JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Structural Requirements of Angiotensin I-Converting Enzyme Inhibitory Peptides: Quantitative Structure−Activity Relationship Study of Di- and Tripeptides

JIANPING WU,† ROTIMI E. ALUKO,*,† AND SHURYO NAKAI‡

Department of Human Nutritional Sciences, University of Manitoba, Winnipeg, Manitoba, R3T 2N2, Canada, and Food, Nutrition and Health, Faculty of Agricultural Sciences, The University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada

A database consisting of 168 dipeptides and 140 tripeptides was constructed from published literature to study the quantitative structure−activity relationships of angiotensin I-converting enzyme (ACE) inhibitory peptides. Two models were computed using partial least squares regression based on the three *z*-scores of 20 coded amino acids and further validated by cross-validation and permutation tests. The two-component model could explain 73.2% of the *Y*-variance (inhibitor concentration that reduced enzyme activity by 50%, IC_{50}) with the predictive ability of 71.1% for dipeptides, while the single-component model could explain 47.1% of the *Y*-variance with the predictive ability of 43.3% for tripeptides. Amino acid residues with bulky side chains as well as hydrophobic side chains were preferred for dipeptides. For tripeptides, the most favorable residues for the carboxyl terminus were aromatic amino acids, while positively charged amino acids were preferred for the middle position, and hydrophobic amino acids were preferred for the amino terminus. According to the models, the IC_{50} values of seven new peptides with matchable primary sequences within pea protein, bovine milk protein, and soybean were predicted. The predicted peptides were synthesized, and their IC_{50} values were validated through laboratory determination of inhibition of ACE activity.

KEYWORDS: Angiotensin converting enzyme; bioactive peptides; quantitative structure−activity relationship (QSAR); partial least squares regression

INTRODUCTION

Bioactive peptides are increasingly becoming important as starting points for the development of drugs and drug-related compounds (*1*, *2*). It is also a fact that bioactive peptides contribute greatly to the content of functional foods, as food scientists realized that some specific sequences within the parent food proteins can provide physiological benefits after they have been released through in vitro processing or in vivo digestion (*3*, *4*). Among various bioactive peptides, antihypertensive peptides have been studied extensively and the mechanism of activity involves inhibition of angiotensin I-converting enzyme (ACE), the key enzyme responsible for the regulation of blood pressure via the renin−angiotensin system. ACE converts angiotensin I to angiotensin II, a potent vasoconstrictor; ACE also hydrolyzes and inactivates bradykinin, a potent vasodilator. Therefore, excessive activity of ACE leads to an increased rate of vasoconstriction and development of high blood pressure. Inhibitory peptides block the ACE-mediated production of angiotensin II, and the reduction in ACE activity results in enhanced levels of bradykinin (*5*, *6*). Some antihypertensive peptides, especially short peptides such as VPP and IPP that are present in sour milk, have been shown to be resistant to in vivo degradation and were able to exert antihypertensive activity through the inhibition of ACE in aorta (*7*, *8*). In a human trial, the administration of this type of sour milk was shown to be effective in reducing the blood pressure of hypertensive patients; a product containing VPP and IPP has since been commercialized (*9*−*11*). Productions of ACE inhibitory peptides from several sources have also been reported (*3*, *4*, *8*, *12*−*34*).

In contrast to the vast amount of information that is available on production and characterization of antihypertensive peptides, there is very limited information on the relationships between structure and activity of food protein-derived antihypertensive peptides, especially the effect of primary structure on potency. Limited works on structure−activity relationships have been reported mostly from the early studies of snake venom peptide analogues and synthetic dipeptides (*35*−*37*). Although structure and activity relationships have also been a subject of research for food-derived ACE inhibitory peptides, all knowledge obtained so far was derived primarily from qualitative analysis of chemically synthesized peptides or peptide analogues having similar structures to those of known peptides (*12*, *13*). Consequently, current understanding of the structure and activity relationships of ACE inhibitory peptides is experimenter-

* To whom correspondence should be addressed. Tel: 204-474-9555. Fax: 204-474-7593. E-mail: alukor@cc.umanitoba.ca.
† University of Manitoba.
‡ The University of British Columbia.

10.1021/jf051263l CCC: $33.50 © 2006 American Chemical Society
Published on Web 01/04/2006

dependent, which is insufficient and unable to provide predictive power. For example, the common applied COST approach (change one structural factor at a time) was reported to be inefficient and unrealistic for the study of bioactive peptides (*38*). In contrast, quantitative structure−activity relationships (QSAR), an important area of chemometrics, search information relating chemical structure to biological and other activities. QSAR has become increasingly helpful in understanding many aspects of chemical−biological interactions in drug and pesticide research as well as many other areas (*39*). The basic assumption in QSAR is that the biological activity (BA) within a set of compounds is related to the structural variation of the compounds; that is, the BA can be modeled as a function of molecular structure (*40*). For example, the three *z*-approach has been successfully applied to develop sequence-dependent models of small peptides' biological properties (*38*, *41*, *42*) and functional behavior of polypeptides (*43*, *44*). Therefore, it is possible to use QSAR to model ACE inhibitory peptides and predict the most potent peptides and thus study and develop new ACE inhibitory peptides from food proteins driven by the predicted structure and activity outcomes.

Therefore, the objectives of this study were to (i) construct an ACE inhibitory peptide database and apply it for modeling of activities using structural descriptors of individual amino acids and (ii) further predict and validate potentially potent ACE inhibitory peptides based on the models.

MATERIALS AND METHODS

Preparation of Data Set. An ACE inhibitory peptide database consisting of 168 dipeptides and 140 tripeptides that were collected from previously published works was constructed. These peptides were selected based on the fact that each one has a reported (in vitro assay) IC_{50} (inhibitory concentration that reduced ACE activity by 50%) value that is less than 20 mM. Peptides reported from different sources with the same sequence were included provided their determined IC_{50} values were different; otherwise, only one peptide result was used. Data sets for the dipeptides and tripetides are presented in **Tables 1** and **2**, respectively.

Analysis of Data Set. The characterization of each individual amino acid by the three *z*-scores (**Table 3**), namely, z_1, z_2, and z_3, was previously calculated by principal component analysis (PCA) from a matrix consisting of 29 physicochemical variables (*41*). These three resulting principal components, so-called principal properties, are linear combinations of the primary data and were tentatively interpreted to represent largely lipophilicity (z_1), steric properties or side chain bulk/molecular size (z_2), and electronic properties (z_3). Because the activities of peptides ranged over almost 5 orders, from −0.68 to 4.23 in log units, they were log-transformed prior to modeling. The amino acid at the amino terminus was designed as n_1, and its properties were described as n_1z_1, n_1z_2, and n_1z_3; the amino acid adjacent to the amino terminus was designed as n_2, and its properties were described as n_2z_1, n_2z_2, and n_2z_3, etc. Partial least squares regression (PLS) analysis between amino acid descriptors (predictors, *X*) and log-transformed IC_{50} values (dependent, *Y*) was carried out using SIMCA-P version 10 (Umetrics Inc., Kinnelon, NJ). All variables were centered and scaled to unit variance prior to the analyses except with specification, and thereby, all variables had an equal participation in the model. In SIMCA-P, the number of significant PLS components was chosen automatically by using various rules based on a statistic called Q^2. Q^2 is the cross-validation correlation coefficient that is calculated from predicted residual sum of squares (PRESS) and refers to model's predictive ability. Another important parameter in PLS analysis is the multiple correlation coefficient (R^2), which provides an estimate of the model fit. The optimal model was made where a reasonable balance between the model's fit and predictive ability was achieved (*44*, *45*).

Validation of Predicted Peptides. Predicted ACE inhibitory peptides with matchable sequences within food proteins (milk protein, soybean protein, and pea protein) were synthesized and used as an external validation data set. Three predicted dipeptides (FW, WW, and YW), four predicted tripeptides (VRF, IKP, LRF, and LRW), and two famous milk protein-derived tripeptides (VPP and IPP) were synthesized by Genscript Corp. (Piscataway, NJ). The purity (95−99.9%) of each peptide was measured by high-performance liquid chromatography (HPLC), and the structure was verified by mass spectrometry (Genscript Corp.). The ACE inhibitory activity of each peptide was analyzed by a spectrophotometric assay according to the method of Cushman and Cheung (*46*).

Table 1. Dipeptides and Their IC_{50} Values for the in Vitro Inhibition of Angiotensin Converting Enzyme[a]

peptide	log IC_{50}	peptide	log IC_{50}	peptide	log IC_{50}	peptide	log IC_{50}
IY	**0.57**	RP	1.32	LF	3.52	GT	3.76
YP	2.95	GP	2.56	IR	2.84	GE	3.85
DY	2.00	GP	3.08	RL	3.39	GG	3.86
YG	3.04	TP	2.46	KP	1.71	GD	3.96
QK	2.95	VP	2.76	FL	1.20	VG	3.04
IY	**0.32**	GI	3.11	VY	1.25	IG	3.08
LY	**0.83**	DF	2.56	IL	1.74	RG	3.08
MF	1.65	NP	3.36	VY	1.55	YG	3.30
RY	1.71	DM	2.78	**IY**	**0.79**	AG	3.40
MY	2.29	DL	3.30	AW	1.27	KG	3.51
LY	1.59	GY	2.42	**FY**	**1.63**	FG	3.57
KW	0.21	**VY**	**1.20**	**VW**	**0.52**	MG	3.68
TF	1.25	GF	2.85	**IW**	**0.18**	WG	3.77
LY	0.81	VW	0.20	**LW**	**1.37**	HG	3.80
YL	1.21	DG	1.09	**FY**	**0.22**	EG	3.87
AF	1.18	FQ	1.71	KF	2.06	SG	3.93
VF	0.96	**VY**	**1.41**	IF	2.97	LG	3.94
AY	1.28	TF	1.95	VY	1.76	TG	4.00
FP	**2.50**	LY	1.51	GQ	3.75	QG	4.00
VW	**0.15**	YL	1.91	TP	3.32	DG	4.15
YW	**1.02**	AF	1.88	TK	3.21	PG	4.23
HY	**1.42**	IY	1.02	**YH**	**0.71**	**VW**	**0.20**
RF	**1.97**	VK	1.11	KW	1.03	IY	0.30
GG	3.94	FY	1.40	**KY**	**0.89**	AW	1.00
GY	2.41	**AY**	**2.00**	KF	1.45	RW	1.20
IY	**0.38**	LF	2.10	**FY**	**0.57**	**VY**	**1.34**
PR	0.61	YV	2.76	VW	1.03	VF	1.72
RY	1.02	YE	2.80	VY	1.64	AY	1.94
AW	1.08	GW	1.48	IW	1.09	IP	2.11
IY	**0.36**	GY	2.32	VY	1.05	RP	2.26
LW	**0.83**	GP	2.65	IA	2.18	AF	2.28
SY	**1.82**	GF	2.80	WL	1.48	AP	2.36
GY	**1.86**	GI	3.08	WA	2.44	RF	2.36
NY	**1.51**	GM	3.15	VW	0.40	VP	2.62
SF	**2.11**	GA	3.30	WM	1.98	AP	2.43
GF	**2.44**	GL	3.40	MW	1.00	IR	2.92
NF	**1.67**	GH	3.49	IW	0.67	VQ	3.11
LW	**1.70**	GR	3.51	LW	1.24	**IY**	**0.43**
YP	**2.86**	GS	3.58	RP	1.96	VW	0.23
YG	3.18	GV	3.66	A P	1.46	MW	0.58
LF	2.54	GK	3.73	**KP**	**1.34**	**RW**	**1.34**
YL	2.09	GQ	3.73	FY	0.81	**KP**	**1.48**

[a] Peptides in bold have also been tested in vivo for antihypertensive activity. Repeated peptides indicate results from different laboratories.

RESULTS AND DISCUSSION

QSAR of Di- and Tripeptides. It is believed that peptides with two or three amino acid residue lengths could be absorbed directly from the digestive tract into the blood circulatory system (*47*) and be able to reach the action sites to exert physiological functions. Therefore, the results could provide very potent peptides that have physiological relevance. Modeling of these di- and tripeptides was conducted according to the data sets in **Tables 1** and **2**, respectively. Initial modeling of the three *z*-scale descriptors with activities resulted in a two-component PLS model explaining 65.1% of the sum of squares in *Y*-variance with a predictive ability of 62% for the dipeptide set, and the

Table 2. Tripeptides and Their IC_{50} Values for the in Vitro Inhibition of Angiotensin Converting Enzyme[a]

peptide	log IC_{50}	peptide	log IC_{50}	peptide	log IC_{50}	peptide	log IC_{50}
FEP	1.08	MNP	1.82	GPM	1.23	YEY	0.60
IKP	**0.23**	NPP	2.46	**GKP**	**2.55**	PSY	1.20
LNY	1.91	PPK	3.00	**IPA**	**2.15**	LGI	1.46
HQG	2.87	ITT	2.83	**VYP**	**2.46**	ITF	1.69
HHT	2.90	TTN	2.83	VWY	0.97	**IPP**	**1.92**
ALP	2.38	TNP	2.32	FYN	1.26	**IAP**	**1.40**
LKP	0.20	**GQP**	**0.51**	YGG	4.00	**EAP**	**2.61**
LYP	0.82	GSH	1.51	GGY	0.11	FAP	0.62
DYG	3.43	RML	3.01	YPR	1.22	**SVY**	**3.23**
AQK	3.26	YVA	0.15	PRY	0.40	LEK	2.90
IEP	**0.20**	GKV	0.59	YGL	2.61	GVY	2.60
IKY	**–0.68**	SVY	0.91	VFK	3.01	IRP	0.26
LAP	**0.54**	FFL	1.57	**IKW**	**–0.68**	**LPP**	**0.98**
LKP	**–0.49**	IFL	1.65	**IKP**	**0.84**	LVL	1.09
GRP	1.30	LPF	1.60	**IWH**	**0.54**	LRP	–0.57
RFH	2.52	GPP	1.25	VAP	0.30	LSP	0.23
AKK	0.50	AGP	2.75	FAP	0.58	LEP	0.28
RVY	2.31	VIY	0.88	**PLW**	**1.56**	VSP	1.00
LKL	2.27	**RIY**	**1.45**	FGK	2.20	LEP	1.63
HIR	2.98	AFL	1.80	AVP	2.53	LNP	1.76
HHL	1.73	FAL	1.42	PYP	2.34	LLP	1.20
HLL	1.76	IAQ	1.54	LVR	1.15	VLP	0.59
HHL	0.73	VVF	1.55	TAP	0.54	LAY	1.40
LIY	**–0.09**	IVQ	1.98	VRP	0.34	IRA	1.11
LAY	0.59	VQV	0.94	MPP	0.98	LAA	1.11
LLP	1.76	AQL	1.76	LKP	0.60	LEE	2.00
LEE	2.00	LVQ	1.15	TVY	1.18	MKY	0.86
FNE	2.53	FDK	2.59	**IVY**	**–0.32**	LRY	0.70
GPL	0.35	IVY	0.38	IMY	0.26	VSW	1.37
GPV	0.67	VLP	1.91	DGL	0.33	LWA	1.10
IPP	0.70	VLY	1.49	TKY	0.36	VTR	2.13
GPL	0.41	ILP	1.51	LTF	0.44	IKW	–0.27
DLP	0.68	**VPP**	**0.95**	AGP	1.95	VGP	1.42
GLY	0.95	RPP	1.78	FNF	0.84	**MRW**	**–0.22**
LLF	1.90	RPK	3.27	AVL	0.85	**IAY**	**1.10**
IAP	**0.43**						

[a] Peptides in bold have also been tested in vivo for antihypertensive activity. Repeated peptides indicate results from different laboratories.

Table 3. Descriptor (*Z*) Scores for Amino Acids (*41*)

amino acid	code	z_1	z_2	z_3	amino acid	code	z_1	z_2	z_3
Ala	A	0.07	–1.73	0.09	His	H	2.41	1.74	1.11
Val	V	–2.69	–2.53	–1.29	Gly	G	2.23	–5.36	0.30
Leu	L	–4.19	–1.03	–0.98	Ser	S	1.96	–1.63	0.57
Ile	I	–4.44	–1.68	–1.03	Thr	T	0.92	–2.09	–1.40
Pro	P	–1.22	0.88	2.23	Cys	C	0.71	–0.97	4.13
Phe	F	–4.92	1.30	0.45	Tyr	Y	–1.39	2.32	0.01
Trp	W	–4.75	3.65	0.85	Asn	N	3.22	1.45	0.84
Met	M	–2.49	–0.27	–0.41	Gln	Q	2.18	0.53	–1.14
Lys	K	2.84	1.41	–3.14	Asp	D	3.64	1.13	2.36
Arg	R	2.88	2.52	–3.44	Glu	E	3.08	0.39	–0.07

one-component PLS model explained 36.1% of the sum of squares in *Y*-variance with a predictive ability of 28.5% for the tripeptide set. In an attempt to improve the predictive ability of the model, outliers were excluded as previously described (*48*, *49*) and calculations were continued to make new models. Outliers were identified if the absolute residual value of predicted log (IC_{50}) − measured log (IC_{50}) exceeded 1.5 units. The *tu* plots (**Figure 1**) displayed the relationship between *X* and *Y*; outliers, which showed a much worse fit than others, were numbered and excluded in the second round of modeling (*48*, *49*). Modeling of the data set that excluded outliers resulted in substantially improved models, a two-component model explaining 73.2% of the *Y*-variance with the predictive ability

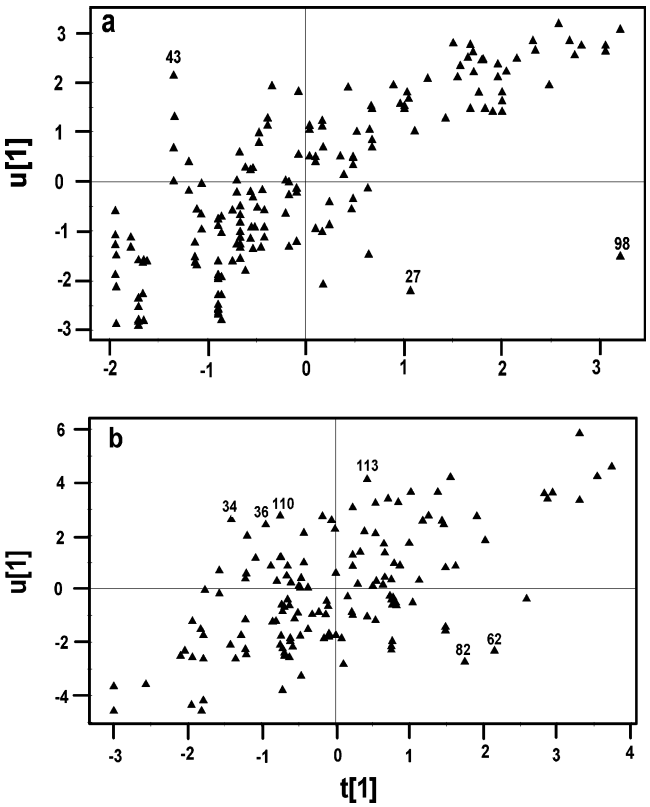

Figure 1. PLS scores, *u*1 and *t*1, of the ACE peptides (**a**) for dipeptides and (**b**) for tripeptides. The scores, such as *u* and *t*, are new variables, are orthogonal, and are summaries of the *X*-variables and *Y*-variables, respectively. Outliers [predicted log (IC_{50}) – measured log (IC_{50}) is greater than 1.5], which are identified with numbers, were excluded from the second round of modeling.

of 71.1% for dipeptides and a single-component model explaining 47.1% of the *Y*-variance with the predictive ability of 43.3% for tripeptides.

The resulting models are illustrated in **Figure 2**, where the relationship between the predicted vs the observed values is presented. The multiple correlation coefficient (R^2) for dipeptides was comparable to the previous results that used a 58 dipeptide set, such as 70.0% in Collantes and Dunn's work (*50*) and 70.8% in Zaliani and Gancia's work (*51*). However, the R^2 value was low for the tripeptide model when compared with the value for the dipeptide model. The lower R^2 for the tripeptide model could be as a result of greater interlaboratory variations and unknown systemic differences (*41*). The two models were validated initially by cross-validation during modeling for a total of seven times. The predictive power of these two ACE models was further validated by response permutation, where the *Y*-response data vectors (log IC_{50}) were each randomly reordered and permuted a number of times but with unperturbed *X* data followed by computation of a QSAR model, which was used to refit the model (*52*). Twenty times permutation and cross-validation rounds computed the resulting R^2 and Q^2 intercepts, which are 0.003 and −0.145 for the dipeptide set and 0.039 and −0.091 for the tripeptide set (**Figure 3**), respectively. SIMCA-P displays the plot of the correlation coefficient between the original *Y* and the permutted *Y* vs the cumulative R^2 and Q^2 and draws the regression line. The intercept (R^2 and Q^2 when the correlation coefficient is zero) is a measure of the over fit (*45*). It was suggested that the desirable limits for a valid model should be an R^2 intercept < 0.3 and a Q^2 intercept < 0.05 (*53*); therefore, the two models developed here were regarded as valid.

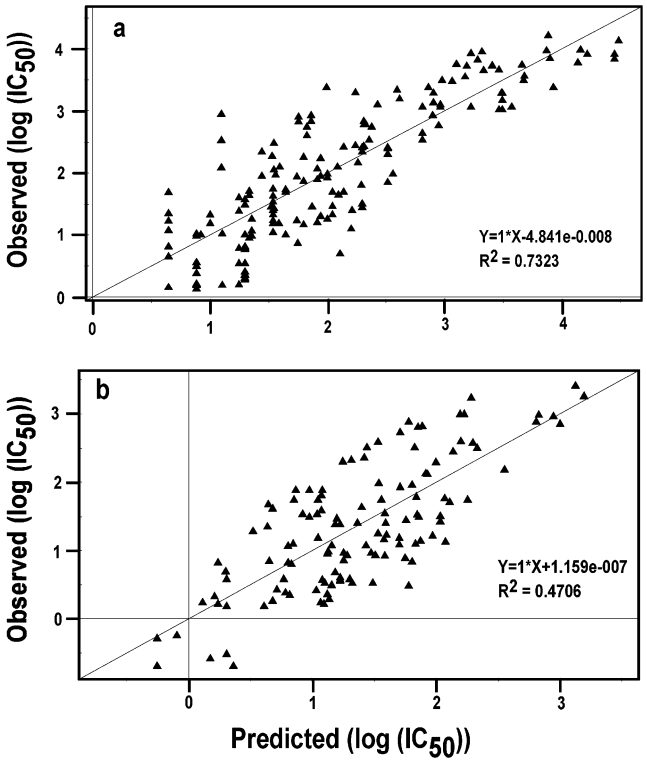

Figure 2. Relationship between the observed and calculated values of log (IC_{50}) for (**a**) dipeptides and (**b**) tripeptides, after exclusion of outliers.

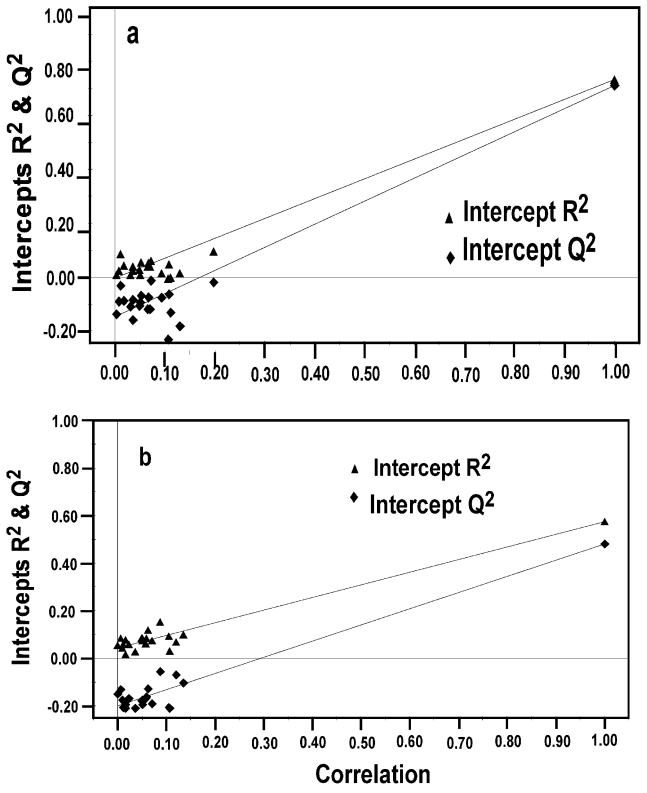

Figure 3. Permutation validation of the PLS model (**a**) for dipeptides and (**b**) for tripeptides. The intercepts (R^2 and Q^2 when the correlation coefficient is zero) are a measure of the over fit.

Prediction and Validation of Potent ACE Inhibitory Peptide Sequences. The importance of a given *X*-variable for *Y* is proportional to its distance from the origin in the loading space and corresponds to the PLS regression coefficients (*49*). The expected amino acid properties in each position are

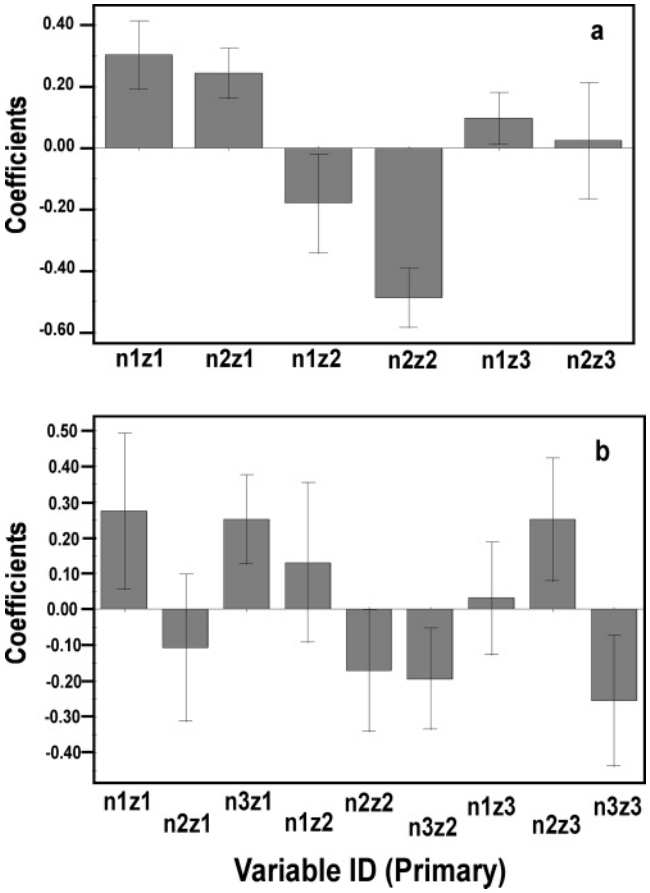

Figure 4. PLS regression coefficients (**a**) for dipeptides and (**b**) for tripeptides. The importance of a given *X*-variable is proportional to its distance (coefficient value) from the origin (zero). The bars indicate 95% confidence intervals based on jack-knifing.

evaluated according to their importance to the *Y*-variable. For dipeptides (**Figure 4a**), n_1z_1, n_1z_3, n_2z_1, and n_2z_3 are positively related to the log values, while n_1z_2 and n_2z_2 are negatively related to the log values. Looking at the coefficients values, it is evident that position n_2, which corresponds to the carboxyl terminus for a dipeptide, is more important than position n_1. The importance of the amino acid residue in position n_2 is mainly decided by z_2 (steric properties) followed by z_1 (lipophilicity). For both positions, amino acid residues with large bulk chain as well as hydrophobic side chains (high z_2 and low z_1) are preferred, such as phenylalanine, tryptophan, and tyrosine. It has been suggested that aromatic side chains and proline are favored in position n_2 and branched aliphatic side amino acids are preferred in position n_1 (*37*), which is similar to our results. However, proline that is well-documented as the most favorable amino acid for binding to ACE (*1*, *2*) (most commercially existing inhibitors bear this residue) was not the most favorable amino acid in this study. Hellberg et al. (*38*) indicated that a highly active peptide should have an amino acid with a hydrophobic side chain, possibly with a positively charged functional group in position 1 and a large hydrophobic amino acid in position 2 based on QSAR model using ACE inhibiting dipeptides reported by Cheung et al. (*37*); our results are in good agreement with this previous modeling. It should be noted that the model of Hellberg et al. (*38*) was constructed based on ideal circumstances where all of the tests were performed in one single laboratory at the same time; however, they applied quadratic and cross-terms of the *z*-scales to achieve the model, which

Table 4. Prediction, Experimental Validation, and Location of Potent ACE Inhibitory Peptides

peptide	position	parent protein[a]	log IC_{50} predicted	log IC_{50} observed	error
FW	f150–151	legumin A2 precursor; *Pisum sativum* (garden pea)	0.60	0.77	0.17
	f149–150	legumin A precursor; *Pisum sativum* (garden pea)			
WW	f150–151	glycinin G1 precursor; *Glycine max* (soybean)	0.52	1.91	1.39
	f147–148	glycinin G2 precursor; *G. max* (soybean)			
YW	f122–123	α-lactalbumin precursor; *Bos taurus* (bovine milk)	0.92	1.64	0.72
	f113–114	albumin 2 (PA2); *Pisum sativum* (garden pea)			
	f219–220	legumin J precursor; *P. sativum* (garden pea)			
	f153–154	legumin K; *P. sativum* (garden pea)			
	f3–4	β-conglycinin, α-chain precursor; *G. max* (soybean)			
	f27–28	glycinin G3 precursor; *G. max* (soybean)			
	f147–148	glycinin G4 precursor; *G. max* (soybean)			
	f156–157	glycinin precursor; *G. max* (soybean)			
VRF	f4–6	β-conglycinin, β-chain precursor; *G. max* (soybean)	0.14	1.38	1.24
IKP	f265–267	glycinin G1 precursor; *G. max* (soybean)	0.37	0.44	0.07
	f279–281	legumin J precursor; *P. sativum* (garden pea)			
	f6–8	vicilin 47 kDa protein; *P. sativum* (garden pea)			
LRW	f377–379	legumin A2 precursor; *P. sativum* (garden pea)	–0.11	–0.64	–0.53
	f374–376	legumin A precursor; *P. sativum* (garden pea)			
LRF	f36–38	α-S_1 casein precursor; *B. taurus* (bovine milk)	0.22	1.33	1.11
VPP	f99–101	β-casein precursor; *B. taurus* (bovine milk)	1.28	1.42	0.14
IPP	f89–91	β-casein precursor; *B. taurus* (bovine milk)	1.18	1.75	0.57
	f129–131	κ-casein precursor; *B. taurus* (bovine milk)			

[a] Protein primary sequences (positions) were obtained from the ExPASy (expert protein analysis system) proteomics server of the Swiss Institute of Bioinformatics. Predicted activity refers to the values obtained from PLS modeling while observed activity refers to the experimentally determined activity using synthetic peptides.

requires a more complex explanation of the effects on peptide activity.

For tripeptides (**Figure 4b**), n_1z_1, n_1z_2, n_1z_3, n_2z_3, and n_3z_1 are positively related to the log values, while n_3z_3, n_3z_2, n_2z_2, and n_2z_1 are negatively related to the values. For active tripeptides, low z_1 and z_2 values such as valine, leucine, and isoleucine are preferred for position n_1, low z_3 with high z_1 and z_2 values such as lysine and arginine are expected for position n_2, and low z_1 with high z_2 and z_3 values such as proline, phenylalanine, and tryptophan are favored for position n_3. The most favorable amino acids for the carboxyl terminus were aromatic acids, for the second position from the carboxyl terminus were positively charged amino acids and hydrophobic amino acids for the amino terminus.

It is well-known that milk protein is a good source for the production of ACE inhibitory peptides (*11*, *12*), such as VPP and IPP. In addition, ACE inhibitory peptides have also been reported in soybean and pea proteins (*54*, *55*). Three predicted dipeptides and four predicted tripeptides were located within the primary structure of food proteins and were then synthesized for validation (**Table 4**). FW was located in pea protein sequence, WW in soybean protein, and YW in bovine milk, soybean, and pea proteins. VRF was located in the primary structure of soybean protein, IKP in soybean and pea proteins, LRW in pea protein, and LRF in milk protein. All of the predictive error of these ACE inhibitory peptides in our study is within the absolute limit of 1.5 units between predicted log (IC_{50}) − measured log (IC_{50}) (**Table 4**). The biggest predictive errors of ACE inhibitory peptides in this study are 1.39 for WW, 1.24 for VRF, and 1.11 for LRF. Among these predicted peptides, LRW is one of the most potent ACE inhibitory peptides ever reported and is over 100 times more potent than the well-known milk tripeptides of VPP and IPP as determined in our study. IKP and FW are also more potent than VPP and IPP. These results confirm the validity of the prediction models, which could provide a useful tool for future prediction of potentially more potent ACE inhibitory peptides.

Conclusions. We have shown that the activities of food-derived ACE inhibitory peptides, namely, di- and tripeptides, can be modeled from the three *z*-scores of amino acids using PLS regression. Potential new ACE inhibitory peptide sequences with higher potency than most previously reported peptides were proposed based on these models. It is anticipated that future research on ACE inhibitory peptides would be driven by these newly developed models. Further research to model peptides with varied amino acid residue lengths will be carried out, and we will expand this approach to include a number of other bioactive peptides. It is likely that the predicted short peptides, e.g., LRW, IKP, and FW, would show stronger antihypertensive activities than the famous milk peptides, VPP and IPP. Evaluation of in vivo activity using spontaneously hypertensive rats and development of technology releasing these peptides from parent proteins is in progress.

LITERATURE CITED

(1) Wyvratt, M. J. Evolution of angiotensin-converting enzyme inhibitors. *Clin. Physiol. Biochem.* **1988**, *6*, 217−219.

(2) Wimart, M. C.; Komajda, M. Angiotensin converting enzyme inhibition: From viper to patient. *Heart* **2000**, *84*, i11−i14.

(3) Ariyoshi, Y. Angiotensin-converting enzyme inhibitor derived from food protein. *Trends Food Sci. Technol.* **1993**, *4*, 139−144.

(4) Yoshikawa, M.; Fujita, H.; Matoba, N.; Takenaka, Y.; Yamamoto, T.; Yamauchi, R.; Tsuruki, H.; Takahata, K. Bioactive peptides derived from food proteins preventing lifestyle-related diseases. *BioFactors* **2000**, *12*, 143−146.

(5) Skeggs, L. T.; Kahn, J. R.; Shumway, N. P. The preparation and function of the angiotensin-converting enzyme. *J. Exp. Med.* **1956**, *103*, 259−299.

(6) Yang, H. Y. T.; Erdos, E. G.; Levin, Y. A dipeptidyl carboxypeptidase that converts angiotensin II and inactivates bradykinin. *Biochim. Biophys. Acta* **1970**, *214*, 374−376.

(7) Nakamura, Y.; Masuda, O.; Tanako, T. Decrease of tissue angiotensin I-converting enzyme activity upon feeding sour milk spontaneously to hypertensive rats. *Biosci., Biotechnol., Biochem.* **1996**, *60*, 488−489.

(8) Nakamura, Y.; Yamamoto, N.; Sakai, K.; Okubo, A.; Yamazaki, S.; Takano, T. Purification and characterization of angiotensin I-converting enzyme inhibitors from a sour milk. *J. Dairy Sci.* **1995**, *78*, 777−783.

(9) Hata, Y.; Yamamoto, M.; Ohni, M.; Nakajima, K.; Nakamura, Y.; Takano, T. A. placebo-controlled study of the effect of sour milk on blood pressure in hypertensive subjects. *Am. J. Clin. Nutr.* **1996**, *64*, 767−771.
(10) Seppo, L.; Jauhiainen, T.; Poussa, T.; Korpela, R. A fermented milk, high in bioactive peptides, has a blood pressure lowering effect in hypertensive subjects. *Am. J. Clin. Nutr.* **2003**, *77*, 326−330.
(11) FitzGerald, R. J.; Murray, B. A.; Walsh, D. J. Hypotensive peptides from milk proteins. *J. Nutr.* **2004**, *134*, 980S−988S.
(12) Maruyama, S.; Mitachi, H.; Tanaka, H.; Tomizuka, N.; Suzuki, H. Studies on the active site and antihypertensive activity of angiotensin I-converting enzyme inhibitory derived from casein. *Agric. Biol. Chem.* **1987**, *51*, 1581−1586.
(13) Yano, S.; Suzuki, K.; Funatsu, G. Isolation from α-zein hydrolysate of thermolysin peptides with angiotensin converting inhibitory activity. *Biosci., Biotechnol., Biochem.* **1996**, *60*, 661−663.
(14) Maeno, M.; Yamamoto, N.; Takano, T. Identification of an antihypertensive peptide from casein hydrolysate produced by a protease from *Lactobacillus helveticus* CP790. *J. Dairy Sci.* **1996**, *79*, 1316−1321.
(15) Yamamoto, N.; Akino, A.; Takano, T. Antihypertensive effects of peptides derived from casein by an extractcelluar proteinase from *Lactobacillus helvetivus* CP790. *J. Dairy Sci.* **1994**, *77*, 917−922.
(16) Fujita, H.; Yokoyama, K.; Yoshikawa, M. Classification and antihypertensive activity of angiotensin I-converting enzyme inhibitory peptides derived from food protein. *J. Food Sci.* **2000**, *5*, 564−569.
(17) Fujita, H.; Yoshikawa, M. LKPNM: A prodrug-type ACE-inhibitory peptide derived from fish protein. *Immunopharmacology* **1999**, *44*, 123−127.
(18) Miyoshi, S.; Ishikawa, H.; Kaneko, T.; Fukui, F.; Tanaka, H.; Marayama, S. Structures and activity of angiotensin converting enzyme inhibitors in α-zein hydrolysate. *Agric. Biol. Chem.* **1991**, *55*, 1314−1318.
(19) Yang, Y.; Marczak, E. D.; Yokoo, M.; Usui, H.; Kawamura, Y.; Yoshikawa, M. Isolation and antihypertensive effect of angiotensin I-converting enzyme (ACE) inhibitory peptides from spinach rubisco. *J. Agric. Food Chem.* **2003**, *51*, 4897−4902.
(20) Kohama, Y.; Oka, H.; Kayamori, Y.; Tsujikawa, K.; Mimura, T.; Nagase, Y.; Satake, M. Potent synthetic analogues of angiotensin-converting enzyme inhibitor derived from tuna muscle. *Agric. Biol. Chem.* **1991**, *55*, 2169−2170.
(21) Kohama, Y.; Oka, H.; Matsumoto, S.; Teramoto, T.; Okaba, M.; Mimura, T.; Nagase, Y.; Chiba, Y.; Fujita, T. Induction of angiotensin-converting enzyme inhibitory activity by acid-limited proteolysis of glyceraldehyde 3-phosphate dehydrogenase. *Biochem. Biophys. Res. Commun.* **1989**, *161*, 456−460.
(22) Matsui, T.; Li, C. H.; Osajima, Y. Preparation and characterization of novel bioactive peptides responsible for angiotensin I-converting enzyme inhibition from wheat germ. *J. Pept. Sci.* **1999**, *5*, 289−297.
(23) Matsui, T.; Yukiyoshi, A.; Doi, S.; Sugimoto, H.; Yamada, H.; Matsumoto, K. Gastrointestinal enzyme production of bioactive peptides from royal jelly protein and their antihypertensive ability in SHR. *J. Nutr. Biochem.* **2002**, *13*, 80−86.
(24) Nomura, A.; Noda, N.; Maruyama, S. Purification of angiotensin I-converting enzyme inhibitors in pelagic thresher *Alopias pelagicus* muscle hydrolysate and viscera extracts. *Fish. Sci.* **2002**, *68*, 954−956.
(25) Mimura, T.; Kohama, Y.; Satake, M.; Nagase, Y. Octopeptide exhibiting antihypertensive activity. U.S. Patent 5,071,955, 1991.
(26) Nakagomi, K.; Yamada, R.; Ebisu, H.; Sadakane, Y.; Akizawa, T.; Tanimura, T. Isolation of Acein-2, a novel angiotensin-I-converting enzyme inhibitory peptide derived from a tryptic hydrolysate of human plasma. *FEBS Lett.* **2000**, *467*, 235−238.
(27) Suetsuna, K.; Maekawa, K.; Chen, J. R. Antihypertensive effects of *Undaria pinnatifida* (wakame) peptide on blood pressure in spontaneously hypertensive rats. *J. Nutr. Biochem.* **2004**, *15*, 267−272.
(28) Sato, M.; Hosokawa, T.; Yamaguchi, T.; Nakano, T.; Muramoto, K.; Kahara, T.; Funayama, K.; Kobayashi, A.; Nakano, K. Angiotensin I-converting enzyme inhibitory peptides derived from wakame (*Undaria pinnatifida*) and their antihypertensive effect in spontaneously hypertensive rats. *J. Agric. Food Chem.* **2002**, *50*, 6245−6252.
(29) Yokoyama, K.; Chiba, H.; Yoshikawa, M. Peptides inhibitors for angiotensin I-converting enzyme from thermolysin digest of dried bonito. *Biosci., Biotechnol., Biochem.* **1992**, *56*, 1541−1545.
(30) Saito, S.; Nakamura, K.; Kawato, A.; Imayasu, S. Structure and activity of angiotensin converting inhibitory peptides from sake and sake lees. *Biosci., Biotechnol., Biochem.* **1994**, *58*, 1767−1771.
(31) Kohmura, M.; Nio, N.; Ariyoshi, Y. Inhibition of angiotensin I-converting enzyme by synthetic peptides of human k-casein. *Agric. Biol. Chem.* **1990**, *54*, 835−836.
(32) Matsumura, N.; Fujii, M.; Takeda, Y.; Shimizu, T. Isolation and characterization of angiotensin I-converting enzyme inhibitory peptides derived from bonito bowels. *Biosci., Biotechnol., Biochem.* **1993**, *57*, 1743−1744.
(33) Li, C. H.; Matsui, T.; Matsumoto, K.; Yamasaki, R.; Kawasaki, T. Latent production of angiotensin I-converting enzyme inhibitors from buckwheat protein. *J. Pept. Sci.* **2002**, *8*, 267−274.
(34) Kohama, Y.; Matsumoto, S.; Oka, H.; Teramoto, T.; Okaba, M.; Mimura, T. Isolation of angiotensin converting enzyme inhibitor from tuna muscle. *Biochem. Biophys. Res. Commun.* **1988**, *155*, 332−337.
(35) Cushman, D. W.; Pluscec, J.; Williams, N. J.; Weaver, E. R.; Sabo, E. F.; Kocy, O.; Cheung, H. S.; Ondetti, M. A. Inhibition of angiotensin-converting enzyme by analogues of peptides from *Bothrops jararaca* venom. *Experientia* **1973**, *29*, 1032−1035.
(36) Cheung, H. S.; Cushman, D. W. Inhibition of homogeneous Angiotensin I-converting enzyme of rabbit lung by synthetic venom peptides of *Bothrops Jararaca*. *Biochim. Biophy. Acta* **1973**, *293*, 451−463.
(37) Cheung, H. S.; Wang, F. L.; Ondetti, M. A.; Sabo, E. F.; Cushman, D. W. Binding of peptide substrates and inhibitors of angiotensin-cnverting enzyme: importance of the COOH-terminal dipeptide sequence. *J. Biol. Chem.* **1980**, *255*, 401−407.
(38) Hellberg, S.; Eriksson, L.; Jonsson, J.; Lindgren, F.; Sjostrom, M.; Skagerberg, B.; Wold, S.; Andrews, P. Minimum analogue peptide sets (MAPS) for quantitative structure−activity relationshios. *Int. J. Pept. Protein Res.* **1991**, *37*, 414−424.
(39) van de Waterbeemd, H. Introduction. In *Chemometric Methods in Molecular Design*; van de Waterbeemd, H., Ed.; VCH Publishers: New York, 1995; pp 1−13.
(40) Sandberg, M.; Eriksson, L.; Jonsson, J.; Sjöström, M.; Wold, W. New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids. *J. Med. Chem.* **1998**, *41*, 2481−2491.
(41) Hellberg, S.; Sjostrom, M.; Skagerberg, B.; Wold, S. Peptide quantitative structure-activity relationships, a multivariate approach. *J. Med. Chem.* **1987**, *30*, 1126−1135.
(42) Jonsson, J.; Eriksson, L.; Hellberg, S.; Sjöström, M.; Wold, S. Multivariate parametrization of 55 coded and noncoded amino acids. *Quant. Struct.-Act. Relat.* **1989**, *8*, 204−209.
(43) Strøm, M. B.; Stensen, W.; Svendsen, J. S.; Rekdal, Ø. Increased antibacterial activity of 15-residue murine lactoferricin derivatives. *J. Pept. Res.* **2001**, *57*, 127−139.
(44) Siebert, K. J. Modeling protein functional properties from amino acid composition. *J. Agric. Food Chem.* **2003**, *51*, 7792−7797.
(45) *User Guide. SIMCA-P and SIMCA-P+ 10.0*; Umetrics AB: Umeå, Sweden, 2002.

(46) Cushman, D. W.; Cheung, H. S. Spectrophotometric assay and properties of the angiotensin-converting enzyme of rabbit lung. *Biochem. Pharmacol*. **1971**, *20*, 1637−1648.
(47) Mathews, D. M.; Adibi, S. A. Peptide absorption. *Gastroenterology* **1976**, *71*, 151−161.
(48) Buydens, L. M. C.; Reijmers, T. H.; Beckers, M. L. M.; Wehrens, R. Molecular data-mining: A challenge for chemometrics. *Chemom. Intell. Lab. Syst.* **1999**, *49*, 121−133.
(49) Wold, S.; Sjostrom, M.; Ericksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109−130.
(50) Collantes, E. R.; Dunn, W. J. Amino acid side chain descriptors for quantitative structure−activity relationship studies of peptide analogues. *J. Med. Chem*. **1995**, *38*, 2705−2713.
(51) Zaliani, A.; Gancia, E. MS-WHIM scores for amino acids: A new 3D-description for peptide QSAR and QSPR studies. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 525−533.
(52) Wold, S.; Eriksson, L. Statistical validation of QSAR results. In *Chemometrics Methods in Molecular Design*; van de Waterbeemd, H., Ed.; VCH Publishers: New York, 1995; Vol. 2, pp 309−318.
(53) Andersson, P. M.; Sjöström, M.; Lundstedt, T. Preprocessing peptide sequences for multivariate sequence-property analysis. *Chemom. Intell. Lab. Syst.* **1998**, *42*, 41−50.
(54) Wu, J. P.; Ding, X. L. Characterization of inhibition and stability of soy-protein-derived agiotensin I-converting enzyme inhibtiory peptides. *Food Res. Int.* **2002**, *35*, 367−375.
(55) Yust, M. M.; Pedroche, J.; Girón-Calle, J.; Alaiz, M.; Millán, F.; Vioque, J. Production of ACE inhibitory peptides by digestion of chickpea legumin with alcalase. *Food Chem.* **2003**, *81*, 363−369.

Received for review May 30, 2005. Revised manuscript received September 30, 2005. Accepted December 1, 2005. Financial support for this work was provided by a grant to R.E.A. and S.N. from the Advanced Foods and Materials Network (AFMnet), part of Canada's Networks of Centres of Excellence.

JF051263L