JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Quantitative Structure–Activity Relationship Study of Bitter Peptides

HYUN-OCK KIM AND EUNICE C. Y. LI-CHAN*

The University of British Columbia Faculty of Land and Food Systems Food, Nutrition and Health Program, FNH Building, 2205 East Mall, Vancouver, British Columbia V6T 1Z4, Canada

A database consisting of 224 di- to tetradecapeptides and five amino acids was compiled to study quantitative structure−activity relationships of bitter peptides. Partial least-squares regression-1 analysis was conducted using the amino acid three *z*-scores and/or three parameters (total hydrophobicity, residue number, and log mass values) as *X*-variables and bitterness values (log 1/*T* where *T* is the bitterness threshold) as *Y*-variables. Using the three parameters only, significant models ($p < 0.001$) were obtained describing the entire data set as well as data subsets, except that comprised only of octa- to tetradecapeptides. For data sets comprising different peptide lengths, the models were improved by including the three *z*-scores at the N-terminal and C-terminal positions. Correlation coefficients for bitterness prediction of 48 dipeptides and 12 pentapeptides were 0.75 (RMSEP = 0.53) and 0.90 (RMSEP = 0.48), respectively. Bulky hydrophobic amino acids at the C terminus and bulky basic amino acids at the N terminus were highly correlated to bitterness.

KEYWORDS: Bitterness; bitter peptides; QSAR; hydrophobicity; mass; residue number; *z*-scores; PLS regression

INTRODUCTION

It is widely known that bitterness is an undesirable outcome that is frequently generated during the enzymatic process to produce functional, bioactive protein hydrolyzates or during the aging process in fermented products such as cheese. Because bitterness decreases the value of these products, there have been many attempts to minimize bitterness (*1*−*4*). The relationships between the bitterness potency and the chemical structure of bitter peptides have been studied extensively by Japanese researchers. These studies have suggested that the hydrophobicity, primary sequence, spatial structure, peptide length, and bulkiness of the molecule are important in bitter taste perception (*5*−*12*).

In general, bitterness was reported to increase as the overall hydrophobicity of the peptide molecules increased (*6*−*8*, *11*, *13*, *14*). Ishibashi et al. (*14*) reported that for the bitter taste to be exhibited, the side chain skeleton of peptides containing the amino acids such as Gly, Ala, Val, and Ile should consist of at least three carbons. Bitter taste was observed in peptides containing Leu (*7*), Tyr, and Phe (*8*). Furthermore, the bitterness was more intense when the hydrophobic amino acid with the L-configuration was located at the C terminus (*7*, *8*, *15*) and with an increase in the number of hydrophobic amino acids in the C-terminal (*8*, *16*, *17*). Nosho et al. (*6*) reported that oligopeptides (Arg-Pro-Phe-Phe) having hydrophobic Phe-Phe at the C terminus exhibited bitterness that was 25 times greater than caffeine, but the bitterness completely vanished when the Phe-Phe was substituted by Gly-Gly. For the intense bitter taste of BPIa (Arg-Gly-Pro-Pro-Phe-Ile-Val), at least two hydrophobic amino acid residues at the C terminus were necessary (*18*).

In addition to overall hydrophobicity, the involvement of basic side chains and the location of basic and hydrophobic groups in the amino acid sequence of peptides are important parameters influencing the binding of peptides with the bitter taste receptors. It was reported that a basic moiety at the N-terminal and a hydrophobic moiety at the C-terminal were necessary for the bitterness of BPIa (Arg-Gly-Pro-Pro-Phe-Ile-Val) peptide (*19*, *20*) and the octapeptide Arg-Gly-Pro-Phe-Pro-Ile-Ile-Val (*13*) isolated from casein hydrolyzate. Similarly, Otagiri et al. (*5*) reported that hydrophobic amino acids located at the C-terminal as well as basic amino acids at the N-terminal are necessary for the bitterness, and furthermore, a strong bitter taste was observed when Arg was contiguous to Pro. On the other hand, a basic moiety at the C-terminal and a hydrophobic moiety at the N-terminal were important for the bitterness of BPIc (Val-Tyr-Pro-Phe-Pro-Pro-Gly-Ile-Asn-His) from casein hydrolyzate (*19*). Kim et al. (*12*) reported that many small bitter peptide fractions (<1000 Da) obtained from soybean proglycinin were composed of uncharged polar as well as hydrophobic amino acids, with a charged residue often being present at either end. Many bitter peptides isolated from soybean 11S glycinin were identified as basic mimics of the common structure, indicating the significance of the primary structure of the peptides in the bitter taste perception (*21*).

In addition to the presence of both basic and hydrophobic amino acids in the molecule, the spatial structure of the whole molecule is considered important for the bitterness of the

* To whom correspondence should be addressed. Tel: 1-604 822 6182. Fax: 1-604 822 5143. E-mail: Eunice.li-chan@ubc.ca.

10.1021/jf062422j CCC: $33.50
Published on Web 11/22/2006

peptides (*13*, *16*, *18*, *22*, *23*). It has been reported that bifunctional units, namely, a bulky basic or hydrophobic group as the stimulating unit and a hydrophobic group as the binding unit, are necessary participants in the mechanism for the bitter taste of peptides (*9*, *24*). Adjacency of these two sites in the steric conformation of peptides was essential (*25*), and the steric distance between two sites was estimated as 4.1 Å (*9*), with the pocket size as 15 Å (*11*). The bitter potency did not increase greatly if the peptides were larger than 15 Å (*11*).

Spatial configuration for the adjacency of bifunctional sites in the amino acid sequence is provided in some peptides by the presence of Pro. The imino ring of the L-Pro molecule induced bitterness of Pro-containing peptides through a conformational alteration leading to folding of the peptide backbone (*25*, *26*) and formation of a ball-like shape instead of a helix conformation (*11*). For example, the bitter taste of BPIa (Arg-Gly-Pro-Pro-Phe-Ile-Val) from casein hydrolyzate was due to the spatial structure attributed to the L-Pro at the 3-position (*26*). For the octapeptide Arg-Gly-Pro-Phe-Pro-Ile-Ile-Val, the location of a hydrophobic amino acid in the L-configuration between the two Pro residues was important to maintain the folded structure of the peptides to produce a strong bitterness (*27*). For the decapeptide BPIc (Val-Tyr-Pro-Phe-Pro-Pro-Gly-Ile-Asn-His), it was suggested that the Pro residues at positions 5 and 6 and the basic charge at the C-terminal, in addition to the hydrophobicity at the N-terminal, were necessary for the strong bitterness of this peptide (*28*).

Regarding the molecular size of the bitter peptides, the bitterness of the peptides was increased with an increasing number of amino acids up to eight (*5*, *11*), and there was no major difference of bitter potency when the peptides were composed of more than seven amino acids (*11*). The most bitter-tasting fractions from soybean proglycinin contained peptides with average molecular masses lower than 1700 Da (*12*), and a molecular range of 200−1400 Da, corresponding to a sequence of 2−12 amino acid residues, was obtained for the bitter peptides from soybean 11S glycinin (*21*).

Quantitative structure−activity relationship (QSAR) analysis has been used as a modeling and predictive tool for the functional activity of food proteins and peptides (*29*) and to find mathematical expressions to describe the structure−activity relationships of antimicrobial, ACE-inhibitory, and bitter-tasting peptides (*30*). Hellberg et al. (*31*) conducted pioneering research for QSAR of peptides by establishing a system to describe the 20 coded amino acids as three principal properties derived by principal components analysis of a matrix of 29 physicochemical variables. These three principal properties, often referred to as the three *z*-scores, represent mainly hydrophilicity/hydrophobicity (z_1), molecular size/bulkiness (z_2), and electronic properties/charge (z_3) of the amino acids. Hellberg et al. (*32*) applied partial least squares (PLS) regression to construct QSAR models of the data set of 48 bitter dipeptides compiled by Asao et al. (*10*). Application of new physiochemical descriptors of the amino acids for QSAR analysis of this data set of 48 bitter dipeptides has been examined extensively by researchers (*33*−*37*), while Asao et al. (*10*) used hydrophobicity and steric parameters in a QSAR study of 93 bitter amino acids, di- and tripeptides, and their derivatives. To our knowledge, the QSAR of bitter peptides including tetrapeptides or longer peptides has not been reported in the literature.

The objective of this study was therefore to elucidate the relationships between structure and bitterness of 224 peptides (di- to tetradecapeptides) and five amino acids whose bitterness values have been reported in the literature. PLS regression analysis was conducted to construct the QSAR models by using the three *z*-scores of amino acids, with or without three additional parameters, namely, total hydrophobicity, residue number (peptide length), and mass values. Models for subsets of the database comprising bitter peptides of the same peptide length as well as bitter peptides of different peptide lengths were validated, and the bitterness potency was predicted. Relationships between the type and the position of the amino acids in the primary sequence of the bitter peptides with the bitterness potency were examined.

MATERIALS AND METHODS

Preparation of Data Set. A database composed of 224 peptides and five amino acids with bitterness values determined by sensory evaluations was compiled from the published literature (**Table 1A**−**D**). The bitterness values were expressed as log $1/T$, where T is the bitter threshold concentration (M).

In addition to the whole data set (229 samples), subsets were evaluated, including di-, tri-, tetra-, penta-, hexa-, hepta-, octa-, and decapeptides and different combinations of the peptides with different peptide lengths (di- to tetradeca-, tetra- to octa-, tetra- to deca-, tetra- to tetradeca-, and octa- to tetradecapeptides). Peptides with high bitterness values (log $1/T \geq 3.7$) and peptides with R at the N terminus (n_1) (R peptides) were also selected for analysis.

In the case of the dipeptides, bitterness values were available for a total of 77 dipeptides, including the 58 listed in **Table 1A** and the 19 shown in **Table 1B**. Two sets of bitterness values were obtained from the literature for 19 of these dipeptides (**Table 1B**), and the mean values of log $1/T$ were calculated for QSAR of the data sets composed of 224 peptides. In order to predict the 48 dipeptides data set compiled by Asao et al. (*10*) (**Table 1A**,**B**), these 48 dipeptides were excluded from the data set (leaving a total of 176 samples) for the construction of the calibration model.

PLS Regression Analysis. PLS-1 regression was used to examine the correlations between the properties or the position of the amino acids in the peptides and the bitterness values of the peptides using the software The Unscrambler (version 9.0, CAMO Inc., Corvallis, OR).

The total hydrophobicity values were calculated using the amino acid hydrophobicity coefficient scale 1 of Wilce et al. (*38*). Because the scale of mass (M) values (115.12−1660.98 Da) of the samples were much larger as compared to the other variables, the log-transformed values (log M) were used as the X-variables. The three *z*-scores, namely, z_1 (hydrophobicity), z_2 (bulkiness/molecular size), and z_3 (electronic property) scores from Hellberg et al. (*31*), were applied to the description of the amino acids. The amino acid at the first position from the N terminus was designated as n_1, and its three *z*-score properties were described as $n_1\ z_1$, $n_1\ z_2$, and $n_1\ z_3$. Amino acid residues at the second, third, fourth, and fifth positions from the N terminus were designated as n_2, n_3, n_4, and n_5. Similarly, amino acid residues at the first, second, third, fourth, and fifth positions from the C terminus were designated as c_1, c_2, c_3, c_4, and c_5.

The *z*-scores and/or the total hydrophobicity, residue number of the peptides (peptide length), and log M were used as X-variables, and the bitterness values (log $1/T$) were used as Y-variables. The models were constructed for the data subsets composed of peptides of the same length or of different peptide lengths and were validated using full cross-validation. For the peptide sets composed of dipeptides and peptides with longer lengths (tri- to tetradecapeptides), the amino acid *z*-scores were applied to only the n_1 and c_1 positions of the peptides. For the peptide sets composed of tetrapeptides and peptides with longer lengths (penta- to tetradecapeptides), the amino acid *z*-scores were applied to n_1, n_2, c_2, and c_1 positions of the peptides. For the peptide sets composed of octapeptides and peptides with longer lengths (nona- to tetradecapeptides), amino acid *z*-scores were applied to n_1, n_2, n_3, n_4, c_4, c_3, c_2, and c_1 positions of the peptides. For peptides (≥tetrapeptides) with high bitterness values (log $1/T \geq 3.7$), amino acid *z*-scores were applied to n_1, n_2, c_2, and c_1 positions and for R peptides (≥tripeptides and R at n_1), amino acid *z*-scores were applied to c_2, c_1 positions of the peptides. Variables were used without scaling for the PLS regression analyses.

Table 1. Bitter Amino Acids and Peptides Used for QSAR Analysis

(**A**) Amino Acids and Dipeptides

sample	log 1/*T*	literature ref	sample	log 1/*T*	literature ref	sample	log 1/*T*	literature ref
R	1.6	*5*	YG	2.52	*8*	IQ	1.49	*10*
F	1.7	*5*	VY	2.52	*8*	SL	1.49	*10*
L	1.7	*7*	VF	2.52	*8*	IN	1.49	*10*
V	1.7	*14*	PR	2.52	*5*	WE	1.56	*10*
P	1.9	*5*	LE	2.52	*44*	IK	1.65	*10*
GR	1	*5*	KP	2.52	*5*	IA	1.68	*10*
YP	1.7	*25*	RF	2.6	*5*	AL	1.7	*10*
IV	1.9	*14*	YY	2.63	*8*	VV	1.71	*10*
VD	1.9	*14*	IF	2.83	*8*	LA	1.72	*10*
KF	2.04	*8*	GE	2.83	*44*	PY	1.8	*10*
RG	2.11	*5*	FI	2.83	*8*	GW	1.89	*10*
RR	2.11	*5*	RP	3.1	*5*	IV	2.05	*10*
LD	2.23	*44*	YF	3.1	*8*	PL	2.22	*10*
VI	2.23	*14*	AV	1.16	*10*	PI	2.33	*10*
VE	2.23	*14*	VA	1.16	*10*	IP	2.4	*10*
PK	2.23	*25*	VG	1.19	*10*	YL	2.4	*10*
LV	2.23	*14*	PA	1.32	*10*	LY	2.46	*10*
AD	2.23	*44*	ID	1.37	*10*	IW	3.05	*10*
FV	2.23	*14*	IE	1.37	*10*	FY	3.13	*10*
PP	2.34	*5*	IS	1.49	*10*	LW	3.4	*10*
LI	2.4	*7*	IT	1.49	*10*	WW	3.6	*10*

(**B**) Dipeptides from Different Refs

dipeptide sample	log 1/*T* (mean)[a]	log 1/*T*	literature refs	dipeptide sample	log 1/*T* (mean)[a]	log 1/*T*	literature refs
GL	1.64	1.68[b] (1.6)	*10* (*9*)	GI	2.17	1.7 (2.64)	*10* (*9*)
LG	1.71	1.72 (1.7)	*10* (*9*)	GF	2.36	1.8 (2.92)	*10* (*5*)
GV	1.74	1.13 (2.34)	*10* (*14*)	LL	2.47	2.35 (2.6)	*10* (*9*)
GP	1.79	1.35 (2.23)	*10* (*5*)	II	2.54	2.26 (2.83)	*10* (*9*)
AF	1.81	1.72 (1.9)	*10* (*8*)	IL	2.54	2.26 (2.83)	*10* (*7*)
FG	2.0	1.77 (2.23)	*10* (*5*)	FP	2.77	2.7 (2.83)	*10* (*25*)
IG	2.01	1.68 (2.34)	*10* (*9*)	LF	2.82	2.75 (2.89)	*10* (*8*)
VL	2.11	2.0 (2.23)	*10* (*14*)	FL	2.85	2.87 (2.83)	*10* (*8*)
PF	2.14	2.8 (1.48)	*10* (*5*)	FF	3.01	3.1 (2.92)	*10* (*5*)
GY	2.15	1.77 (2.52)	*10* (*8*)				

(**C**) Tripeptides and Tetrapeptides

sample	log 1/*T*	literature ref	sample	log 1/*T*	literature ref	sample	log 1/*T*	literature ref
LGG	1	*7*	PGI	2.63	*25*	YYY	3.7	*8*
GGV	1.48	*14*	FFG	2.65	*5*	FFF	3.7	*5*
PGR	1.6	*5*	PPP	2.7	*5*	GGLG	1.6	*7*
GPG	1.7	*25*	RPF	2.83	*5*	GLGG	1.7	*7*
GYG	1.7	*8*	EGG	2.83	*44*	LGGG	1.9	*7*
RGP	1.9	*5*	FIV	2.83	*14*	GGGL	2.34	*7*
GLG	2	*7*	GGF	2.83	*5*	PFPP	2.34	*25*
GGL	2	*7*	GGY	2.83	*8*	FFGG	2.52	*24*
GGP	2.04	*25*	GLL	2.83	*7*	FFPP	2.52	*8*
PGP	2.04	*25*	PIP	2.85	*45*	GPPF	2.52	*8*
PPG	2.04	*25*	VIF	2.89	*8*	RRPP	2.7	*24*
LLG	2.3	*7*	LLL	2.92	*7*	FFPE	2.76	*24*
LGL	2.3	*7*	FGF	2.92	*5*	GGFF	2.85	*24*
FGG	2.34	*5*	YGY	3.1	*8*	FFPG	2.9	*24*
FPP	2.34	*8*	GRP	3.1	*5*	LLLL	3.23	*7*
GVV	2.34	*14*	DLL	3.1	*44*	RPFG	3.41	*6*
PGG	2.34	*8*	RPG	3.1	*5*	FGFG	3.52	*11*
VVV	2.34	*24*	YYG	3.2	*8*	VYPF	3.52	*8*
RRR	2.4	*5*	GFF	3.23	*5*	PFIV	3.52	*18*
VYP	2.52	*25*	PFP	3.4	*25*	GPFF	3.8	*6*
KPK	2.52	*25*	KPF	3.4	*8*	RGFF	3.8	*6*
GFG	2.52	*5*	GYY	3.4	*8*	RPGF	3.8	*6*
FPK	2.52	*8*	FPF	3.4	*8*	FGGF	3.92	*11*
YGG	2.63	*8*	ELL	3.4	*44*	FFPR	4	*6*
PPF	2.63	*8*	YPF	3.52	*25*	RPFF	4.4	*8*

Table 1. Continued

(**D**) Penta-, Hexa-, Hepta-, Octa-, Nona-, Deca-, Undeca-, Dodeca-, and Tetradecapeptides

sample	log 1/*T*	literature ref	sample	log 1/*T*	literature ref	sample	log 1/*T*	literature ref
GGGLG	1.9	*7*	RPGGFF	4.04	*6*	VYPFPPGI	3.82	*28*
GGLGG	1.9	*7*	GGRPFF	4.04	*24*	VIIPFPGR	3.85	*13*
LGGGG	1.9	*7*	RPPFIV	4.1	*18*	RGPKPIIV	4.08	*27*
GLGGG	1.9	*7*	RGPPFF	4.23	*17*	RGPPGGFF	4.11	*17*
GGVVV	2.11	*24*	RGPFIV	4.3	*18*	RGPPFIIV	4.3	*27*
RGPPF	2.63	*18*	RRPPGF	4.4	*15*	GGRPFFGG	4.4	*24*
GGGGL	2.65	*7*	RGPPFI	4.6	*18*	RGPEPIIV	4.51	*27*
FFPGG	2.83	*24*	RRPPFF	5.15	*5*	RGPGPIIV	4.81	*11*
PPFIV	2.92	*18*	RGPPGGV	2.48	*17*	RPFFRPFF	5	*11*
PGPIP	3.11	*45*	RGPPGIG	2.78	*17*	RGPFPIIV	5.4	*27*
RPGFF	3.51	*6*	RGPPGGF	3.08	*17*	RRPPPFFF	5.7	*5*
RRPFF	4.7	*8*	RGPPFGG	3.23	*17*	RGPPGGGFF	3.95	*17*
PGPGPG	2.6	*45*	VYPFPPG	3.52	*28*	GGRGPPFIV	4.1	*22*
VIFPPG	2.68	*19*	VIIPFPG	3.6	*13*	RGPPFIVGG	4.31	*22*
GPPFIV	2.92	*18*	PFPGPIP	3.6	*45*	VYPFPPGIGG	3.52	*28*
RGGFIV	3.1	*26*	RGPPGFG	3.68	*17*	VYPFGGGINH	3.64	*28*
PVLGPV	3.3	*13*	YPFPGPI	3.8	*45*	VYPFPPIGNH	4.3	*8*
PFPGPI	3.36	*45*	RGPFPIV	3.95	*13*	VYPFPPGINH	4.3	*5*
RGPPGF	3.52	*15*	VIFPPGR	4.1	*19*	FFRPFFRPFF	5.15	*6*
FPPFIV	3.52	*20*	VIPFPGR	4.15	*13*	PVRGPFPIIV	5.4	*11*
GGFFGG	3.7	*24*	RGPPFIV	4.3	*5*	GGRGPPFIVGG	4.4	*22*
KPPFIV	3.82	*20*	RGPPGFF	4.4	*17*	RPFFRPFFRPFF	5	*11*
PFPIIV	3.9	*13*	RPPPFFF	4.7	*5*	RGPPFIVRGPPFIV	4.4	*11*
RPFFGG	3.92	*24*	RGPPFFF	5	*17*	PVLGPVRGPFPIIV	4.83	*11*

[a] Mean log 1/*T* values used for sample sets 1 and 2 (**Table 2**) and data sets of dipeptides B and C (**Table 3**). [b] log 1/*T* values used for data set of dipeptides A.

However, all *X*- and *Y*-variables were weighted (standardized to the same scale by dividing with the standard deviation) in order to study the relative influence on bitterness of the three *z*-scores or properties of the amino acids at specific positions of the peptide sequences.

RESULTS AND DISCUSSION

Characteristics of Peptides in the Database. Figure 1 shows the histograms of the 224 peptides and five amino acids in the database, according to bitterness values as log 1/*T*, total hydrophobicity, log *M*, and residue number. In this database, the bitterness values expressed as log 1/*T* ranged from 1.0 for GR and LGG to 5.7 for RRPPPFFF. The total hydrophobicity values ranged from −2.55 for RRR to 28.52 for FFRPFFRPFF. The values of log *M* varied from 2.06 (M = 115.12 Da) for the imino acid P to 3.22 (M = 1660.98 Da) for the dodecapeptide RPFFRPFFRPFF. The residue number (length) of the peptides varied from two to 14, and 56% of the peptides in the database consisted of di- and tripeptides. Furthermore, the database included 95 bitter peptides composed of tetra- to tetradecapeptides, contrary to the suggestion of Asao et al. (*10*) that there are few bitter compounds that are equal to or larger in size than tetrapeptides. In fact, as shown in **Table 1**, the eight peptides having highest bitterness intensity (log 1/*T* values of 5.0 or higher) were a hexapeptide, a heptapeptide, three octapeptides, two decapeptides, and one dodecapeptide (**Table 1D**).

Relationships of Bitterness with Total Hydrophobicity, Residue Number, and Mass. Figure 2 shows highly significant correlations ($p < 0.001$) of the bitterness values with total hydrophobicity ($R^2 = 0.56$), residue number ($R^2 = 0.59$), log *M* ($R^2 = 0.75$), and mass ($R^2 = 0.75$) by using polynomial models.

The positive correlations of these parameters with the bitterness indicated that total hydrophobicity, residue number, and mass (log *M*) contribute to the bitterness values of the peptides. This result was in agreement with previous findings that total hydrophobicity and length of the peptides were important factors for bitterness (*10*, *11*, *39*). For example, a positive correlation ($R^2 = 0.791$) was observed between the level of hydrophobic peptides in pasteurized milk cheese and the mean panel bitterness scores (*40*). Gulyaeva et al. (*41*) reported that the peptide bitterness threshold was quantitatively related to the peptide structure described as a combination of the relative hydrophobicity and lipophilicity of peptides. As shown in **Figure 2**, the bitterness of the peptides was increased largely with increased residue number (peptide length) up to 8−10, and there was little effect of the longer peptides. Tamura et al. (*11*) reported that bitterness increased largely when peptides are composed of less than eight amino acids. Although positive correlations were observed between the bitterness values with both log *M* and mass values, **Figure 2** showed that the correlation was primarily for the mass values with molecular masses up to 1000 Da (8−10 residues). Kukman et al. (*42*) also reported that the bitterness of the peptides produced from soybean protein was mainly caused by hydrophobic bitter peptides of molecular masses less than 1000 Da.

PLS regression analysis was conducted using total hydrophobicity, residue number, and log *M* values for each sample, correlated with their bitterness intensity values. Initially, the PLS regressions were conducted using the bitterness values expressed as R_{caf} (bitterness intensity as compared to the threshold concentration for 1 mM caffeine standard, which is assigned a R_{caf} value of 1.0). However, correlation coefficients for calibrations and validations obtained using R_{caf} as bitterness values were lower than those obtained using log 1/*T* values, where *T* is the threshold molar concentration.

PLS regression results for the different sample sets (composed of varying subsets of the 224 peptides and five amino acids) using total hydrophobicity, residue number, and log *M* values as *X*-variables and the bitterness values, log 1/*T*, as *Y*-variables are shown in **Table 2A**. When mass values instead of log *M*

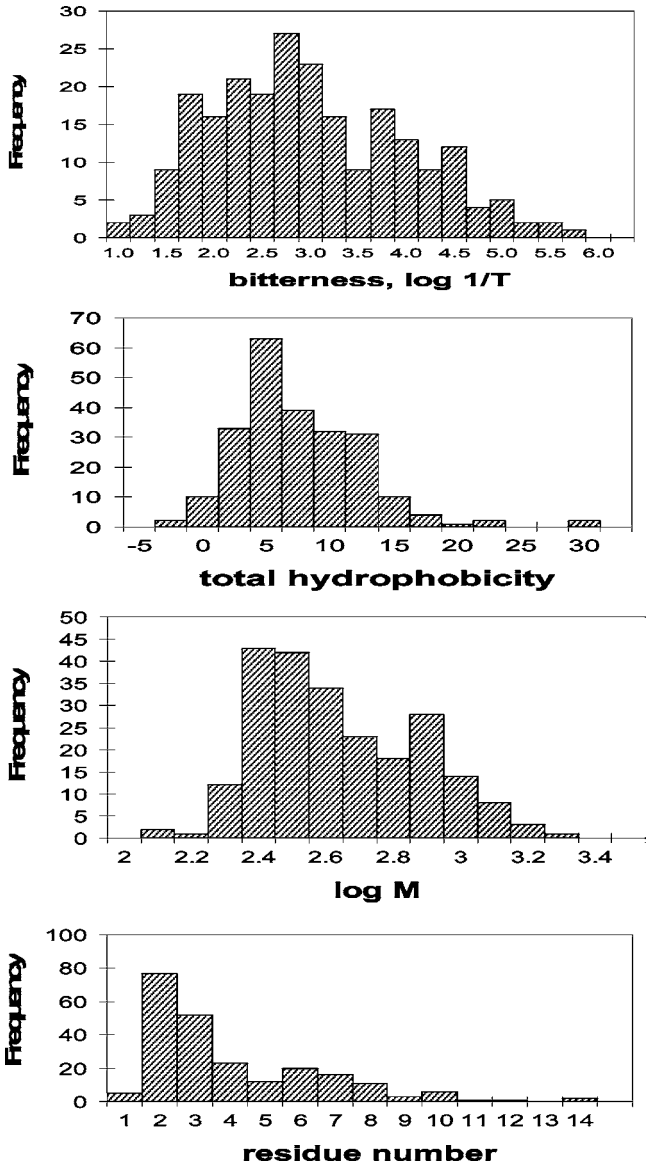

Figure 1. Histograms of bitterness values (log 1/*T*), total hydrophobicity, log *M*, and residue number for bitter peptides data base (229 samples).

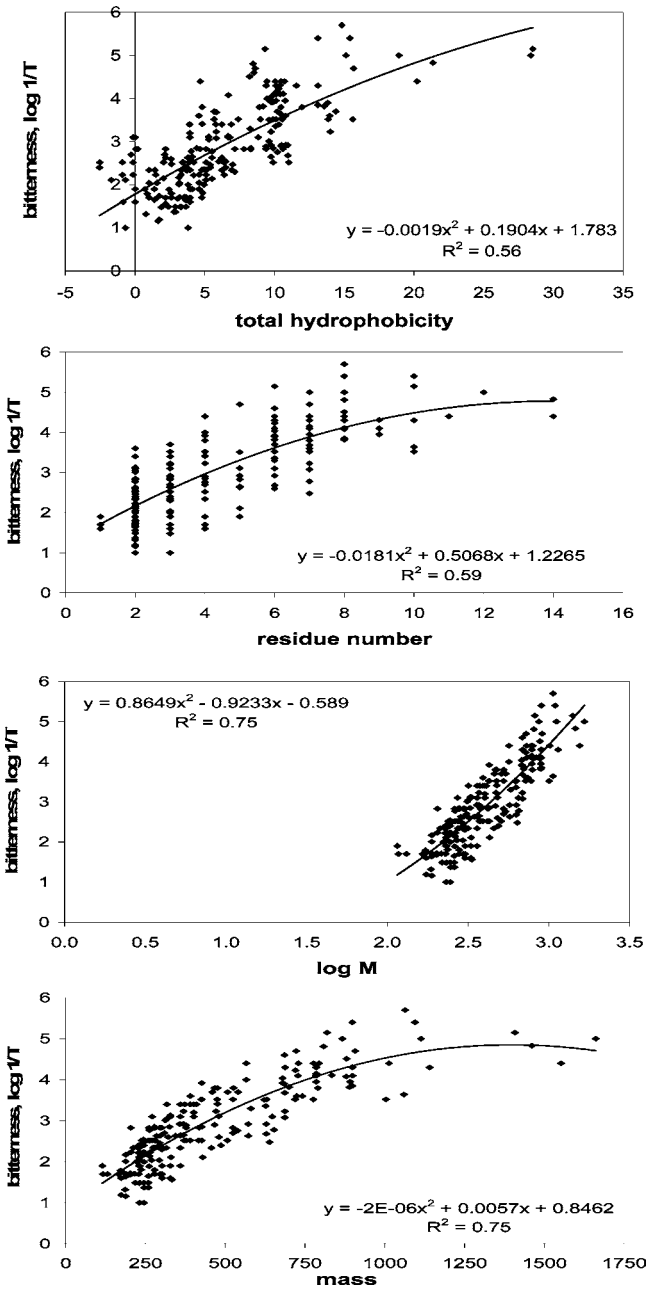

Figure 2. Correlations of the bitterness values with total hydrophobicity, residue number, log *M*, and mass values for bitter peptides data base (229 samples).

were used for regression analysis, with the exception of R peptides (sample set 9), in general, slightly higher correlation coefficients for the calibrations and validations were obtained (data not shown).

Highly significant ($p < 0.001$) correlation coefficients were obtained for both calibration ($R = 0.68-0.81$) and cross-validation (RCV $= 0.65-0.80$) using total hydrophobicity, residue number, and log *M* values for all these sample sets, with the exception of the subset comprising octa- and longer peptides (set 7), and the highly bitter peptides data set with log $1/T \geq 3.7$ (set 8). The lower correlation coefficients of calibration and validation for the octapeptides and longer peptides (set 7) may be explained by the findings that bitterness potency did not increase with increasing number of amino acids beyond 8−10 residues (**Figure 2**). For the highly bitter peptide set (set 8, with log $1/T$ of 3.7−5.7), which was composed of peptides varying in length from tri- to tetradecapeptides, other structural parameters including those related to the position of the hydrophobic residues may be necessary for explaining the high bitterness intensity.

QSAR Analyses Using *z*-Scores. Table 2B shows PLS regression results using *z*-scores only or *z*-scores with total hydrophobicity, log *M*, and residue number values as *X*-variables for the peptide data sets with different peptide lengths. Because the peptides in the data sets had different peptide lengths, amino acid *z*-scores at the specified N-terminal and C-terminal positions in the peptides (as described in the Materials and Methods and in the footnote to **Table 2B**), along with total hydrophobicity, log *M*, and residue number values of the whole peptide molecules were used as *X*-variables.

By using *z*-scores together with total hydrophobicity, residue number, and log *M* values as *X*-variables, all of the peptide data sets showed improved correlation coefficients for the calibrations and the validations as compared to those obtained by using *z*-scores only. The improvements of the correlation coefficients for the calibrations and validations were more pronounced for the peptide data sets with a large range of peptide lengths, such

Table 2. Results of PLS Regression of Bitter Peptides Using (**A**) Total Hydrophobicity, Residue Number, and log *M* or (**B**) *z*-Scores without or with Total Hydrophobicity, Residue Number, and log *M* as *X*-Variables[a]

sample set	sample no.	PLS regression results[b,c] PCs[d]	*R*[e]	RCV[f]
	part **A**[b]			
(#1) peptides (224) + amino acids (5)	229	2	0.81***	0.80***
(#2) peptides (224) (di- to tetradecapeptides)	224	2	0.80***	0.79***
(#3) peptides (176) (di- to tetradecapeptides)	176	2	0.78***	0.77***
(#4) tetra- to octapeptides	82	2	0.71***	0.69***
(#5) tetra- to decapeptides	91	2	0.70***	0.67***
(#6) tetra- to tetradecapeptides	95	2	0.68***	0.65***
(#7) octa- to tetradecapeptides	24	1	0.44^{NS}	0.31^{NS}
(#8) highly bitter peptides (log $1/T \geq 3.7$)	51	1	0.51***	0.46***
(#9) R peptides (R at n_1)	49	1	0.79***	0.76***
	part **B**[c]			
(#2)	224	1 (2)	0.53*** (0.87)***	0.51*** (0.86)***
(#3)	176	1 (2)	0.51*** (0.86)***	0.47*** (0.84)***
(#4)	82	5 (3)	0.74*** (0.88)***	0.64*** (0.84)***
(#5)	91	2 (3)	0.63*** (0.87)***	0.54*** (0.84)***
(#6)	95	1 (3)	0.56*** (0.84)***	0.49*** (0.79)***
(#7)	24	1 (1)	0.68*** (0.69)***	0.42* (0.50)*
(#8) ≥tetrapeptides	49	1 (3)	0.43** (0.76)***	0.27^{NS} (0.61)***
(#9) ≥tripeptides	44	3 (1)	0.66*** (0.77)***	0.53*** (0.72)***

[a] NS, not significant; significant at $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$. [b] PLS regression results using total hydrophobicity, residue number, and log *M* values (**A**). [c] PLS regression results using *z*-scores without or with (in parentheses) total hydrophobicity, residues number, and log *M* values (**B**). The *z*-scores were at n_1, c_1 for sets 2 and 3; n_1, n_2, c_2, and c_1 for sets 4–6 and 8; n_1–n_4, c_4–c_1 for set 7; and c_2, c_1 for set 9. [d] Number of PLS components used in regression analyses. [e] Multivariate correlation coefficients for calibration set. [f] Multivariate correlation coefficients of the cross-validation.

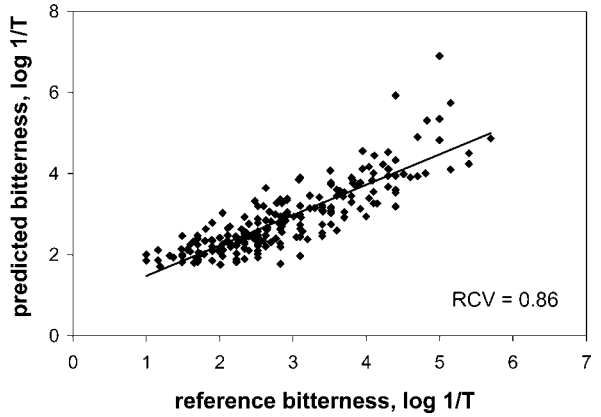

Figure 3. Plot of predicted and reference bitterness values for 224 bitter peptides including di- to tetradecapeptides analyzed by PLS regression with full cross-validation using amino acid *z*-scores at n_1, c_1 positions, total hydrophobicity, log *M*, and residue number values as *X*-variables.

as for peptide data sets 2 and 3 (di- to tetradecapeptides). Correlation of cross-validation between the predicted bitterness values and the reference bitterness values using sample set 2 composed of 224 peptides is shown in **Figure 3**. With the exception of the R peptides (sample set 9), the correlation coefficients for the calibrations and validations obtained by using all of these parameters (**Table 2B**) for all of these data sets were also higher than those obtained by using total hydrophobicity, residue number, and log *M* values only as *X*-variables (**Table 2A**). This result showed that total hydrophobicity, residue number, and log *M* (or mass) values can be used in addition to the amino acid *z*-scores for better QSAR modeling when the data sets include peptides with different peptide lengths.

PLS regression results for the data subsets comprised of bitter peptides with the same lengths, using *z*-scores only or *z*-scores with total hydrophobicity and log *M* values as *X*-variables, are shown in **Table 3**. Using the three *z*-scores only as *X*-variables, the correlation coefficients for calibrations (*R*) ranged from 0.63 to 0.95, and the cross-validations (RCV) ranged from 0.52 to

Table 3. PLS Regression Results for Data Sets of Bitter Peptides of the Same Length, Using the Three *z*-Scores without or with Total Hydrophobicity and log *M* Values as *X*-Variables[a]

sample set	sample no.	PLS regression results[b] PCs[c]	*R*[d]	RCV[e]
dipeptides A[f]	48	2 (1)	0.91*** (0.92)***	0.88*** (0.90)***
dipeptides B[g] (average)	48	3 (1)	0.85*** (0.85)***	0.80*** (0.82)***
dipeptides C[h]	77	1 (2)	0.63*** (0.68)***	0.57*** (0.60)***
tripeptides	52	1 (1)	0.71*** (0.75)***	0.62*** (0.65)***
tetrapeptides	23	4 (4)	0.90*** (0.92)***	0.71*** (0.75)***
pentapeptides	12	1 (1)	0.88*** (0.89)***	0.74** (0.76)**
hexapeptides	20	1 (1)	0.75*** (0.76)***	0.52* (0.49)*
heptapeptides	16	3 (3)	0.95*** (0.95)***	0.77*** (0.82)***
octapeptides	11	1 (1)	0.69* (0.73)*	0.15^{NS} (0.15)^{NS}
decapeptides	6	1 (1)	0.94** (0.89)*	0.76^{NS} (0.64)^{NS}

[a] NS, not significant; significant at $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$. [b] PLS regression results using three *z*-scores only or three *z*-scores with total hydrophobicity and log *M* values (in parentheses). [c] Number of PLS components used in regression analyses. [d] Multivariate correlation coefficients for calibration set. [e] Multivariate correlation coefficients of the cross-validation. [f] Dipeptides (*48*) data set compiled by Asao et al. (*10*), using bitterness values reported by these authors. [g] Dipeptides (*48*) data set compiled by Asao et al. (*10*), using averaged bitterness values for the 19 samples shown in **Table 1B**. [h] Dipeptides (*77*) data set, using averaged bitterness values for the 19 samples shown in **Table 1B**.

0.88 for all of the data sets comprised of peptides with same length, with the exception of the data sets for octapeptides and decapeptides. In general, inclusion of total hydrophobicity and log *M* values to the *z*-scores as *X*-variables led to little improvements in the correlation coefficients for the calibrations ($R = 0.68-0.95$) and cross-validations (RCV = 0.49−0.90). This result indicated that the *z*-scores used for QSAR analyses for these peptide sets were sufficient to represent the hydrophobicity and bulkiness properties of the peptides.

Using the amino acid three *z*-scores of Hellberg et al. (*31*), correlation coefficients for the calibration ($R = 0.91$) and cross-validation (RCV = 0.88) were obtained for the sample set of 48 dipeptides compiled by Asao et al. (*10*) (**Table 3**, dipeptides

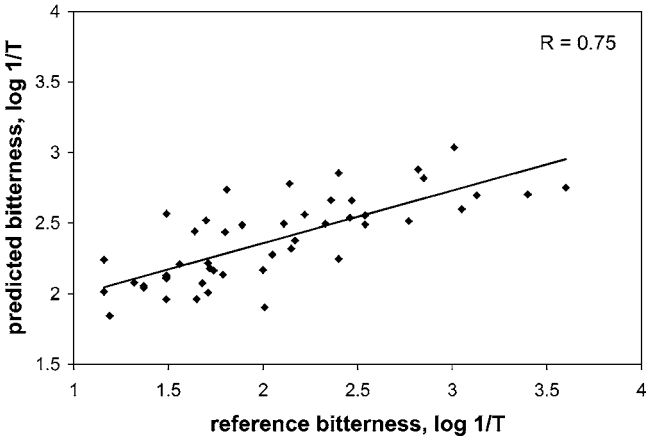

Figure 4. Correlation between the predicted and the reference bitterness values for 48 dipeptides B data set by using the calibration model constructed using the 176 peptides data set.

A). These results are similar to the findings of Jonsson et al. (*33*), who applied three extended *z*-scales (*z*′) to this 48 dipeptide data set and reported a correlation of 0.88 (RCV) between observed bitterness and calculated bitterness. In the present study, the 48 dipeptide data set of Asao et al. (*10*) was compared to dipeptide data reported by other researchers, and the mean bitterness values were calculated when different bitterness values were reported for the same dipeptide (**Table 1B**). By using these averaged bitterness values for the 48 dipeptides data set, *R* and RCV of 0.85 and 0.80, respectively, were obtained (**Table 3**, dipeptides B), as compared to *R* and RCV of 0.63 and 0.57 for the whole 77 dipeptide data set (**Table 3**, dipeptides C). Despite possible variability in bitterness values resulting from sensory evaluation conducted by different research groups, the correlation coefficients for all three dipeptide sample sets were highly significant (**Table 3**), and the inclusion of the 48 dipeptides data to the whole peptides data set did not affect the correlation coefficients for calibration nor the validation, as shown by the results for sample sets 2 and 3 (**Table 2**).

The low and nonsignificant ($p > 0.05$) correlation coefficients of cross-validations obtained for both the octapeptides and the decapeptides data sets, may have been due to the small number of samples with high bitterness range (log 1/*T*, 3.82−5.7 for octapeptides and log 1/*T*; 3.52−5.4 for decapeptides) in the sample set.

Prediction of Bitterness by QSAR Models. External validation was conducted by constructing QSAR models to predict bitterness of two sets of peptides, which were not previously included for the calibration. The dipeptides B data set (48 samples, **Table 3**) was excluded from the peptide data set (224 samples), and the remaining 176 peptides were used to develop a calibration model using amino acid *z*-scores at n_1, c_1 positions together with total hydrophobicity, log *M*, and residue number values as *X*-variables (set 3, **Table 2B**). The resulting model with two PLS components explained 74% of the variance in the *Y*-variable (bitterness intensity) of the 176 samples and was used to predict the bitterness values of the excluded 48 dipeptides B data set. The correlation coefficient for the prediction was 0.75 ($p < 0.001$) with RMSEP of 0.53 (**Figure 4**).

Pentapeptides (12 samples) were excluded from the data set 4 (82 samples) composed of tetra- to octapeptides, and the remaining 70 peptides were used to construct a calibration model using *z*-scores for n_1, n_2, c_1, and c_2 together with total hydrophobicity, log *M*, and residue number values as *X*-variables. The correlation coefficients of calibration and cross-validation for the model made from the 70 peptides were 0.86 and 0.81, respectively ($p < 0.001$), and the two PLS components of this model explained 72% of the variance in the *Y*-variable (bitterness intensity). When this model was used to predict bitterness of the excluded pentapeptides, the correlation coefficient for prediction was 0.90 ($p < 0.001$), with a RMSEP of 0.48 (data not shown).

As shown in **Tables 1D** and **3**, there is limited data available on longer peptides with higher bitterness values for construction of QSAR models. In the present study, using the currently available data, the QSAR model derived from 224 peptides (**Table 2**, set 2) could be useful especially for the prediction of bitterness in peptides up to 8−10 residues in length (**Figure 2**) with the expected bitterness values (log 1/*T*) lower than 4.5 (**Figure 3**). The QSAR model is given in the following equation:

$$\text{bitterness (log } 1/T) = 1.87 + 0.08_{n_1z_1} + 0.07_{n_1z_2} - 0.04_{n_1z_3} - 0.02_{c_1z_1} + 0.03_{c_1z_2} + 0.01_{\log M} + 0.11_{\text{totalHP}} + 0.09_{\text{residuenumber}}$$

where n_1z_1, n_1z_2, and n_1z_3 are the *z*-scores for the amino acid in the N-terminal position and c_1z_1 and c_1z_2 are the *z*-scores for the amino acid position in the C-terminal position, respectively. Using the approach reported in this study, better QSAR models could be constructed in the future for the prediction of bitterness in peptides by incorporating additional data that may be generated through further research on the longer peptides with higher bitterness values.

Relationship of Bitterness with Amino Acids in the Peptide Sequences. Regression coefficients obtained by PLS regression analysis of the weighted *X*- and *Y*-variables (standardized to the same scale by dividing with the standard deviation) were examined for relative importance of the *X*-variables in the PLS regression models. Typical PLS weighted regression coefficients plots for di-, tri- tetra-, penta-, hexa-, and heptapeptides data sets by using *z*-scores only as *X*-variables are shown in **Figure 5A−F**.

For the sample set composed of 77 dipeptides, the hydrophobicity (z_1) at $c_1 > n_1$ and size/bulkiness (z_2) at n_1 and c_1 positions were important for the prediction of the bitterness. There was also some contribution of electronic effects/charge (z_3) for the c_1 position. This result was in agreement with the finding by Hellberg et al. (*32*) that the most important factors in the model were hydrophobicity (z_1) and size (z_2) for both amino positions of the 48 dipeptides data set compiled by Asao et al. (*10*). Using other descriptors such as the isotropic surface area and electronic charge index (*34*) or MS-weighted holistic invariant molecular (WHIM) scores (*35*) or VHSE (principal components score vectors of hydrophobic, steric, and electronic properties) (*37*), it was also reported that highly bitter dipeptides should have hydrophobic amino acids at both positions (*34*, *37*) or bulkiness at the c_1 position (*35*) as well as polar/charged amino acids at the n_1 position (*34*, *35*, *37*).

For the sample set of tripeptides, bulky amino acids at n_2, $c_1 > n_1$ and hydrophobic amino acids at $c_1 > n_2$ were important for the prediction of the bitterness. As for tetrapeptides, a basic, bulky, hydrophobic amino acid at c_1 and a bulky basic amino acid at n_1 were important, whereas for the prediction of the bitterness of pentapeptides, bulky hydrophobic amino acids at $c_1 > c_2$ and bulky basic amino acids at $n_1 > n_2$ were important. For the prediction of hexapeptides, bulky hydrophobic amino acids at c_1 and bulky basic and hydrophilic amino acid at n_1

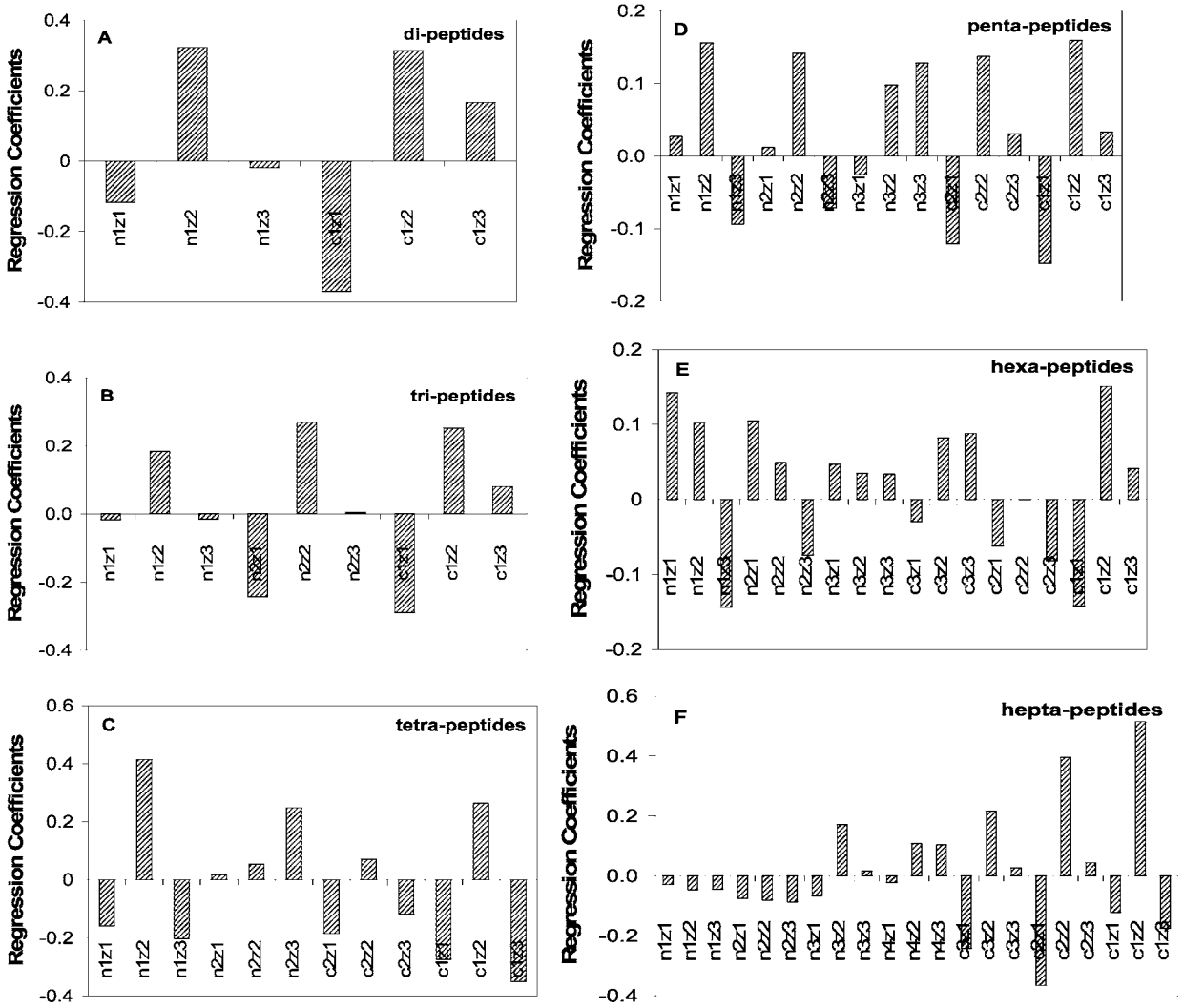

Figure 5. PLS regression coefficient (weighted) plots for (**A**) di-, (**B**) tri-, (**C**) tetra-, (**D**) penta-, (**E**) hexa-, and (**F**) heptapeptides data sets by using *z*-scores.

were important, while for heptapeptides, bulky basic amino acids at c_1 with bulky hydrophobic amino acids at $c_2 > c_3$ were important.

For the prediction of bitterness of R peptides (44 samples, set 9, **Table 2B**), hydrophobic amino acids at $c_2 > c_1$ were important (data not shown).

Typical PLS weighted regression coefficient plots using amino acid *z*-scores only (**A**) and *z*-scores with total hydrophobicity, log *M*, and residue number values (**B**) as *X*-variables are shown in **Figures 6** and **7** for the 224 sample set composed of di- to tetradecapeptides and the 95 sample set composed of tetra- to tetradecapeptides, respectively.

For the 224 peptide sample set, hydrophobic amino acid at c_1 and bulky basic and hydrophilic amino acid at n_1 were important for the prediction of the bitterness. For the 95 peptide sample set, bulky basic amino acids with hydrophobicity at c_1 and bulky basic amino acids at n_1 were important. It was interesting that hydrophilic amino acids were found at the N terminus of the bitter peptides. Although it has been reported that the bitterness was principally proportional to the content of hydrophobic amino acids, bitterness also was correlated with the presence of uncharged polar amino acid residues (*12*). Hydrophilic amino acids may affect the overall taste of the peptides as suggested by Belitz and Wieser (*43*), and the polar amino acids probably affect taste quality of the bitter peptides (*10*).

When total hydrophobicity, log *M*, and residue number were used together with *z*-scores as *X*-variables for the above two data sets, PLS regression results showed high positive values for the weighted regression coefficients of these three parameters, which were greater in magnitude than any of the *z*-score coefficients (**Figures 6B** and **7B**). These results suggest that log *M*, total hydrophobicity, and residue number, i.e., parameters describing the overall rather than sequence-specific properties of the samples, may be dominant factors in prediction of bitterness values. The results also are consistent with the finding that these three parameters were sufficient for establishing the PLS models (**Table 2A**). Nonetheless, analysis of the weighted regression coefficients for the *z*-scores provides further information on the relative importance of properties at specific locations of the sequence of the peptides. In general, bulky hydrophobic amino acid at the C-terminal with bulky amino acid at the N-terminal were important for the bitterness of small peptides (di- and tripeptides). For large peptides (≥tetrapeptides), bulky hydrophobic amino acids with or without basic properties at the C-terminal and bulky basic amino acids at the N-terminal were related with the bitterness of the peptides.

ACKNOWLEDGMENT

We gratefully acknowledge constructive discussions with Dr. S. Nakai at the University of British Columbia.

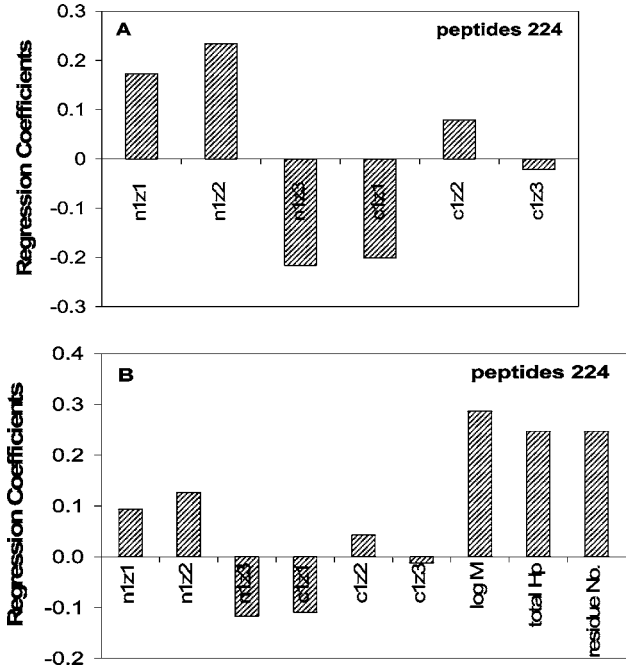

Figure 6. PLS regression coefficient (weighted) plots for 224 peptide sample set (di- to tetradecapeptides) using (**A**) *z*-scores only and (**B**) *z*-scores with total hydrophobicity, log *M*, and residue number values.

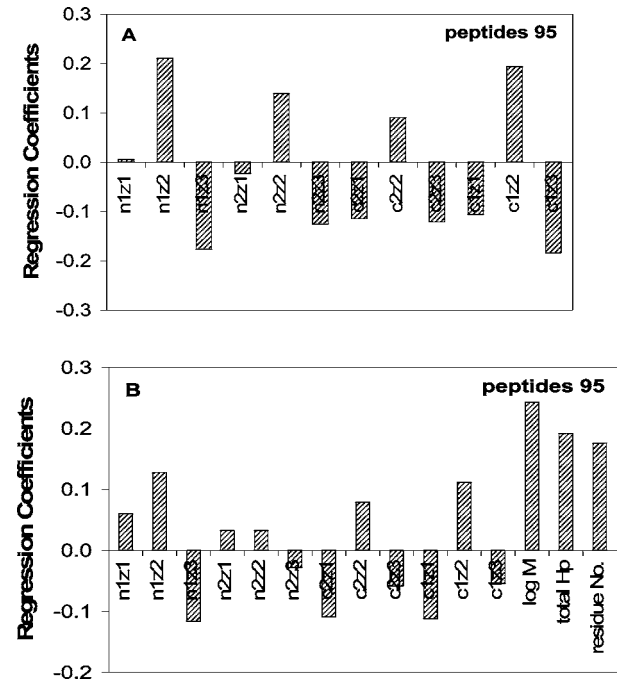

Figure 7. PLS regression coefficient (weighted) plots for 95 peptide sample set (tetra- to tetradecapeptides) using (**A**) *z*-scores only and (**B**) *z*-scores with total hydrophobicity, log *M*, and residue number values.

LITERATURE CITED

(1) Tamura, M.; Mori, N.; Miyoshi, T.; Koyoma, S.; Kohri, H.; Okai, H. Practical debittering using model peptides and related compounds. *Agric. Biol. Chem.* **1990**, *54*, 41−51.

(2) Saha, B. C.; Hayashi, K. Debittering of protein hydrolyzates. *Biotechnol. Adv.* **2001**, *19*, 355−370.

(3) Nishiwaki, T.; Yoshimizu, S.; Furuta, M.; Hayashi, K. Debittering of enzymatic hydrolysates using an aminopeptidase from the edible Basidiomycete *Grifola frondosa*. *J. Biosci. Bioeng.* **2002**, *93*, 60−63.

(4) Raksakulthai, R.; Haard, N. F. Exopeptidases and their application to reduce bitterness in food: A review. *Crit. Rev. Food Sci. Nutr.* **2003**, *43*, 401−445.

(5) Otagiri, K.; Nosho, Y.; Shinoda, I.; Fukui, H.; Okai, H. Studies on model of bitter peptides including arginine, proline and phenylalanine residues. I. Bitter taste of di- and tripeptides, and bitterness increase of the model peptides by extension of the peptide chain. *Agric. Biol. Chem.* **1985**, *49*, 1019−1026.

(6) Nosho, Y.; Otagiri, K.; Shinoda, I.; Okai, H. Studies on a model of bitter peptides including arginine, proline and phenylalanine residues. II. Bitterness behavior of a tetrapeptide (Arg-Pro-Phe-Phe) and its derivatives. *Agric. Biol. Chem.* **1985**, *49*, 1829−1837.

(7) Ishibashi, N.; Arita, Y.; Kanehisa, H.; Kouge, K.; Okai, H.; Fukui, S. Bitterness of leucine-containing peptides. *Agric. Biol. Chem.* **1987**, *51*, 2389−2394.

(8) Ishibashi, N.; Sadamori, K.; Yamamoto, O.; Kanehisa, H.; Kouge, K.; Kikuchi, E.; Okai, H.; Fukui, S. Bitterness of phenylalanine- and tyrosine-containing peptides. *Agric. Biol. Chem.* **1987**, *51*, 3309−3313.

(9) Ishibashi, N.; Kouge, K.; Shinoda, I.; Kanehisa, H.; Okai, H. A mechanism for bitter taste sensibility in peptides. *Agric. Biol. Chem.* **1988**, *52*, 819−827.

(10) Asao, M.; Iwamura, H.; Akamatsu, M.; Fujita, T. Quantitative structure-activity relationships of the bitter thresholds of amino acids, peptides, and their derivatives. *J. Med. Chem.* **1987**, *30*, 1873−1879.

(11) Tamura, M.; Miyoshi, T.; Mori, N.; Kinomura, K.; Kawaguchi, M.; Ishibashi, N.; Okai, H. Mechanism for the bitter tasting potency of peptides using *O- aminoacyl sugars as model compounds. Agric. Biol. Chem.* **1990**, *54*, 1401−1409.

(12) Kim, M.-R.; Choi, S.-Y.; Kim, C.-S.; Kim, C.-W.; Utsumi, S.; Lee, C.-H. Amino acid sequence analysis of bitter peptides from a soybean proglycinin subunit synthesized in *Escherichia coli*. *Biosci., Biotechnol., Biochem.* **1999**, *63*, 2069−2074.

(13) Shinoda, I.; Fushimi, A.; Kato, H.; Okai, H.; Fukui, S. Bitter taste of synthetic C-terminal tetradecapeptide of bovine β-casein, H-Pro^{196}-Val-Leu-Gly-Pro-Val-Arg-Gly-Pro-Phe-Pro-Ile-Ile-Val^{209}-OH, and its related peptides. *Agric. Biol. Chem.* **1985**, *49*, 2587−2596.

(14) Ishibashi, N.; Ono, I.; Kato, K.; Shigenaga, T.; Shinoda, I.; Okai, H.; Fukui, S. Role of the hydrophobic amino acid residue in the bitterness of peptides. *Agric. Biol. Chem.* **1988**, *52*, 91−94.

(15) Shinoda, I.; Nosho, Y.; Otagiri, K.; Okai, H.; Fukui, S. Bitterness of diastereomers of a hexapeptide (Arg-Arg-Pro-Pro-Phe-Phe) containing D-phenylalanine in place of L-phenylalanine. *Agric. Biol. Chem.* **1986**, *50*, 1785−1790.

(16) Kanehisa, H.; Miyake, I.; Okai, H.; Aoyagi, H.; Izumiya, N. Studies of bitter peptides from casein hydrolyzate. X. Synthesis and bitter taste of H-Arg-Gly-Pro-Phe-Pro-Ile-Ile-Val-OH corresponding to C-terminal portion of β-casein. *Bull. Chem. Soc. Jpn.* **1984**, *57*, 819−822.

(17) Otagiri, K.; Shigenaga, T.; Kanehisa, H.; Okai, H. Studies of bitter peptides from casein hydrolyzate. IV. Relationship between bitterness and hydrophobic amino acids moiety in the C-terminal of BPIa (Arg-Gly-Pro-Pro-Phe-Ile-Val). *Bull. Chem. Soc. Jpn.* **1984**, *57*, 90−96.

(18) Otagiri, K.; Miyake, I.; Ishibashi, N.; Fukui, H.; Kanehisa, H.; Okai, H. Studies of bitter peptides from casein hydrolyzate. II. Syntheses of bitter peptide fragments and analogs of BPIa (Arg-Gly-Pro-Pro-Phe-Ile-Val) from casein hydrolyzate. *Bull. Chem. Soc. Jpn.*, **1983**, *56*, 1116−1119.

(19) Shigenaga, T.; Otagiri, K.; Kanehisa, H.; Okai, H. Studies of bitter peptides from casein hydrolyzate. VII. Bitterness of the retro-BPIa (Val-Ile-Phe-Pro-Pro-Gly-Arg) and its fragments. *Bull. Chem. Soc. Jpn.* **1984**, *57*, 103−107.

(20) Kanehisa, H.; Okai, H. Studies of bitter peptides from casein hydrolyzate. V. Bitterness of the synthetic N-terminal analogs of des-Gly^2-BPIa (Arg-Pro-Pro-Phe-Ile-Val). *Bull. Chem. Soc. Jpn.* **1984**, *57*, 301−302.

(21) Kim, M.-R.; Kawamura, Y.; Lee, C.-H. Isolation and identification of bitter peptides of tryptic hydrolysate of soybean 11S glycinin by reverse-phase high performance liquid chromatography. *J. Food Sci.* **2003**, *68*, 2416−2422.

(22) Miyake, I.; Kouge, K.; Kanehisa, H.; Okai, H. Studies of bitter peptides from casein hydrolyzate. IX. Syntheses and bitter taste of bitter peptide BPIa dimer. $(Arg\text{-}Gly\text{-}Pro\text{-}Pro\text{-}Phe\text{-}Ile\text{-}Val)_2$, and Gly-Gly BPIa derivatives. *Bull. Chem. Soc. Jpn.* **1984**, *57*, 815−818.

(23) Miyake, I.; Kouge, K.; Kanehisa, H.; Okai, H. Studies of bitter peptides from casein hydrolyzate. VIII. Bitter taste of cyclic analog of BPIa (Arg-Gly-Pro-Pro-Phe-Ile-Val). *Bull. Chem. Soc. Jpn.* **1984**, *57*, 1163−1164.

(24) Shinoda, I.; Nosho, Y.; Kouge, K.; Ishibashi, N.; Okai, H.; Tatsumi, K.; Kikuchi, E. Variation in bitterness potency when introducing Gly-Gly residue into bitter peptides. *Agric. Biol. Chem.* **1987**, *51*, 2103−2110.

(25) Ishibashi, N.; Kubo, T.; Chino, M.; Fukui, H.; Shinoda, I.; Kikuchi, E.; Okai, H.; Fukui, S. Taste of proline-containing peptides. *Agric. Biol. Chem.* **1988**, *52*, 95−98.

(26) Miyake, I.; Kouge, K.; Kanehisa, H.; Okai, H. Studies of bitter peptides from casein hydrolyzate. III. Bitter taste of synthetic analogs of BPIa (Arg-Gly-Pro-Pro-Phe-Ile-Val) containing D-proline or glycine in place of L-proline. *Bull. Chem. Soc. Jpn.* **1983**, *56*, 1678−1681.

(27) Nakatani, M.; Nakata, T.; Kouge, K.; Okai, H. Studies on bitter peptides from casein hydrolyzate. XIV. Bitter taste of synthetic analogs of octapeptide, Arg-Gly-Pro-Phe-Pro-Ile-Ile-Val, corresponding to the C-terminal portion of β-casein. *Bull. Chem. Soc. Jpn.* **1994**, *67*, 438−444.

(28) Kanehisa, H. Studies of bitter peptides from casein hydrolyzate. VI. Syntheses and bitter taste of BPIc (Val-Tyr-Pro-Phe-Pro-Pro-Gly-Ile-Asn-His) and its analogs and fragments. *Bull. Chem. Soc. Jpn.* **1984**, *57*, 97−102.

(29) Nakai, S.; Li-Chan, E. Recent advances in structure and function of food proteins: QASR approach. *Crit. Rev. Food Sci. Nutr.* **1993**, *33*, 477−499.

(30) Pripp, A. H.; Isaksson, T.; Stepaniak, L.; Sorhaug, T.; Ardo, Y. Quantitative structure activity relationship modelling of peptides and proteins as a tool in food science. *Trends Food Sci. Technol.* **2005**, *16*, 484−494.

(31) Hellberg, S.; Sjostrom, M.; Skagerberg, B.; Wold, S. Peptide quantitative structure-activity relationships, a multivariate approach. *J. Med. Chem.* **1987**, *30*, 1126−1135.

(32) 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 relationships. *Int. J. Pept. Protein Res.* **1991**, *37*, 414−424.

(33) Jonsson, J.; Eriksson, L.; Hellberg, S.; Sjostrom, M.; Wold, S. Multivariate parametrization of 55 coded and non-coded amino acids. *Quant. Struct.-Act. Relat.* **1989**, *8*, 204−209.

(34) Collantes, E. R.; Dunn, W. J., III. Amino acid side chain descriptors for quantitative structure-activity relationship studies of peptide analogues. *J. Med. Chem.* **1995**, *38*, 2705−2713.

(35) 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.

(36) De Armas, R. R.; Diaz, H. G.; Molina, R.; Gonzalez, M. P.; Uriarte, E. Stochastic-based descriptors studying peptides biological properties: Modeling the bitter tasting threshold of dipeptides. *Bioorg. Med. Chem.* **2004**, *12*, 4815−4822.

(37) Mei, H.; Liao, Z. H.; Zhou, Y.; Li, S. Z. A new set of amino acid descriptors and its application in peptide QSARs. *Biopolymers* **2005**, *80*, 775−786.

(38) Wilce, M. C. J.; Agullar, M.-I.; Hearn, M. T. W. Physicochemical basis of amino acid hydrophobicity scales: Evaluation of four new scales of amino acid hydrophobicity coefficients derived from RP-HPLC of peptides. *Anal. Chem.* **1995**, *67*, 1210−1219.

(39) Akamatsu, M.; Yoshida, Y.; Nakamura, H.; Asao, M.; Iwamura, H.; Fujita, T. Hydrophobicity of di- and tripeptides having unionizable side chains and correlation with substituent and structural parameters. *Quant. Struct.-Act. Relat.* **1989**, *8*, 195−203.

(40) Gomez, M. J.; Garde, S.; Gaya, P.; Medina, M.; Nunez, M. Relationship between level of hydrophobic peptides and bitterness in cheese made from pasteurized and raw milk. *J. Dairy Res.* **1997**, *64*, 289−297.

(41) Gulyaeva, N.; Zaslavsky, A.; Chait, A.; Zaslavsky, B. Relative hydrophobicity of di- to hexapeptides as measured by aqueous two-phase partitioning. *J. Pept. Res.* **2003**, *61*, 129−139.

(42) Kukman, I. L.; Zelenik-Blatnik, M.; Abram, V. Isolation of low-molecular-mass hydrophobic bitter peptides in soybean protein hydrolysates by reversed-phase high-performance liquid chromatography. *J. Chromatogr. A* **1995**, *704*, 113−120.

(43) Belitz, H. D.; Wieser, H. Bitter compounds: Occurrence and structure-activity relationships. *Food Rev. Int.* **1985**, *1*, 271−354.

(44) Ohyama, S.; Ishibashi, N.; Tamura, M.; Nishizaki, H.; Okai, H. Synthesis of bitter peptides composed of aspartic acid and glutamic acid. *Agric. Biol. Chem.* **1988**, *52*, 871−872.

(45) Shinoda, I.; Tada, M.; Okai, H.; Fukui, S. Bitter taste of H-Pro-Phe-Pro-Gly-Pro-Ile-Pro-OH corresponding to the partial sequence (positions 61−67) of bovine β-casein, and related peptides. *Agric. Biol. Chem.* **1986**, *50*, 1247−1254.

Received for review August 23, 2006. Revised manuscript received October 11, 2006. Accepted October 12, 2006. This research was funded by a grant from the Natural Sciences and Engineering Research Council of Canada.

JF062422J