models.

Although we tried to optimize the network architecture and to minimize overtraining and other related problems during the course of development, our models have some weak points. First, we assumed that randomly-selected heptapeptide sequences could be used as negative controls. Because the random sequences are not the ones identified by experiments, they are not true negatives and our model might cause prediction errors. However, as mentioned in our previous study [13], heptapeptides with random sequences are very likely to be organ-non-targeting peptides because the organ-targeting peptides identified by the in vivo experiment covered only a very small portion of the entire 'heptapeptide space'.

Second, our models were focused on the prediction of organ-targeting peptides rather than organ-specific-targeting peptides because there are some common sequences found in the positive controls for each organ.

Finally, in order for a QSAR model to be used for computational screening, it must have high Positive Predictive Value (PPV). In the previous study for the prediction of intestinal permeability [13], our models produced high sensitivity values and PPVs. In this study, however, the sensitivity values are high but the PPVs are low, presumably because the number of organ-targeting peptide sequences used as positive controls is relatively small. Thus, we could produce more robust models if we were able to obtain more organ-targeting peptides and organ-non-targeting peptide from the in vivo experiments.

We tried to develop QSAR models using diverse descriptors as well. In our previous study [13], binary descriptors tended to produce slightly better training models than VHSE descriptors in prediction of intestinal permeability on the basis of peptide sequence information. In this study, however,

Table 3 The results of validation for models with zero neuron in hidden layer and one in output layer

Descriptor	Kidney		Liver		Lung		Spleen	
	Training	Test	Training	Test	Training	Test	Training	Test
VHSE	0.864 ± 0.006	0.695 ± 0.091	0.851 ± 0.004	0.669 ± 0.087	0.853 ± 0.004	0.703 ± 0.095	0.870 ± 0.004	0.701 ± 0.112
Z3	0.771 ± 0.005	0.637 ± 0.099	0.717 ± 0.007	0.607 ± 0.115	0.695 ± 0.006	0.566 ± 0.110	0.704 ± 0.009	0.574 ± 0.140
Z5	0.864 ± 0.006	0.695 ± 0.091	0.778 ± 0.007	0.638 ± 0.137	0.784 ± 0.007	0.643 ± 0.123	0.791 ± 0.007	0.640 ± 0.130

The results of rigorous test using leave-5%-out method. The results of 20 rigorous tests are averaged and expressed as mean ± SD

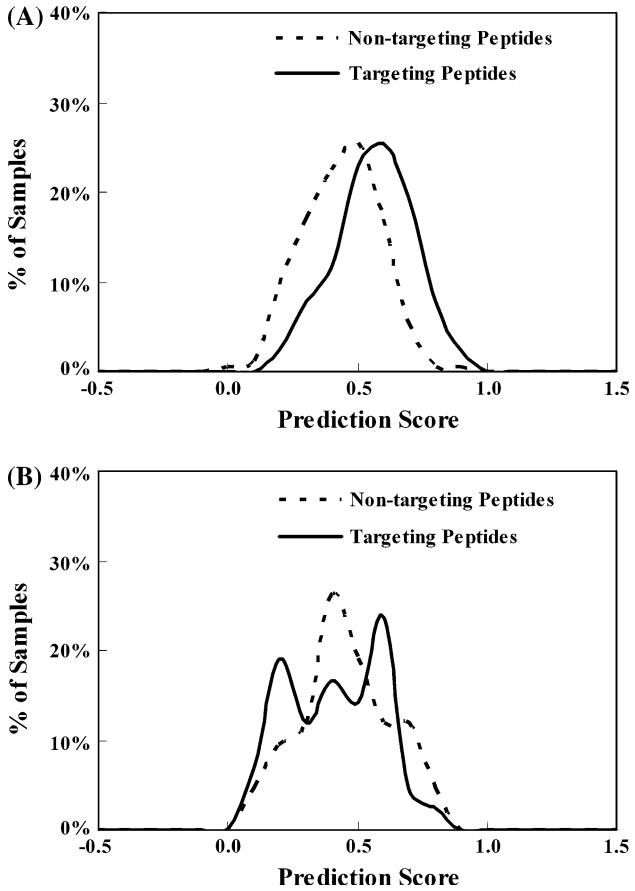

Fig. 3 The features of the model constructed with the decoy set. The models were constructed with zero neuron in a hidden layer and one in an output layer using VHSE descriptor. **a** Training set and **b** test set

Table 4 Comparison of ROC scores between real and decoy set for model without the hidden node

Organ	Descriptor	Real set		Decoy set	
		Training	Test	Training	Test
Kidney	VHSE	0.88	0.75	0.73	0.55
	Z3	0.78	0.71	0.65	0.54
	Z5	0.85	0.70	0.70	0.51
Liver	VHSE	0.86	0.75	0.75	0.54
	Z3	0.73	0.63	0.64	0.52
	Z5	0.79	0.67	0.68	0.50
Lung	VHSE	0.86	0.76	0.77	0.55
	Z3	0.71	0.60	0.63	0.54
	Z5	0.77	0.80	0.71	0.50
Spleen	VHSE	0.88	0.76	0.73	0.52
	Z3	0.69	0.68	0.66	0.56
	Z5	0.80	0.69	0.70	0.52

because the number of sequences in the data sets is relatively small and risk of overtraining a neural network exists, binary descriptors did not produce better models and therefore the

Table 5 Comparison of truth table statistics for the test sets for various models

Organ	Descriptor	SE[a]	SP[b]	PPV[c]	NPV[d]	Acc[e]
Kidney	VHSE	83	57	66	77	70
	Z3	74	62	66	70	68
	Z5	67	60	62	64	63
Liver	VHSE	82	52	63	74	67
	Z3	64	59	61	62	61
	Z5	66	45	55	57	56
Lung	VHSE	73	47	64	57	62
	Z3	55	59	57	57	57
	Z5	77	64	68	74	70
Spleen	VHSE	63	63	63	63	63
	Z3	63	63	63	63	63
	Z5	49	61	56	54	55

[a] *SE*, Sensitivity: the proportion of all organ-targeting peptides correctly predicted, SE = TP/(TP + FN) where TP is the number of organ-targeting peptides correctly predicted and FN is the number of organ-targeting peptides incorrectly predicted as non-targeting peptides

[b] *SP*, Specificity: the proportion of organ non-targeting peptides correctly predicted, SP = TN/(TN + FP) where TN is the number of organ non-targeting peptides correctly predicted and FP is the number of organ non-targeting peptides incorrectly predicted as organ-targeting peptides

[c] *PPV*, Positive predictive value: the probability that a predicted organ-targeting peptide is in fact a organ-targeting peptide, PPV = TP/(TP + FP)

[d] *NPV*, Negative predictive value: the probability that a predicted organ non-targeting peptide is in fact organ non-targeting peptide, NPV = TN/(TN + FN)

[e] *Acc*, Accuracy: the percentage of all predictions that are correct, Acc = (TP + TN)/Total

result using binary descriptors is not reported in this paper. VHSE descriptors tended to produce slightly better training models than z scale descriptors. Mei et al. [14] noted that better results are obtained from the models using VHSE descriptors because VHSE scales possess relatively definite physicochemical meanings and more information in comparison with z scales. As shown in Table 2 and 3, our models produced similar results.

Tissue-specific targeting peptides have been shown to be promising materials in the delivery of drugs and other therapies to designated sites. For example, peptide ligands selectively homing to certain tissues have been linked to therapeutic compounds to target them to the sites of interest, thus enhancing their therapeutic efficacy and decreasing side effects. Such peptides were linked to the chemotherapeutic drug doxorubicine [6], proapoptotic peptides [8, 21, 22], cytotoxic agents [23], and cytokines [24, 25], resulting in a marked increase in therapeutic efficiency compared with that of the untargeted drug. Our QSAR study on the selection of organ-targeting peptides would be applicable to the development a strategy for the selective delivery of drugs to target organs.

Conclusions

We used artificial neural networks to screen organ-targeting peptides on the basis of sequence information. The VHSE descriptors tend to produce a slightly better training model than the z series descriptors. Though the models have limited robustness because of the small size of positive data set, they have the capabilities to make reliable predictions. The prediction power and the robustness are likely to be enhanced when more experimental data are available. These models are expected to find applications in the selection of organ-targeting peptides from large peptide libraries, and the selected peptides might be used to increase the efficacy and to reduce the toxicity associated with systemic administration of therapy.

Acknowledgments This work was supported by the Korea Science and Engineering Foundation (KOSEF) NRL Program grant funded by the Korea government (MEST) (No. R0A-2008-000-20024-1).

References

1. Pasqualini R, Ruoslahti E (1996) Organ targeting in vivo using phage display peptide libraries. Nature 380:364–366
2. Rajotte D, Arap W, Hagedorn M, Koivunen E, Pasqualini R, Ruoslahti E (1998) Molecular heterogeneity of the vascular endothelium revealed by in vivo phage display. J Clin Invest 102:430–437
3. Trepel M, Arap W, Pasqualini R (2000) Exploring vascular heterogeneity for gene therapy targeting. Gene Ther 7:2059–2060
4. Rajotte D, Ruoslahti E (1999) Membrane dipeptidase is the receptor for a lung-targeting peptide identified by in vivo phage display. J Biol Chem 274:11593–11598
5. Pasqualini R, Koivunen E, Kain R, Lahdenranta J, Sakamoto M, Stryhn A, Ashmun RA, Shapiro LH, Arap W, Ruoslahti E (2000) Aminopeptidase N is a receptor for tumor-homing peptides and a target for inhibiting angiogenesis. Cancer Res 60:722–727
6. Kolonin MG, Pasqualini R, Arap W (2001) Molecular addresses in blood vessels as targets for therapy. Curr Opin Chem Biol 5:308–313
7. Arap W, Pasqualini R, Ruoslahti E (1998) Cancer treatment by targeted drug delivery to tumor vasculature in a mouse model. Science 279:377–380
8. Arap W, Haedicke W, Bernasconi M, Kain R, Rajotte D, Krajewski S, Ellerby HM, Bredesen DE, Pasqualini R, Ruoslahti E (2002) Targeting the prostate for destruction through a vascular address. Proc Natl Acad Sci USA 99:1527–1531
9. Durr E, Yu J, Krasinska KM, Carver LA, Yates JR, Testa JE, Oh P, Schnitzer JE (2004) Direct proteomic mapping of the lung microvascular endothelial cell surface in vivo and in cell culture. Nat Biotechnol 22:985–992
10. Oh P, Li Y, Yu J, Durr E, Krasinska K, Carver LA, Testa JE, Schnitzer JE (2004) Subtractive proteomic mapping of the endothelial surface in lung and solid tumours for tissue-specific therapy. Nature 429:629–635

11. Trepel M, Arap W, Pasqualini R (2002) In vivo phage display and vascular heterogeneity: implications for targeted medicine. Curr Opin Chem Biol 6:399–404
12. Arap W, Kolonin MG, Trepel M, Lahdenranta J, Cardó-Vila M, Giordano RJ, Mintz PJ, Ardelt PU, Yao VJ, Vidal CI, Chen L, Flamm A, Valtanen H, Weavind LM, Hicks ME, Pollock RE, Botz GH, Bucana CD, Koivunen E, Cahill D, Troncoso P, Baggerly KA, Pentz RD, Do KA, Logothetis CJ, Pasqualini R (2002) Steps toward mapping the human vasculature by phage display. Nat Med 8:121–127
13. Jung E, Kim J, Kim M, Jung DH, Rhee H, Shin JM, Choi K, Kang SK, Kim MK, Yun CH, Choi YJ, Choi SH (2007) Artificial neural network models for prediction of intestinal permeability of oligopeptides. BMC Bioinformatics 8:245
14. Mei H, Lian ZH, Zhou Y, Li SZ (2005) A new set of amino acid descriptors and its application in peptide QSARs. Biopolymers (Peptide Science) 80:775–786
15. Hellberg S, Sjöström M, Skagerberg B, Wold S (1987) Peptide quantitative structure-activity relationships, a multivariate approach. J Med Chem 30:1126–1135
16. Sandberg M, Eriksson L, Jonsson J, Sjöströ M, Wold S (1998) New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids. J Med Chem 41:2481–2491
17. The nnet of VR 7.2 package (2004) http://www.r-project.org/, Version R 2.0.1
18. Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29–36
19. Springer C, Adalsteinsson H, Young MM, Kegelmeyer PW, Roe DC (2005) PostDock: a structural, empirical approach to scoring protein ligand complexes. J Med Chem 48:6821–6831
20. Cramer RD, Bunce JD, Patterson DE, Frank IE (1988) Cross-validation, bootstrapping, and partial least squares compared with multiple regression in conventional QSAR studies. Quant Struct-Act Relat 7:18–25
21. Ellerby HM, Arap W, Ellerby LM, Kain R, Andrusiak R, Rio GD, Krajewski S, Lombardo CR, Rao R, Ruoslahti E, Bredesen DE, Pasqualini R (1999) Anti-cancer activity of targeted pro-apoptotic peptides. Nat Med 5:1032–1038
22. Chen Y, Xu X, Hong S, Chen J, Liu N, Underhill CB, Creswell K, Zhang L (2001) RGD-tachyplesin inhibits tumor growth. Cancer Res 61:2434–2438
23. Yoneda Y, Steiniger SC, Capková K, Mee JM, Liu Y, Kaufmann GF, Janda KD (2008) A cell-penetrating peptidic GRP78 ligand for tumor cell-specific prodrug therapy. Bioorg Med Chem Lett 18:1632–1636
24. Curnis F, Sacchi A, Borgna L, Magni F, Gasparri A, Corti A (2000) Enhancement of tumor necrosis factor α antitumor immunotherapeutic properties by targeted delivery to aminopeptidase N (CD13). Nat Biotechnol 18:1185–1190
25. Curnis F, Arrigoni G, Sacchi A, Fischetti L, Arap W, Pasqualini R, Corti A (2002) Differential binding of drugs containing the NGR motif to CD13 isoforms in tumor vessels, epithelia, and myeloid cells. Cancer Res 62:867–874