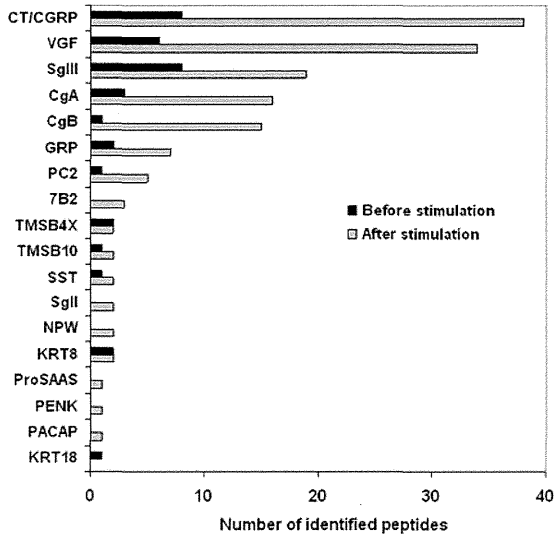


FIG. 2. **The numbers of peptides identified before and after stimulation sorted by precursor names.** CT and CGRP are grouped as they arise from alternatively spliced exons. Peptide sequences are indicated in Table I. *PACAP*, pituitary adenylate cyclase-activating polypeptide; *NPW*, neuropeptide W; *PENK*, proenkephalin A; *KRT*, cytokeratin; *TMSB4X*, thymosin β-4 X-linked; *TMSB10*, thymosin β-10.

polypeptide, neuropeptide W, and proenkephalin A) and three granin-like proteins (7B2, SgII, and pro-SAAS). In total, 146 of 152 peptides arose from the 15 RSP precursors. The remaining six peptides were derived from thymosins and cytokeratin 8. According to the stimulus-induced increase in total peptide amounts released to culture supernatant, more peptides were identified from the former nine precursor proteins (Figs. 1 and 2). Across the CT and CGRP precursor sequences, known major processing products (17) were identified, namely CT N-terminal propeptide (6217.04 Da), bioactive CT (3415.58 Da), katacalcin (2435.07 Da), CGRP N-terminal propeptide (6056.04 Da), and bioactive CGRP (3786.96 Da) (Fig. 3).

Investigation of Cleavage Sites through Identified Peptides—Regarding the processing of peptide hormone precursors, it has long been known that PC1/3 or PC2 cleaves the precursors at sites containing consecutive basic amino acids following N-terminal signal peptide cleavage (1, 2). Cameron *et al.* (18) recently studied the specificity of PCs and drew a conclusion that the PC-mediated cleavage occurs at sites containing pairs of basic amino acids separated by 0, 2, 4, or 6 residues. However, some peptide hormone precursors are processed at monobasic residues although at much less frequency (2). In any case, the resultant C-terminal basic residues are subsequently removed by carboxypeptidase E. If the exposed C-terminal residue is glycine, peptidyl-glycine α-amidating monooxygenase catalyzes peptide α-amidation, a common post-translational modification often required for a peptide to be fully bioactive (19).

Having confirmed that almost all the sequenced peptides arose from RSP precursors (Table I), we extracted 10 amino acids N- or C-terminally flanking the sequenced peptides to analyze their cleavage sites. Any glycine immediately followed by a basic residue(s) that creates an amidation site was also counted as a PC consensus site. Monobasic sites were defined as those that do not harbor basic residues except for P1 position and considered potential processing sites as well. PC consensus cleavage sites, signal sequence cleavage sites, and monobasic sites, referred to as informative cleavage sites in the present study, were found in 120, 26, and 29 peptides, respectively (72 peptides were counted twice under this definition). Overall 142 of 152 had such informative cleavage sites at either or both ends.

Identity of Major Peptides in the TT Secretopeptidome—In the LC-MSMS setting used throughout this study, any signal that transcended a given intensity threshold was automatically subjected to MSMS and ignored thereafter if it persisted within a precursor mass tolerance of 5 ppm with the aim of sequencing as many peptides as possible. The signal is not always subjected to MSMS at its maximum intensity in LC-MS profiles, and therefore this setting could return relatively low scores even for abundant peptides, which may not be included in Table I. To identify intense signals in LC-MSMS base peak chromatograms, we examined peptide peak intensities in all MS spectra. Table II provides the list of 35 peptides that were detected at the indicated monoisotopic *m/z* and charge state with a base peak intensity beyond 2*e*+06 (see also Fig. 3 and supplemental Fig. 1). Three peptide sequences had expectation values (above 0.05) that did not exceed the Mascot significance threshold. First, the 1278.67-Da peptide was qualified as a CGRP-derived peptide based on the observation that CGRP-derived peptides are most abundantly expressed in the TT secretopeptidome (Fig. 3 and supplemental Fig. 3). Second, the 5687.91-Da peptide also yielded suboptimum MSMS spectra but was identified as a PC2-derived peptide because of matches for eight consecutive b-ions (supplemental Fig. 3). Third, the 8559.61-Da peptide was qualified using MSMS spectral comparison with commercially available human ubiquitin (supplemental Fig. 2).

This list covered all the major processing products of CT and CGRP. Intact CGRP continued to be observed at multiple charged ions (+3 to +6) over a retention time of 10 min in the mass chromatogram, suggesting that it represents one of the most abundant peptides in this secretopeptidome (Fig. 3). The CGRP-derived peptide (ACDTATCVTHRLAGLLSRSGGVVKN) appeared to be generated from endoproteolytic cleavage of intact CGRP because the C-terminal cleaved half (1278.67 Da) was detected as well (Table II). Except for this peptide, all the N-terminal cleavage sites of CT/CGRP-derived peptides are known as major processing sites (17). Similar findings were obtained with C-terminal cleavage sites except for four non-basic sites. It remains to be clarified whether these non-basic cleavage sites point to the processing that actually occurred in the RSP.

TABLE I

Peptides identified before and after stimulation

Data were obtained without gel filtration chromatography and summarized from three runs using an identical LC-MSMS setting and peptides whose scores (column 7) exceeded homology thresholds (HT, column 8) in at least two runs are listed. For peptides identified in multiple runs, a higher score is listed. Mr(Calc) represents the theoretical monoisotopic molecular mass (Da) based on the peptide sequence. The score value beyond an identity threshold (IT, column 9) is indicated in bold. In column 1, "CT/CGRP" indicates that the peptides are shared by CT and CGRP precursors. If the peptide was identified at different charges, the charge states are also shown in column 4. Expectation values are indicated in column 10. The N- (N-term) and C-terminal (C-term) flanking 10 amino acids (columns 11 and 13) are shown and marked as follows: closed boxes with white letters, typical PC cleavage sites containing consecutive basic residues and sites having C-terminal amidation motifs; dark gray boxes, cleavage sites containing basic residues at P4, P6 or P8 position; pale gray boxes, sites having basic residues at P1 but not at P2, P4, P6 or P8 position. "Signal" indicates that the peptide flanks its signal sequence. "C-term" indicates that the peptide C-terminus is the end of the precursor protein. Ac-, N-terminal acetylation; -NH2, C-terminal amidation; <Q, pyroglutamic acid. Oxidized methionine residues are underlined in column 12. Sequences are based on the following IPI accession numbers: CgA, 00746813; CgB, 00006601, CGRP, 00027855; CT, 00000914; GRP, 00011722; KRT18, 00554788; KRT8, 00554648; PC2, 00029131; SgIII, 00292071; SST, 00000130; TMSB10, 00220827; TMSB4X, 00220828; VGF, 00069058; 7B2, 00008944; NPW, 00853190; PACAP, 00000027; PENK, 00000828; SgII, 00009362. PACAP, pituitary adenylate cyclase-activating polypeptide; NPW, neuropeptide W; PENK, proenkephalin A; KRT, cytokeratin; TMSB4X, thymosin β-4 X-linked; TMSB10, thymosin β-10.

Before stimulation

Precursor	m/z (obsd)	z	Also ID at	Mr(calc) (Da)	Mass error (Da)	Score	HT	IT	Expect. value	N-term	Peptide	C-term
CgA	581.8036	2		1161.5931	-0.0005	**60**	29	45	0.0016	SSMKLSFRAR	AYGFRGPGPQL	RRGWRPSSRE
CgA	725.3738	2		1448.7333	-0.0003	30	16	45	1.8	LSKEWEDSKR	WSKMDQLAKELT	AEKRLEGQEE
CgA	825.4139	2		1648.8130	0.0002	**54**		45	0.0078	LSKEWEDSKR	WSKMDQLAKELTAE	KRLEGQEEEE
CgB	587.9692	3		1760.8846	0.0012	**51**	20	47	0.022	ARVPKLDLKR	QYDRVAQLDQLLHY	RKKSAEFPDF
CGRP	877.8292	3		2630.4657	0.0000	**74**		43	4.40E-05	CDTATCVTHR	LAGLLSRSGGVVKNNFVPTNVGSKAF-NH2	GRRRRDLQA
CGRP	778.0571	3		2331.1488	0.0007	**51**	35	47	0.023	LAALVQDYVQ	MKASELEQEQEREGSRIIAQ	KRACDTATCV
CGRP	961.0179	2		1920.0218	-0.0005	**56**	20	46	0.0065	THRLAGLLSR	SGGVVKNNFVPTNVGSKAF-NH2	GRRRRDLQA
CGRP	509.7875	2		1017.5607	-0.0003	35	21	46	0.71	RSGGVVKNNF	VPTNVGSKAF-NH2	GRRRRDLQA
CT	751.0047	3		2249.9906	0.0017	**44**	17	42	0.035	IGVGAPGKKR	DMSSDLERDHRPHVSMPQN	AN
CT	774.6831	3		2321.0277	-0.0003	**50**	16	43	0.011	IGVGAPGKKR	DMSSDLERDHRPHVSMPQNA	N
CT	812.6974	3		2435.0706	-0.0003	**52**	17	42	0.0054	IGVGAPGKKR	DMSSDLERDHRPHVSMPQNAN	C-term
CT/CGRP	881.7805	3		2642.3187	0.0009	**122**	25	48	1.90E-09	Signal	APFRSALESSPADPATLSEDEARLL	LAALVQDYVQ
GRP	800.9424	2		1599.8695	0.0009	41	20	45	0.15	Signal	VPLPAGGGTVLTKMYP	RGNHWAVGHL
GRP	586.3308	3		1755.9706	0.0001	**76**	24	44	3.50E-05	Signal	VPLPAGGGTVLTKMYPR	GNHWAVGHLM
KRT18	485.4749	5	4	2422.3405	-0.0021	40		44	0.12	NSMQTIQKTT	TRRIVDGKVVSETNDTKVLRH	C-term
KRT8	884.2651	5		4416.2885	0.0007	**93**		49	2.10E-06	TSPGLSYSLG	SSFGSGAGSSSFSRTSSSRAVVVKKIETRDGKLVSESSDVLPK	C-term
KRT8	742.1170	4		2964.4438	-0.0048	23	16	47	16	RVTQKSYKVS	TSGPRAFSSRSYTSGPGSRISSSSFSRVG	SSNFRGGLGG
PC2	633.3122	2		1264.6122	-0.0023	**49**	23	46	0.032	HKQQLERDPR	VKMALQQEGFD	RKKRGYRDIN
SgIII	735.3826	2		1468.7521	-0.0014	37	33	46	0.44	GSQDKSLHNR	ELSAERPLNEQIA	EAEEDKIKKT
SgIII	799.9040	2		1597.7947	-0.0012	**48**	30	46	0.035	GSQDKSLHNR	ELSAERPLNEQIAE	AEEDKIKKTY
SgIII	899.9442	2		1797.8744	-0.0005	**67**	22	46	0.00045	GSQDKSLHNR	ELSAERPLNEQIAEAE	EDKIKKTYPP
SgIII	1021.9787	2		2041.9439	-0.0011	**101**	22	45	1.50E-07	GSQDKSLHNR	ELSAERPLNEQIAEAEED	KIKKTYPPEN
SgIII	762.0485	3		2283.1230	0.0007	**54**	27	47	0.011	GSQDKSLHNR	ELSAERPLNEQIAEAEEDKI	KKTYPPENKP
SgIII	756.3841	2	3	1510.7528	0.0008	**50**	29	46	0.022	Signal	FPKPGGSQDKSLHN	RELSAERPLN
SgIII	787.4010	5		3931.9663	0.0021	29	16	49	5.4	Signal	FPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKI	KKTYPPENKP
SgIII	866.9151	2		1731.8163	-0.0007	45	22	45	0.06	KIEKERQSIR	SSPLDNKLNVEDVDST	KNRK
SST	622.7880	2		1243.5615	0.0000	**66**	22	41	0.0002	QDEMRLELQR	SANSNPAMAPRE	RKAGCKNFFW
TMSB10	823.2613	6	4, 5, 7	4933.5229	0.0012	**148**		49	8.50E-12	M	Ac-ADKPDMGEIASFDKAKLKKTETQEKNTLPTKETIEQEKRSEIS	C-term
TMSB4X	827.7546	6	4, 5, 7	4960.4862	-0.0020	**95**		50	1.40E-06	M	Ac-SDKPDMAEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES	C-term
TMSB4X	996.3035	5		4976.4811	-0.0002	**77**		49	9.90E-05	M	Ac-SDKPDMAEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES	C-term
VGF	917.7143	4		3666.8278	0.0004	**68**	23	48	0.00067	Signal	APPGRPEAQPPPLSSEHKEPVAGDAVPGPKDGSAPEV	RGARNSEPQD
VGF	988.7540	4	5, 6	3950.9875	-0.0006	**90**	23	49	3.90E-06	Signal	APPGRPEAQPPPLSSEHKEPVAGDAVPGPKDGSAPEVRGA	RNSEPQDEGE
VGF	742.1746	5		3705.8346	0.0020	32	18	48	2.5	YPGREAQARR	AQEEAEAEERRLQEQEELENYIEHVLLRRP	C-term
VGF	639.0110	3		1914.0112	0.0001	32	20	46	1.3	EVEEKRKRKK	NAPPEPVPPPRAAPAPTHV	RSPQPPPPAP
VGF	848.4221	4		3389.6600	-0.0007	**109**	20	48	4.00E-08	RDFSPSSAKR	<QQETAAAETETRTHTLTRVNLESPGPERVW	RASWGEFQAR
VGF	852.6792	4		3406.6865	0.0010	**102**		48	2.50E-07	RDFSPSSAKR	QQETAAAETETRTHTLTRVNLESPGPERVW	RASWGEFQAR

After stimulation

7B2	851.4489	2		1700.8846	-0.0012	44	25	46	0.083	MKGGERRKRR	SVNPYLQGQRLDNVV	AKKSVPHFSD
7B2	886.9685	2		1771.9217	0.0008	**51**	26	47	0.022	MKGGERRKRR	SVNPYLQGQRLDNVVA	KKSVPHFSDE
7B2	751.3260	2		1500.6369	0.0006	**64**	22	40	0.00022	QRLDNVVAKK	SVPHFSDEDKDPE	C-term
CgA	525.2614	2		1048.5090	-0.0007	**69**	32	45	0.0002	SSMKLSFRAR	AYGFRGPGPQ	LRRGWRPSSR
CgA	581.8039	2		1161.5931	0.0001	**65**	29	45	0.00055	SSMKLSFRAR	AYGFRGPGPQL	RRGWRPSSRE
CgA	464.7536	2		927.4927	-0.0001	40	37	46	0.19	MKLSFRARAY	GFRGPGPQL	RRGWRPSSRE
CgA	666.6731	3	2	1996.9966	0.0010	43	25	47	0.12	FRGPGPQLRR	GWRPSSREDSLEAGLPLQ	VRGYPEEKKE
CgA	699.6958	3		2096.0650	0.0004	**54**	26	47	0.012	FRGPGPQLRR	GWRPSSREDSLEAGLPLQV	RGYPEEKKEE
CgA	1302.9266	3		3905.7636	-0.0056	**51**	16	43	0.0092	AKERAHQQKK	HSGFEDELSEVLENQSSQAELKEAVEEPSSKDVME	KREDSKEAEK
CgA	699.3080	3		2094.9011	0.0010	**59**	18	41	0.00082	LAKELTAEKR	LEGQEEEEDNRDSSMKLS	FRARAYGFRG
CgA	748.3303	3		2241.9695	-0.0003	**83**	23	41	3.20E-06	LAKELTAEKR	LEGQEEEEDNRDSSMKLSF	RARAYGFRGP
CgA	824.0432	3		2469.1077	-0.0001	**94**	18	43	4.70E-07	LAKELTAEKR	LEGQEEEEDNRDSSMKLSFRA	RAYGFRGPGP
CgA	754.7653	5		3768.7914	-0.0013	37	19	47	0.51	LAKELTAEKR	LEGQEEEEDNRDSSMKLSFRARAYGFRGPGPQL	RRGWRPSSRE
CgA	727.3265	3	2	2178.9552	0.0023	**109**	24	42	1.10E-08	VPQGLFRGGK	SGELEQEEERLSKEWEDS	KRWSKMDQLA
CgA	674.8492	2		1347.6856	-0.0017	**64**	32	47	0.0011	LSKEWEDSKR	WSKMDQLAKEL	TAEKRLEGQE
CgA	725.3741	2		1448.7333	0.0003	**63**	26	44	0.00066	LSKEWEDSKR	WSKMDQLAKELT	AEKRLEGQEE
CgA	760.8932	2		1519.7704	0.0015	**54**	26	46	0.0093	LSKEWEDSKR	WSKMDQLAKELTA	EKRLEGQEEE
CgA	825.4145	2	3	1648.8130	0.0014	**92**	27	46	1.50E-06	LSKEWEDSKR	WSKMDQLAKELTAE	KRLEGQEEEE
CgA	546.2849	2		1090.5560	-0.0007	27	16	44	3.4	SMKLSFRARA	YGFRGPGPQL	RRGWRPSSRE
CgB	624.7640	2		1247.5128	0.0006	**56**	29	40	0.0013	HHRGRGGEPR	AYFMSDTREE	KRFLGEGHHR
CgB	448.6857	2		895.3573	-0.0005	15	14	41	20	VSMASLGEKR	DHHSTHY	RASEEEPEYG
CgB	689.3519	3		2065.0327	0.0010	**78**	22	47	4.60E-05	VLKTSRKDVK	DKETTENENTKFEVRLL	RDPADASEAH
CgB	599.6091	3		1795.8060	-0.0004	21	17	44	9.8	FMSDTREEKR	FLGEGHHRVQENQMD	KARRHPQGAW
CgB	665.9865	3		1994.9381	-0.0004	**87**	21	46	3.70E-06	FMSDTREEKR	FLGEGHHRVQENQMDKA	RRHPQGAWKE
CgB	653.2990	3		1956.8748	0.0004	**58**	19	43	0.0019	LEPGKGRHHR	GRGGEPRAYFMSDTREE	KRFLGEGHHR
CgB	689.3477	4		2753.3602	0.0013	**86**	23	48	8.10E-06	YSSHHTAEKR	KRLGELFNPYYDPLQWKSSHFE	RRDNMNDNFL
CgB	824.0616	3		2469.1641	-0.0012	35	19	46	0.65	YSSHHTAEKR	LGELFNPYYDPLQWKSSHFE	RRDNMNDNFL
CgB	569.8506	2		1137.6870	-0.0003	35	32	41	0.24	EDVNWGYEKR	NLARVPKLDL	KRQYDRVAQL

TABLE I—*continued*

CgB	731.1431	5		3650.6808	-0.0015	**50**		45	0.016	EDVNWGYEKR	NYPSLELDKMAHGYGEESEEERGLEPGKGRHH	RGRGGEPRAY
CgB	671.7838	2		1341.5513	0.0016	22	16	41	4.4	EYNYDWWEKK	PFSEDVNWGYE	KRNLARVPKL
CgB	731.3881	2		1460.7623	-0.0007	27	27	46	4.3	ARVPKLDLKR	QYDRVAQLDQLL	HYRKKSAEFP
CgB	872.9364	2		1743.8580	0.0003	**61**	32	46	0.0016	ARVPKLDLKR	<QYDRVAQLDQLLHY	RKKSAEFPDF
CgB	587.9688	3		1760.8846	0.0000	**64**	21	47	0.001	ARVPKLDLKR	QYDRVAQLDQLLHY	RKKSAEFPDF
CgB	1068.4681	3		3202.3844	-0.0019	**100**	21	40	5.80E-08	HHHGRSRPDR	SSQGGSLPSEEKGHPQEESEESNVSMASLGE	KRDHHSTHYR
CGRP	1186.6031	5	6	5927.9850	-0.0059	**136**		50	1.50E-10	Signal	APFRSALESSPADPATLSEDEARLLLAALVQDYVQMKASELEQEQEREGSRIIA	QKRACDTATC
CGRP	1010.3481	6	5	6056.0436	0.0014	**156**	24	50	1.60E-12	Signal	APFRSALESSPADPATLSEDEARLLLAALVQDYVQMKASELEQEQEREGSRIIAQ	KRACDTATCV
CGRP	734.3766	3		2200.1083	-0.0004	**52**	20	46	0.014	AALVQDYVQM	KASELEQEQEREGSRIIAQ	KRACDTATCV
CGRP	877.8294	3		2630.4657	0.0006	**118**		43	2.10E-09	CDTATCVTHR	LAGLLSRSGGVVKNNFVPTNVGSKAF-NH2	GRRRRDLQA
CGRP	797.4484	3		2389.3230	0.0002	**81**		44	1.20E-05	ATCVTHRLAG	LLSRSGGVVKNNFVPTNVGSKAF-NH2	GRRRRDLQA
CGRP	735.3707	3		2203.0902	0.0000	**55**	34	47	0.0087	LAALVQDYVQ	MKASELEQEQEREGSRIIA	QKRACDTATC
CGRP	778.0572	3		2331.1488	0.0010	**100**	33	47	3.30E-07	LAALVQDYVQ	MKASELEQEQEREGSRIIAQ	KRACDTATCV
CGRP	788.4104	2		1574.8053	0.0010	38	21	47	0.43	THRLAGLLSR	SGGVVKNNFVPTNVGS	KAFGRRRR
CGRP	961.0179	2		1920.0218	-0.0005	**96**	27	46	5.90E-07	THRLAGLLSR	SGGVVKNNFVPTNVGSKAF-NH2	GRRRRDLQA
CGRP	509.7875	2		1017.5607	-0.0003	38	22	46	0.3	RSGGVVKNNF	VPTNVGSKAF-NH2	GRRRRDLQA
CGRP	853.7657	3		2558.2758	-0.0005	40	18	47	0.28	LLLAALVQDY	VQMKASELEQEQEREGSRIIAQ	KRACDTATCV
CGRP	649.3277	4		2593.2805	0.0010	**71**	17	47	0.00025	RLLLAALVQD	YVQMKASELEQEQEREGSRIIA	QKRACDTATC
CGRP	908.1203	3	4	2721.3391	-0.0001	**120**	33	47	3.30E-09	RLLLAALVQD	YVQMKASELEQEQEREGSRIIAQ	KRACDTATCV
CT	1037.1800	6	5	6217.0396	-0.0033	**143**	23	49	2.50E-11	Signal	APFRSALESSPADPATLSEDEARLLLAALVQDYVQMKASELEQEQEREGSSLDSPRS	KRCGNLSTCM
CT	1247.6136	5		6233.0345	-0.0029	**58**		49	0.0077	Signal	APFRSALESSPADPATLSEDEARLLLAALVQDYVQMKASELEQEQEREGSSLDSPRS	KRCGNLSTCM
CT	602.7485	2		1203.4826	-0.0002	**74**	21	37	1.30E-05	IGVGAPGKKR	DMSSDLERDH	RPHVSMPQNA
CT	610.7458	2		1219.4775	-0.0005	**54**	21	36	0.00099	IGVGAPGKKR	DMSSDLERDH	RPHVSMPQNA
CT	532.2388	3		1593.6954	-0.0009	29	16	42	1.1	IGVGAPGKKR	DMSSDLERDHRPH	VSMPQNAN
CT	756.3361	3		2265.9855	0.0008	40	16	41	0.079	IGVGAPGKKR	DMSSDLERDHRPHVSMPQN	AN
CT	761.6672	3		2281.9804	-0.0007	24	15	41	2.6	IGVGAPGKKR	DMSSDLERDHRPHVSMPQN	AN
CT	780.0152	3		2337.0226	0.0013	**49**	16	42	0.011	IGVGAPGKKR	DMSSDLERDHRPHVSMPQNA	N
CT	774.6831	3		2321.0277	-0.0002	**65**	20	43	0.00033	IGVGAPGKKR	DMSSDLERDHRPHVSMPQNA	N
CT	812.6975	3	2, 4	2435.0706	0.0000	**65**	21	42	0.00031	IGVGAPGKKR	DMSSDLERDHRPHVSMPQNAN	C-term
CT	818.0293	3		2451.0655	0.0005	**59**	17	41	0.0009	IGVGAPGKKR	DMSSDLERDHRPHVSMPQNAN	C-term
CT	721.3836	2		1440.7514	0.0012	**53**	29	45	0.0086	LGTYTQDFNK	FHTFPQTAIGVGAP-NH2	GKKRDMSSDL
CT	831.7223	3	4	2492.1449	0.0002	**134**	27	45	6.40E-11	LAALVQDYVQ	MKASELEQEQEREGSSLDSPRS	KRCGNLSTCM
CT	490.7348	2		979.4545	0.0005	41	32	43	0.09	MSSDLERDHR	PHVSMPQNA	N
CT	547.7556	2		1093.4975	-0.0007	37	23	43	0.26	MSSDLERDHR	PHVSMPQNAN	C-term
CT	961.7856	3		2882.3352	-0.0003	**134**	22	45	7.70E-11	RLLLAALVQD	YVQMKASELEQEQEREGSSLDSPRS	KRCGNLSTCM
CT/CGRP	674.3307	2		1346.6466	0.0002	40	27	44	0.14	Signal	APFRSALESSPAD	PATLSEDEAR
CT/CGRP	881.7808	3		2642.3187	0.0020	**106**	23	48	8.60E-08	Signal	APFRSALESSPADPATLSEDEARLL	LAALVQDYVQ
CT/CGRP	919.4757	3		2755.4028	0.0025	28	22	48	5.1	Signal	APFRSALESSPADPATLSEDEARLLL	AALVQDYVQM
CT/CGRP	1118.5786	3		3352.7150	-0.0010	**61**	23	48	0.003	Signal	APFRSALESSPADPATLSEDEARLLLAALVQD	YVQMKASELE
CT/CGRP	560.3060	2		1118.5972	0.0002	39	27	46	0.24	TLSEDEARLL	LAALVQDYVQ	MKASELEQEQ
GRP	560.7768	2		1119.5396	-0.0005	28	21	45	2.8	GTVLTKMYPR	GNHWAVGHLM-NH2	GKKSTGESSS
GRP	670.8837	2		1339.7534	-0.0004	**62**	28	41	0.00039	Signal	VPLPAGGGTVLTKM	YPRGNHWAVG
GRP	752.4157	2		1502.8167	0.0001	**64**	22	43	0.0004	Signal	VPLPAGGGTVLTKMY	PRGNHWAVGH
GRP	800.9420	2	3	1599.8695	-0.0001	**50**	21	43	0.011	Signal	VPLPAGGGTVLTKMYP	RGNHWAVGHL
GRP	808.9396	2		1615.8644	0.0003	**63**		44	0.00074	Signal	VPLPAGGGTVLTKMYP	RGNHWAVGHL
GRP	878.9924	2	3	1755.9706	-0.0003	**139**	26	43	1.40E-11	Signal	VPLPAGGGTVLTKMYPR	GNHWAVGHLM
GRP	591.6626	3		1771.9655	0.0003	**54**	23	45	0.007	Signal	VPLPAGGGTVLTKMYPR	GNHWAVGHLM
KRT8	479.9398	3		1436.7987	-0.0010	**60**	27	44	0.0014	M	Ac-SIRVTQKSYKVS	TSGPRAFSSR
KRT8	513.6225	3		1537.8464	-0.0008	39	18	44	0.17	M	Ac-SIRVTQKSYKVST	SGPRAFSSRS
NPW	541.8009	2		1081.5880	-0.0007	31	22	45	1.2	VQELWETRRR	SSQAGIPVRAP	RSPRAPEPAL
NPW	413.5702	3		1237.6891	-0.0003	28	17	41	1.2	VQELWETRRR	SSQAGIPVRAPR	SPRAPEPALE
PACAP	774.3423	2		1546.6688	0.0012	31	18	42	0.75	GDDAEPLSKR	HSDGIFTDSYSRY	RKQMAVKKYL
PC2	715.7713	5		3573.8229	-0.0030	**88**	21	48	5.50E-06	Signal	ERPVFTNHFLVELHKGGEDKARQVAAEHGFGV	RKLPFAEGLY
PC2	813.5659	7		5687.9079	0.0027	44		50	0.19	Signal	ERPVFTNHFLVELHKGGEDKARQVAAEHGFGVRKLPFAEGLYHFYHNGLA	KAKRRRSLHH
PC2	496.5950	3		1486.7641	-0.0009	**51**	24	46	0.016	NGLAKAKRRR	SLHHKQQLERDP	RVKMALQQEG
PC2	633.3134	2		1264.6122	0.0000	**69**	25	45	0.00022	HKQQLERDPR	VKMALQQEGFD	RKKRGYRDIN
PC2	641.3111	2		1280.6071	0.0006	**72**	27	44	9.20E-05	HKQQLERDPR	VKMALQQEGFD	RKKRGYRDIN
PENK	693.8411	2		1385.6674	0.0002	**62**	28	46	0.0013	YGGFMRGLKR	SPQLEDEAKELQ	KRYGGFMRRV
ProSAAS	743.6491	4		2970.5635	0.0037	**53**	21	47	0.013	ETGAPRRFRR	SVPRGEAAGAVQELARALAHLLEAERQE	RARAEAQEAE
SgII	831.4189	6		4982.4661	0.0039	37		49	1	TDKLAPVSKR	FPVGPPKNDDTPNRQYWDEDLLMKVLEYLNQEKAEKGREHIA	KRAMENM
SgII	1088.2673	4		4349.0419	-0.0018	**112**		46	1.50E-08	EIINSNQVKR	VPGQGSSEDDLQEEEQIEQAIKEHLNQGSSQETDKLAPVS	KRFPVGPPKN
SgIII	643.3237	2		1284.6310	0.0019	**49**	29	46	0.029	GSQDKSLHNR	ELSAERPLNEQ	IAEAEEDKIK
SgIII	699.8649	2		1397.7150	0.0002	41	26	46	0.18	GSQDKSLHNR	ELSAERPLNEQI	AEAEEDKIKK
SgIII	735.3837	2		1468.7521	0.0007	44	41	44	0.066	GSQDKSLHNR	ELSAERPLNEQIA	EEDKIKKTYP
SgIII	799.9048	2		1597.7947	0.0003	**49**	36	46	0.03	GSQDKSLHNR	ELSAERPLNEQIAE	AEEDKIKKTY
SgIII	899.9447	2		1797.8744	0.0005	**67**	22	46	0.00047	GSQDKSLHNR	ELSAERPLNEQIAEAE	EDKIKKTYPP
SgIII	964.4659	2		1926.9170	0.0002	**99**	32	45	2.30E-07	GSQDKSLHNR	ELSAERPLNEQIAEAEE	DKIKKTYPPE
SgIII	1021.9788	2		2041.9439	-0.0009	**98**	21	45	2.80E-07	GSQDKSLHNR	ELSAERPLNEQIAEAEED	KIKKTYPPEN
SgIII	1142.5676	2	3	2283.1230	-0.0023	**95**	32	47	8.50E-07	GSQDKSLHNR	ELSAERPLNEQIAEAEEDKI	KKTYPPENKP
SgIII	466.7270	2		931.4400	-0.0005	32	30	45	1.1	Signal	FPKPGGSQD	KSLHNRELSA
SgIII	630.8329	2		1259.6510	0.0002	27	25	45	3.4	Signal	FPKPGGSQDKSL	HNRELSAERP
SgIII	504.5913	3	2	1510.7528	-0.0007	**48**	22	45	0.025	Signal	FPKPGGSQDKSLHN	RELSAERPLN
SgIII	*862.6868*	4		3446.7178	0.0004	*22*	*19*	48	*24*	*Signal*	*FPKPGGSQDKSLHNRELSAERPLNEQIAEAE*	*EDKIKKTYPP*
SgIII	923.7038	4		3690.7873	-0.0013	**50**	19	48	0.033	Signal	FPKPGGSQDKSLHNRELSAERPLNEQIAEAEED	KIKKTYPPEN
SgIII	787.4007	5	6	3931.9663	0.0008	**59**	18	49	0.0055	Signal	FPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKI	KKTYPPENKP
SgIII	661.1581	6		3960.9050	0.0000	**71**		48	0.00027	IEWLKKHDKK	GNKEDYDLSKMRDFINKQADAYVEKGILDKEEAE	AIKRIYSSL
SgIII	830.0125	5	6	4145.0262	-0.0002	**95**	23	48	1.40E-06	IEWLKKHDKK	GNKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAI	KRIYSSL
SgIII	999.5105	5	6	4992.5178	-0.0018	**177**		50	1.10E-14	IEWLKKHDKK	GNKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL	C-term
SgIII	591.8170	2		1181.6193	0.0001	**47**	23	44	0.026	TEAYLEAIRK	NIEWLKKHD	KKGNKEDYDL
SgIII	866.9154	2		1731.8163	-0.0001	**52**	27	45	0.014	KIEKERQSIR	SSPLDNKLNVEDVDST	KNRKLIDDYD
SST	480.2160	2		958.4178	-0.0004	31	26	42	0.82	QDEMRLELQR	SANSNPAMAP	RERKAGCKNF
SST	622.7879	2		1243.5615	-0.0002	**74**	20	42	3.30E-05	QDEMRLELQR	SANSNPAMAPRE	RKAGCKNFFW
TMSB10	987.7114	5	6, 7	4933.5229	-0.0022	**131**		49	3.80E-10	M	Ac-ADKPDMGEIASFDKAKLKKTETQEKNTLPTKETIEQEKRSEIS	C-term
TMSB10	825.9273	6		4949.5179	0.0021	47		49	0.089	M	Ac-ADKPDMGEIASFDKAKLKKTETQEKNTLPTKETIEQEKRSEIS	C-term
TMSB4X	1241.1269	4	5, 6, 7	4960.4862	-0.0077	**100**		49	5.10E-07	M	Ac-SDKPDMAEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES	C-term
TMSB4X	996.3027	5		4976.4811	-0.0041	**93**		49	2.20E-06	M	Ac-SDKPDMAEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES	C-term
VGF	892.9472	4		3567.7594	0.0003	47	23	48	0.074	Signal	APPGRPEAQPPPLSSEHKEPVAGDAVPGPKDGSAPE	VRGARNSEPQ
VGF	1223.2823	3	4, 5	3666.8278	-0.0027	**72**		48	0.00024	Signal	APPGRPEAQPPPLSSEHKEPVAGDAVPGPKDGSAPEV	RGARNSEPQD
VGF	776.9976	5		3879.9504	0.0014	23	18	49	22	Signal	APPGRPEAQPPPLSSEHKEPVAGDAVPGPKDGSAPEVRG	ARNSEPQDEG
VGF	988.7539	4	5, 6	3950.9875	-0.0012	**119**	23	49	5.40E-09	Signal	APPGRPEAQPPPLSSEHKEPVAGDAVPGPKDGSAPEVRGA	RNSEPQDEGE
VGF	779.8765	2		1557.7383	0.0002	**53**	23	45	0.0078	YPGREAQARR	AQEEAEAEERRLQ	EQEELENYIE
VGF	927.4661	4	5, 6	3705.8346	0.0009	**112**		48	2.40E-08	YPGREAQARR	AQEEAEAEERRLQEQEELENYIEHVLLRRP	C-term
VGF	542.5338	4		2166.1069	-0.0010	39	16	46	0.3	EAEAEERRLQ	EQEELENYIEHVLLRRP	C-term
VGF	807.0214	3	2	2418.0418	0.0005	**101**	28	41	4.80E-08	GARQRGLGGR	GLQEAAEERESARE EEEEAEQE	RRGGEERVGE
VGF	507.0039	4	3	2023.9877	-0.0011	44	27	46	0.098	LQPPSALRRR	HYHHALPPSRHYPGREA	QARRAQEEAE
VGF	718.3561	3	4	2152.0463	0.0002	**56**	20	46	0.005	LQPPSALRRR	HYHHALPPSRHYPGREAQ	ARRAQEEAEA
VGF	742.0357	3	4	2223.0834	0.0018	**63**	21	47	0.0014	LQPPSALRRR	HYHHALPPSRHYPGREAQA	RRAQEEAEAE
VGF	537.6075	3		1609.8014	-0.0006	19	15	46	28	LQPPSALRRR	HYHHALPPSRHYP-NH2	GREAQARRAQ

TABLE I—*continued*

VGF	543.5572	4	3	2170.2011	-0.0015	69	21	44	0.00019	IEEVEEKRKR	KKNAPPEPVPPPRAAPAPTHV	RSPQPPPPAP
VGF	602.8200	4		2407.2495	0.0015	68	19	47	0.00043	QEEAEAEERR	LQEQEELENYIEHVLLRRP	C-term
VGF	909.6714	5	6	4543.3210	-0.0002	145		49	1.50E-11	PPAPSQFQAR	MPDSGPLPETHKFGEGVSSPKTHLGEALAPLSKAYQGVAAPFPK	ARRPESALLG
VGF	923.8788	5	6	4614.3581	-0.0002	118	22	49	7.60E-09	PPAPSQFQAR	MPDSGPLPETHKFGEGVSSPKTHLGEALAPLSKAYQGVAAPFPKA	RRPESALLGG
VGF	585.8179	2		1169.6193	0.0019	43	23	46	0.11	EVEEKRKRKK	NAPPEPVPPPR	AAPAPTHVRS
VGF	839.9493	2		1677.8838	0.0002	83	28	46	1.10E-05	EVEEKRKRKK	NAPPEPVPPPRAAPAPT	HVRSPQPPPP
VGF	958.0125	2	3	1914.0112	-0.0008	89	22	46	3.30E-06	EVEEKRKRKK	NAPPEPVPPPRAAPAPTHV	RSPQPPPPAP
VGF	958.9349	2		1915.8548	0.0005	35	18	43	0.35	GSAPEVRGAR	NSEPQDEGELFQGVDPR	ALAAVLLQAL
VGF	828.8768	2		1655.7387	0.0002	77	23	43	2.10E-05	RDFSPSSAKR	<QQETAAAETETRTHT	LTRVNLESPG
VGF	747.3683	3		2239.0829	0.0001	48	17	47	0.043	RDFSPSSAKR	<QQETAAAETETRTHTLTRVN	LESPGPERVW
VGF	848.4226	4	3	3389.6600	0.0011	93	19	48	1.70E-06	RDFSPSSAKR	<QQETAAAETETRTHTLTRVNLESPGPERVW	RASWGEFQAR
VGF	629.9729	3		1886.8970	-0.0002	23	14	46	9.3	RDFSPSSAKR	QQETAAAETETRTHTLT	RVNLESPGPE
VGF	989.4860	3		2965.4377	-0.0017	72	17	47	0.00021	RDFSPSSAKR	QQETAAAETETRTHTLTRVNLESPGPE	RVWRASWGEF
VGF	852.6789	4	3, 5	3406.6865	-0.0002	118	22	48	5.70E-09	RDFSPSSAKR	QQETAAAETETRTHTLTRVNLESPGPERVW	RASWGEFQAR
VGF	572.2870	3		1713.8394	-0.0002	54	25	46	0.009	HYPGREAQAR	RAQEEAEAEERRLQ	EQEELENYIE
VGF	773.3947	5		3861.9357	0.0013	56		49	0.0099	HYPGREAQAR	RAQEEAEAEERRLQEQEELENYIEHVLLRRP	C-term
VGF	641.8454	4		2563.3506	0.0018	59	22	46	0.0032	AQEEAEAEER	RLQEQEELENYIEHVLLRRP	C-term
VGF	442.6910	2		883.3672	0.0003	32	25	41	0.41	DGEAGAEDKR	SQEETPGH	RRKEAEGTEE
VGF	727.3609	2		1452.7069	0.0003	17	16	44	32	DGEAGAEDKR	SQEETPGHRRKE	AEGTEEGGEE
VGF	698.0365	3		2091.0861	0.0016	40	17	46	0.24	ETAAAETETR	THTLTRVNLESPGPERVW	RASWGEFQAR
VGF	691.8575	2		1381.6990	0.0014	35	24	46	0.88	TETRTHTLTR	VNLESPGPERVW	RASWGEFQAR
VGF	577.6458	3		1729.9151	0.0003	37	21	45	0.44	RASWGEFQAR	VPERAPLPPPAPSQFQ	ARMPDSGPLP

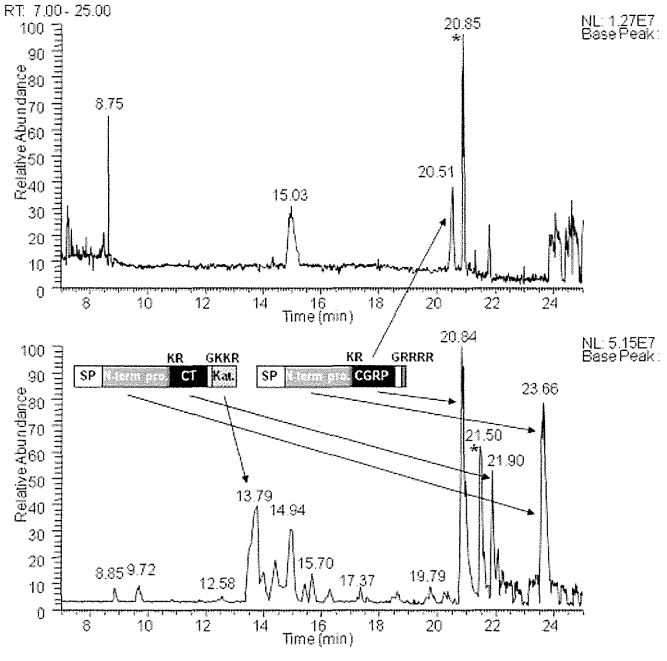


FIG. 3. **Representative base peak chromatograms of the secretopeptidome from unstimulated (*top*) and stimulated (*bottom*) cells.** Samples without gel filtration chromatography were analyzed. Major processing products of CT and CGRP precursors are illustrated along with *arrows* pointing at their peaks in the chromatogram. The base peaks marked with *asterisks* at 20.85 min (unstimulated) and 21.50 min (stimulated) are intact ubiquitin. *SP*, signal peptide; *N-term pro.*, N-terminal propeptide; *Kat.*, katacalcin. Base peaks at 8.75 and 15.03 min in the *top panel* were unrelated to peptide signals. *RT*, retention time; *NL*, normalized ion intensity.

With regard to the SST precursor, the 1243.56-Da peptide corresponds to the first 12 amino acids of SST-28. The dibasic sites RK (position 78–79) located downstream of this 12-residue peptide (Table II) is the known cleavage site for SST-14 (AGCKNFFWKTFTSC) (20). The single arginine (position 65) flanking the 1243.56-Da peptide and SST-28 is also an established processing site (20). At position −4 relative to this scissible bond (referred to as the P4 position), a single arginine (position 62) exists and forms a consensus PC cleavage site. As for the GRP precursor, the 1599.87-Da peptide is C-terminally flanked by an atypical single arginine, which is followed by intact neuromedin C (21). This arginine is known as an established processing site, although responsible enzymes remain to be identified. Thus, a majority of the cleavage sites of peptide hormone precursors were consistent with the previously identified processing sites. Except for precursor C termini, all the peptides derived from granin-like precursors (SgII, SgIII, and VGF) retained informative cleavage sites defined in this study. The 2677.41-Da VGF-derived amidated peptide was recently identified and designated NERP-1 (16).

Integrity of the Secretopeptidome Demonstrated by an In-depth Analysis—The secretopeptidome was separated into four fractions using gel filtration HPLC to perform an in-depth analysis. This analysis contributed to a substantial increase in sequenced peptides, and thus we were able to identify a total of 400 peptides from 23 precursors (Fig. 4 and the supplemental table). Some peptides arose from precursors not identified by the one-dimensional analysis; these included peptide hormone precursors neuromedin U and ghrelin, processing enzymes PC1 and peptidyl-glycine α-amidating monooxygenase, and the calcium-binding protein calnuc. The identification of PC1- and peptidyl-glycine α-amidating monooxygenase-derived peptides confirmed the integrity of this secretopeptidome as they are enzymes involved in the RSP proteolytic processing. Four calnuc-derived peptides had typical cleavage sites suggestive of the PC function (supplemental table). Recently calnuc was identified in a soluble fraction of bovine adrenal secretory granules (13). Altogether it is likely that nearly 99% of sequenced peptides were released upon exocytosis, which again demonstrates that this secretopeptidome is extremely rich in peptides stored in secretory granules.

We examined a total of 400 sequenced peptides to see whether they meet the criteria for the PC consensus sites (supplemental table). PC consensus sites were found in 299 peptides, and signal sequence cleavage sites were found in 43

TABLE II

Identity of the major peptides released on exocytosis (base peak intensities beyond 2e+06 in the Fig. 3 base peak chromatogram)

Values in brackets represent those obtained by peptides after reductive alkylation. For the ubiquitin MSMS spectrum, see supplemental Fig. 2. Note that the MSMS spectra of the 1278.67- and 5687.91-Da peptides (supplemental Fig. 3) did not meet the homology threshold criteria but were considered identified as described in the text. Grayscale boxes are defined in Table I legend. Oxidized methionine residues are underlined in column 10. TMSB4X, thymosin β-4 X-linked; TMSB10, thymosin β-10; RT, retention time; obsd, observed; Mr(Calc), theoretical monoisotopic molecular mass (Da) based on the peptide sequence; Expect., Expectation; N-term, the N-terminal flanking 10 amino acids; C-term, the C-terminal flanking 10 amino acids.

RT (min)	m/z (obsd)	z	Mr(Calc)	Mass error (ppm)	Base peak intensity	Score	Expect. value	N-term	Sequence	C-term	Pre-cursor	Validated by
8.85	622.7882	2	1243.5615	0.27	4.33E+06	74	3.30E-05	QDEMRLELQR	SANSNPAMAPRE	RKAGCKNFFW	SST	
9.72	504.5915	3	1510.7528	-0.09	4.74E+06	48	0.025	Signal	FPKPGGSQDKSLHN	RELSAERPLN	SgII	
12.55	613.7739	4	2451.0655	0.40	2.78E+06	59	0.0009	IGVGAPGKKR	DMSSDLERDHRPHVSMPQNAN	C-term	CT	
13.51	639.0115	3	1914.0112	0.76	1.17E+07	89	0.0000033	EVEEKRKRKK	NAPPEPVPPPRAAPAPTHV	RSPQPPPPAP	VGF	
13.59	609.7757	4	2435.0706	1.26	1.29E+07	65	0.00031	IGVGAPGKKR	DMSSDLERDHRPHVSMPQNAN	C-term	CT	
13.68	751.0048	3	2249.9906	0.87	4.65E+06	45	0.03	IGVGAPGKKR	DMSSDLERDHRPHVSMPQN	AN	CT	
13.79	659.5054	6	3950.9875	0.31	1.96E+07	119	5.40E-09	Signal	APPGRPEAQPPPLSSEHKEPVAGDAVPGPKDGSAPEVRGA	RNSEPQDEGE	VGF	
14.01	774.6837	3	2321.0277	0.67	7.48E+06	65	0.00033	IGVGAPGKKR	DMSSDLERDHRPHVSMPQNA	N	CT	
14.42	734.3739	5	3666.8278	1.45	9.82E+06	72	0.00024	Signal	APPGRPEAQPPPLSSEHKEPVAGDAVPGPKDGSAPEV	RGARNSEPQD	VGF	
14.94	827.7560	6	4960.4862	1.23	1.58E+07	100	5.10E-07	M	SDKPDMAEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES	C-term	TMSB4X	
14.94	633.3134	2	1264.6122	0.03	2.66E+06	69	0.00022	HKQQLERDPR	VKMALQQEGFD	RKKRGYRDIN	PC2	
15.46	823.2621	6	4933.5229	1.22	5.12E+06	131	3.80E-10	M	ADKPDMGEIASFDKAKLKKTETQEKNTLPTKETIEQEKRSEIS	C-term	TMSB10	
15.70	640.3438	2	1278.6721	0.74	7.21E+06	28	0.86	LSRSGGVVKN	NFVPTNVGSKAF-NH2	GRRRDLQA	CGRP	manual
16.14	787.4012	5	3931.9663	0.84	2.37E+06	59	0.0055	Signal	FPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKI	KKTYPPENKP	SgII	
16.30	682.3453	5	3406.6865	1.06	4.07E+06	118	5.70E-09	RDFSPSSAKR	QQETAAAETETRTHTLTRVNLESPGPERVW	RASWGEFQAR	VGF	
17.37	848.4229	4	3389.6600	0.73	4.51E+06	93	1.70E-06	RDFSPSSAKR	<QQETAAAETETRTHTLTRVNLESPGPERVW	RASWGEFQAR	VGF	
17.60	762.0490	3	2283.1230	0.95	2.56E+06	95	8.50E-07	GSQDKSLHNR	ELSAERPLNEQIAEAEEDKI	KKTYPPENKP	SgII	
18.64	800.9424	2	1599.8695	0.46	3.70E+06	50	0.011	Signal	VPLPAGGGTVLTKMYP	RGNHWAVGHL	GRP	
19.65	632.5773	4	[2642.3381]	[-1.20]	2.39E+06	[98]	[2.50E-07]	EGSRIIAQKR	ACDTATCVTHRLAGLLSRSGGVVKN	NFVPTNVGSK	CGRP	
19.79	819.3663	2	[1752.7753]	[-0.79]	4.71E+06	[70]	[4.90E-05]	NPAMAPRERK	AGCKNFFWKTFTSC	C-term	SST	
19.91	881.7797	3	2642.3187	-0.55	2.47E+06	106	8.60E-08	Signal	APFRSALESSPADPATLSEDEARLL	LAALVQDYVQ	CT/CGRP	
20.25	577.2836	3	[1844.8873]	[-1.10]	3.57E+06	[98]	[1.20E-07]	EGSRIIAQKR	ACDTATCVTHRLAGLLS	RSGGVVKNNF	CGRP	
20.33	711.9964	8	5687.9079	0.89	2.13E+06	44	0.19	Signal	ERPVFTNHFLVELHKGGEDKARQVAAEHGFGVRKLPFAEGLYHFYHNGLA	KAKRRRSLHH	PC2	manual
20.37	1088.2665	4	4349.0419	-1.15	3.73E+06	112	1.50E-08	EIINSNQVKR	VPGQGSSEDDLQEEEQIEQAIKEHLNQGSSQETDKLAPVS	KRFPVGPPKN	SgI	
20.60	1144.8639	3	[3547.6323]	[-1.68]	2.29E+06	[62]	[0.00049]	SSLDSPRSKR	CGNLSTCMLGTYTQDFNKFHTFPQTAIGVGAP-NH2	GKKRDMSSDL	CT	
20.84	758.3990	5	[3902.9996]	[-1.10]	5.05E+07	[117]	[2.80E-09]	EGSRIIAQKR	ACDTATCVTHRLAGLLSRSGGVVKNNFVPTNVGSKAF-NH2	GRRRDLQA	CGRP	
21.16	830.0130	5	4145.0262	0.58	2.61E+06	95	1.40E-06	IEWLKKHDKK	GNKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAI	KRIYSSL	SgII	
21.50	714.3080	12	8559.6167	-0.94	3.09E+07	8	2.20E+02		intact ubiquitin	C-term	Ubq	standard
21.52	742.1743	5	3705.8346	0.13	2.51E+07	112	2.40E-08	YPGREAQARR	AQEEAEAEERRLQEQEELENYIEHVLLRRP	C-term	VGF	
21.55	893.4781	3	2677.4147	-0.82	4.27E+06	69	0.00031	GVAAPFPKAR	RPESALLGGSEAGERLLQQGLAQVEA-NH2	GRRQAEATRQ	VGF	
21.90	1139.5341	3	[3531.6374]	[-1.57]	2.68E+07	[78]	[1.30E-05]	SSLDSPRSKR	CGNLSTCMLGTYTQDFNKFHTFPQTAIGVGAP-NH2	GKKRDMSSDL	CT	
21.93	1158.8690	3	[3589.6429]	[-1.88]	9.68E+06	[64]	[0.0003]	SSLDSPRSKR	CGNLSTCMLGTYTQDFNKFHTFPQTAIGVGAPG	KKRDMSSDLE	CT	
22.25	833.0936	6	4992.5178	0.02	7.83E+06	177	1.10E-14	IEWLKKHDKK	GNKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL	C-term	SgII	
23.66	1037.1808	6	6217.0396	0.24	3.95E+07	143	2.50E-11	Signal	APFRSALESSPADPATLSEDEARLLLAALVQDYVQMKASELEQEQEREGSSLDSPRS	KRCGNLSTCM	CT	
23.69	1010.3493	6	6056.0436	1.41	3.91E+07	156	1.60E-12	Signal	APFRSALESSPADPATLSEDEARLLLAALVQDYVQMKASELEQEQEREGSRIIAQ	KRACDTATCV	CGRP	

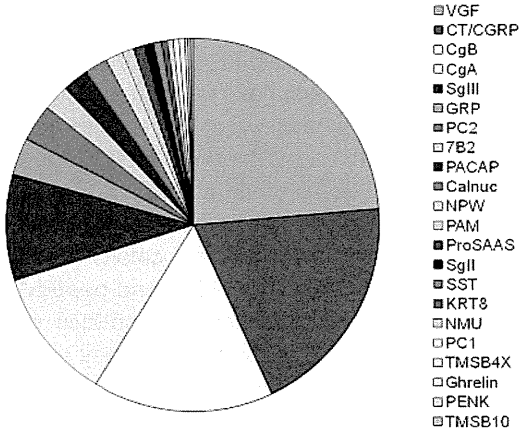


FIG. 4. **Pie representation of the 400 peptides (sorted by names of 23 precursors) secreted after stimulation.** *PACAP*, pituitary adenylate cyclase-activating polypeptide; *NPW*, neuropeptide W; *PENK*, proenkephalin A; *KRT*, cytokeratin; *TMSB4X*, thymosin β-4 X-linked; *TMSB10*, thymosin β-10; *PAM*, peptidyl-glycine α-amidating monooxygenase; *NMU*, neuromedin U.

peptides. Monobasic sites were found in 98 peptides. Overall a total of 373 peptides were shown to have informative cleavage sites to predict the precursor processing in the RSP. As an example, the sequences of CT-derived peptides listed in Tables I and II were mapped to the precursor sequence (Fig. 5). The map consists of the known major processing products and intermediate products sharing N or C termini, which correspond to the established PC1/3 or PC2 cleavage sites. Consistent with the reported length of its signal peptide (17), no peptide was identified in the first 25 amino acids. Overall this map is reminiscent of CT precursor processing previously elucidated in parafollicular cells of the thyroid gland (17).

DISCUSSION

The most outstanding finding in the present study is that an exocytotic stimulus applied to cultured endocrine cells is highly effective in identifying secretory peptides. Peptide profiles identified with this protocol strongly suggest that most peptides were released from secretory granules on exocytosis (Fig. 4). It should be noted that in a total of 400 identified peptides nearly 97% arose from previously known RSP precursor proteins. This non-invasive approach dispenses with time-consuming procedures such as subcellular fractionation and can be extended to different cell culture models. To the best of our knowledge, this is the first study ever to conduct a comprehensive analysis focused on peptides from the RSP proteins.

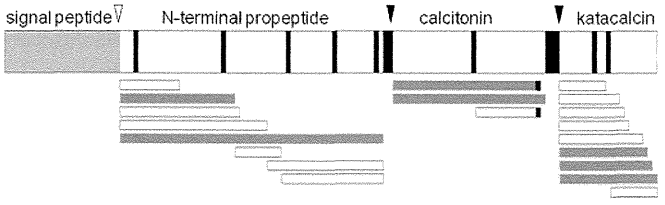


FIG. 5. **CT precursor processing deduced by a panel of identified peptides.** The established signal peptide cleavage site (*open arrowhead*) and known processing sites (*closed arrowheads*) are shown across the *top* of the CT precursor with basic residues (*thin black boxes*) and the signal peptide (*hatched box*) indicated. Sequences of identified peptides (*bars*) are detailed in Tables I and II. Major peptides, defined in Table II, are indicated by *gray bars*. *Black boxes* denote C-terminally amidated residues.

Most peptidomics studies have dealt with endocrine organs or brains, which are considered major sources of peptide hormones or neuropeptides (4, 7–11). It is now recognized that endogenous proteases must be inactivated before tissue extraction to prevent massive production of peptide fragments caused by degradation of abundant intracellular proteins, which hampers MS detection of endogenous peptides (7, 8). For tissue peptidome studies, microwave irradiation before or after decapitation has been proposed to prevent protease activation that is thought to occur immediately after sacrifice (7, 8). Despite these efforts, a limited number of secretory peptides were identified (4, 7–11).

Regarding MSMS of naturally occurring peptides, precursor mass acquisition with conventional mass spectrometers (mass accuracy, ~50 ppm) and subsequent filtering in the setting of "no enzyme" very often lead to ambiguous peptide identification (22). To cope with this issue, many peptidomics studies have used in-house databases containing a limited number of entries to identify peptides that would otherwise remain elusive (11, 23). On the other hand, we used the public IPI human database and took into account four variable modifications simultaneously for MSMS interpretation. Because of the mass accuracy of Orbitrap (~2 ppm), a false discovery rate using a decoy database was minimized to 0% for peptide matches above identity thresholds in the Mascot MSMS ion search. This MSMS identification scheme may inevitably miss many peptides, which could be considered identified if a specific database with limited entries was used as in previous peptidomics studies (11, 23). The peptides listed in the supplemental table are all beyond identity thresholds; accepting peptides beyond homology thresholds will allow more than 200 additional secretory peptides to enter the table, with only two different peptides from keratin 8 turning up (data not shown). These examples include neuromedin U 25 (3018.52 Da) and pancreastatin (5076.36 Da) with a Mascot expectation value of 0.063 and 0.43, respectively. In the supplemental table, they are not considered identified because an expectation value for accepting MSMS spectra is the Mascot default significant threshold value of 0.05. Nonetheless we used the stringent setting (described under "Experimental Procedures") to preclude misleading assignments and to demonstrate that this secretopeptidome shows little contamination by non-secretory components.

In the present study, we made every effort to prevent peptide degradation or chemical modifications (deamidation, methionine oxidation, and pyroglutamination) that may occur during sample preparation. As described under "Experimental Procedures," peptides released during 2 min were immediately subjected to solid phase extraction. Peptide extraction was performed at 4 °C and completed within 20 min after harvesting the supernatant. In addition, lyophilized samples were analyzed by LC-MSMS immediately after reconstitution. Even with these attentive procedures, several clusters of N- or C-terminally truncated peptides that share cleavage sites at the other end were sequenced as reported in previous peptidomics studies on tissue peptidomes (4, 7–11). However, most peptides (30 of 35) dominantly detected in LC-MSMS (Table II, Fig. 3, and supplemental Fig. 1) did not have cleavage sites suggestive of exopeptidase digestion aside from the five CT- or CGRP-derived peptides mentioned under "Results." These five peptides appeared to be N- or C-terminally truncated peptides of major processing products. They have not been reported as major processing products in previous biochemical studies to the best of our knowledge. However, the possibility that their C termini or N termini were generated by unknown endopeptidases (cutting within a peptide) cannot be excluded.

It was unexpected that thymosins and ubiquitin represented major peptides in the TT secretopeptidome. Given a previous report of its storage in adrenal secretory granules and exocytosis-induced secretion (24), ubiquitin may be localized in TT secretory granules and secreted upon exocytosis. The secretory nature of thymosin β-4 has also been reported (25); however, thymosins are not regarded as peptides localized in secretory granules. In any case, the successful identification of intact peptide forms indicates that a majority of peptides may not be affected by exopeptidase digestion.

Gel filtration-based separation caused an increase in the number of sequenced peptides among which several peptides suggestive of the PC-mediated cleavage were identified, such as those from the calcium-binding protein calnuc (supplemental table). Calnuc was identified in a soluble fraction of bovine adrenal secretory vesicles (13), and hence our finding suggests that it is a precursor to unknown bioactive peptides.

To identify unique cleavage sites of RSP precursor proteins, we examined N- and C-terminal flanking sequences of the 400 peptides identified. Overall 152 unique cleavage sites that match PC consensus sites were elucidated of which 105 cleavage sites were conserved consecutive dibasic residues. This finding appears to support the contention revealed by previous studies that the most often encountered PC cleavage sites are conserved paired dibasic sites (1, 2, 19). The

observation that a majority of cleavage sites were consistent with PC consensus sites ((R/K)X_n(R/K) where n = 0, 2, 4, or 6) should not be overestimated. For instance, the C-terminal cleavage site of the CGRP precursor (RLAGLLS ↓ RSG) matches this rule but is not known as a processing site. At present its abundance relative to the longer major product of intact CGRP was unavailable, and therefore it remains to be clarified whether the peptide would represent a major processing product secreted by TT cells. We should also consider that shorter peptides tend to be better ionized and readily detected in mass spectrometry schemes.

Conversely cleavages at non-consensus monobasic sites could represent *bona fide* initial endoproteolytic cleavage sites. In the present study, CgA-derived peptides (1048.51 and 1161.59 Da) sharing the N-terminal cleavage site (FRAR ↓ AYGF) were identified (supplemental table), suggesting that the single arginine (position 356) is a processing site. The C-terminal cleavage site of the 1914.01-Da peptide (VGF residues 463–481) had an arginine at the P10 position. Indeed they have both been shown to be an authentic recognition site for PC2 using PC2 knock-out mice (26). Another example is the N-terminal cleavage site (LSFR ↓ ARAY) of the CgA 1275.65- and 1388.73-Da peptides known as the processing site for CgA LF-19 peptide (27). Thus atypical monobasic sites should also be considered potential processing sites for a precursor whose processing remains largely unknown. In this context, the single arginine of VGF (position 212) in FQAR ↓ MPDS may represent a cleavage site as shown by six peptides (supplemental table). It is envisaged that further detailed analysis of the secretopeptidome could identify processing sites with higher confidence. Although the peptide repertoire from a cancer cell line does not necessarily reflect the *in vivo* processing pattern of RSP precursor proteins, this will not detract from the significance of our study. Indeed the peptides identified with our approach retain cleavage sites created in the RSP to a degree that allows the accurate prediction of processing sites in known peptide hormone precursors (Fig. 5). In summary, we showed that peptidomics has the potential to identify processing sites of precursors processed in the RSP. By dissecting the secretopeptidome we should have a clearer picture of the precursor processing that actually occurs in the RSP that would also facilitate the discovery of bioactive peptides.

Acknowledgment—We thank Gary S. Goldberg (University of Medicine and Dentistry of New Jersey) for critical reading of the manuscript.

* This work was supported in part by the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation and by grants-in-aid for scientific research from the Japanese Society for the Promotion of Science, Japan.

[S] The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

§ To whom correspondence may be addressed. Tel.: 81-6-6833-5004; Fax: 81-6-6835-5349; E-mail: ksasaki@ri.ncvc.go.jp.

‖ Present address: Pharmaceutical Research Division, Discovery Research Center, Takeda Pharmaceutical Co. Ltd., 17-85 Jusohonmachi 2-chome, Yodogawa-ku, Osaka 532-8686, Japan.

** To whom correspondence may be addressed. Tel.: 81-6-6833-5004; Fax: 81-6-6835-5349; E-mail: minamino@ri.ncvc.go.jp.

REFERENCES

1. Zhou, A., Webb, G., Zhu, X., and Steiner, D. F. (1999) Proteolytic processing in the secretory pathway. *J. Biol. Chem.* **274,** 20745–20748
2. Fricker, L. D. (2005) Neuropeptide-processing enzymes: applications for drug discovery. *AAPS J.* **7,** E449–455
3. Brakch, N., Rholam, M., Boussetta, H., and Cohen, P. (1993) Role of beta-turn in proteolytic processing of peptide hormone precursors at dibasic sites. *Biochemistry* **32,** 4925–4930
4. Clynen, E., Baggerman, G., Veelaert, D., Cerstiaens, A., Van der Horst, D., Harthoorn, L., Derua, R., Waelkens, E., De Loof, A., and Schoofs, L. (2001) Peptidomics of the pars intercerebralis-corpus cardiacum complex of the migratory locust, Locusta migratoria. *Eur. J. Biochem.* **268,** 1929–1939
5. Schrader, M., and Schulz-Knappe, P. (2001) Peptidomics technologies for human body fluids. *Trends Biotechnol.* **19,** S55–60
6. Sasaki, K., Sato, K., Akiyama, Y., Yanagihara, K., Oka, M., and Yamaguchi, K. (2002) Peptidomics-based approach reveals the secretion of the 29-residue COOH-terminal fragment of the putative tumor suppressor protein DMBT1 from pancreatic adenocarcinoma cell lines. *Cancer Res.* **62,** 4894–4898
7. Svensson, M., Sköld, K., Svenningsson, P., and Andren, P. E. (2003) Peptidomics-based discovery of novel neuropeptides. *J. Proteome Res.* **2,** 213–219
8. Che, F. Y., Lim, J., Pan, H., Biswas, R., and Fricker, L. D. (2005) Quantitative neuropeptidomics of microwave-irradiated mouse brain and pituitary. *Mol. Cell. Proteomics* **4,** 1391–1405
9. Boonen, K., Baggerman, G., D'Hertog, W., Husson, S. J., Overbergh, L., Mathieu, C., and Schoofs, L. (2007) Neuropeptides of the islets of Langerhans: a peptidomics study. *Gen. Comp. Endocrinol.* **152,** 231–241
10. Cape, S. S., Rehm, K. J., Ma, M., Marder, E., and Li, L. (2008) Mass spectral comparison of the neuropeptide complement of the stomatogastric ganglion and brain in the adult and embryonic lobster, Homarus americanus. *J. Neurochem.* **105,** 690–702
11. Bora, A., Annangudi, S. P., Millet, L. J., Rubakhin, S. S., Forbes, A. J., Kelleher, N. L., Gillette, M. U., and Sweedler, J. V. (2008) Neuropeptidomics of the supraoptic rat nucleus. *J. Proteome Res.* **7,** 4992–5003
12. Wegrzyn, J., Lee, J., Neveu, J. M., Lane, W. S., and Hook, V. (2007) Proteomics of neuroendocrine secretory vesicles reveal distinct functional systems for biosynthesis and exocytosis of peptide hormones and neurotransmitters. *J. Proteome Res.* **6,** 1652–1665
13. Brunner, Y., Couté, Y., Iezzi, M., Foti, M., Fukuda, M., Hochstrasser, D. F., Wollheim, C. B., and Sanchez, J. C. (2007) Proteomics analysis of insulin secretory granules. *Mol. Cell. Proteomics* **6,** 1007–1017
14. Gkonos, P. J., Born, W., Jones, B. N., Petermann, J. B., Keutmann, H. T., Birnbaum, R. S., Fischer, J. A., and Roos, B. A. (1986) Biosynthesis of calcitonin gene-related peptide and calcitonin by a human medullary thyroid carcinoma cell line. *J. Biol. Chem.* **261,** 14386–14391
15. Zabel, M., Seidel, J., Kaczmarek, A., Surdyk-Zasada, J., Grzeszkowiak, J., and Górny, A. (1994) Hybridocytochemical and immuno-ultrastructural study of calcitonin gene expression in cultured medullary carcinoma cells. *Histochemistry* **102,** 323–327
16. Yamaguchi, H., Sasaki, K., Satomi, Y., Shimbara, T., Kageyama, H., Mondal, M. S., Toshinai, K., Date, Y., González, L. J., Shioda, S., Takao, T., Nakazato, M., and Minamino, N. (2007) Peptidomic identification and biological characterization of neuroendocrine regulatory peptide-1 and -2. *J. Biol. Chem.* **282,** 26354–26360
17. Jacobs, J. W., Goodman, R. H., Chin, W. W., Dee, P. C., Habener, J. F., Bell, N. H., and Potts, J. T., Jr. (1981) Calcitonin messenger RNA encodes multiple polypeptides in a single precursor. *Science* **213,** 457–459
18. Cameron, A., Apletalina, E. V., and Lindberg, I. (2002) *The enzymology of PC1 and PC2*. In *The Enzymes* (Dalby, R. E., and Sigman, D. S., eds) 3rd Ed., pp. 291–328, Academic Press, NY
19. Eipper, B. A., Stoffers, D. A., and Mains, R. E. (1992) The biosynthesis of neuropeptides: peptide alpha-amidation. *Annu. Rev. Neurosci.* **15,** 57–85

20. Shen, L. P., Pictet, R. L., and Rutter, W. J. (1982) Human somatostatin I: sequence of the cDNA. *Proc. Natl. Acad. Sci. U. S. A.* **79,** 4575–4579
21. Minamino, N., Kangawa, K., and Matsuo, H. (1984) Neuromedin C: a bombesin-like peptide identified in porcine spinal cord. *Biochem. Biophys. Res. Commun.* **119,** 14–20
22. Brosch, M., Swamy, S., Hubbard, T., and Choudhary, J. (2008) Comparison of Mascot and X!Tandem performance for low and high accuracy mass spectrometry and the development of an adjusted mascot threshold. *Mol. Cell. Proteomics* **7,** 962–970
23. Fälth, M., Sköld, K., Norrman, M., Svensson, M., Fenyö, D., and Andren, P. E. (2006) SwePep, a database designed for endogenous peptides and mass spectrometry. *Mol. Cell. Proteomics* **5,** 998–1005
24. Kieffer, A. E., Goumon, Y., Ruh, O., Chasserot-Golaz, S., Nullans, G., Gasnier, C., Aunis, D., and Metz-Boutigue, M. H. (2003) The N- and C-terminal fragments of ubiquitin are important for the antimicrobial activities. *FASEB J.* **17,** 776–778
25. Bock-Marquette, I., Saxena, A., White, M. D., Dimaio, J. M., and Srivastava, D. (2004) Thymosin beta4 activates integrin-linked kinase and promotes cardiac cell migration, survival and cardiac repair. *Nature* **432,** 466–472
26. Pan, H., Che, F. Y., Peng, B., Steiner, D. F., Pintar, J. E., and Fricker, L. D. (2006) The role of prohormone convertase-2 in hypothalamic neuropeptide processing: a quantitative neuropeptidomic study. *J. Neurochem.* **98,** 1763–1777
27. Orr, D. F., Chen, T., Johnsen, A. H., Chalk, R., Buchanan, K. D., Sloan, J. M., Rao, P., and Shaw, C. (2002) The spectrum of endogenous human chromogranin A-derived peptides identified using a modified proteomic strategy. *Proteomics* **2,** 1586–1600

research articles Journal of proteome research

A Peptidomics Strategy for Discovering Endogenous Bioactive Peptides

Kazuki Sasaki,*,† Noriyuki Takahashi,‡ Mitsuo Satoh,§ Motoo Yamasaki,‡ and Naoto Minamino*,†

Department of Molecular Pharmacology, National Cerebral and Cardiovascular Center Research Institute, Suita, Osaka 565-8565, Japan, Innovative Drug Research Laboratories, Kyowa Hakko Kirin Co., Ltd., Machida, Tokyo 194-8533, Japan, and Antibody Research Laboratories, Kyowa Hakko Kirin Co., Ltd., Machida, Tokyo 194-8533, Japan

Received April 15, 2010

Peptide hormones and neuropeptides constitute an important class of naturally occurring peptides that are generated from precursor proteins by limited proteolytic processing. An important but unaddressed issue in peptidomics is to pin down novel bioactive peptides in a bulk of peptide sequences provided by tandem mass spectrometry. Here, we describe an approach to simultaneously screen for bioactive peptides and their target tissues. The principle behind this approach is to identify intact secretory peptides that have the ability to raise intracellular calcium levels. In practice, we used nanoflow liquid chromatography–tandem mass spectrometry to analyze peptides released by exocytosis from cultured cells. Peptide sequence information was utilized to deduce intact peptide forms, among which those highly conserved between species are selected and tested on an ex vivo calcium assay using tissue pieces from transgenic mice that systemically express the calcium indicator apoaequorin. The calcium assay can be applied to various cell types, including those not amenable to in vitro culture. We used this approach to identify novel bioactive neuropeptides derived from the neurosecretory protein VGF, which evoke a calcium response in the pituitary and hypothalamus.

Keywords: peptidomics • proteolytic processing • apoaequorin • secretory peptides • bioactive peptides

Introduction

Peptide hormones and neuropeptides function as cell-to-cell signaling molecules that mediate many physiological effects. The generation of these bioactive peptides involves a series of cleavage events orderly executed by specific proteases, starting with the endoproteolytic cleavage of precursor proteins to produce peptides with defined lengths.[1,2] These biosynthetic cleavages are often accompanied by post-translational modifications on specific residues, such as N-terminal acetylation and C-terminal amidation.[3] Resultant peptides, termed major processing products or "intact" peptides, are subsequently degraded and inactivated by a variety of proteases.

Peptidomics has been advocated to comprehensively study these naturally cleaved peptides that are beyond the reach of conventional proteomics.[4,5] Tandem mass spectrometry techniques enable the sequence determination of peptides present in complex mixtures for an organism whose genome sequencing nears completion. Efforts have been made to discover neuropeptides by analyzing the total peptide complement of the hypothalamus and pituitary, which are considered treasure troves of bioactive peptides.[6–9] However, despite initial enthusiasm, it has become clear that intact secretory peptides account for a relatively small proportion in a tissue peptidome. Even with known neuropeptides and peptide hormones, most of the identified peptides are N-terminally or C-terminally trimmed from intact peptides, implying that bioactive peptides, present in trace amounts, are difficult to identify in their native molecular forms.[6–9] Hence, delineating intact peptides emerges as a key factor in mass spectrometry-based peptidomic approaches to discovering bioactive peptides.

This "peptide first" approach, however, only leaves behind many peptide sequences and does not offer information about bioactivity or target tissues. One solution to the problem is to focus on post-translational modifications characteristic of known bioactive peptides, for instance C-terminal amidation.[3] Using tandem mass spectrometry, we profiled peptides secreted by cultured endocrine cells and identified novel C-terminally amidated peptides, designated NERP-1 and NERP-2.[10] Since C-terminal amidation does not guarantee that the peptides are bioactive, we were forced to conduct many hit-or-miss experiments to demonstrate that they are bioactive.[10] A more efficient method for identifying bioactive peptides would therefore be desired.

Besides the importance of investigating intact secretory peptides, we have considered the following common features

* To whom correspondence should be addressed. Department of Molecular Pharmacology, National Cerebral and Cardiovascular Center Research Institute, Fujishirodai 5-7-1, Suita, Osaka 565-8565, Japan. Phone: +81 6 6833 5004ex. 2600; Fax: +81 6 6835 5349. E-mail: ksasaki@ri.ncvc.go.jp or minamino@ri.ncvc.go.jp.

† National Cerebral and Cardiovascular Center Research Institute.

‡ Innovative Drug Research Laboratories, Kyowa Hakko Kirin Co., Ltd.

§ Antibody Research Laboratories, Kyowa Hakko Kirin Co., Ltd.

10.1021/pr1003455

Published on Web 08/04/2010

of known bioactive peptides: (1) strong interspecies homology at the amino acid level and (2) frequent use of calcium as a second messenger. In the present study, we describe an approach to screen for bioactive peptides, independent of specific modifications. The core part of this approach is to identify intact peptides released by exocytosis from cultured cells with secretory granules, followed by testing the ability of a candidate to raise intracellular calcium levels. We hypothesized that our transgenic mice systemically expressing apoaequorin, a calcium-sensitive photoprotein,[11] could be used to simultaneously screen for bioactive peptides and their target tissues. In the present study, we examined peptides derived from the neurosecretory protein VGF, the processing of which has not been elucidated, aside from the C-terminal region.[12] We used this peptidomic approach to identify two neuropeptides, hidden in the precursor sequence, that elicited a calcium transient in the pituitary and hypothalamus.

Materials and Methods

Peptide Preparation. Monolayer cultures of TT cells[13] were rinsed three times with Hanks medium (Invitrogen). Culture supernatants of the cells incubated for 15 min after stimulation with 10 μM forskolin plus 10 μM carbachol were harvested and rapidly extracted at 4C using an RP-1 solid-phase extraction cartridge (GL Sciences). Peptide fractions were obtained by HPLC on a gel filtration column (G2000SWXL, 21.5 × 300 mm, TOSOH) at 1.5 mL/min to obtain two fractions named G7−8 and G9−12, which contains peptides around 3000−7000 Da and smaller peptides (around 1000−3000 Da), respectively. Both fractions were reductive alkylated and desalted with Empore C18 cartridges (3M). The resultant samples were individually reconstituted in solvent A (10 mM ammonium formate, pH 3.8: acetonitrile (ACN) = 9:1 (v/v)) and applied to a TSK gel SP-2SW cation-exchange column (1.0 × 50 mm, TOSOH) and eluted at 50 uL/min with a gradient of 0−100% solvent B (1 M ammonium formate, pH 3.8: ACN = 9:1 (v/v)) in 30 min. Fraction G7−8 was divided into five fractions (collected every 5 min, starting 5 min after injection) and fraction G9−12 was into nine fractions (collected every 3 min, starting immediately after injection) using this cation-exchange HPLC. A total of 14 cation-exchange fractions were individually desalted with Empore C18 cartridges and analyzed by LC−MS/MS.

LC−MS/MS. NanoLC−MS/MS experiments were performed with a Chorus nanoflow system (CTC Analytics) connected to an LTQ-Orbitrap mass spectrometer (ThermoFisher Scientific) equipped with a nanoelectrospray emitter (MonoSpray C18 Nano, 100 um × 50 mm, GL Sciences).[14] Sample dissolved in 2% ACN and 0.1% formic acid was loaded via a PAL autosampler (CTC Analytics) onto an L column (Chemicals Evaluation and Research Institute, Japan) and eluted with a linear gradient from 5% ACN, 0.1% formic acid to 60% ACN, 0.1% formic acid over 40 min at 500 nL/min. A protonated ion of polycyclodimethylsiloxane with m/z 445.120025 was used for internal calibration throughout. The mass spectrometer was operated in a data-dependent mode to automatically switch between MS and MS/MS acquisitions. After survey full scan (400−1500 m/z range), five most intense ions (intensity threshold 2×10^5) were isolated for MS/MS in the linear ion trap using collision induced dissociation, with dynamic exclusion onward throughout the following scans. The resultant product ions were recorded in the Orbitrap. For each fraction, any precursor m/z (except for singly charged ions) was subjected to MS/MS and excluded thereafter if detected within a mass width of 5 ppm. Multiple identifications of the same sequence were allowed given that each identification was derived from a different precursor charge state. In addition, redundant identification of the same peptide, which was in most cases caused by distribution over multiple chromatographic fractions, was also allowed in this secretopeptidome analysis (Supplemental Table, Supporting Information).

Data Analysis and Peptide Identification. Peak picking, deisotoping, and deconvolution of MSMS spectra were performed using Mascot Distiller (ver. 2.1.1.0) with the default parameters for OrbiTrap. Peak lists were searched against IPI Human (76539 entries on March 4, 2009) using Mascot (ver. 2.2.04), with no enzyme specification. Mascot was used with monoisotopic mass selected, a precursor mass tolerance of 2 ppm, and a fragment mass tolerance of 25 mmu. Pyroglutamination and C-terminal amidation were simultaneously allowed as variable modifications. The significance threshold was the Mascot default setting of 5%. Each MSMS spectrum was checked manually to confirm or contradict the Mascot assignment. The false discovery rate for the identity threshold was in all cases 0% as estimated by using the Mascot decoy database function. Only peptides identified with a score above the Mascot identity threshold were considered in the present study.

Peptide Synthesis. All peptides were synthesized using Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) strategy, purified by reversed phase HPLC (SigmaGenosys), and verified for correct synthesis by mass spectrometry and amino acid analysis.

Antibody Preparation. Cysteinyl C-terminal peptides of rat NERP-3 (CESPGPERVW) and AQEE-30 (CEHVLLHRP) were synthesized and each coupled with maleimide activated keyhole limpet hemocyanin (Pierce). Rabbits were immunized with each conjugate and antisera were characterized as described.[15]

Mass Spectrometric Characterization of VGF-Derived Peptides. Rat whole brain excluding cerebellum was extracted and condensed with a SepPak C18 cartridge as previously described.[10] Sephadex G-50 gel filtrated fractions of the rat brain extract (2.38 and 1.40 g equivalent for NERP-3 and AQEE-30, respectively) were immunoprecipitated with the antibodies and analyzed on a surface-enhanced laser desorption ionization mass spectrometer (Ciphergen) as described.[16]

Calcium Assay Using Tissues from Apoaequorin Transgenic Mice. Assays were performed as previously described with a few modifications.[11] Tissues or organs from the apoaequorin transgenic mice were cut out as 1−2 mm^3 pieces and incubated with coelenterazine for 3 h. They were sequentially treated with RPMI medium containing 0.01% (w/v) BSA, test peptide (1 μM), a standard and Triton X-100 (final 2.5%). Triton X-100 was used to lyse cells and confirm the expression of functionally active aequorin. The standard refers to one of the following positive controls: angiotensin II (1 μM), bradykinin (1 μM) and ATP (100 μM). A positive control that constantly exhibited higher signals than any other control was used as standard for a given organ, where bradykinin served as standard for pituitary and uterus, ATP for hypothalamus, liver and heart, and angiotensin II for adrenal gland. The mean relative luminescent units (RLU) elicited by the standards from 20 assays (values in parentheses indicate RLU ranges) for the pituitary, hypothalamus, liver, adrenal gland, uterus, and heart were 40,000 (23 000−51 500), 880 (200−3950), 160 (100−300), 2020 (120−9200), 11 600 (1700−32 000), and 820 (120−3000), respectively. Signals induced by test peptides were defined as positive in case RLUs were above 100 for the pituitary or 50 for the others, or above

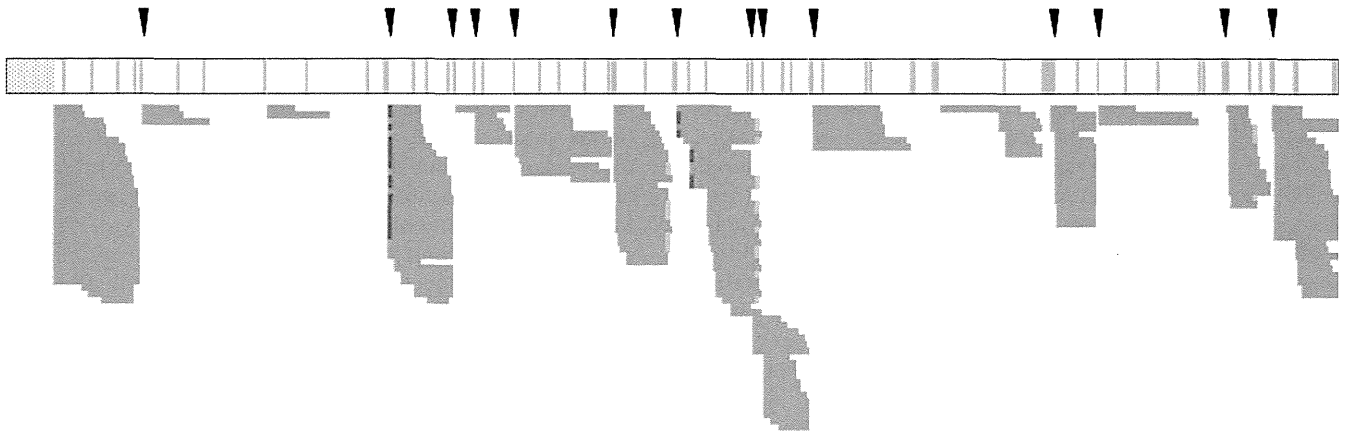


Figure 1. VGF processing deduced from identified peptides. Sequenced peptides are shown by gray boxes and detailed in Supplemental Table (Supporting Information). Closed arrowheads denote major cleavage sites across the top of the precursor, with basic residues (pale magenta boxes) and the signal sequence (a stippled box) indicated. Orange boxes, C-terminal amidation; red boxes, pyroglutamination.

5% of the RLUs induced by each standard in a given experiment. Peptides were dissolved in RPMI containing 0.01% (w/v) fatty acid-free BSA (Sigma) treated with activated charcoal. Data for test peptides were obtained from three or four independent experiments. A given peptide was considered positive on a tested tissue if positive signals were observed in at least two experiments. All the peptides were used at 1 μM throughout due to limited sensitivity inherent in this assay; endothelin-1, known as a potent vasoconstrictive peptide, did not produce a positive signal below 0.1 μM (data not shown).

Results and Discussion

The human thyroid cell line TT was stimulated with carbachol plus forskolin for 15 min to recover peptides secreted in the medium. The TT secretopeptidome was separated by gel filtration and subsequent cation exchange chromatography into 14 fractions for LC–MS/MS. Among the MS/MS-identified peptides, we studied peptides derived from VGF, a neurosecretory protein.[12,17] *Vgf*, expressed in various neurons and endocrine cells, encodes a 68-kDa protein of 617 amino acids in rodents and of 615 in human that bears several consensus motifs for processing enzymes PC1/3 and PC2 to yield peptides secreted via the regulated secretory pathway.[12] Studies of *Vgf*-deficient mice received much attention since they are lean, hypermetabolic, and resistant to various types of obesity.[18] These findings support the view that VGF is a precursor to multiple bioactive peptides. Major bioactive peptides derived from VGF are AQEE-30,[19] TLQP-62,[19] TLQP-21,[20] NERP-1 and NERP-2.[10] It is hence possible that additional bioactive peptides are embedded in this precursor sequence.

Consequently, we were able to identify 240 redundant peptides (155 distinct sequences) that arise from the human VGF sequence. These peptide sequences were aligned on the precursor sequence to construct a "peptide alignment map" (Figure 1 and Supplemental Figure 1, Supporting Information). Of note, this map is conceptually different from the peptide coverage map commonly used in proteomics studies. The peptide alignment map delineated 15 distinct clusters of the redundant peptides across the entire sequence. Consistent with the reported signal peptide,[12] the N-terminal region of 22 amino acids was not occupied by any peptide sequence (Figure 1). Each cluster consisted of peptides having N- or C-terminal trimming. The peptide ladders observed are not unique to VGF and common among different precursors of the regulatory secretion pathway, such as neuropeptide precursors and chromogranins identified in secretion media of endocrine cells, presumably caused by the action of exopeptidases.[14,21]

In 12 of 15 clusters, the longest peptide in each cluster was flanked by basic residue(s) at both ends (Figure 1 and Supplemental Figure 1, Supporting Information). Except for 232QARI MPD237, the boundaries that separate each cluster fit the consensus motifs for PC1/3 and PC2 ((R/K)Xn(R/K) where $n =$ 0, 2, 4, or 6).[2] While not experimentally verified, the stretch of six basic residues 479KRKRKK484 may be a bona fide target of PC1/3, since PC1/3 actually cleaves the corresponding rat residues 483KRKRKK488.[22] Judging from the number of human peptides identified, APPG-40 (aa 23–62), QQET-30 (aa 177–206), MPDS-45 (aa 235–279), NERP-1 (aa 282–306, C-terminal amidation), NERP-2 (aa 310–347, C-terminal amidation), NAPP-19 (aa 485–503), and AQEE-30 (aa 586–615) were considered intact peptides released from TT cells. As noted earlier, NERP-1, NERP-2,[10] and AQEE-30[19] have been reported as bioactive. Although not illustrated by redundant identification (Supplemental Figure 1, Supporting Information), GGEE-45 (aa 373–417) also appeared to be an intact peptide.

The limited number of human VGF-derived peptides reported in cerebrospinal fluid peptitomics studies[23–25] does not suffice to elucidate an overall processing pattern of the precursor (Supplemental Figure 2, Supporting Information). In the present study, however, we were able to identify most, if not all, of potential processing sites of the precursor (Figures 1 and 2). While some of the clusters are represented by relatively few peptides (Figure 1), this might be caused by poor presentation of acidic peptides in positive ion mode, as exemplified by GGEE-45 with a calculated pI of 3.8. Otherwise, some regions may not have been appreciably converted to peptides in TT cells. For instance, the cleavage site 553RPR|TLQ558 (corresponding human sequence 551RPR|TLQ556) for the recently reported rat TLQP-21 as well as TLQP-62, an abundant C-terminal peptide in rat brain,[20] did not emerge in the TT secretopeptidome analysis. Since TT cells are of thyroid C-cell origin, the lack of identification might reflect the difference in cell type-specific precursor processing that involves multiple processing enzymes.[26] In two human pancreatic neuroendocrine cell lines we examined, we got evidence that 551RPR|TLQ556 represented a major processing site (data not shown). To the

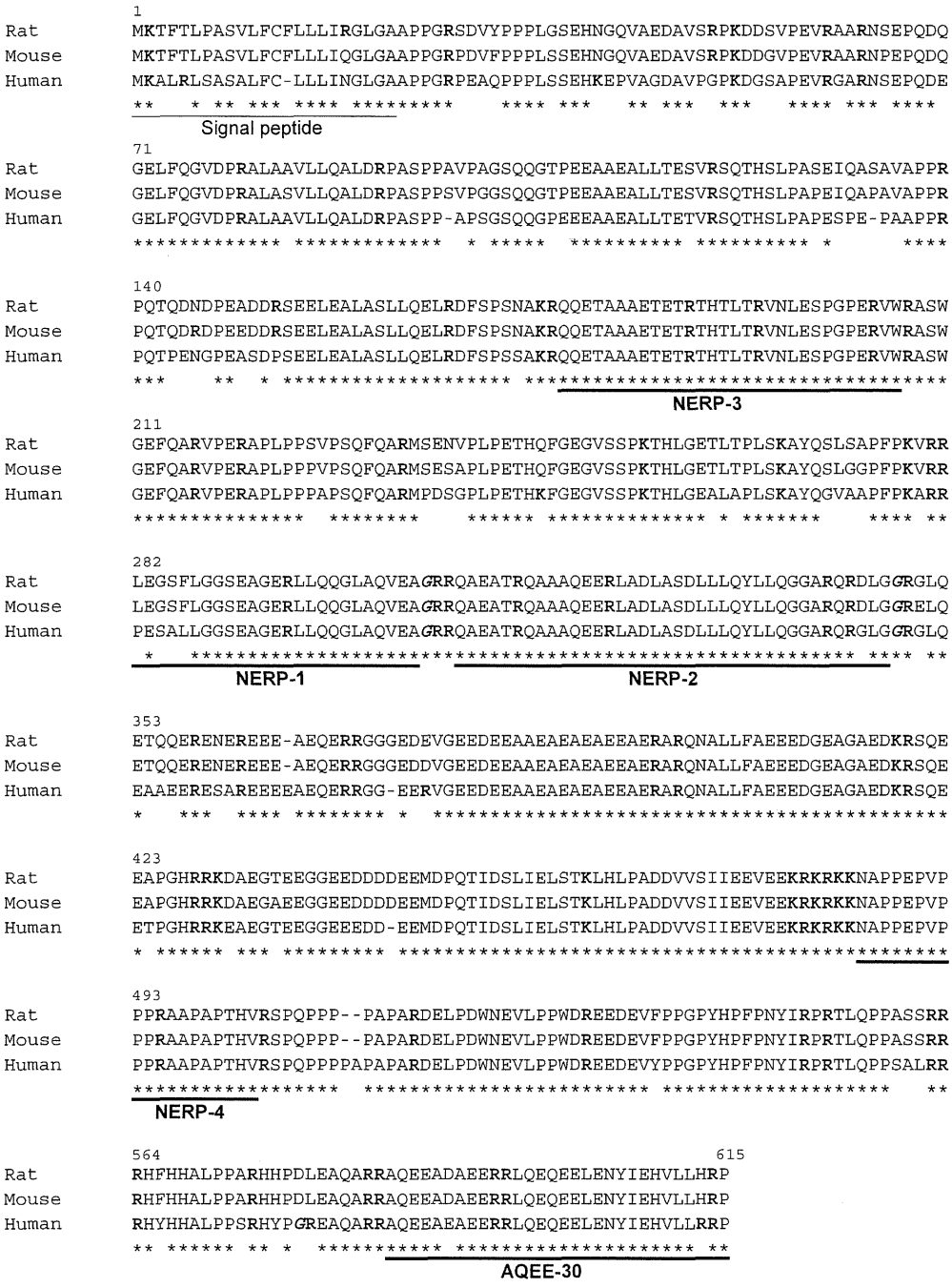


Figure 2. Alignment of rat (IPI00324864), mouse (IPI00378764), and human (IPI00069058) VGF sequences. Asterisks denote identical residues. Basic residues and amidation motifs are marked in bold and bold italic, respectively. The calcium assay was performed with bold underlined peptides. The residue numbering is for the human sequence.

best of our knowledge, this is the first study to demonstrate that accumulated peptide information contributes to elucidation of the overall processing of a precursor protein.

Bioactivity of a peptide cannot be inferred just from its sequence. It would be impossible to synthesize all of the peptides of various lengths revealed by LC−MS/MS. We therefore need to narrow down the number of candidate peptides for assessing bioactivity. In the present study, we took into consideration that peptides with definite biological activity are structurally conserved between species. Given the overall processing pattern (Figure 1), as well as interspecies homology between human and rodent VGF sequences (Figure 2), we selected QQET-30, hereafter designated NERP-3, and NAPP-19, designated NERP-4 (Figure 2). At the amino acid level, they are completely identical between human and rodents. APPG-40 and HYHH-19 (aa 565−583) were excluded since the location of internal basic residues is not conserved between rodent and human sequences (Figure 2). MPDS-45 was disregarded since rodent corresponding peptides were not so identical. HYHH-13 (aa 565−577, C-terminal amidation)[14] was also excluded because rodent VGF sequences do not harbor corresponding amidation sites. The stretch of acidic residues (seven aspartic and/or glutamic acids in human and eight in rodents) led us to disregard human EAEG-48 (aa 431−478. calculated p*I* 3.5). Also, the presence of multiple proline residues in SPQP-49 (aa

Table 1. Calcium Mobilization Induced by Test Peptides in Tissues from Apoaequorin Transgenic Mice[a]

	pituitary	hypothalamus	liver	adrenal gland	uterus	heart
NERP-3	+ (2/3)	+ (2/3)	− (0/3)	− (0/3)	− (0/3)	− (0/3)
NERP-4	+ (3/4)	+ (2/3)	− (0/3)	− (1/4)	− (0/3)	− (0/3)
AQEE-30	+ (2/4)	+ (2/3)	− (0/3)	− (1/4)	− (0/3)	− (0/3)
NERP-1	+ (3/3)	+ (2/3)	− (0/3)	ND	ND	ND
NERP-1-Gly	− (0/3)	− (0/3)	− (0/3)	ND	ND	ND
NERP-2	+ (3/3)	+ (2/3)	− (0/3)	ND	ND	ND
NERP-2-Gly	− (0/3)	− (0/3)	− (0/3)	ND	ND	ND

[a] Peptides were tested at a final concentration of 1 μM throughout. Data were summarized from at least three separate experiments. The peptide was considered active on the indicated tissue if positive signals were observed in at least two experiments. In parentheses, (X/Y) indicates the occurrence of positive signals (X) in total experiments (Y). ND, not determined.

505−553) was a reason for not including this peptide in subsequent assays.

Screening methods for bioactive peptides should be as comprehensive as possible. Given the presence of numerous cell types in an organism, a cell-based assay using a limited number of permanent cell lines or primary cultured cells is not practical. Since calcium is the most common second messenger, we used transgenic mice systemically expressing apoaequorin[11] to develop a tissue-based calcium assay. A candidate peptide is tested for its ability to raise intracellular calcium levels on selected tissues from the mice in a standardized analytical platform. Namely, the evoked calcium transient serves as a marker to report that the peptide is active on a tested tissue. The assay system was validated by bradykinin (pituitary) and ATP (hypothalamus, liver, and heart) and angiotensin II (adrenal gland) in advance. After an induced calcium transient returned to basal levels, expression of functional aequorin in the tested tissues was confirmed by lysing cells with Triton X-100. This treatment liberated intracellular apoaequorin to allow binding to extracellular calcium in incubation medium. At first, the utility of this assay system was assessed with the recently reported bioactive peptides NERP-1 and NERP-2[10] on pieces excised from the pituitary, hypothalamus, and liver. Consistent with their central actions reported in recent studies,[10] they elicited positive responses in the pituitary and hypothalamus (Figure 3 and Table 1). The inability of glycine-extended NERP-1 and NERP-2 to evoke the calcium response was also consistent with the previous report that NERPs' function is dependent on C-terminal amidation.[10] In this assay, the liver was considered a negative control since in most cases it does not respond to known bioactive peptides (data not shown).

NERP-3 and NERP-4, in addition to AQEE-30, were tested on the tissues including pituitary, hypothalamus, liver, adrenal gland, uterus, and heart. Aside from the liver, they are established targets for known neuropeptides and peptide hormones, including previously reported VGF-derived neuropeptides.[10,20] Positive signals were observed in the pituitary and hypothalamus in response to NERP-3 and NERP-4 (Figure 3 and Table 1). The observed calcium transient in the hypothalamus triggered by AQEE-30 may support a previous view that the peptide is a neuropeptide whose activity was first reported in isolated hypothalamic tissue cultures.[19] Overall, we identified two novel VGF-derived peptides that are functional toward the pituitary and hypothalamus.

As for human/mouse/rat NERP-3 and mouse/rat AQEE-30, we were able to prepare antibodies that specifically recognize

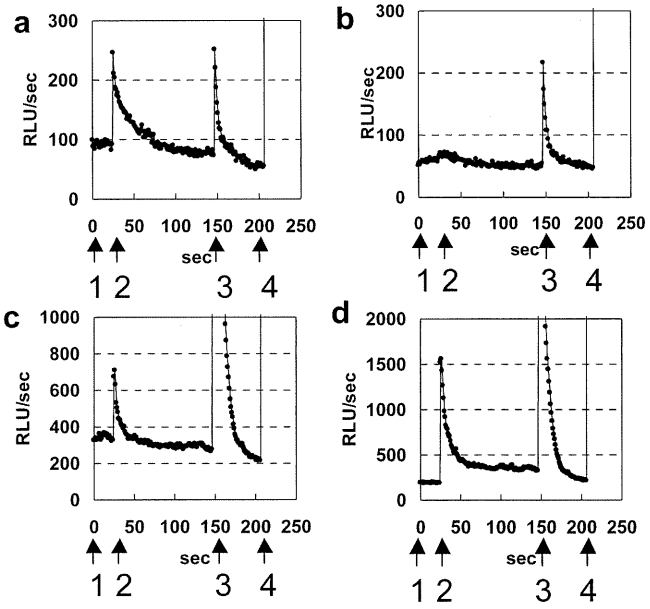


Figure 3. Typical traces of the relative luminescence units (RLU) evoked in hypothalamic tissues from apoaequorin transgenic mice. (a) NERP-1, (b) NERP-1-Gly, (c) NERP-3, and (d) AQEE-30. Peptides were tested at a final concentration of 1 μM. Medium (1), peptide (2), ATP (3), and Triton X-100 (4) were sequentially administered where indicated by arrows.

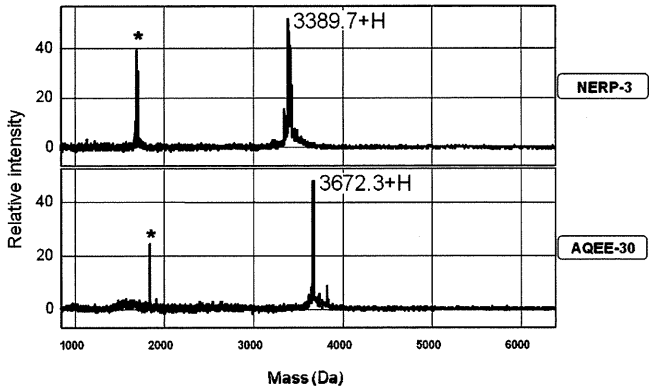


Figure 4. Mass characterization of peptides recognized by NERP-3 and AQEE-30 antibodies. Brain extract was immunoprecipitated with each antibody and analyzed on a surface-enhanced laser desorption ionization mass spectrometer. Asterisks indicate doubly charged ions.

the C-terminus of respective peptide as described in Materials and Methods. Peptides immunoprecipitated from whole rat brain extract were analyzed on a surface-enhanced laser desorption ionization mass spectrometer. Major peaks of the immunoprecipitate coincided with the calculated mass of 3389 Da (NERP-3) and 3672 Da (mouse/rat AQEE-30) (Figure 4), indicating that NERP-3 and AQEE-30 are actual major processing products also in rat brain. Although antibody was not prepared, NERP-4 may be a major peptide also in rodents as deduced by the multiple identifications in the human peptidomic data shown in Figure 1. Overall, this assay resulted in the identification of two VGF-derived peptides, designated NERP-3 and NERP-4, which are active on the pituitary and hypothalamus.

Conclusions

We described an approach to discover bioactive peptides and their target tissues. This approach can be extended to different

cell culture models. Identification of target organs or tissues will help to characterize responding cells and receptors.

Acknowledgment. We thank Emi Mishiro-Sato, Masako Matsubara (National Cerebral and Cardiovascular Center) and Miho Inoue, Machi Kusunoki, Teruyoshi Imada and Toyoko Kashiwagi (Kyowa Hakko Kirin) for expert technical assistance. This work was supported in part by the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation, and by Grants-in-Aid for Scientific Research from the Japanese Society for the Promotion of Science, Japan. We also thank Junko Kimata, Makoto Takahata and Morihiko Yoshida (Thermo Fisher Scientific, Japan) for LTQ-Orbitrap operation.

Supporting Information Available: Supplemental Table, peptides identified by LC–MS/MS; Supplemental Figure 1, detailed presentation of Figure 1; Supplemental Figure 2, peptides reported in previous cerebrospinal peptidomics studies; Supplemental Figure 3, MS/MS spectra of peptides with modified residues listed in Supplemental Table. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Zhou, A.; Webb, G.; Zhu, X.; Steiner, D. F. Proteolytic processing in the secretory pathway. *J. Biol. Chem.* **1999**, *274*, 20745–8.
(2) Fricker, L. D. Neuropeptide-processing enzymes: applications for drug discovery. *AAPS J.* **2005**, *7*, E449–455.
(3) Eipper, B. A.; Stoffers, D. A.; Mains, R. E. The biosynthesis of neuropeptides: peptide alpha-amidation. *Annu. Rev. Neurosci.* **1992**, *15*, 57–85.
(4) Clynen, E.; Baggerman, G.; Veelaert, D.; Cerstiaens, A.; Van der Horst, D.; Harthoom, L.; Derua, R.; Waelkens, E.; De Loof, A.; Schoofs, L. Peptidomics of the pars intercerebralis-corpus cardiacum complex of the migratory locust, Locusta migratoria. *Eur. J. Biochem.* **2001**, *268*, 1929–39.
(5) Schrader, M.; Schultz-Knappe, P. Peptidomics technologies for human body fluids. *Trends Biotechnol.* **2001**, *10*, S55–60.
(6) Svensson, M.; Sköld, K.; Svenningsson, P.; Andren, P. E. Peptidomics-based discovery of novel neuropeptides. *J. Proteome Res.* **2003**, *2*, 213–9.
(7) Che, F. Y.; Lim, J.; Pan, H.; Biswas, R.; Fricker, L. D. Quantitative neuropeptidomics of microwave-irradiated mouse brain and pituitary. *Mol. Cell. Proteomics* **2005**, *4*, 1391–1405.
(8) Dowell, J. A.; Heyden, W. V.; Li, L. Rat neuropeptidomics by LC-MS/MS and MALDI-FTMS: enhanced dissection and extraction techniques coupled with 2D RP-RP HPLC. *J. Proteome Res.* **2006**, *5*, 3368–75.
(9) Bora, A.; Annangudi, S. P.; Millet, L. J.; Rubakhin, S. S.; Forbes, A. J.; Kelleher, N. L.; Gillette, M. U.; Sweedler, J. V. Neuropeptidomics of the rat supraoptic nucleus. *J. Proteome Res.* **2008**, *7*, 4992–5003.
(10) Yamaguchi, H.; Sasaki, K.; Satomi, Y.; Shimbara, T.; Kageyama, H.; Mondal, M. S.; Toshinai, K.; Date, Y.; Gonzalez, L. J.; Shioda, S.; Takao, T.; Nakazato, M.; Minamino, N. Peptidomic identification and biological characterization of neuroendocrine regulatory peptide-1 and -2. *J. Biol. Chem.* **2007**, *282*, 26354–60.
(11) Yamano, K.; Mori, K.; Nakano, R.; Kusunoki, M.; Inoue, M.; Satoh, M. Identification of the functional expression of adenosine A3 receptor in pancreas using transgenic mice expressing jellyfish apoaequorin. *Transgenic Res.* **2007**, *16*, 429–35.
(12) Levi, A.; Ferri, G. L.; Watson, E.; Possenti, R.; Salton, S. R. Processing, distribution, and function of VGF, a neuronal and endocrine peptide precursor. *Cell Mol. Neurobiol.* **2004**, *24*, 517–33.
(13) Gkonos, P. J.; Born, W.; Jones, B. N.; Petermann, J. B.; Keutmann, H. T.; Bimbaum, R. S.; Fischer, J. A.; Roos, B. A. Biosynthesis of calcitonin gene-related peptide and calcitonin by a human medullary thyroid carcinoma cell line. *J. Biol. Chem.* **1986**, *261*, 14386–91.
(14) Sasaki, K.; Satomi, Y.; Takao, T.; Minamino, N. Snapshot peptidomics of the regulated secretory pathway. *Mol. Cell. Proteomics* **2009**, *8*, 1638–47.
(15) Mishiro-Sato, E.; Sasaki, K.; Matsuo, Y.; Kageyama, H.; Yamaguchi, H.; Date, Y.; Matsubara, M.; Ishizu, T.; Yoshizawa-Kumagaye, K.; Satomi, Y.; Takao, T.; Shioda, S.; Nakazato, M.; Minamino, N. Distribution of neuroendocrine regulatory peptide-1 and–2, and proteolytic processing of their precursor VGF protein in the rat. *J. Neurochem.* **2010**, *114*, 1097–106.
(16) Sasaki, K.; Sato, K.; Akiyama, Y.; Yanagihara, K.; Oka, M.; Yamaguchi, K. Peptidomics-based approach reveals the secretion of the 29-residue COOH-terminal fragment of the putative tumor suppressor protein DMBT1 from pancreatic adenocarcinoma cell lines. *Cancer Res.* **2002**, *62*, 4894–8.
(17) Levi, A.; Eldridge, J. D.; Paterson, B. M. Molecular cloning of a gene sequence regulated by nerve growth factor. *Science* **1985**, *229*, 393–5.
(18) Hahm, S.; Mizuno, T. M.; Wu, T. J.; Wisor, J. P.; Priest, C. A.; Kozak, C. A.; Boozer, C. N.; Peng, B.; McEvoy, R. C.; Good, P.; Kelley, K. A.; Takahashi, J. S.; Pintar, J. E.; Roberts, J. L.; Mobbs, C. V.; Salton, S. R. Targeted deletion of the Vgf gene indicates that the encoded secretory peptide precursor plays a novel role in the regulation of energy balance. *Neuron* **1999**, *23*, 537–48.
(19) Alder, J.; Thakker-Varia, S.; Bangasser, D. A.; Kuroiwa, M.; Plummer, M. R.; Shors, T. J.; Black, I. B. Brain-derived neurotrophic factor-induced gene expression reveals novel actions of VGF in hippocampal synaptic plasticity. *J. Neurosci.* **2003**, *23*, 10800–8.
(20) Bartolomucci, A.; Possenti, R.; Levi, A.; Pavone, F.; Moles, A. The role of the vgf gene and VGF-derived peptides in nutrition and metabolism. *Genes Nutr.* **2007**, *2*, 169–80.
(21) Nikoulina, S. E.; Andon, N. L.; McCowen, K. M.; Hendricks, M. D.; Lowe, C.; Taylor, S. W. A primary colonic crypt model enriched in enteroendocrine cells facilitates a peptidomic survey of regulated hormone secretion. *Mol. Cell. Proteomics* **2010**, *9*, 728–41.
(22) Trani, E.; Giorgi, A.; Canu, N.; Amadoro, G.; Rinaldi, A. M.; Halban, P. A.; Ferri, G. L.; Possenti, R.; Schinina, M. E.; Levi, A. Isolation and characterization of VGF peptides in rat brain. Role of PC1/3 and PC2 in the maturation of VGF precursor. *J. Neurochem.* **2002**, *81*, 565–74.
(23) Selle, H.; Lamerz, J.; Buerger, K.; Dessauer, A.; Hager, K.; Hampel, H.; Karl, J.; Kellmann, M.; Lannfelt, L.; Louhija, J.; Riepe, M.; Rollinger, W.; Tumani, H.; Schrader, M.; Zucht, H.-D. Identification of novel biomarker candidates by differential peptidomics analysis of cerebrospinal fluid in Alzheimer's disease. *Comb. Chem. High Throughput Screen.* **2005**, *8*, 801–6.
(24) Huang, J. T.; Leweke, F. M.; Oxley, D.; Wang, L.; Harris, N.; Koethe, D.; Gerth, C. W.; Nolden, B. M.; Gross, S.; Schreiber, D.; Reed, B.; Bahn, S. Disease biomarkers in cerebrospinal fluid of patients with first-onset psychosis. *PLoS Med.* **2006**, *3*, e428.
(25) Zougman, A.; Pilch, B.; Podtelejnikov, A.; Kiehntopf, M.; Schnabel, C.; Kumar, C.; Mann, M. Integrated analysis of the cerebrospinal fluid peptidome and proteome. *J. Proteome Res.* **2008**, *7*, 386–99.
(26) Rehfeld, J. F.; Bundgaard, J. R.; Hannibal, J.; Zhu, X.; Norrbom, C.; Steiner, D. F.; Friis-Hansen, L. The cell-specific pattern of cholecystokinin peptides in endocrine cells versus neurons is governed by the expression of prohormone convertases 1/3, 2, and 5/6. *Endocrinology* **2008**, *149*, 1600–8.

PR1003455

Journal of proteome research

ARTICLE

pubs.acs.org/jpr

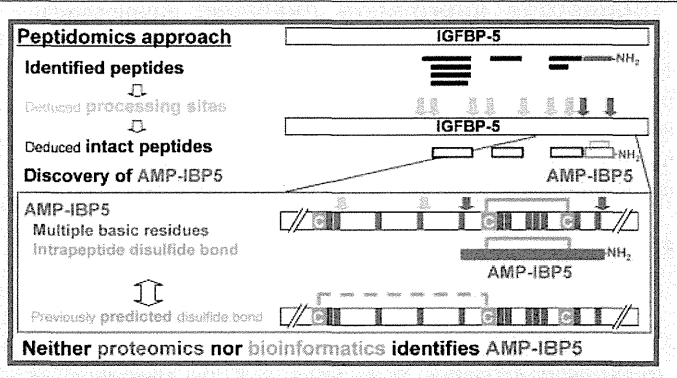
Peptidomics-Based Discovery of an Antimicrobial Peptide Derived from Insulin-Like Growth Factor-Binding Protein 5

Tsukasa Osaki, Kazuki Sasaki,* and Naoto Minamino*

Department of Molecular Pharmacology, National Cerebral and Cardiovascular Center Research Institute, Suita, Osaka 565-8565, Japan

Supporting Information

ABSTRACT: Antimicrobial peptides (AMPs) are effector molecules that are able to kill or inactivate microbial pathogens. However, most AMPs harbor multiple basic amino acids that hamper current proteomic identification. In our peptidomic survey of endogenous peptides, we identified a novel intramolecular disulfide-linked 22-residue amidated peptide. This peptide, designated AMP-IBP5 (antimicrobial peptide derived from insulin-like growth factor-binding protein 5), showed antimicrobial activity against six of the eight microorganisms tested at concentrations comparable to or lower than those for well-characterized AMPs cathelicidin and β-defensin-2. AMP-IBP5 is identical at the amino acid level between human, mouse, rat, pig, and cow. Natural occurrence of this peptide as the originally isolated form was demonstrated in the rat brain and intestine, using mass spectrometric characterization of major immunoreactivity. The peptide is flanked N-terminally by a single arginine and C-terminally by a common amidation signal, indicating that insulin-like growth factor-binding protein 5 (IGFBP-5) undergoes specific cleavage by a defined set of processing proteases. Furthermore, the intramolecular linkage C199-C210 reveals itself as a correct disulfide pairing in the precursor protein, the finding not inferred from closely related family members IGFBP-4 and -6. In principle, neither conventional proteomics nor bioinformatics would achieve the identification of this AMP. Our study exemplifies the impact of peptidomics to study naturally occurring peptides.



KEYWORDS: antimicrobial peptide, bioactive peptide, insulin-like growth factor-binding protein 5, mass spectrometry, peptidomics, proteolytic processing, secretopeptidome

INTRODUCTION

Bioactive peptides, such as peptide hormones and antimicrobial peptides (AMPs), have been identified through activity-guided biochemical purification that starts with a bulk of biological samples. Recently, technological advances in mass spectrometry (MS) enable us to identify naturally occurring peptides present in mixtures.[1−3] Nevertheless, we are faced with a daunting task of identifying bioactive peptides of a secretory nature owing to their relative low abundance. Once extracted from biological samples, secreted peptides are not discriminated from nonsecreted peptides or peptide fragments caused by degradation of cytosolic proteins. Since relatively abundant molecules are preferentially detected in MS schemes, we need to work on samples rich in secreted peptides for facilitating the discovery of bioactive peptides. We recently used tandem mass spectrometry (MS/MS) techniques to characterize a total pool of naturally occurring peptides that are released by exocytosis from cells in culture. This study, referred to as secretopeptidome analysis, allows us to identify peptides localized in secretory granules in a noninvasive as well as efficient manner.[4,5]

In the present study, we focused on secreted peptides with a highly basic nature. Some bioactive peptides, especially AMPs, harbor multiple basic amino acids. In fact, well-characterized mammalian AMPs β-defensin-2[6] and cathelicidin[7] both bear a net charge of +6 at pH 7.0. These peptides should be analyzed in their native forms, and therefore cannot be studied by current proteomics that needs an enzymatic digestion step for MS-based identification. We characterized highly basic fractions of the secretopeptidome from cultured human pancreatic neuroendocrine tumor cells and identified a previously unknown peptide that arises from insulin-like growth factor-binding protein 5 (IGFBP-5). With regard to antimicrobial activity and spectrum, this peptide is almost as potent as cathelicidin and even superior to β-defensin-2. The peptide was thus designated AMP-IBP5 (antimicrobial peptide derived from IGFBP-5). We provide evidence that AMP-IBP5 is generated through site-specific cleavages in the brain and small intestine. In addition, the identification of this intramolecular disulfide-linked peptide led us to conclude that IGFBP-5 protein possesses a disulfide pairing different from that previously predicted on the basis of analogy to IGFBP-4 and -6.[8] While bioinformatics is currently used for in silico prediction of bioactive peptides, this peptide could not be identified in a

Received: November 5, 2010
Published: January 06, 2011

situation where we had relied on the previous incorrect disulfide pairing. Our data demonstrated that peptidomics is a promising tool to uncover previously uncharacterized peptides.

EXPERIMENTAL PROCEDURES

Peptide Preparation

Monolayer cultures of the human pancreatic neuroendocrine tumor cell line QGP-1[9] were rinsed three times with Hanks' medium (Invitrogen, Carlsbad, CA). Culture supernatant of the cells (ca. 3.2×10^7 cells) stimulated with 10 μM forskolin and 10 μM carbachol for 15 min was harvested. Peptides were extracted and gel-filtrated as previously described.[4] Cysteine residues of the gel-filtrated samples were converted to carbamidomethyl cysteines (CmC) using dithiothreitol and iodoacetamide, followed by desalting and lyophilization. Resultant products were dissolved in solvent A (10 mM ammonium formate, pH 3.8: acetonitrile (ACN) = 9:1 (v/v)) and applied to a TSK gel SP-2SW cation-exchange column (1.0 × 50 mm; TOSOH, Tokyo, Japan) equilibrated with solvent A, and eluted at a flow rate of 50 μL/min with a gradient of 0−100% solvent B (1 M ammonium formate, pH 3.8: ACN = 9:1 (v/v)) in 30 min and then maintained at 100% B for 20 min. Highly basic fractions eluted after 30 min and the preceding fractions eluted between 0 and 30 min were separately desalted and applied to a C18 PepMap column (0.075 × 150 mm; Dionex, Sunnyvale, CA) using an Ultimate liquid chromatography (LC) system (Dionex) equilibrated with solvent A (5% ACN, 0.1% trifluoroacetic acid (TFA)) at a flow rate of 300 nL/min. Peptides were eluted with four steps, 5% B (95% ACN, 0.1% TFA) in 5 min, a linear gradient of 5−60% B in 55 min, a linear gradient of 60−100% B in 5 min, and then 100% B in 20 min. Fractions were collected every 20 s from 10 min after sample injection and spotted onto a 384-well matrix-assisted laser desorption/ionization (MALDI) plate with an infusion of 1.75 mg/mL α-cyano-4-hydroxy cinnamic acid (Nacalai tesque, Kyoto, Japan) in 50% ACN, 0.1% TFA at a flow rate of 1.5 μL/min using a Probot microfraction collector (Dionex).

MS Analysis

Samples spotted on a MALDI plate were analyzed on a MALDI-TOF/TOF mass spectrometer (4800 Proteomics Analyzer, Applied Biosystems, Foster City, CA). Each spot was first analyzed in MS positive ion reflector mode in the mass range from 1000 to 5000 Da by accumulating signals of 1000 laser shots. The 15 most abundant parent ions with a signal-to-noise ratio >20 were selected for top-down MS/MS scans, excluding identical parent ions contained in adjacent spots from a given LC-MALDI run. MS/MS was conducted using medium collision energy in positive ion mode.

Data Analysis and Peptide Identification

Peak lists were generated by the "Launch Peaks to Mascot" function of 4000 Series Explorer software (ver. 3.5, Applied Biosystems) using the default parameters supplied by the manufacturer. Peak lists were searched against IPI Human (80128 entries on June 18, 2009) using Mascot (ver. 2.2), with no enzyme specification. Carbamidomethylation of cysteine was set as a fixed modification, and pyroglutamination and C-terminal amidation were simultaneously allowed as variable modifications. Peptide tolerance was set to 125 ppm and MS/MS tolerance was 0.4 Da. The significance threshold was the Mascot default setting of 5%. In Tables 1 and 2, peptides with a score above the identity threshold (corresponding to an expectation value below 0.05) were listed and considered identified.

Peptide Synthesis

Peptides derived from IGFBP-5 were synthesized on an Abacus peptide synthesizer (Sigma Genosys, Sigma Aldrich Japan, Hokkaido, Japan) using Fmoc (*N*-(9-fluorenyl) methoxycarbonyl) strategy, purified by reverse phase high performance liquid chromatography (HPLC), and verified for correct synthesis by MS and amino acid analysis. Purity of the peptides was confirmed on separate HPLC systems.

Heparin-Binding Assay

Synthetic peptides (1 nmol) were dissolved in 450 μL of phosphate-buffered saline (PBS, 10 mM phosphate buffer, pH 7.0 containing 0.15 M NaCl) and incubated with 50 μL of heparin-Sepharose CL-6B (GE Healthcare, Piscataway, NJ) at room temperature for 1 h. Supernatant was obtained by centrifugation at 800 × g for 3 min. Beads were washed three times with PBS (1 mL each). All the supernatants were combined as unbound fraction. After washing, bound peptides were eluted off the beads in 1 mL of 20 mM Tris-HCl, pH 8.0, containing 1.5 M NaCl (bound fractions). One-tenth volume of the bound and unbound fractions was each subjected to a reverse phase HPLC system. The linear gradient consisted of 10−60% ACN in 0.1% TFA for 40 min at a flow rate of 0.05 mL/min on a C18 reverse phase HPLC column (Vydac 218TP5115, 1.0 × 150 mm; Hesperia, CA). For a positive control of heparin-binding activity, cathelicidin (LL37; Bachem, Bubendorf, Switzerland) was used.[10]

Measurement of Antibacterial and Antifungal Activity

AlamarBlue (BioSource International, Camarillo, CA) was used to determine antimicrobial activity. As a consequence of bacterial growth, this redox indicator turns from blue to red in color. Antimicrobial activities of test peptides (up to 10 μM) or IGFBP-5 protein (1 μM; Ray Biotech, Norcross, GA) were assessed for the target microbes *Enterococcus hirae* (*E. hirae*), *Micrococcus luteus* (*M. luteus*), *Staphylococcus aureus* (*S. aureus*) 209P, *S. saprophyticus* KD, *Escherichia coli* (*E. coli*) B, *E. coli* K12, *E. coli* kp and *Pichia pastoris* (*P. pastoris*) GS115. The optimal growth temperature of *M. luteus*, *S. saprophyticus* KD and *P. pastoris* GS115 was 30 °C and that of the other microbes was 37 °C. After grown in 3% tryptosoy broth (Eiken Chemical, Tokyo, Japan) for 16 h with shaking at each optimal temperature, cells were washed twice with 10 mM phosphate buffer, pH 7.0, and diluted to 8×10^5 colony-forming units/ml in the same buffer. Twenty-five microliters of bacteria suspension was mixed with an equal volume of sample in the absence or presence of peptides, and incubated for 1 h. After incubation, 200 μL of 3% tryptosoy broth containing 10% alamarBlue was added to the reaction mixture and further incubated as follows: 4 h for *E. hirae*, *S. aureus* 209P and *E. coli* B; 6 h for *E. coli* K12 and *E. coli* kp; 7 h for *M. luteus*; 7.5 h for *S. saprophyticus* KD; and 20 h for *P. pastoris* GS115. Aliquots containing all assay reagents but microbes were used as blank. After incubation, the reactions were monitored by absorbance at 570 and 600 nm. Molar extinction coefficients of OD_{570} and OD_{600} in the oxidative condition are 80586 and 117216. Therefore, viability (%) was expressed using the following formula: viability (%) = $(117216 \times OD_{570} - 80586 \times OD_{600}$ in the presence of peptides$/117216 \times OD_{570} - 80586 \times OD_{600}$ in the absence of peptides$) \times 100$. The classical colony formation assay was also performed as described[11,12] using *S. aureus* 209P, *E. coli* K12 and *P. pastoris* GS115. For positive controls of

Table 1. Peptides Identified in Highly Basic Fractions[a]

precursor	IPI accession	m/z (obsd)	M_r (calc) (Da)	mass error (Da)	MASCOT score	expect.	N-term	peptide	C-term	net charge (pH 7.0)
CgA	00383975	1779.87	1778.93	−0.07	63	0.026	L	SFRARAYGFRGPGPQL	R	+3
CgB	00006601	2460.91	2460.12	−0.21	125	6.90×10^{-9}	M	AHGYGEESEEERGLEPGKGRHH	R	−3
CgB	00006601	1995.79	1994.94	−0.16	83	0.00018	R	FLGEGHHRVQENQMDKA	R	−1
CgB	00006601	1952.87	1952.03	−0.17	65	0.0044	R	GLEPGKGRHHRGRGGEPR	A	+3
CgB	00006601	2106.84	2105.98	−0.15	59	0.011	R	SETHAAGHSQEKTHSREKS	S	−1
CGRP	00027855	2390.12	2389.32	−0.21	98	9.70×10^{-6}	G	LLSRSGGVVKNNFVPTNVGSKAF-NH2	G	+4
CT	00000914	2435.85	2435.07	−0.23	69	0.0013	R	DMSSDLERDHRPHVSMPQNAN	C-term	−2
DSG2	00028931	1822.88	1822.03	−0.16	99	1.70×10^{-6}	R	NENKLLPKHPHLVRQ	K	+2
IGFBP5	00029236	2770.21	2769.44	−0.24	61	0.046	R	AVYLPNCDRKGFYKRKQCKPSR-NH2	G	+7
NUCB1	00893068	1580.70	1579.85	−0.16	72	0.0031	R	ELDFVSHHVRTKL	D	0
PC2	00643663	3111.38	3110.55	−0.18	73	0.0031	R	<QVAAEHGFGVRKLPFAEGLYHFYHNGLA	K	−1
PC2	00643663	2423.07	2422.24	−0.18	116	3.20×10^{-8}	A	ERPVFTNHFLVELHKGGEDKA	R	−1
PC2	00643663	1976.90	1975.99	−0.10	109	7.20×10^{-7}	R	KLPFAEGLYHFYHNGLA	K	0
PC2	00643663	1390.57	1389.71	−0.15	71	0.0037	R	SLHHKQQLERD	P	−1
PC2	00643663	1643.71	1642.87	−0.16	87	9.30×10^{-5}	R	SLHHKQQLERDPR	V	+1
SST	00000130	1157.44	1156.55	−0.11	70	0.0024	K	AGCKNFFWK	T	+2
SST	00000130	1258.48	1257.60	−0.12	67	0.0054	K	AGCKNFFWKT	F	+2
SST	00000130	1593.64	1592.74	−0.11	59	0.043	K	AGCKNFFWKTFTS	C	+2
SST	00000130	1753.62	1752.78	−0.17	77	7.10×10^{-5}	K	AGCKNFFWKT FTSC	C-term	+2
SST	00000130	1625.61	1624.72	−0.11	86	6.40×10^{-5}	G	CKNFFWKTFTSC	C-term	+2
SST	00000130	1334.54	1333.63	0.10	74	0.0016	A	GCKNFFWKTF	T	+2
SST	00000130	1682.63	1681.74	−0.12	91	1.90×10^{-5}	A	GCKNFFWKTFTSC	C-term	+2
VGF	00289501	3706.59	3705.83	−0.25	69	0.0071	R	AQEEAEAEERRLQEQEELENYIEHVLLRRP	C-term	−6
VGF	00289501	1786.80	1785.90	−0.11	70	0.0062	H	HALPPSRHYPGREAQA	R	+1
VGF	00289501	1724.74	1723.87	−0.13	70	0.0052	Y	HHALPPSRHYPGREA	Q	+1
VGF	00289501	1923.82	1922.96	−0.15	88	1.80×10^{-5}	Y	HHALPPSRHYPGREAQA	R	+1
VGF	00289501	1351.55	1350.67	−0.13	68	0.0065	R	HYHHALPPSRH	Y	+1
VGF	00289501	1953.80	1952.95	−0.16	53	0.05	R	HYHHALPPSRHYPGRE	A	+1
VGF	00289501	2024.81	2023.99	−0.18	89	1.10×10^{-5}	R	HYHHALPPSRHYPGREA	Q	+1
VGF	00289501	2152.91	2152.05	−0.15	88	1.60×10^{-5}	R	HYHHALPPSRHYPGREAQ	A	+1
VGF	00289501	2223.91	2223.08	−0.18	94	4.20×10^{-6}	R	HYHHALPPSRHYPGREAQA	R	+1
VGF	00289501	1610.68	1609.80	−0.13	69	0.0052	R	HYHHALPPSRHYP-NH2	G	+2
VGF	00289501	2380.01	2379.18	−0.18	54	0.043	R	RHYH HALPPSRHYPGREAQA	R	+2
VGF	00289501	1766.75	1765.90	−0.16	79	0.00065	R	RHYHHALPPSRHYP-NH2	G	+3
VGF	00289501	2564.15	2563.35	−0.21	121	9.50×10^{-9}	R	RLQEQEELENYIEHVLLRRP	C-term	−2
VGF	00289501	1594.76	1593.89	−0.14	66	0.014	R	TLQPPSALRRRHY	H	+3
VGF	00289501	2086.87	2086.02	−0.16	68	0.008	H	YHHALPPSRHYPGREAQA	R	0

[a] Peptides whose expectation values (Expect., column 7) were less than 0.05 are listed. M_r (calc) represents the theoretical monoisotopic molecular mass (Da) of the peptide sequence. MASCOT scores are indicated in column 6. The N-terminal (N-term) and C-terminal (C-term) flanking one amino acid (columns 8 and 10) are shown. Net charge (pH 7.0) was calculated as follows: D and E are −1, K and R are +1, and H is 0. N-terminal amino group is +1, and C-terminal carboxyl group is −1. <Q, pyroglutamic acid; -NH2, C-terminal amidation. CgA, chromogranin A; CgB, chromogranin B; CGRP, calcitonin gene-related peptide; CT, calcitonin; DSG2, desmoglein 2; IGFBP5, insulin-like growth factor-binding protein 5; NUCB1, nucleobindin 1; PC2, prohormone convertase 2; SST, somatostatin.

1872
dx.doi.org/10.1021/pr101114a | J. Proteome Res. 2011, 10, 1870–1880

Table 2. IGFBP-5-Derived Peptides Identified by Mass Spectrometry[a]

m/z (obsd)	M_r (calc) (Da)	mass error (Da)	MASCOT score	identity threshold	expect.	N-term	peptide	C-term
3332.62	3331.54	0.07	134	54	5.50×10^{-10}	VKIER	DSREHEEPTTSEMAEETYSPKIFRPKH	RISEL
2974.46	2973.38	0.07	165	54	4.10×10^{-13}	ERDSR	EHEEPTTSEMAEETYSPKIFRPKHT	RISEL
2956.13	2955.37	−0.25	154	57	1.10×10^{-11}	ERDSR	<EHEEPTTSEMAEETYSPKIFRPKHT	RISEL
2873.41	2872.33	0.07	80	54	0.00013	ERDSR	EHEEPTTSEMAEETYSPKIFRPKH	TRISE
1812.88	1811.93	−0.05	143	54	7.80×10^{-11}	DRRKK	LTQSKFVGGAENTAHPR	IISAP
2080.34	2079.07	0.26	71	50	0.0005	GPCRR	HMEASLQELKASPRMVPR	AVYLP
1596.80	1595.81	−0.01	72	54	0.00097	GPCRR	HMEASLQELKASPR	MVPRA
2770.21	2769.44	−0.24	61	60	0.046	RMVPR	AVYLPNCDRKGFYKRKQCKPSR-NH2	GRKRG

[a] Peptides whose MASCOT scores (column 4) exceeded identity thresholds (column 5) are listed. Mr (calc) represents the theoretical monoisotopic molecular mass (Da) of the peptide sequence. Expectation values (Expect.) are indicated in column 6. The N-terminal (N-term) and C-terminal (C-term) flanking five amino acid sequences (columns 7 and 9) are shown. <E, pyroglutamic acid; -NH2, C-terminal amidation.

antimicrobial activity, cathelicidin and β-defensin-2 (Peptide Institute, Osaka, Japan) were used.

Statistical Analysis

Statistical analysis was performed using Student's t test, with a level of significance set at $p < 0.05$.

Preparation of Cell Culture Supernatant

The culture supernatant from QGP-1 was prepared as described above. Monolayer culture of the human lung neuroendocrine tumor cell line SHP-77[13] (Ca. 1.1×10^7 cells) was stimulated with 50 mM potassium chloride and 10 μM carbachol for 10 min and its culture supernatant was harvested. Peptides were extracted as previously described[4] and used for radioimmunoassay (RIA).

Tissue Collection from Sprague-Dawley Rats

All experimental procedures of tissue collection from rats and immunization of rabbits were approved by our institutional animal experiments and care committee. Brain, pituitary gland, lung, heart, stomach, small intestine, liver, pancreas, kidney and uterus were collected from three female Sprague-Dawley rats (11-week-old, 220−280 g) immediately after decapitation, and used for RIA.

Antibody Preparation

Synthetic AMP-IBP5 (AVYLPNCDRKGFYKRKQCKPSR-NH_2, intramolecularly disulfide-linked) was conjugated with bovine thyroglobulin (Sigma Aldrich, St. Louis, MO) by the action of water-soluble carbodiimide (Peptide Institute). Rabbits were immunized with each conjugate emulsified with an equal volume of Freund's complete adjuvant as reported.[14]

RIA

RIA was carried out as reported[15] using intramolecularly disulfide-linked ^{125}I-radiolabeled YAVYLPNCDRKGFYKRKQ-CKP SR-NH_2 and anti-AMP-IBP5 antibody (#569-5) at a dilution of 1:210 000. A fifty percent inhibitory concentration (IC_{50}) of ligand binding in the RIA was 20 fmol/tube. Specificity of the RIA was examined with C-terminally Gly-extended AMP-IBP5, carbamidomethylated (CAM)-AMP-IBP5, IGFBP-5 protein and seven known bioactive peptides, vasopressin, calcitonin, adrenomedullin, proadrenomedullin N-terminal 20-amino acid peptide (PAMP-20), neurokinin A, angiotensin II and leucine-enkephalin. The six bioactive peptides except PAMP-20 showed no cross-reactivity up to 100 000 fmol/tube, and IGFBP-5 protein had no cross-reactivity up to 10 000 fmol/tube. The IC_{50} values of C-terminally Gly-extended AMP-IBP5, CAM-AMP-IBP5, and PAMP-20 (arginine amide) were 20 000, 20, and 100 000 fmol/tube, respectively. These results indicate that the antiserum strictly recognizes the C-terminal region including amide structure but not the disulfide bond or the intact IGFBP-5 protein.

Immunological Detection of AMP-IBP5

Tissues were collected as described above, extracted and condensed with a Sep-Pak C18 cartridge as described previously.[16] An aliquot of cartridge eluate was examined by RIA to quantify immunoreactive (IR)-AMP-IBP5. Brain and small intestine extracts were loaded onto a gel filtration column (Sephadex G-50 fine, GE Healthcare; 1.8 × 135 cm) equilibrated with 1 M CH_3COOH at a flow rate of 7 mL/h, fractionated every 6 mL/tube, and assessed by RIA to evaluate an IR-AMP-IBP5 level in each fraction. Fractions containing most abundant IR-AMP-IBP5 were pooled, lyophilized, and separated on a reverse phase HPLC column (Symmetry300 C18 5 μ, 4.6 × 250 mm; Waters Co., Milford, MA) equilibrated with solvent A (10% ACN, 0.1% TFA) at a flow rate of 1 mL/min. Adsorbed samples were eluted with a linear gradient of 0−100% B (60% ACN, 0.1% TFA) in 60 min, fractionated every 1 mL/tube, and assessed by RIA.

For MS of immunoprecipitates, the fraction containing a highest level of IR-AMP-IBP5 in the reverse phase HPLC was lyophilized, dissolved in 40 μL of the antibody (#569-5) 10-fold diluted with PBS, and incubated overnight with 10 μL of Protein A-Sepharose CL-4B (GE Healthcare) at 4 °C. Immunocomplexes were washed three times with PBS and twice with distilled water (1 mL each), followed by elution in 20 μL of 1% TFA and desalting using u-C18 Zip Tips (Millipore, Billerica, MA). Recovered AMP-IBP5-IR materials were spotted on target plates with α-cyano-4-hydroxy cinnamic acid matrix and then analyzed in MS positive ion reflector mode in the mass range from 1000 to 5000 Da on a 4800 Proteomics Analyzer.

SDS-PAGE and Immunoblotting

Tissues were collected as described above and extracted with 10 volumes (w/v) of SDS-PAGE loading buffer, followed by separation on a 15% SDS-PAGE gel and transfer to a PVDF membrane (GE Healthcare) as reported.[17] Membranes were probed with the first antibody and then probed with horseradish peroxidase-conjugated goat antirabbit IgG (1:5000; Cell Signaling Technology, Beverly, MA). First antibodies were used as follows: a rabbit polyclonal antibody (H-100) raised against human IGFBP-5[81-180] (1:1000; Santa Cruz Biotechnology, Santa Cruz, CA) for detecting intact IGFBP-5; and a rabbit monoclonal antibody (14C10) raised against human/rat glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (1:5000; Cell

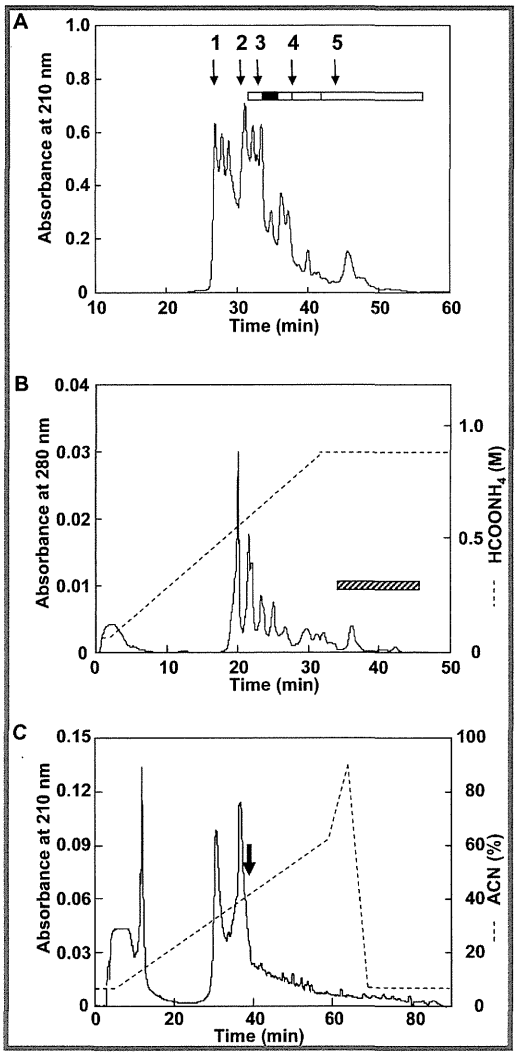


Figure 1. Preparation of highly basic peptides for peptidomic analysis. (A) Gel filtration profile of culture supernatant extracts from QGP-1 cells. Peptide-containing fractions were indicated by five boxes; the closed box indicated the fraction in which AMP-IBP5 originated. Molecular weight makers: 1, bovine serum albumin (66.5 kDa); 2, ribonuclease A (13.5 kDa); 3, neuropeptide Y (4.3 kDa); 4, neurotensin (1.7 kDa) and 5, Leu-enkephalin (0.6 kDa). (B) Cation-exchange HPLC of the AMP-IBP5-containing fraction obtained by the preceding gel filtration. Fractions eluted between 35 and 46 min (hatched box) were pooled for LC-MALDI-MS. (C) Subsequent nanoLC separation of the fraction denoted by the hatched box in (B). AMP-IBP5 was identified at the arrowed microfraction.

Signaling Technology). Membranes were visualized using an ECL-plus kit, according to the protocol provided by the manufacturer (GE Healthcare), and images were recorded with a LAS-1000 plus imager (Fujifilm, Tokyo, Japan) for 1−10 min. Digital images were quantitated using Image J software (National Institutes of Health, Bethesda, MD).

RESULTS

Peptidomic Identification of IGFBP-5-Derived Peptides

We recovered the culture supernatant from QGP-1 cells that received an exocytotic stimulus of 10 μM forskolin plus 10 μM carbachol for 15 min. Substances extracted from the supernatant were separated by gel filtration HPLC to obtain peptides

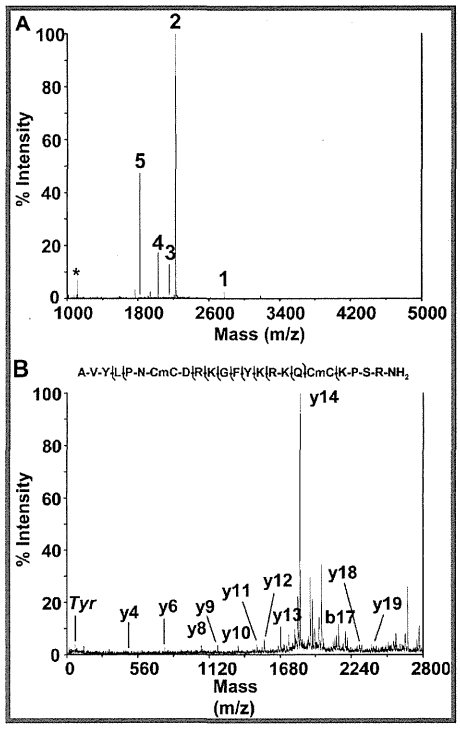


Figure 2. MS identification of AMP-IBP5. (A) MS spectrum from the microfraction indicated by the arrow in Figure 1C. 1, IGFBP-5[193-214]-NH_2 (AMP-IBP5) AVYLPNCDRKGFYKRKQCKPSR-NH_2; 2, VGF[543-561] HYHHALPPSRHYPGREAQA; 3, VGF[543-560]; 4, VGF[543-559]; 5, Desmoglein 2[10-24] NENKLLPKHPHLVRQ; *, double charged ions of 2. (B) MS/MS identification of AMP-IBP5. The peptide was identified with one b-ion and eleven y-ions. *Tyr* denotes the immonium ion of Tyr.

distributed over five fractions (Figure 1A, open and closed boxes). These five fractions were each subjected to reductive alkylation and further separated by cation-exchange HPLC (Figure 1B). To obtain a highly basic fraction, peptides eluted over the ammonium formate concentration of 0.9 M (pH 3.8) (Figure 1B, hatched box) were pooled for subsequent LC-MALDI-MS analysis. Table 1 summarizes a list of MS/MS-identified peptides in these basic fractions obtained by the cation-exchange separations. We identified 37 peptides, of which 35 peptides arose from precursor proteins reported to be enriched in secretory granules,[4,18] including VGF, somatostatin, prohormone convertase 2 (PC2), chromogranin A, chromogranin B, calcitonin gene-related peptide, calcitonin and nucleobindin 1. Among the identified peptides, AVYLPN(CmC)DRKGFY-KRKQ(CmC)KPSR-NH_2 was derived from IGFBP-5 and found to be unique with a net charge of +7 even at pH 7.0 (Table 1). The closed box (Figure 1A), the hatched box (Figure 1B) and the arrow (Figure 1C) indicate the fractions in which this peptide originated. MS profiling of the microfraction indicated by the arrow in Figure 1C showed the signal of this peptide (Figure 2A, peak 1), whose sequence was identified by MS/MS (Figure 2B and Supplementary data).

IGFBP-5 is a secreted protein, but its proteolytic processing and concomitant generation of functional peptides remained unknown. To search for different IGFBP-5-derived peptides, we analyzed other fractions of the secretopeptidome by LC-MALDI-MS/MS. As shown in Table 2 and Figure 3, we identified a total of eight distinct peptides derived from IGFBP-5. This result

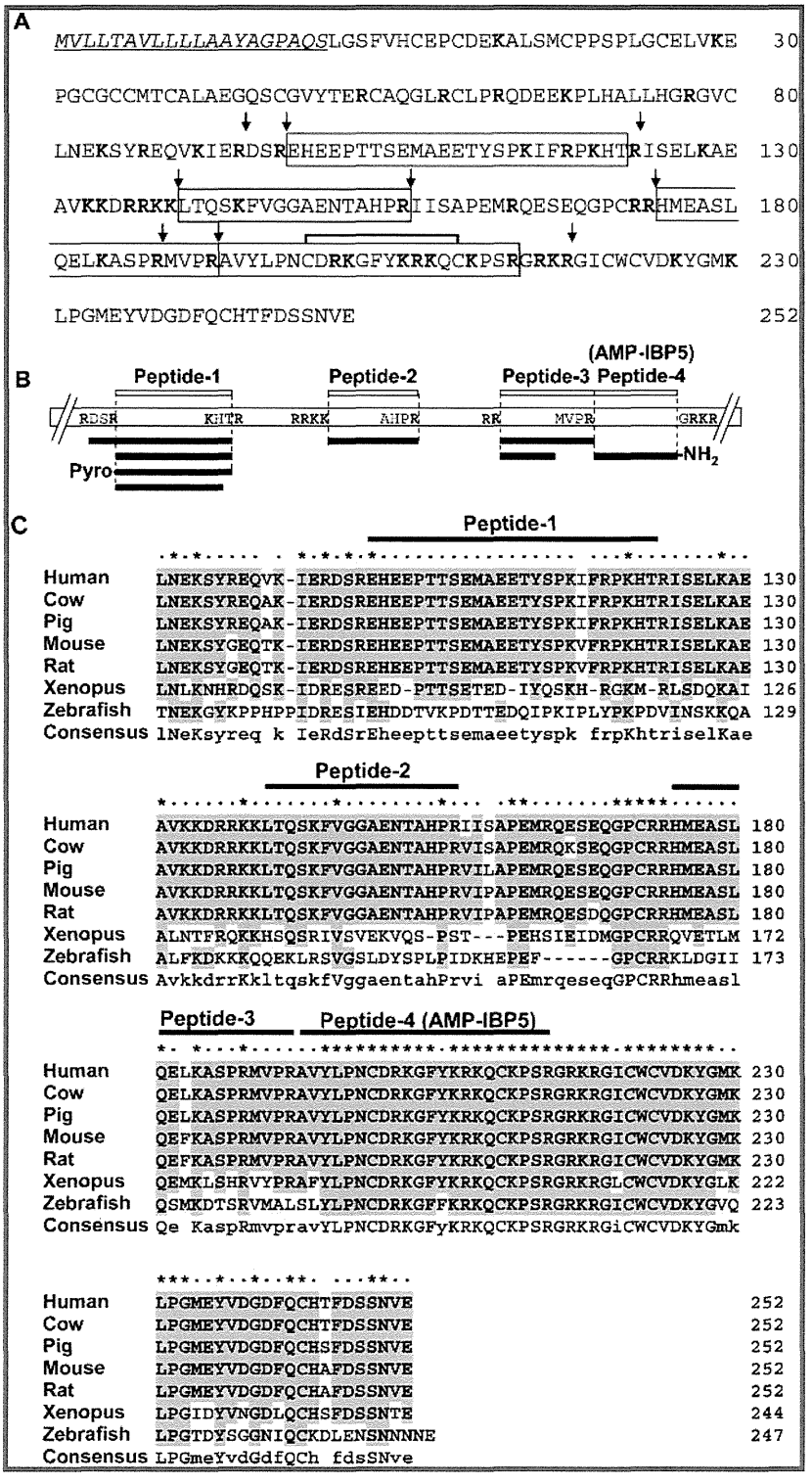


Figure 3. Peptides processed from IGFBP-5 and sequence alignment of IGFBP-5 between various vertebrates. (A) Amino acid sequence of human IGFBP-5. Signal sequence is indicated in italic and underlined. Basic amino acids, K and R, are denoted by bold letters. Four distinct regions from which peptides are processed out are boxed as follows: Peptide-1, IGFBP-5[98-122]; Peptide-2, IGFBP-5[140-156]; Peptide-3, IGFBP-5[175-192]; Peptide-4 (AMP-IBP5), IGFBP-5[193-214]-NH_2. The disulfide-bond (C199-C210) determined in this study is shown by the bold line. Processing sites predicted by identified peptides (Figure 3B) are indicated by arrows. (B) Identified peptides derived from IGFBP-5. Peptides listed in Table 2 are denoted by closed boxes. Above the entire precursor, open boxes indicate synthesized peptides. Basic amino acids in the processing sites are denoted by bold letters. Pyro- and -NH_2 mean N-terminal pyroglutamination and C-terminal amidation, respectively. (C) Sequence alignment of the central and C-terminal domains of various vertebrate IGFBP-5s. Residues conserved in more than four species are shaded. Asterisks indicate the completely conserved residues and dots indicate residues conserved in more than four vertebrates. Consensus amino acid residues conserved completely and in more than four species are indicated by upper and lower cases, respectively. Accession numbers of IGFBP-5: human, NP_000590.1; cow, NP_001098797.1; pig, NP_999264.1; mouse, NP_034648.2; rat, NP_036949.1; *Xenopus*, NP_001083938.1; and zebrafish, NP_991289.1. Residue numbering for zebrafish is based on the signal sequence prediction by SignalIP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP/).