A Database of 660 Peptide Ion Cross Sections: Use of Intrinsic Size Parameters for Bona Fide Predictions of Cross Sections

Stephen J. Valentine, Anne E. Counterman, and David E. Clemmer
Department of Chemistry, Indiana University, Bloomington, Indiana, USA

An ion trap/ion mobility/time-of-flight mass spectrometry technique has been used to measure collision cross sections for 660 peptide ions generated by tryptic digestion of 34 common proteins. Measured cross sections have been compiled into a database that contains peptide molecular weight and sequence information. The database is used to generate average intrinsic contributions to cross section (size parameters) for different amino acid residues by solving systems of equations that relate the unknown contributions of individual residues to the sequences and cross sections of database peptides. Size parameters are combined with information about amino acid composition to calculate cross sections for database peptides. Bona fide cross section predictions (made prior to measurement) for peptides observed in tryptic digests of sperm whale myoglobin and yeast enolase are made. Eight of 10 predicted cross sections are within 2% of the experimental values and all 10 are within 3.2%. The utility of size parameters for cross section prediction is explored and discussed. (J Am Soc Mass Spectrom 1999, 10, 1188–1211) © 1999 American Society for Mass Spectrometry

Recently we have developed a hybrid electrospray ionization (ESI) [1]/ion trap [2, 3]/ion mobility [4]/time-of-flight mass spectrometry (TOFMS) technique [5, 6] that allows mass-to-charge (m/z) ratios of mobility-separated ions to be measured in a single experimental sequence. The method is rapid, sensitive, and especially suited for determination of average cross sections [7] for components of mixtures, such as peptides formed by tryptic digestion of proteins. In this paper we report a database of cross sections and sequences for 660 singly and doubly charged peptide ions obtained from tryptic digestion of common proteins. Prior to measurement of this database, cross sections of only a few dozen peptides were reported [6, 8–10]. Cross sections provide information about the three-dimensional structures of ions; a collection of these measurements should benefit efforts to understand the intrinsic structural properties of peptides in the absence of sovent.

In an initial report, we used 113 cross sections for singly-protonated peptides having the general form $[Xxx_nLys + H]^+$ (Xxx refers to any naturally occurring amino acid except Cys, His, Lys, or Arg and $n = 4–9$) to derive information about the intrinsic sizes of amino acids [11]. This was accomplished by writing a system of 113 equations that relates the occurrence frequency and unknown amino acid sizes (17 in this study) to a reduced cross section of each peptide, as given by

$$\frac{\sum_i n_{ij} p_i}{\sum_i n_{ij}} = \Omega_{j,\text{reduced}} \quad (1)$$

In this equation, n_{ij} corresponds to the number of times an amino acid i occurs in each sequence j, $\Omega_{j,\text{reduced}}$ is the reduced cross section of sequence j (described below), and p_i is the unknown amino acid size parameter. The 113 equations were solved for the average, best-fit intrinsic size parameters (p_i) using a linear algebra regression method described previously [12]. Size parameters were subsequently combined with peptide sequences in order to calculate cross sections for the 113 peptides. More than 90% of the calculated cross sections were within 2% of experiment [11].

There are three objectives in the work presented here: (1) to report the measured cross sections; (2) to examine the utility of intrinsic size parameters for cross section predictions; and (3) to begin to assess the utility of size parameters as structural indicators. This report is primarily aimed at simply making these values available. A consideration in the decision to report an extensive list of tryptic digest peptides is that peptide sequences reflect those associated with elements of structure relevant to proteins. One use of this database may be to restrict structures for gas-phase ions generated by molecular modeling methods. We have recently used this approach to derive volumes for individual amino acids in peptide ions and compare these values

Address reprint requests to David E. Clemmer, Dept. of Chemistry, Indiana University, Bloomington, IN 47405.

1044-0305/99/$20.00
PII S1044-0305(99)00079-3

Received February 11, 1999
Revised May 19, 1999
Accepted May 19, 1999

with those found in protein cores [13]. Because peptides are derived from common proteins, it is often possible to compare model structures of the gas-phase peptide ion sequences to the structures of identical sequences within the anteceding protein core. The data should be directly useful for calculation of collision frequencies during ion formation and the transient time that ions exist in various mass spectrometers.

Insomuch as it is possible, the accomplishment of the second goal—to predict peptide cross sections from sequence—would have a significant impact on the ability to assign peaks in nested ion mobility/time-of-flight mass spectra. The combination of accurate *m/z* measurements and calculated cross sections (for expected peptide sequences) may provide a viable alternative to fragmentation-based strategies for peak assignments in spectra of complicated mixtures (e.g., peptides from digestion of protein mixtures). The mobility separation is an interesting time regime (milliseconds), which is intermediate between the microsecond flight times required for *m/z* determination and seconds-to-minutes timescales for condensed phase separations. Our laboratory is currently interfacing a condensed phase separation strategy with the ion mobility/time-of-flight approach; the data presented here are important in the foundation associated with developing these methods.

Predictions of cross section from sequences would also help define gas-phase ion structures. In the small systems reported here, it appears that structure is largely determined by composition instead of sequence [11]; some dramatic examples, in which position of a charged residue is used to stabilize and destabilize helices, have been presented [10a]. Here, we find that the number of calculated cross sections within 2% of experimental data can be remarkably high (~70%–90% for different regions of the database). Further improvements in accuracy are expected if sequence constraints are added; however, none have been used in this report. Calculations of cross sections for peptides that were used to derive the intrinsic size parameters (i.e., retrodictions) [14] are more accurate than calculations for regions not contained in the parameterization set. Peptide length appears to be an important factor that limits the application of parameters for cross section prediction. Below we provide several parameter sets that should be useful for calculating cross sections for peptides containing 3–15 residues. We also report bona fide predictions of cross section (predictions made prior to measurement) [14]. In these predictions, 8 of 10 cross sections that were predicted for small $[Xxx_nLys + H]^+$ peptides (n = 4 to 9) were within 2% of experimental values; all 10 calculated cross sections were within 3.2%.

The third goal of this work—assessment of intrinsic size factors as structural indicators—is more challenging. Intrinsic size parameters, determined directly from experimental data, arise from a combination of factors including the average contributions of residues to the physical shapes of the peptides, as well as differences in the collision dynamics associated with interaction of the buffer gas with different residues. As we discuss below, a parameter derived for an individual residue (within a parameter set) can vary substantially, depending upon the behavior of other residues; at this stage, even the relatively large database allows only very limited exploration of how parameters vary with length, sequence, and conformer type. However, we begin this assessment here and note that in future work we will expand this database. This should help clarify the meaning of size parameters.

A number of other methods are being developed to investigate the conformations of anhydrous biomolecular ions. Hydrogen–deuterium exchange [15], proton-transfer [16], and molecular adduction experiments [17] provide structural information based on differences in the chemical properties of specific conformations. Kinetic energy release measurements [18], microscopy studies of high-energy impacts of proteins on surfaces [19], and collision cross section measurements utilizing ion scattering [20, 21] and ion mobility [8, 9, 22, 23] methods yield information about the physical sizes of the gas-phase biomolecules. These experimental methods have been complemented by molecular modeling techniques to provide insight into gas-phase conformations [8, 24, 25]. The work presented here is also related to condensed-phase studies involving structural prediction based on sequence information [14, 26, 27].

Experimental

Overview

The ESI/ion trap/ion mobility/mass spectrometry methods used here have been discussed previously [3, 5, 6]; a brief description is presented here. The continuous electrospray ion beam is accumulated for ~100 ms in an ion trap. Concentrated ion packets are injected into a 40.4 cm long drift tube containing ~2–3 torr of helium buffer gas at 300 K. Ions travel across the drift tube under the influence of an attractive uniform electric field (8.66 V cm^{-1}) that is established by a series of evenly spaced BeCu rings. As ions drift through the buffer gas they are separated based on differences in their mobilities; compact ions have higher mobilities than elongated ones. After exiting the drift tube, ions are focussed into the source region of a time-of-flight (TOF) mass spectrometer where high-voltage, high-frequency pulses (synchronous with the initial ion-mobility injection pulse) are used to initiate time-of-flight measurements.

Formation of Mixtures of Peptide Ions

Information about the purities and sources for proteins used in tryptic digestion is provided in Tables 1 and 2. Tryptic fragments for each protein are generated by combining 150 μL of a 0.2 mg/mL trypsin (Sigma, sequencing grade) solution in 0.2 M ammonium bicar-

Table 1. Cross sections for singly protonated peptides from tryptic digests[a]

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
2	AK	gludehy_bov	217.28	87.01	1
2	AR	lactotrans_bov	245.29	93.79	1
2	EK	albu_pig alcodehy_yst BSA	275.31	97.89 (1.14)	9
2	ER	albu_dog albu_horse	303.33	102.02 (0.65)	6
2	FK	albu_pig HSA	293.37	101.74 (1.00)	3
2	FR	canhyd_bov	321.39	111.10 (0.03)	2
2	HK	cytc_horse bcas_bov crephos_rab	283.30	98.96 (0.76)	6
2	IR	conalb_chick	287.37	106.73	1
2	KK	canhyd_bov cytc_horse	274.37	97.40 (0.60)	3
2	LK	albu_horse albu_pig albu_sheep HSA transfr_hum tryps	259.36	100.92 (2.59)	11
2	LR	hb_dog albu_sheep BSA crephos_rab hb_bov hb_hum hb_pig hb_sheep	287.37	110.03 (1.44)	19
2	NR	canhyd_bov gludehy_bov lys_tew	288.31	101.76 (1.49)	5
2	QR	albu_pig albu_sheep BSA	302.34	104.07 (1.92)	7
2	RP	transfr_hum	271.33	96.48 (2.22)	2
2	SK	aldol_rab	233.27	88.57	1
2	TK	blacto_bov	247.30	92.76 (0.86)	2
2	TR	gludehy_bov	275.32	100.59 (1.09)	3
2	VK	bcas_bov enolase_yst gludehy_bov hb_pig hb_sheep	245.33	93.56 (1.37)	7
2	VR	conalb_chick	273.34	104.93 (2.94)	3
2	YH	hb_bov hb_dog hb_hum hb_pig hb_sheep	318.34	105.75 (1.10)	11
2	YK	crephos_rab HSA	309.37	104.66 (2.10)	4
2	YR	hb_bov hb_dog hb_hum hb_pig hb_sheep	337.39	110.63 (1.16)	13
3	ADR	transfr_hum	360.38	123.31 (0.68)	2
3	AFK	albu_dog albu_pig HSA	364.45	121.15 (1.45)	8

(*Continued*)

Table 1. (*Continued*)

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
3	ALK	alcodehy_yst blacto_bov	330.43	116.85 (1.95)	4
3	APR	lactotrans_bov	342.4	113.05 (1.78)	3
3	ATK	HSA	318.38	103.25	1
3	AVK	albu_horse	316.41	110.53 (0.76)	3
3	CEK	kcas_bov	378.45	124.58	1
3	CRR	lactotrans_bov	433.54	137.86 (0.38)	2
3	DER	kcas_bov	418.41	123.18 (3.30)	3
3	EFK	albu_horse	422.49	134.38 (0.88)	3
3	FDK	blacto_bov	408.46	127.82 (0.54)	2
3	FGK	lactotrans_bov	350.43	116.66 (1.41)	2
3	FPK	albu_pig albu_sheep BSA	390.49	123.14 (0.85)	8
3	FYR	aldol_rab	484.56	147.93 (1.34)	3
3	GER	crephos_rab	360.38	116.82 (0.34)	3
3	GKK	cytc_horse	331.40	111.71 (1.19)	3
3	GWK	alcodehy_yst	389.46	123.70 (1.18)	3
3	HFK	albu_horse	430.51	133.96 (1.01)	3
3	HLK	myo_horse	396.50	133.05 (0.71)	2
3	HPK	crephos_rab	380.45	123.77 (2.63)	3
3	IAK	albu_horse kcas_bov	330.43	113.56 (1.50)	6
3	IEK	bcas_bov	388.47	127.42 (0.54)	3
3	INK	bcas_bov	373.46	122.09 (0.19)	3
3	ITK	albu_horse	360.46	118.89 (1.19)	2
3	KLR	crephos_rab	415.54	139.37 (1.45)	3
3	LAK	albu_dog enolase_yst	330.43	118.08 (1.19)	5
3	LFK	myo_sw	406.53	130.29	1
3	LVK	albu_sheep	358.49	127.00 (1.69)	3
3	MEK	alcodehy_yst	406.50	126.51 (3.22)	2
3	NLR	lactotrans_bov transfr_hum	401.47	132.47 (2.00)	6
3	NYK	albu_pig	423.48	133.47 (1.39)	2
3	QIK	albu_horse albu_sheep BSA	387.49	129.74 (1.13)	5
3	QVR	canhyd_bov	401.47	131.12 (1.87)	2
3	SIK	crephos_rab	346.46	116.59 (0.35)	2
3	TGR	crephos_rab	332.37	109.48 (2.51)	3
3	TVR	transfr_hum	374.45	122.49 (0.85)	3
3	VHK	albu_sheep	382.47	123.04 (1.57)	2
3	VTK	albu_sheep	346.43	114.24 (0.79)	2
3	YTK	albu_dog albu_horse albu_pig HSA	410.48	126.96 (1.08)	9
3	YTR	albu_sheep BSA	438.49	131.96 (0.97)	6
3	YVR	alcodehy_yst	436.52	140.16	1
4	ADEK	albu_sheep BSA	461.48	138.92 (1.61)	4
4	ADTR	alcodehy_yst	461.48	141.34	1
4	AGIK	cytc_horse	387.50	130.20 (0.42)	2
4	AHGK	hb_bov hb_dog hb_pig hb_rab hb_sheep hb_hum	411.47	133.57 (1.15)	15
4	ANVK	hb_pig	430.51	136.46 (1.52)	3

(*Continued*)

Table 1. (*Continued*)

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
4	APPK	conalb_chick	411.51	131.96	1
4	AQEK	lactotrans_bov	474.50	141.97 (0.52)	2
4	ATNE	cytc_horse	433.43	128.33	1
4	AVEK	crephos_rab	445.52	141.29 (1.64)	3
4	AVLK	canhyd_bov	429.57	145.63 (0.35)	2
4	AVSK	enolase_yst	403.49	132.83 (1.99)	3
4	DAHK	HSA	469.51	143.45	1
4	DLTK	conalb_chick	475.55	147.56 (1.64)	3
4	DTHK	albu_sheep BSA	499.53	142.74 (1.62)	5
4	DTYK	albu_pig	525.57	150.51 (1.36)	3
4	EAYK	albu_dog	509.57	148.42 (0.47)	3
4	ENLK	aldol_rab	502.58	150.35 (1.73)	2
4	EVFR	crephos_rab	549.63	157.46 (2.83)	3
4	EVVR	acas_bov	501.59	148.81 (0.11)	2
4	FGAK	gludehy_bov	421.50	134.34 (0.96)	3
4	FGDR	albu_dog	493.53	144.12 (0.44)	3
4	FGER	albu_pig	507.55	148.17 (0.18)	3
4	FLGK	albu_horse	463.58	152.63 (0.74)	3
4	FWGK	albu_pig albu_sheep BSA	536.64	160.46 (0.50)	8
4	GFPK	canhyd_bov	447.54	141.49 (0.20)	2
4	GGIR	gludehy_bov	401.47	129.69 (0.50)	3
4	GHGK	hb_hum	397.44	129.51 (1.12)	2
4	GILA	albu_pig	372.47	123.99 (0.99)	2
4	GILR	gludehy_bov	457.58	148.38 (0.66)	3
4	GLVK	alcodehy_yst	415.54	141.97 (1.46)	3
4	GNVK	hb_bov	416.49	134.17 (1.11)	3
4	GTDK	apotransf_bov	419.44	128.01	1
4	GVFR	enolase_yst	477.57	146.75 (0.83)	3
4	HKPK	albu_horse BSA	508.63	149.94 (2.10)	6
4	IPSK	apotransf_bov lactotrans_bov	443.55	139.78 (1.80)	4
4	KIEK	bcas_bov	516.65	155.03 (0.24)	3
4	LEHK	alcodehy_yst	525.61	154.49	1
4	LHDR	transfr_hum	539.60	157.14 (0.75)	3
4	LLSK	lactotrans_bov	459.59	155.26 (0.63)	3
4	LNYK	crephos_rab	536.64	163.88 (1.09)	2
4	LSQK	albu_horse albu_sheep BSA	474.57	149.68 (1.08)	9
4	LSQR	albu_dog albu_pig HSA	502.58	154.85 (1.16)	9
4	LTFK	gludehy_bov	507.64	161.05 (0.27)	3
4	MDAK	transfr_hum	463.56	135.56 (0.42)	3
4	NFGK	hb_bov	464.53	137.95 (0.76)	4
4	NLGK	HSA	430.51	132.68	1
4	QIMR	gludehy_bov	546.69	159.34	1
4	Ac-QRLK	HSA	585.71	163.41	1
4	SETK	lactotrans_bov	463.50	155.20	1
4	SLGK	BSA	403.49	132.39 (1.82)	3
4	SLTK	enolase_yst	447.54	143.99 (0.30)	3
4	SNVK	hb_sheep	446.51	136.52 (0.23)	3
4	SPIK	alcodehy_yst	443.55	136.08 (0.90)	3
4	TNIK	hb_dog hb_rab	474.57	144.91 (0.63)	5
4	TNVK	hb_hum	460.54	141.41 (0.17)	3
4	VGTK	albu_dog	403.49	131.69 (0.66)	3

(*Continued*)

Table 1. (*Continued*)

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
4	VGTR	BSA	431.50	136.71 (2.06)	3
4	VYAR	enolase_yst	507.60	157.06 (0.93)	3
4	WMGK	enolase_yst	520.65	152.87 (0.55)	3
4	WQWR	lactotrans_bov	674.77	176.44 (0.68)	3
4	YLHR	gludehy_bov	587.69	171.66 (3.96)	2
4	YTKK	HSA	538.65	155.28	1
5	AAWGK	hb_bov hb_hum hb_pig hb_sheep	531.62	157.36 (1.02)	13
5	ADLAK	albu_dog albu_sheep BSA	516.60	159.31 (1.06)	9
5	AFDEK	BSA	608.66	168.36 (0.44)	3
5	AIAEK	lactotrans_bov	530.63	160.73 (1.84)	3
5	APNAK	glox_aspgn	499.57	147.31 (0.98)	2
5	AWGGK	aldol_rab	517.59	152.16 (2.53)	3
5	CLVEK	transfr_hum	590.74	170.73	1
5	DIAAK	myo_sw	516.60	155.37 (2.98)	3
5	DLLFK	apotransf_bov conalb_chick lactotrans_bov transfr_hum	634.78	183.11 (3.14)	12
5	FFSDK	kcas_bov	642.72	172.73 (2.52)	3
5	Ac-GDVEK	cytc_horse	588.62	163.18 (1.29)	2
5	GGNMK	crephos_rab	505.60	147.44 (1.24)	2
5	GITWK	cytc_horse	603.70	169.34 (0.58)	3
5	GTFAK	hb_dog hb_pig hb_rab	522.61	153.97 (0.91)	9
5	HPEAK	HSA	580.65	163.89	1
5	IDAMR	albu_sheep	604.73	172.80 (0.75)	3
5	IEHLR	albu_pig	666.79	181.70 (0.16)	3
5	IETMR	BSA	648.78	181.29 (0.81)	3
5	IFLER	gludehy_bov	676.82	191.86 (0.96)	3
5	IFVQK	cytc_horse	633.80	181.96 (0.51)	3
5	IIAEK	blacto_bov	572.71	172.54 (0.50)	3
5	KATNE	cytc_horse	561.60	153.95 (0.28)	3
5	KFWGK	BSA	664.81	185.85	1
5	KNQDK	kcas_bov	631.69	168.54 (3.29)	3
5	KVYGR	lys_tew	621.74	175.06 (2.80)	2
5	LDALK	albu_horse	558.68	172.36 (1.48)	3
5	LDELR	HSA	644.73	181.93 (0.96)	2
5	LHSMK	acas_bov	614.77	169.24 (1.22)	2
5	LNELR	apotransf_bov	643.75	182.27 (1.13)	3
5	NLNEK	transfr_hum	616.68	167.97 (0.94)	3
5	NRQVR	canhyd_bov	671.76	175.89	1
5	NTYEK	transfr_hum	653.70	175.90 (2.66)	3
5	SVDGK	lactotrans_bov	504.55	150.56	1
5	TAWEK	hb_rab	633.71	170.03 (1.25)	3
5	TGQIK	enolase_yst	545.64	157.62 (0.28)	3
5	TLTGK	ubiq	518.62	157.34 (2.11)	2
5	TPGSK	lys_tew	488.55	145.50 (1.17)	2
5	TVGGK	apotransf_bov	460.54	140.18 (3.52)	2
5	VASLR	BSA	544.66	163.70 (0.92)	3
5	VATLR	albu_sheep	558.69	166.85 (0.49)	3
5	YYPLK	crephos_rab	682.82	187.31 (0.38)	2
6	AAAAEK	enolase_yst	559.63	160.38 (1.06)	2
6	AKIQDK	ubiq	701.83	180.72	1
6	ALELFR	myo_horse	747.90	202.99 (0.33)	3
6	ANIDVK	enolase_yst	658.76	176.78 (0.76)	3
6	ASEDLK	myo_sw	661.72	175.16 (0.58)	3

(*Continued*)

Table 1. (*Continued*)

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
6	AVGNLR	transfr_hum	628.74	172.89 (0.27)	3
6	AWSVAR	albu_dog albu_horse albu_sheep BSA	688.79	185.37 (0.81)	11
6	EAMAPK	bcas_bov	645.78	176.19 (1.24)	2
6	EDLIAY	cytc_horse	722.80	191.70 (0.29)	3
6	ELEDFK	gludehy_bov	779.85	198.28	1
6	ELGFQG	myo_horse	649.71	172.29 (0.14)	3
6	EMPFPK	bcas_bov	747.91	193.26 (1.43)	3
6	FVIEIR	albu_pig	775.95	207.16 (0.16)	3
6	GGVHVK	crephos_rab	595.71	167.11	1
6	GQIVGR	alcodehy_yst	628.74	173.63 (0.70)	3
6	GVPTKK	glox_aspgn	628.78	178.28 (0.28)	2
6	HLDDLK	hb_bov hb_sheep	739.83	193.90 (0.71)	7
6	HLDNLK	hb_pig	738.85	196.18 (1.17)	3
6	HPEAKR	HSA	736.84	197.52	1
6	IEEIFK	crephos_rab	777.92	197.00 (2.97)	3
6	ISATGR	gludehy_bov	603.68	169.18 (0.75)	3
6	IVAPGK	aldol_rab	583.74	173.33 (1.17)	3
6	LIFAGK	ubiq	647.83	186.10 (1.58)	2
6	LNQLLR	enolase_yst	755.92	204.97 (0.46)	3
6	LVEDLK	gludehy_bov	715.85	192.17 (1.05)	2
6	MQIFVK	ubiq	764.99	203.90 (0.59)	2
6	NDIAAK	myo_horse	630.70	173.78 (0.84)	3
6	NLDNLK	hb_dog	715.81	192.40 (0.69)	3
6	NVPLYK	enolase_yst	732.89	195.27 (0.90)	3
6	NYAEAK	HSA	694.75	181.87	1
6	NYQEAK	albu_dog albu_sheep BSA	751.80	191.16 (0.92)	8
6	QCSTSK	apotransf_bov	652.73	171.10	1
6	QLEDGR	ubiq	716.76	186.94	1
6	QTIASN	tryps	632.68	170.79	1
6	SEIAHR	albu_horse albu_pig albu_sheep BSA albu_pig albu_sheep	711.78	181.53 (0.71)	11
6	SEVAHR	HSA	697.76	177.66 (2.46)	3
6	SGIQVR	tryps	658.76	178.36	1
6	SKKTAK	conA	661.81	168.94	1
6	SVYDSR	enolase_yst	725.76	184.17	1
6	TEAEMK	myo_horse myo_sw	707.80	182.54 (2.25)	5
6	TGAPAR	enolase_yst	571.64	162.57 (1.27)	3
6	TGPNLH	cytc_horse	637.70	164.29 (0.30)	3
6	TPVSEK	albu_horse BSA	659.74	175.98 (0.73)	5
6	VKAHGK	hb_dog	638.78	165.87	1
6	YDLDFK	enolase_yst	799.89	201.01	1
6	YIPGTK	cytc_horse	677.80	183.75	1
6	YLTTLK	lactotrans_bov	737.90	197.94 (0.70)	3
7	AACLLPK	HSA	714.93	194.76	1
7	AAGHDGK	enolase_yst	654.69	170.18 (2.14)	2
7	ALPMHIR	blacto_bov	837.06	213.70 (1.22)	3
7	APVDAFK	lactotrans_bov	746.87	189.05 (0.76)	3
7	ASEDLKK	myo_horse	789.89	196.81 (0.17)	3
7	ATDEQLK	albu_sheep	803.87	205.89 (0.16)	2
7	ATEEQLK	BSA	817.90	206.40 (1.81)	2

(*Continued*)

Table 1. (*Continued*)

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
7	AVPYPQR	bcas_bov	829.96	202.81 (0.47)	2
7	DGADFAK	aldol_rab	722.76	185.83 (0.90)	3
7	DSAIMLK	conalb_chick	776.95	203.77 (1.32)	3
7	DTLQQER	conalb_chick	888.94	206.76 (0.37)	3
7	ELTEFAK	albu_sheep	836.95	209.40 (0.69)	2
7	EVTEFAK	albu_pig	822.92	202.91 (0.29)	3
7	FGERAFK	HSA	853.98	218.63 (1.78)	3
7	FNDLGEK	albu_horse	821.89	206.73 (1.10)	3
7	GDVAFVK	apotransf_bov conalb_chick transfr_hum	734.86	200.45 (1.77)	8
7	GFFYTPK	ins_bov	859.00	221.88	1
7	GGVVGIK	aldol_rab	628.78	175.87 (0.62)	3
7	GPFPIIV	bcas_bov	741.94	195.69 (0.70)	3
7	GVLHAVK	enolase_yst	722.89	198.60 (0.42)	3
7	HLADLSK	enolase_yst	782.90	201.78 (0.62)	3
7	HNNHMAK	crephos_rab	850.96	207.15 (4.31)	2
7	IATAIEK	enolase_yst	744.89	202.92 (0.37)	3
7	ILLSSAK	albu_horse	730.91	202.97 (0.75)	3
7	ISQRYQK	acas_bov	922.06	225.85	1
7	IVTDLAK	albu_pig	758.92	204.75 (1.40)	3
7	IVTDLTK	albu_horse	788.95	207.02 (0.28)	3
7	LGLVGSR	albu_pig	700.84	194.18 (0.27)	2
7	LMVEMEK	crephos_rab	879.10	220.91	1
7	LSQRFPK	HSA	875.05	225.29 (1.46)	2
7	LVTDLTK	BSA HSA	788.95	205.76 (0.98)	6
7	MIFAGIK	cytc_horse	779.00	207.13 (0.39)	3
7	MLTAEEK	hb_bov hb_sheep	820.96	209.77 (0.59)	6
7	NPDPWAK	transfr_hum	826.91	198.09 (0.38)	3
7	PIANGER	canhyd_bov	755.84	201.77	1
7	Ac-SIPETQK	alcodehy_yst	843.94	205.88 (1.31)	3
7	TFAEALR	enolase_yst	806.92	209.98 (0.52)	3
7	VAAALTK	hb_bov hb_sheep	672.83	190.32 (0.73)	7
7	VADALTK	hb_pig	716.84	194.69 (1.73)	3
7	VDPVNFK	hb_dog hb_bov hb_hum	817.95	208.73 (0.91)	12
7	VLAAVYK	aldol_rab	762.95	206.91 (1.08)	3
7	VLASSAR	albu_sheep	702.82	195.27 (0.36)	3
7	VLPVPQK	bcas_bov	779.99	206.94 (0.57)	3
7	VLSAADK	hb_bov hb_pig hb_sheep	702.81	193.03 (1.28)	10
7	VLSPADK	hb_dog hb_hum hb_rab	728.85	196.51 (1.02)	8
7	VLTSAAK	albu_pig	688.83	194.37 (0.13)	3
7	VSEALTK	hb_rab	746.87	198.59 (2.27)	3
7	VVTDLTK	albu_dog	774.92	202.35 (0.99)	3
7	WNMQNGK	conA	876.99	206.26 (0.19)	3
7	YLYEIAR	albu_pig BSA HSA	927.08	228.03 (0.78)	6
7	YLYEVAR	albu_horse albu_sheep	913.05	226.13 (0.63)	5
8	AADALLLK	enolase_yst	814.00	223.86 (2.78)	3
8	ADFAEISK	albu_dog	879.97	218.06 (0.55)	3
8	ADFAEVSK	albu_horse	865.95	214.17 (0.67)	3
8	ADFTDVTK	albu_sheep	895.97	214.56 (1.13)	3

(*Continued*)

Table 1. (*Continued*)

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
8	ADFTEISK	albu_pig	910.00	219.61 (0.21)	3
8	AEFAEVSK	HSA	879.97	214.97 (4.51)	2
8	AEFVEVTK	BSA	922.05	223.42	1
8	AHELLNTK	glox_aspgn	925.06	227.41	1
8	ALQASALK	aldol_rab	800.96	218.92 (0.62)	3
8	DDHPNLPK	albu_horse	935.01	211.02 (2.10)	3
8	DDNPNLPR	HSA	939.99	218.41 (2.84)	2
8	DERFFSDK	kcas_bov	1043.11	238.70	1
8	DIPVPKPK	alcodehy_yst	893.10	219.20 (1.10)	3
8	DIVGAVLK	alcodehy_yst	814.00	206.07 (0.54)	3
8	DLGEEHFK	BSA	974.05	225.29	1
8	DLGEENFK	HSA	951.01	223.98 (1.41)	3
8	DLGEQYFK	albu_pig	999.10	232.22 (0.30)	3
8	DSADGFLK	apotransf_bov	851.92	209.30 (1.09)	3
8	DSALGFLR	lactotrans_bov	878.00	216.27 (4.91)	2
8	EALDFFAR	alcodehy_yst	968.08	231.05	1
8	ELSDIAHR	aldol_rab	940.03	221.32 (2.53)	3
8	EYEATLEK	albu_dog	982.06	229.28 (0.93)	3
8	FGVNGSEK	conalb_chick	836.91	201.76 (0.59)	3
8	GASIVEDK	gludehy_bov	817.90	205.22 (1.78)	2
8	GFFYTPKA	ins_bov	930.08	239.80	1
8	IDALNENK	blacto_bov	916.01	225.09 (1.45)	3
8	IGDYAGIK	alcodehy_yst	835.96	210.40 (0.63)	3
8	KLGLVGSR	albu_pig	829.02	204.43	1
8	KVSEALTK	hb_rab	875.04	213.55 (1.46)	3
8	LIVTQTMK	blacto_bov	933.18	243.91 (0.71)	3
8	MVEGFFDR	gludehy_bov	1000.14	237.35 (0.51)	3
8	RCQYVTEK	aldol_rab	1026.18	244.74 (0.92)	2
8	SAASLNSR	tryps	804.87	201.95	1
8	SEVAHRFK	HSA	973.11	235.87 (1.31)	3
8	SHHWGYGK	canhyd_bov	971.05	227.58 (0.33)	2
8	TKIPAVFK	blacto_bov	903.14	227.74 (1.26)	3
8	TKLTEEEK	acas_bov	977.09	243.22	1
8	TYETTLEK	HSA	984.08	239.55 (1.30)	2
8	VHLSAEEK	hb_pig	912.02	222.58 (1.10)	3
8	VHLTAEEK	hb_dog	926.04	226.87 (0.45)	3
8	VHLTPEEK	hb_hum	952.08	228.28 (2.05)	2
8	VLTPDLYK	crephos_rab	948.13	230.73 (1.13)	2
8	VYGVQGLR	glox_aspgn	891.05	218.35 (3.65)	2
8	YKELGFQG	myo_horse	941.06	227.91 (0.59)	3
8	YLGEEYVK	transfr_hum	1000.12	238.79 (1.26)	3
8	YNLGLDLR	gludehy_bov	963.11	230.06	1
9	AAVTAFWGK	hb_bov	950.11	237.76 (0.53)	4
9	AAVTGFWGK	hb_sheep	936.09	233.00 (0.18)	3
9	ALMSAVEDR	glox_aspgn	991.13	235.30	1
9	ANELLINVK	alcodehy_yst	1013.21	249.69 (0.86)	3
9	APNHAVVTR	transfr_hum	964.10	219.79 (0.94)	3
9	DFPIANGER	canhyd_bov	1018.10	229.78	1
9	DGKYDLDFK	enolase_yst	1100.20	250.07 (0.42)	2
9	DLLFRDDTK	apotransf_bov	1122.25	248.68	1
9	EAVLGLWGK	hb_pig	972.16	235.37 (1.14)	3
9	EGIPPDQQR	ubiq	1039.12	229.87	1
9	ESTLHLVLR	ubiq	1067.26	247.60 (3.52)	2
9	FMMFESQNK	conalb_chick	1161.36	259.37 (4.27)	2
9	FQPLVDEPK	albu_pig	1072.23	255.89 (0.74)	3
9	GNTHNVYAK	glox_aspgn	1003.09	234.91 (0.68)	2
9	GTDVHAWIR	lys_tew	1054.18	241.75 (0.34)	2
9	HGGFKPTDK	crephos_rab	986.10	226.66 (0.55)	2
9	LDELRDEGK	HSA	1074.16	233.18	1
9	LHVDPENFK	hb_dog	1098.23	248.13 (1.98)	4
9	LHVDPENFR	hb_hum hb_sheep	1126.24	246.51 (1.79)	4

(*Continued*)

Table 1. (*Continued*)

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
9	LLGLFPDAN	conA	959.12	232.59 (0.23)	3
9	LRVDPVNFK	hb_bov hb_sheep	1087.29	250.74 (1.07)	2
9	MFLGFPTTK	hb_pig hb_pig	1041.28	250.02 (1.45)	6
9	MFLSFPTTK	hb_bov hb_hum hb_sheep	1071.31	255.16 (0.70)	10
9	QLLLTADDR	aldol_rab	1044.18	246.37 (2.12)	2
9	QSALAELVK	albu_horse	958.13	234.39 (1.35)	3
9	QTALVELLK	albu_sheep	1014.24	245.01 (1.53)	3
9	QTALVELVK	HSA	1000.21	242.43 (0.50)	3
9	RPAQPLKNR	canhyd_bov	1079.30	254.31 (0.44)	2
9	SAVTALWGK	hb_hum hb_rab	932.09	231.09 (1.09)	6
9	SKGGVVGIK	aldol_rab	844.03	211.01 (0.83)	2
9	SLVSGLWGK	hb_dog	946.12	236.29 (0.69)	3
9	TDLNHENLK	crephos_rab	1083.17	245.50 (0.58)	2
9	TFQSFPTTK	hb_dog	1056.19	245.22 (0.25)	2
9	TGAPARSER	enolase_yst	944.02	249.59 (3.78)	2
9	TLSDYNIQK	ubiq	1081.20	248.81	1
9	VAAHAVVAR	conalb_chick	893.06	228.54 (1.39)	3
9	YFGYTGALR	conalb_chick	1047.19	238.95 (2.80)	2
9	YYGYTGAFR	apotransf_bov	1097.20	242.91 (1.19)	2
10	AQSDFGVDTK	conalb_chick	1067.13	241.43 (1.36)	3
10	DGAGDVAFVK	transfr_hum	978.08	229.25 (0.61)	3
10	DLFDPIIQDR	crephos_rab	1231.38	266.68 (0.41)	2
10	DSNVNWNNLK	conalb_chick	1203.28	261.59	1
10	EAYKSEIAHR	albu_dog	1203.33	267.64 (0.91)	3
10	ECCDKPLLEK	BSA	1177.40	268.27	1
10	EKDIVGAVLK	alcodehy_yst	1071.29	254.31 (1.07)	3
10	FKDLGEENFK	HSA	1226.36	266.70 (2.94)	2
10	FKDLGEQYFK	albu_pig	1274.45	269.57 (0.71)	2
10	GVIFYESHGK	alcodehy_yst	1136.28	254.90 (0.56)	3
10	IGSEVYHNLK	enolase_yst	1159.31	269.70 (0.37)	3
10	KQSALAELVK	albu_horse	1086.30	255.37 (2.87)	3
10	KQTALVELLK	albu_pig albu_sheep BSA	1142.41	272.66 (2.62)	8
10	LLVVYPWTQR	hb_hum hb_pig hb_rab hb_sheep	1274.54	283.07 (1.87)	10
10	LVASSQLALA	albu_horse	972.16	234.24 (0.40)	3
10	LVASTQAALA	albu_sheep	944.10	225.98 (0.62)	3
10	LVKELTEFAK	albu_sheep	1177.41	265.20 (1.14)	3
10	LVNELTEFAK	BSA	1163.34	267.51 (1.38)	3
10	LVNEVTEFAK	albu_horse	1149.32	262.66 (0.85)	4
10	LVVSTQTALA	BSA	1002.18	239.27 (1.48)	3
10	SEEEYPDLSK	crephos_rab	1196.24	257.78 (0.49)	2
10	TAAYVNAIEK	gludehy_bov	1079.23	246.52 (2.11)	2
10	VLDSFSNGMK	hb_bov hb_sheep	1097.26	252.25 (1.41)	4
10	VLNSFSDGLK	hb_dog	1079.23	255.16 (0.96)	3
10	VLQSFSDGLK	hb_pig	1093.25	255.95 (1.41)	3
10	YNDLGEEHFR	albu_dog	1279.34	271.02 (1.40)	3
11	EMGGVVDNAAR	glox_aspgn	1118.24	244.16 (0.66)	2
11	HKTDLNHENLK	crephos_rab	1348.49	288.37	1
11	HLVDEPQNLIK	albu_sheep BSA	1305.50	282.63 (0.86)	5
11	HPDYSVSLLLR	albu_horse	1299.50	296.03 (0.62)	3
11	HPEYAVSVLLR	albu_sheep	1283.50	281.06 (2.47)	3

(*Continued*)

Table 1. (*Continued*)

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
11	KVLDSFSNGMK	hb_bov hb_sheep	1225.43	266.78 (1.13)	4
11	KVLQSFSDGLK	hb_pig	1221.43	269.66	1
11	LEQWAEEAVAR	glox_aspgn	1301.43	285.03 (2.29)	2
11	LFTGHPETLEK	myo_horse	1271.44	277.00	1
11	LQHGTILGFPK	gludehy_bov	1210.45	284.39 (0.59)	2
11	SASDLTWDNLK	transfr_hum	1249.35	275.99 (2.94)	2
11	SCQAQPTTMAR	kcas_bov	1193.40	260.11 (4.55)	2
11	SSGTSYPDVLK	Tryps	1153.26	251.05	1
11	TGPNLHGLFGR	cytc_horse	1168.33	257.73	1
11	TPEVDDEALEK	blacto_bov	1245.31	249.83 (1.34)	3
11	VDEVGAEALGR	hb_sheep	1115.22	251.98 (0.49)	3
11	VGLSASTGLYK	conA	1095.27	270.04 (0.79)	3
11	VLSPADKTNIK	hb_rab	1185.39	259.52 (2.72)	3
12	AAFTECCQAADK	HSA	1257.41	271.68	1
12	EETLMEYLENPK	cytc_horse	1495.67	309.03	1
12	EFTPVLQADFQK	hb_bov	1422.61	305.06 (0.90)	4
12	FLANVSTVLTSK	hb_rab hb_sheep	1279.51	297.19 (1.30)	5
12	FLASVSTVLTSK	hb_hum	1252.48	298.80 (0.41)	2
12	GQSIDDMIPAQK	crephos_rab	1302.47	276.29 (2.26)	2
12	GTEFTVNDLQGK	conalb_chick	1308.42	276.26 (0.93)	2
12	SIGGEVFIDFTK	alcodehy_yst	1312.49	296.14 (0.34)	2
12	SISIVGSYVGNR	alcodehy_yst	1251.41	267.82 (1.51)	3
12	TVDMESTEVFTK	acas_bov	1386.54	304.91	1
12	TVLGNFSAFVAK	albu_horse	1253.47	290.06 (1.14)	3
12	VNQIGTLSESIK	enolase_yst	1288.47	278.77 (0.59)	3
12	VVAGVANALAHK	hb_dog hb_hum hb_pig hb_rab hb_sheep	1149.37	267.76 (1.71)	14
13	GNPTVEVELTTEK	enolase_yst	1416.56	302.96 (0.35)	2
13	LGEYGFQNALIVR	albu_pig	1479.71	305.73	1
13	VKVDEVGAEALGR	hb_sheep	1342.52	286.34 (1.24)	2
13	VKVDEVGGEALGR	hb_bov	1328.50	281.42 (2.25)	2
13	VNVDEVGGEALGR	hb_dog hb_hum hb_pig	1314.43	278.91 (1.10)	9
13	VNVEEVGGEALGR	hb_rab	1328.45	278.43 (2.38)	2
14	APQVSTPTLVEIGR	albu_horse	1467.69	305.13	1
14	GILAADESTGSIAK	aldol_rab	1332.48	300.14 (0.72)	3
14	TAGWNIPMGLLYSK	apotransf_bov	1550.84	323.84	1
14	TGQAPGFTYTDANK	cytc_horse	1470.60	286.74	1
15	ANGTTVLVGMPAGAK	alcodehy_yst	1386.64	286.04 (2.23)	2
15	AVDDFLISLDGTANK	enolase_yst	1578.75	316.92 (2.52)	2
15	LGANAILGVSLAASR	enolase_yst	1412.66	332.98	1
15	VGGNAGAYGAEALER	hb_sheep	1434.54	287.23 (0.59)	2
15	VGGQAGAHGAEALER	hb_pig	1422.53	284.89 (2.35)	2

[a]All values were obtained using an injected-ion mobility/time-of-flight technique. For a description, see: Hoaglund, C. S.; Valentine, S. J.; Sporleder, C. R.; Reilly, J. P.; Clemmer, D. E. *Anal. Chem.* **1998,** *70,* 2236; Henderson, S. C.; Valentine, S. J.; Counterman, A. E.; Clemmer, D. E. *Anal. Chem.* **1999,** *71,* 291.
[b]Peptide sequences correspond to fragments expected from tryptic digests as obtained from peptide_mass.pl (http://expasy.hcuge.ch/sprot/peptide-mass.html). N-terminal acetylation is indicated by the prefix Ac.
[c]All proteins were obtained from Sigma and used without further purification. Purities were typically ≥70%. Tryptic digests were performed by addition of 150 μL of a 0.2-mg/mL trypsin (Sigma, sequencing grade) solution in 0.2 M ammonium bicarbonate (EM Science) to 0.5 mL of a 20-mg/mL solution of each protein. Protein names are abbreviated as follows: albumin (albu), alcohol dehydrogenase (alcodehy), alpha-casein (acas), aldolase (aldol), apotransferrin (apotransf), beta-casein (bcas), beta-lactoglobulin (blacto), bovine serum albumin (BSA), carbonic anhydrase (canhyd), conalbumin (conalb), concanavalin A (conA), creatine phosphokinase (crephos), cytochrome c (cytc), glucose dehydrogenase (gludehy), glucose oxidase (glox), hemoglobin (hb), human serum albumin (HSA), kappa-casein (kcas), lactotransferrin (lactotrans), lysozyme (lys), myoglobin (myo), transferrin (transfr), and ubiquitin (ubiq). The sources of the proteins are separated from the protein name by an underscore, and are abbreviated as follows: aspergillus niger (aspgn), bovine (bov), chicken (chick), human (hum), rabbit (rab), sperm whale (sw), turkey egg white (tew), and yeast (yst).
[d]Molecular weights are reported as an isotopic average and were checked by comparison with http://expasy.hcuge.ch/sprot/peptide-mass.html.
[e]Cross sections correspond to the average of multiple data sets. Uncertainties are given in parentheses and correspond to one standard deviation when three or more measurements were made, or as the range when only two measurements were made.
[f]Total number of separate observations and cross section measurements for each peptide sequence.

Table 2. Cross sections for doubly protonated peptides from tryptic digests[a]

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
4	IQDK	ubiq	502.58	174.97	1
4	YTKK	HSA	538.65	188.84	1
5	GITWK	cytc_horse	603.70	172.82	1
5	IDAMR	albu_sheep	604.73	205.17 (0.08)	2
5	IEHLR	albu_pig	666.79	199.48 (1.23)	2
5	KNQDK	kcas_bov	631.69	179.66	1
5	LRLQK	crephos_rab	656.83	188.79	1
5	TLTGK	ubiq	518.62	174.97	1
6	AKIQDK	ubiq	701.83	212.99	1
6	AWSVAR	albu_sheep	688.79	188.27	1
6	DWPDAR	crephos_rab	758.80	208.84	1
6	FVIEIR	albu_pig	775.95	211.25	1
6	GGVHVK	crephos_rab	595.71	179.46	1
6	HLDDLK	hb_bov hb_sheep	739.83	212.92 (1.16)	4
6	HLDNLK	hb_pig	738.85	212.80 (0.72)	3
6	IEEIFK	crephos_rab	777.92	226.04	1
6	LIFAGK	ubiq	647.83	199.02	1
6	LNQLLR	enolase_yst	755.92	204.90	1
6	MQIFVK	ubiq	764.99	222.87	1
6	QLEDGR	ubiq	716.76	223.07	1
6	SEIAHR	albu_pig albu_sheep	711.78	193.99 (2.55)	5
6	YIPGTK	cytc_horse	677.80	189.88 (3.56)	2
7	ASEDLKK	myo_horse	789.89	238.09	1
7	ATDEQLK	albu_sheep	803.87	205.29	1
7	ATKEQLK	HSA	816.96	215.40	1
7	AVPYPQR	bcas_bov	829.96	211.55	1
7	GDVAFVK	transfr_hum	734.86	194.38	1
7	GFFYTPK	ins_bov	859.00	225.36	1
7	GVLHAVK	enolase_yst	722.89	202.50	1
7	IATAIEK	enolase_yst	744.89	196.62	1
7	IVTDLAK	albu_pig	758.92	201.59 (1.94)	2
7	LGLVGSR	albu_pig	700.84	192.82 (2.40)	3
7	LMVEMEK	crephos_rab	879.10	225.09	1
7	LVTDLTK	BSA	788.95	210.68	1
7	MIFAGIK	cytc_horse	779.00	207.03 (1.25)	3
7	NPDPWAK	transfr_hum	826.91	213.72	1
7	Ac-SIPETQK	alcodehy_yst	843.94	214.09	1
7	VDPVNFK	hb_dog hb_bov hb_hum hb_pig	817.95	203.64 (3.57)	7
7	VKAHGKK	hb_sheep	766.95	203.95	1
7	VLPVPQK	bcas_bov	779.99	215.38	1
7	VVTDLTK	albu_dog	774.92	205.05	1
7	WNMQNGK	conA	876.99	211.48 (0.37)	2
7	YLYEIAR	albu_pig	927.08	236.05 (0.46)	3
7	YLYEVAR	albu_sheep	913.05	230.94 (0.30)	3
8	ADFAEISK	albu_dog	879.97	213.59	1
8	ADFTDVTK	albu_sheep	895.97	213.10 (1.15)	2
8	ADFTEISK	albu_pig	910.00	219.23	1
8	DDHPNLPK	albu_horse	935.01	225.23	1
8	DDNPNLPR	HSA	939.99	229.08	1
8	DDTVCLAK	transfr_hum	863.99	213.30	1
8	DIPVPKPK	alcodehy_yst	893.10	231.34	1
8	DLGEENFK	HSA	951.01	238.58	1
8	DLGEQYFK	albu_pig	999.10	243.83 (0.55)	2
8	ELSDIAHR	aldol_rab	940.03	232.56 (3.58)	2
8	EYEATLEK	albu_dog	982.06	230.57 (2.63)	3
8	FGVNGSEK	conalb_chick	836.91	211.12 (1.97)	3
8	IDALNENK	blacto_bov	916.01	226.20 (1.73)	3
8	KLGLVGSR	albu_pig	829.02	227.76 (0.13)	3

(*Continued*)

Table 2. (*Continued*)

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
8	KVSEALTK	hb_rab	875.04	211.32 (0.42)	2
8	LIVTQTMK	blacto_bov	933.18	234.72	1
8	SEVAHRFK	HSA	973.11	231.57	1
8	VHLSAEEK	hb_pig	912.02	226.77 (2.70)	3
8	VHLTAEEK	hb_dog	926.04	230.04 (2.15)	2
8	VLTPDLYK	crephos_rab	948.13	241.28 (2.73)	3
8	YLGEEYVK	transfr_hum	1000.12	244.85 (4.55)	2
8	YLYEIARR	HSA	1083.26	252.64	1
9	AAVTAFWGK	hb_bov	950.11	229.12 (3.11)	3
9	AAVTGFWGK	hb_sheep	936.09	229.67	1
9	ANELLINVK	alcodehy_yst	1013.21	252.18 (2.33)	3
9	APNHAVVTR	transfr_hum	964.10	231.13 (3.65)	3
9	EAVLGLWGK	hb_pig	972.16	232.31 (0.88)	2
9	EGIPPDQQR	ubiq	1039.12	252.51	1
9	EKVLASSAR	albu_sheep	960.11	229.84 (0.42)	3
9	LHDRNTYEK	transfr_hum	1175.27	261.21 (0.90)	2
9	LHVDPENFK	hb_bov hb_dog	1098.23	252.51 (3.57)	3
9	LHVDPENFR	hb_hum hb_sheep	1126.24	255.74 (4.51)	2
9	LRVDPVNFK	hb_bov hb_hum hb_sheep	1087.29	261.06 (3.73)	5
9	MFLGFPTTK	hb_pig hb_rab	1041.28	248.32 (1.69)	5
9	MFLSFPTTK	hb_bov hb_hum hb_sheep	1071.31	254.66 (2.89)	8
9	QLLLTADDR	aldol_rab	1044.18	264.51	1
9	QSALAELVK	albu_horse	958.13	238.41	1
9	QTALVELLK	albu_sheep	1014.24	246.39 (2.50)	2
9	QTALVELVK	HSA	1000.21	247.90	1
9	SAVTALWGK	hb_hum hb_rab	932.09	229.36 (1.54)	2
9	SKGGVVGIK	aldol_rab	844.03	227.29	1
9	SLVSGLWGK	hb_dog	946.12	236.87	1
9	TDLNHENLK	crephos_rab	1083.17	258.47 (2.16)	3
9	TFQSFPTTK	hb_dog	1056.19	244.10 (1.36)	2
9	TGAPARSER	enolase_yst	944.02	239.16	1
9	TLSDYNIQK	ubiq	1081.20	258.99	1
9	VAAHAVVAR	conalb_chick	893.06	224.58 (2.42)	2
9	YFGYTGALR	conalb_chick	1047.19	250.59 (4.81)	3
10	AWGGKKENLK	aldol_rab	1130.32	265.36	1
10	DLFDPIIQDR	crephos_rab	1231.38	275.32 (2.20)	3
10	EAYKSEIAHR	albu_dog	1203.33	270.89 (0.00)	2
10	ECCEKPLLEK	HSA	1191.43	267.09	1
10	EKDIVGAVLK	alcodehy_yst	1071.29	253.92 (2.96)	3
10	FKDLGEEHFK	BSA	1249.40	282.95	1
10	FKDLGEENFK	HSA	1226.36	275.66 (0.71)	3
10	FKDLGEQYFK	albu_pig	1274.45	286.80 (0.18)	3
10	GVIFYESHGK	alcodehy_yst	1136.28	259.73 (2.27)	3
10	IGSEVYHNLK	enolase_yst	1159.31	264.63 (1.78)	3
10	KQSALAELVK	albu_horse	1086.30	252.69	1
10	KQTALVELLK	albu_pig albu_sheep BSA	1142.41	267.73 (2.24)	8
10	KQTALVELVK	HSA	1128.38	264.46	1
10	LLVVYPWTQR	hb_hum hb_pig hb_rab	1274.54	287.97 (2.67)	4
10	LVNELTEFAK	BSA	1163.34	266.65	1
10	LVNEVTEFAK	albu_horse HSA	1149.32	261.00 (1.56)	3

(*Continued*)

Table 2. (*Continued*)

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
10	SEEEYPDLSK	crephos_rab	1196.24	261.41 (1.98)	2
10	VLNSFSDGLK	hb_dog	1079.23	249.32 (0.48)	3
10	VLQSFSDGLK	hb_pig	1093.25	255.01 (2.35)	3
10	YNDLGEEHFR	albu_dog	1279.34	270.41 (0.71)	3
11	HKPHATEEQLR	albu_pig	1345.49	284.51 (1.83)	3
11	HKTDLNHENLK	crephos_rab	1348.49	295.16	1
11	HLVDEPQNLIK	albu_sheep BSA	1305.50	290.36 (0.83)	4
11	HPDYSVSLLLR	albu_horse	1299.50	287.44 (0.31)	3
11	HPEYAVSVLLR	albu_sheep	1283.50	287.75 (3.00)	2
11	HQTVPQNTGGK	transfr_hum	1166.27	256.68 (4.86)	2
11	KVLDSFSNGMK	hb_bov hb_sheep	1225.43	269.78 (4.85)	3
11	KVLQSFSDGLK	hb_pig	1221.43	290.65	1
11	LAKEYEATLEK	albu_dog	1294.47	255.81	1
11	LEQWAEEAVAR	glox_aspgn	1301.43	279.51 (3.87)	2
11	LFTGHPETLEK	myo_horse	1271.44	286.20 (4.02)	2
11	SASDLTWDNLK	transfr_hum	1249.35	277.98	1
11	SCQAQPTTMAR	kcas_bov	1193.40	267.63	1
11	TGPNLHGLFGR	cytc_horse	1168.33	267.56 (0.72)	3
11	TPEVDDEALEK	blacto_bov	1245.31	269.64 (4.98)	2
11	VGLSASTGLYK	conA	1095.27	253.96 (0.20)	2
11	VLSPADKTNIK	hb_rab	1185.39	263.79 (2.50)	2
12	AAFTECCQAADK	HSA	1257.41	269.89 (5.22)	2
12	AVMDDFAAFVEK	HSA	1342.54	275.89	1
12	DTDFKLNELRGK	apotransf_bov	1435.61	308.04 (0.23)	2
12	EETLMEYLENPK	cytc_horse	1495.67	312.71 (0.80)	3
12	EFTPPVQAAYQK	hb_hum	1378.55	289.31 (0.95)	2
12	EFTPVLQADFQK	hb_bov	1422.61	300.31 (2.37)	4
12	ENLKAAQEEYVK	aldol_rab	1421.58	308.18	1
12	FLANVSTVLTSK	hb_rab hb_sheep	1279.51	287.78 (0.33)	3
12	FLASVSTVLTSK	hb_hum	1252.48	289.60 (2.14)	2
12	GQSIDDMIPAQK	crephos_rab	1302.47	279.24 (2.24)	3
12	GTEFTVNDLQGK	conalb_chick	1308.42	284.14 (1.72)	3
12	RHPEYAVSVLLR	albu_sheep BSA	1439.69	313.44 (2.35)	4
12	SIGGEVFIDFTK	alcodehy_yst	1312.49	282.95 (4.08)	3
12	SISIVGSYVGNR	alcodehy_yst	1251.41	278.59 (1.60)	2
12	TVDMESTEVFTK	acas_bov	1386.54	294.42 (2.64)	2
12	TVLGNFSAFVAK	albu_horse	1253.47	282.48	1
12	VNQIGTLSESIK	enolase_yst	1288.47	278.95	1
12	VVAGVANALAHK	hb_dog hb_hum hb_pig hb_rab hb_sheep	1149.37	272.31 (2.65)	14
12	VYGRCELAAAMK	lys_tew	1311.59	283.85	1
13	CASIQKFGERALK	BSA	1450.73	319.03	1
13	CLQDGAGDVAFVK	lactotrans_bov	1322.51	284.08	1
13	FDKALKALPMHIR	blacto_bov	1539.91	297.70	1
13	GNPTVEVELTTEK	enolase_yst	1416.56	301.98 (2.86)	2
13	KGTEFTVNDLQGK	conalb_chick	1436.59	298.50 (3.64)	3
13	KVYGRCELAAAMK	lys_tew	1439.76	305.24	1
13	LGEYGFQNALIVR	albu_pig BSA	1479.71	313.94 (3.30)	2
13	RFYRQLLLTADDR	aldol_rab	1666.91	325.91	1
13	SFLVWVNEEDHLR	crephos_rab	1643.83	341.65 (2.24)	3
13	TCVADESHAGCEK	BSA	1349.46	283.88	1
13	VKVDEVGAEALGR	hb_sheep	1342.52	285.25 (0.03)	2
13	VKVDEVGGEALGR	hb_bov	1328.50	285.58 (1.90)	4

(*Continued*)

Table 2. (*Continued*)

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
13	VNVDEVGGEALGR	hb_dog hb_hum hb_pig	1314.43	287.12 (1.80)	7
13	VNVEEVGGEALGR	hb_rab	1328.45	286.28 (3.03)	3
13	VVGLSTLPEIYEK	alcodehy_yst	1447.70	310.43 (1.85)	2
14	APQVSTPTLVEIGR	albu_horse	1467.69	306.85 (1.58)	2
14	GILAADESTGSIAK	aldol_rab	1332.48	289.61 (3.75)	2
14	HGTVVLTALGGILK	myo_horse	1378.79	312.22 (0.49)	3
14	LDELRDEGKASSAK	HSA	1518.65	309.42	1
14	LSVEALNSLTGEFK	crephos_rab	1507.71	318.15 (2.52)	3
14	SAGWNIPIGTLLHR	conalb_chick	1534.79	315.82	1
14	TGQAPGFTYTDANK	cytc_horse	1470.60	297.61 (2.24)	3
14	TPEVDDEALEKFDK	blacto_bov	1635.75	333.91 (1.43)	2
14	VGTAHIIYNSVDKR	conA	1572.79	343.52 (1.21)	3
14	VVAGVANALAHRYH	hb_bov	1477.70	325.95	1
15	ANGTTVLVGMPAGAK	alcodehy_yst	1386.64	302.28	1
15	AVDDFLISLDGTANK	enolase_yst	1578.75	316.45	1
15	HQGLPQEVLNENLLR	acas_bov	1759.99	353.86	1
15	IGGHAGDYGGEALDR	hb_dog	1487.56	298.65 (0.89)	2
15	IGSHGGEYGAEAVER	hb_rab	1531.61	299.00 (0.81)	2
15	KAPQVSTPTLVEIGR	albu_horse	1595.87	330.12 (1.17)	2
15	KAPQVSTPTLVEISR	albu_sheep	1625.89	331.49 (1.37)	3
15	LCMGSGLNLCEPNNK	transfr_hum	1592.87	320.80 (0.05)	3
15	LGANAILGVSLAASR	enolase_yst	1412.66	310.74 (6.42)	2
15	LGHDFNPNVQAAFQK	hb_pig	1685.87	333.24 (0.91)	2
15	VEADIAGHGQEVLIR	myo_horse	1606.81	340.95 (1.00)	3
15	VGAHAGEYGAEALER	hb_hum	1529.64	306.47 (1.16)	2
15	VGGNAGAYGAEALER	hb_sheep	1434.54	288.98 (1.80)	3
15	VGGQAGAHGAEALER	hb_pig	1422.53	300.48 (2.81)	3
15	VLGIDGGEGKEELFR	alcodehy_yst	1618.81	354.90 (1.40)	3
15	YSHEEIAMATVTALR	aldol_rab	1691.93	335.42 (0.90)	3
16	GLSDGEWQQVLNVWGK	myo_horse	1816.01	352.48 (0.52)	2
16	HGGTIPIVPTAEFQDR	gludehy_bov	1737.94	335.66	1
16	HHGNEFTPVLQADFQK	hb_sheep	1868.05	353.50 (1.14)	3
16	LLGNVLVCVLAHHFGK	hb_hum	1720.12	318.03	1
16	NTDGSTDYGILQINSR	lys_tew	1753.85	330.83	1
16	SISIVGSYVGNRADTR	alcodehy_yst	1694.87	334.59 (1.85)	3
16	TITLEVEPSDTIENVK	ubiq	1787.99	360.05 (0.47)	2
16	TYFPHFDFTHGSEQIK	hb_rab	1954.14	363.50 (2.18)	3
16	TYFPHFDLSHGSAQVK	hb_bov hb_hum hb_sheep	1834.03	364.16 (1.49)	9
16	TYFPHFDLSPGSAQVK	hb_dog	1794.00	350.32 (1.56)	3
16	TYFPHFNLSHGSDQVK	hb_pig	1877.05	369.15	1
16	VLAAFSEGLSHLDNLK	hb_rab	1713.96	375.98 (3.62)	3
16	VLGAFSDGLAHLDNLK	hb_hum	1669.91	365.62 (1.48)	2
16	WLTGPQLADLYHSLMK	enolase_yst	1873.21	372.93 (1.77)	2
16	YLEFISDAIIHVLHSK	myo_horse	1885.20	404.45 (0.04)	2
17	AAQDSFAAGWGVMVSHR	enolase_yst	1789.99	371.96 (3.71)	2
17	GHHEAELKPLAQSHATK	myo_horse	1854.06	361.52	1
17	HPYFYGPELLFHAEEYK	albu_horse	2140.39	397.72 (1.79)	3
17	SGETEDTFIADLVVGLR	enolase_yst	1822.01	354.32 (5.38)	2
17	STHETNALHFMFNQFSK	conA	2039.26	384.39 (3.28)	3
17	TAGIQIVADDLTVTNPK	enolase_yst	1755.99	357.02	1
18	AKIIAEGANGPTTPQADK	gludehy_bov	1781.99	341.97 (4.73)	2
18	KGHHEAELKPLAQSHATK	myo_horse	1982.24	374.41	1
18	MLTAEEKAAVTAFWGKVK	hb_bov	1980.36	379.02 (4.54)	2
18	MLTAEEKAAVTGFWGKVK	hb_sheep	1966.33	357.38	1
18	SIVPSGASTGVHEALEMR	enolase_yst	1841.08	394.66 (1.36)	3
19	AHELLNTKLEQWAEEAVAR	glox_aspgn	2208.47	390.65	1
19	DMPIQAFLLYQEPVLGPVR	bcas_bov	2186.60	400.67	1
19	EPMIGVNQELAYFYPELFR	acas_bov	2316.66	405.63	1

(*Continued*)

Table 2. (*Continued*)

Number of residues	Assigned sequence[b]	Peptide source[c]	MW[d]	Cross section (Å^2)[e]	Number of measurements[f]
19	FFDSFGDLSTPDAVMSNAK	hb_dog	2049.25	385.30 (2.98)	2
19	FFEHFGDLSNADAVMNNPK	hb_sheep	2153.36	393.23 (0.93)	3
19	FFESFGDLSNADAVMGNPK	hb_pig	2046.25	377.85 (0.96)	2
19	FFESFGDLSSANAVMNNPK	hb_rab	2075.29	381.63 (5.56)	3
19	FFESFGDLSTADAVMNNPK	hb_bov	2090.30	388.79 (1.76)	4
19	FFESFGDLSTPDAVMGNPK	hb_hum	2059.29	378.16	1
19	TVDYIIAGGGLTGLTTAAR	glox_aspgn	1850.11	366.03 (2.52)	2
20	IGEHTPSALAIMENANVLAR	aldol_rab	2107.42	399.25	2
20	VYVEELKPTPEGDLEILLQK	blacto_bov	2313.68	426.75 (2.14)	2
21	AAANFFSASCVPCADQSSFPK	apotransf_bov	2148.40	389.55	1
21	GTGGVDTAAVGSVFDISNADR	crephos_rab	2009.12	363.68 (1.95)	2
22	AVEHLDDLPGALSELSDLHAHK	hb_bov	2367.61	437.19 (2.50)	4
22	AVGHLDDLPGALSALSDLHAHK	hb_pig	2237.51	418.64 (0.31)	2
22	AVGHLDDLPGALSTLSDLHAHK	hb_rab	2267.53	422.37 (2.90)	2
22	AVGHLDDLPGTLSDLSDLHAHK	hb_sheep	2311.54	438.43 (3.25)	3
22	RGTGGVDTAAVGSVFDISNADR	crephos_rab	2165.31	387.93 (4.68)	2
22	TVGGKEDVIWELLNHAQEHFGK	apotransf_bov	2507.80	432.00	1
22	YTPSGQAGAAASESLFISNHAY	aldol_rab	2242.39	404.01	1
23	GVVPLAGTNGETTTQGLDGLSER	aldol_rab	2272.46	396.30 (3.46)	3
23	WSGFSGGAIECETAENTEECIAK	apotransf_bov	2432.62	412.48	1
24	ATDGGAHGVINVSVSEAAIEASTR	alcodehy_yst	2312.49	415.18 (1.48)	3
24	NLCNIPCSALLSSDITASVNCAKK	lys_tew	2465.89	412.61	1

[a]All values were obtained using an injected-ion mobility/time-of-flight technique. For a description, see: Hoaglund, C. S.; Valentine, S. J.; Sporleder, C. R.; Reilly, J. P.; Clemmer, D. E. *Anal. Chem.* **1998,** *70,* 2236; Henderson, S. C.; Valentine, S. J.; Counterman, A. E.; Clemmer, D. E. *Anal. Chem.* **1999,** *71,* 291.
[b]Peptide sequences correspond to fragments expected from tryptic digests as obtained from peptide-mass.pl (http://expasy.hcuge.ch/sprot/peptide-mass.html). N-terminal acetylation is indicated by the prefix Ac.
[c]All proteins were obtained from Sigma and used without further purification. Purities were typically ≥70%. Tryptic digests were performed by addition of 150 μL of a 0.2-mg/mL trypsin (Sigma, sequencing grade) solution in 0.2 M ammonium bicarbonate (EM Science) to 0.5 mL of a 20-mg/mL solution of each protein. Protein names are abbreviated as follows: albumin (albu), alcohol dehydrogenase (alcodehy), alpha-casein (acas), aldolase (aldol), apotransferrin (apotransf), beta-casein (bcas), beta-lactoglobulin (blacto), bovine serum albumin (BSA), carbonic anhydrase (canhyd), conalbumin (conalb), concanavalin A (conA), creatine phosphokinase (crephos), cytochrome c (cytc), glucose dehydrogenase (gludehy), glucose oxidase (glox), hemoglobin (hb), human serum albumin (HSA), kappa-casein (kcas), lactotransferrin (lactotrans), lysozyme (lys), myoglobin (myo), transferrin (transfr), and ubiquitin (ubiq). The sources of the proteins are separated from the protein name by an underscore, and are abbreviated as follows: aspergillus niger (aspgn), bovine (bov), chicken (chick), human (hum), rabbit (rab), sperm whale (sw), turkey egg white (tew), and yeast (yst).
[d]Molecular weights are reported as an isotopic average and were checked by comparison with http://expasy.hcuge.ch/sprot/peptide-mass.html.
[e]Cross sections correspond to the average of multiple data sets. Uncertainties are given in parentheses and correspond to one standard deviation when three or more measurements were made, or as the range when only two measurements were made.
[f]Total number of separate observations and cross section measurements for each peptide sequence.

bonate (EM Science) with 0.5 mL of a 20 mg/mL solution of each protein. The digest solution is incubated for 20 h at 37 °C and is filtered to remove trypsin using a microconcentrator (microcon 10, Amicon, Inc.). The remaining fragment peptides are lyophilized. Tryptic digest solutions containing 0.5 mg/mL of the lyophilized product in 49:49:2 (% volume) water:acetonitrile: acetic acid are electrosprayed to produce positively charged (protonated) peptide ions. Solution flow rates of 0.05–0.10 mL/h are utilized in these studies. Solutions are electrosprayed at atmospheric pressure with a needle that is biased at +3200 V relative to the entrance of the desolvation region.

Ion Trap/Drift Tube Interface

Ions that exit the source are focused into an ion trap (R. M. Jordan, model C-1251) [3, 5, 6], where they are influenced by a 1.1-MHz rf field (2600 V peak to peak) and collisions with residual gas molecules (10^{-4} to 10^{-3} torr). Here, the continuous beam is accumulated in a small volume in the center of the trap that is aligned with the entrance aperture of the drift tube. Ions are ejected from the trap into the drift tube by turning off the rf voltage and applying a 0.5 μs pulse (−100 to −200 V) to the exit endcap. Injection voltages of 70 to 110 V were used. Cross section measurements over a range of trapping potentials (from 1500 to 2600 V peak to peak) and times (50 to 150 ms) are identical. Subsequent measurements of many cross sections have been made using a high-pressure ion mobility instrument that is suited for observation of fragile structure [28]; all of the high pressure measurements agree with the average values reported here to within the reported uncertainty. We interpret the reproducibility of these values over the range of conditions as an indication that the values correspond to favorable gas-phase ion conformations rather than metastable states that are preserved during ESI, as has been observed for larger protein systems [23c, 29–31].

Mass Spectrometer

Ions that exit the drift tube are focussed into the shape of a ribbon using an einzel/dc-quadrupole lens system. A four-grid system is used to extract ions into an orthogonal TOF mass spectrometer. The field-free flight tube is 17.5 cm long. Ions are detected by a pair of microchannel plates mounted directly to the back of the flight tube. High-frequency pulses (3 μs, +2220 V) are supplied by a high-voltage pulser (Directed Energy Inc., model GRX-3.0K-H). Voltages in the time-of-flight instrument were determined from a space- and velocity-focusing algorithm [32]. Typically the resolving power ($m/\Delta m$) for singly-charged peptides is ~300.

Acquisition of Nested Drift(Flight) Time Data

Because flight times in the evacuated mass spectrometer are much shorter than the drift times (in the high-pressure drift tube) hundreds of mass spectra can be recorded for each pulse of ions that is injected into the drift tube. This is referred to as a "nested" drift(flight) time measurement [5]. An initial pulse is used to inject ions into the drift tube and simultaneously activate a programmable delay generator (PDG, Lecroy 4222). The PDG triggers the high-voltage TOF pulser at specified delay times and a time-to-digital converter (TDC, Lecroy 4208) is used to record ion flight times. The initial injection (drift) pulse, PDG pulse sequence, TDC, and high-voltage TOF pulser are synchronized and under computer control. Contour plots of the data were created using the MATLAB software [33].

Experimental Collision Cross Sections

The arrival time of an ion at the detector is the sum of its drift time, flight time, and the time required to travel through interface regions of the instrument. To determine drift times it is necessary to account for the flight time and a small correction associated with transport of the ions from the exit of the drift tube to the entrance of the time-of-flight region. Correction times are small (80 to 220 μs) compared with drift times (2 to 10 ms) in these experiments. Experimental collision cross sections (or collision integrals) [7] are obtained directly from the drift time distributions using the relation [4e, 4f]

$$\Omega = \frac{(18\pi)^{1/2}}{16}\frac{ze}{(k_bT)^{1/2}}\left[\frac{1}{m_I}+\frac{1}{m_B}\right]^{1/2} \times \frac{t_DE_D}{L}\frac{760}{P}\frac{T}{273.2}\frac{1}{N} \quad (2)$$

where t_D, E_D, L, P, and T correspond to the average drift time, the drift electric field strength, the drift tube length, buffer gas pressure (in torr) and temperature, respectively. The other terms ze, N, k_b, m_I, and m_B are the charge of the ion, the neutral number density, Boltzmann's constant, and the masses of the ion and buffer gas, respectively. Reproducibility of recorded cross sections is high as noted by the small uncertainties in Tables 1 and 2.

Drift times are determined from peak maxima in the three-dimensional spectra, which are usually located in the center of the peak. Most peaks can be accurately represented by calculated peak shapes for single conformations obtained from the transport equation [4e, 34]. Although this is consistent with each peak corresponding to single conformations, it is more likely that related structures, which interconvert rapidly (compared with experimental timescales) or have indistinguishable drift times, are present. Molecular dynamics simulations for peptide ions at 300 K suggest that peptides displaying compact conformations are comprised of rapidly interconverting (compared with the millisecond experimental timescales) structures [35]. We currently believe that in most cases measured cross sections correspond to a distribution of related states centered about the favored conformers. Interconversion of conformations over longer timescales would lead to peak skewing. This is observed for some peptides, and suggests that these peptides have conformations that interconvert on a timescale similar to (or longer than) the experimental timescale. For the peptides studied here, ion mobility data display only a single resolved peak.

Results and Discussion

Nested Measurements

Figure 1 shows a two-dimensional contour plot of the nested drift(flight) times for a peptide mixture obtained by electrospraying a tryptic digest of yeast enolase (mol. wt. = 46.7 kDa). A time-of-flight mass spectrum obtained by compression of the drift axis is shown on the left. Typically, peaks in the two-dimensional spectra for tryptic digests fall into singly- ($[M + H]^+$) and doubly- ($[M + 2H]^{2+}$) charged families [5b]. The observation of singly- and doubly-charged ions is expected because trypsin cleaves peptide bonds on the C-terminal side of basic residues [36]; this cleavage pattern results in the generation of two protonatable sites (the N-terminus and the C-terminal basic residue). Generally, for peptides with similar m/z ratios, the $[M + 2H]^{2+}$ ions have higher mobilities (shorter drift times) than $[M + H]^+$ ions because they experience twice the drift force. Solid lines are used to distinguish the $[M + H]^+$ and $[M + 2H]^{2+}$ families in Figure 1. The combined mobility/TOF approach allows rapid determination of the charge state and reduces spectral congestion produced by the multiple-charging phenomenon of ESI. These advantages facilitate peak identification.

Complete lists of singly and doubly charged peptides identified in the digest spectra of the 34 proteins are given in Tables 1 and 2, respectively. Only those peptide sequences that have been unambiguously assigned have been included in this database. Average

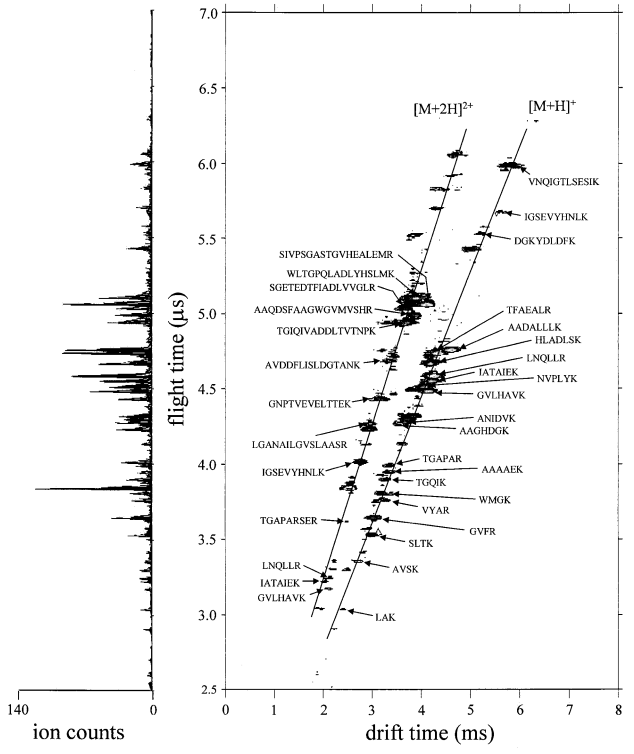

Figure 1. A contour plot of "nested" drift time (bottom axis) and flight time (left axis) data for the electrosprayed mixture of peptide ions obtained from a tryptic digest of yeast enolase. The drift time axis has been scaled to a He pressure of 2.000 torr. Two distinct mobility families are observed; the higher and lower mobility families have been determined to contain $[M + 2H]^{2+}$ and $[M + H]^{+}$ peptides, respectively. Labels have been assigned to peaks that correspond with expected tryptic digest fragments based on comparison of measured *m/z* ratios (from ion flight times) to calculated molecular weights. In all cases the maximum peak height is used to obtain a drift time from which collision cross sections are derived. The time-of-flight distribution shown on the left is obtained by compressing the drift time axis.

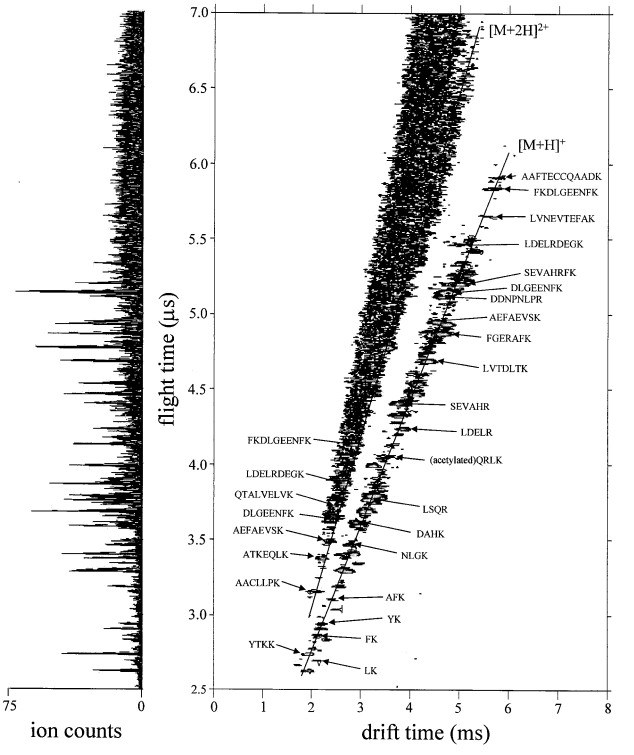

Figure 2. Contour plot of drift and flight times for a mixture of peptides obtained from tryptic digestion of human serum albumin. Labels are shown for peaks that correspond to expected tryptic digest fragments. Unresolved features arriving at longer flight times and shorter drift times are believed to correspond to multiply charged, disulfide-linked peptide fragments or multiply charged peptide fragments resulting from incomplete tryptic digestion. The time-of-flight distribution shown on the left is obtained by compressing the drift time axis.

cross section, peptide molecular weight, protein digest source, and the number of separate cross section measurements for 420 singly charged peptides (containing 2 to 15 residues) and 240 doubly charged peptides (containing 4 to 24 residues) are presented in these tables. Most of the database peptides, 290 (+1) and 164 (+2), contain a C-terminal lysine residue. Of the remaining peptides, 114 (+1) and 74 (+2) contain a C-terminal arginine residue. A significant percentage of these, ~12% (+1) and ~25% (+2), results from incomplete tryptic digestion. Peptides containing cysteine residues comprise only ~2% (+1) and 6% (+2) of the data.

For some tryptic digest systems, two-dimensional spectra are extremely complicated. Figure 2 shows data obtained upon ESI of tryptic fragments of human serum albumin (mol. wt. = 69.4 kDa). The $[M + H]^{+}$ and $[M + 2H]^{2+}$ families are identified with solid lines; however, at longer flight times (higher *m/z*), unresolved features with mobilities that are higher than those expected for $[M + 2H]^{2+}$ family are observed. The combined mobility-MS measurements suggest that these features are associated with many unresolved +3 and higher charge state ions. Higher charge state ions probably correspond to sequences with more than two strongly basic (Arg, His, Lys, or N-terminal) residues. Human serum albumin contains 17 disulfide linkages; high-mobility, high *m/z* ions may correspond to disulfide-linked peptides or fragments that arise from incomplete tryptic digestion. Partially solvated peptides and aggregates may also contribute to the many unresolved peaks in this region; however, the experimental conditions that were used favor formation of the naked peptide monomer ions [9].

Cross Sections for Singly and Doubly Protonated Peptide Ions

Figure 3a, b shows cross sections as a function of molecular weight for the singly and doubly charged database peptides, respectively. A strong correlation of increasing cross section with increasing molecular weight is observed; at similar molecular weights, cross sections for different peptides vary by as much as ~10%–15%. In some regions, variations as large as 25%

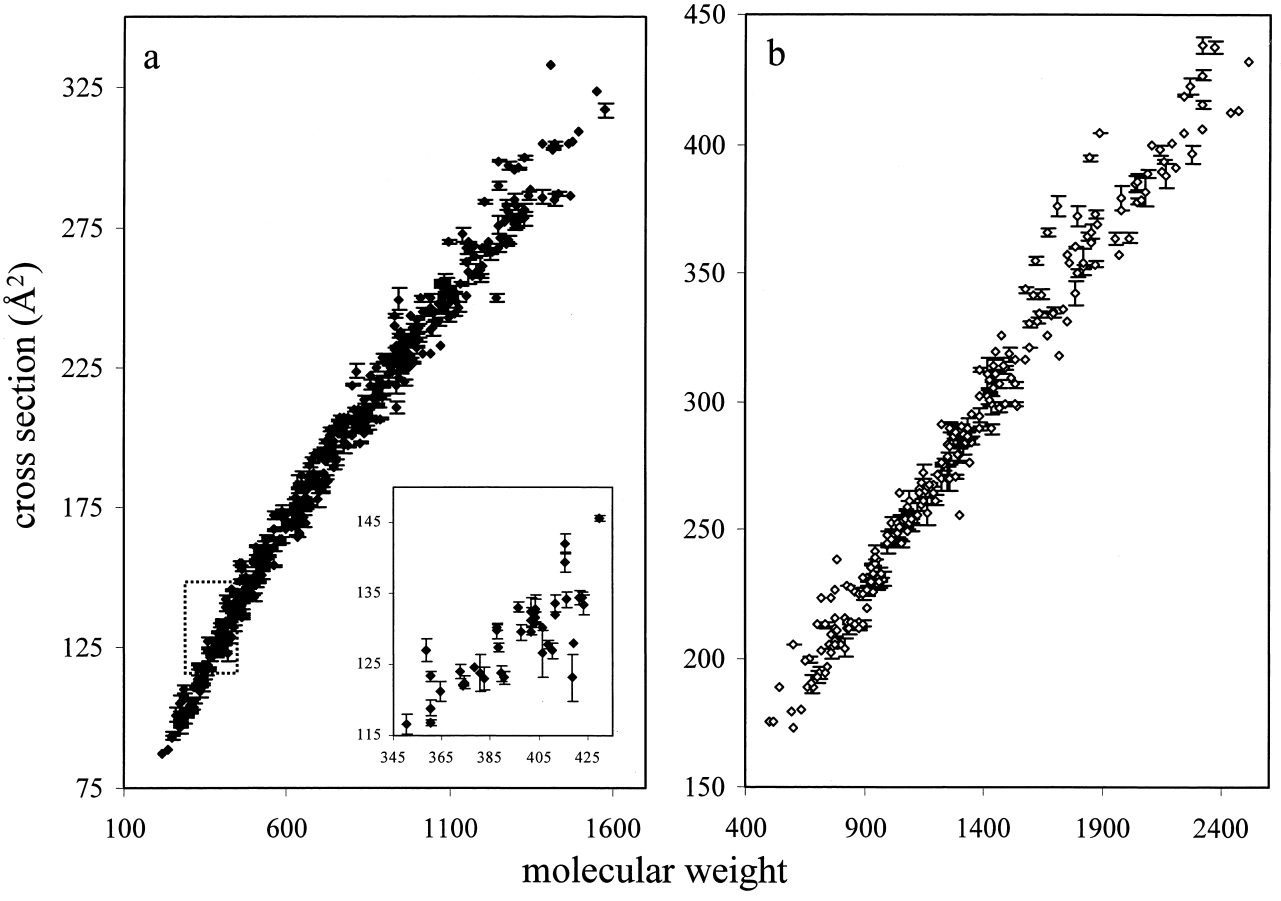

Figure 3. (**a**) Cross sections for 420 $[M + H]^+$ peptides (solid diamonds) as a function of molecular weight. Uncertainties correspond to one standard deviation (when three or more measurements have been made) or to a range (when only two values are available). The inset shows variations in cross sections for $[M + H]^+$ peptides over a smaller molecular weight range (defined by the dashed-line box). (**b**) Cross section measurements for 240 $[M + 2H]^{2+}$ peptides (open diamonds) as a function of molecular weight.

are observed. In order to compare sizes of different peptides it is useful to normalize cross section data with respect to values obtained for a series of singly protonated, polyalanine peptides containing 3 to 19 residues. This removes the variations in cross section that occur because of differences in mass. Polyalanine peptides have roughly spherical (globular) conformations resulting from the self-solvation of the protonated N-terminal amino group through interactions with electronegative backbone carbonyl groups along the peptide chain [10b, 35]. We have previously defined reduced cross section as the cross section of the database peptides divided by the cross section of polyalanine at the same molecular weight (provided by a polynomial fit to the polyalanine data) [11]. Reduced cross sections for $[Xxx_nLys + H]^+$ (n = 1–14) peptides are shown in Figure 4. Values range from 0.96 to 1.18 (a range of ~25%) over the molecular weight range studied. Reduced cross sections for the majority (~96%) of the database peptides are larger than 1.00, indicating that these peptides are larger than polyalanine (for a given molecular weight).

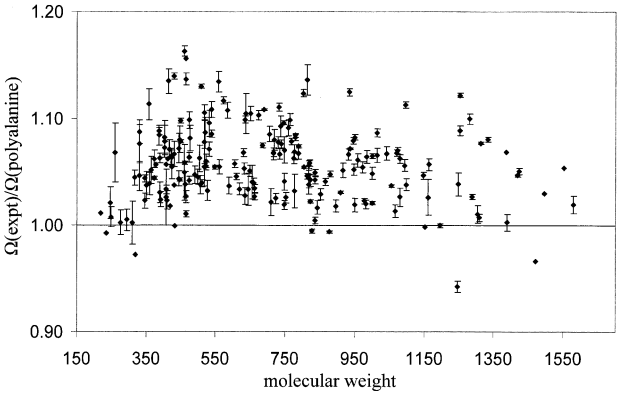

Figure 4. Reduced cross sections and uncertainties for the $[(Xxx)_nLys + H]^+$ (Xxx represents any naturally occurring amino acid and n = 1–14) database peptides as a function of molecular weight. The solid line corresponds to reduced cross sections for peptides having the same cross section as the polyalanine data.

Table 3. Intrinsic size parameters for individual amino acid residues[a]

	Parameterization set			
Residue	5-10-Lys[b]	9-15-Lys[c]	3-5-Lys[d]	5-10-Arg[e]
Gly	0.99 (0.03)	0.95 (0.06)	1.11 (0.04)	0.99 (0.03)
Ala	1.08 (0.01)	1.05 (0.05)	1.11 (0.03)	1.04 (0.03)
Val	1.08 (0.02)	1.24 (0.08)	1.20 (0.08)	1.07 (0.05)
Ile	1.13 (0.04)	1.30 (0.15)	1.20 (0.07)	1.12 (0.05)
Leu	1.19 (0.02)	1.24 (0.05)	1.29 (0.04)	1.16 (0.02)
Met	1.04 (0.08)	1.04 (0.18)	1.07 (0.27)	1.07 (0.16)
Phe	1.05 (0.03)	1.29 (0.09)	1.14 (0.07)	0.97 (0.05)
Tyr	0.99 (0.07)	1.07 (0.36)	1.14 (0.12)	0.91 (0.03)
Trp	0.95 (0.12)	1.25 (0.35)	1.11 (0.16)	0.97 (0.41)
Ser	0.99 (0.04)	1.09 (0.06)	1.13 (0.08)	0.97 (0.04)
Thr	1.00 (0.02)	1.00 (0.04)	1.08 (0.05)	0.96 (0.07)
Asn	0.94 (0.04)	0.87 (0.11)	1.04 (0.06)	0.89 (0.08)
Asp	0.89 (0.03)	0.83 (0.07)	1.04 (0.09)	0.86 (0.04)
Gln	0.98 (0.07)	0.92 (0.13)	1.11 (0.20)	0.86 (0.05)
Glu	0.91 (0.02)	0.98 (0.06)	1.14 (0.07)	1.00 (0.05)
Pro	1.00 (0.05)	1.01 (0.14)	1.09 (0.10)	1.01 (0.06)
Lys	1.23 (0.04)	0.76 (0.27)	0.87 (0.04)	
Arg				1.27 (0.07)

[a]The amino acid size parameters were derived by solving a system of equations that relate the occurrence frequency of each amino acid and unknown size parameters to a reduced cross section of each peptide (see text). Uncertainties, given parenthetically, correspond to one standard deviation about the mean.
[b]Amino acid size parameters derived from 113 peptides having the form $[Xxx_nLys + H]^+$. Here, Xxx is any naturally occurring amino acid except Cys, His, Lys, or Arg and $n = 4$ to 9.
[c]Amino acid size parameters derived from 43 peptides having the form $[Xxx_nLys + H]^+$ ($n = 8$ to 14).
[d]Amino acid size parameters derived from 81 peptides having the form $[Xxx_nLys + H]^+$ ($n = 2$ to 4).
[e]Amino acid size parameters derived from 41 peptides having the form $[Xxx_nArg + H]^+$ ($n = 4$ to 9).

High-density regions of points in this plot may indicate preferred structural types; however, no assignable trends have been determined.

Intrinsic Size Parameters

It is not feasible to measure cross sections for all sequences of small peptides. Consider, for example, five-residue tryptic fragments containing a C-terminal arginine or lysine residue; there will be 320,000 [$2(20)^4$] different sequences. A method for accurate prediction of cross sections would be valuable. From a series of preliminary predictions of cross sections that were carried out while the database was being compiled, we determined that accurate predictions for peptide sequences that have not been measured require the development of different parameterization sets from homologous peptides. A particularly important criterion for a successful prediction based on sequence parameterization appears to be peptide length. In this section we show several systems and examine their ability to accurately predict cross sections for other peptides in the database. A prediction is considered to be accurate if it falls within 2% of the experimental values, because this is often the level of accuracy reported for ion mobility measurements of cross section [4e].

Table 3 lists four parameterization sets that have been derived from the database. The first set was obtained from the 113 $[Xxx_nLys + H]^+$ ($n = 4$–9) peptides mentioned above [11] and is referred to as the 5-10-Lys parameter set. It is possible to calculate cross sections that are within 2% of experimental values for 90% of the 113 peptides that were used to derive the parameters [11]. Table 3 also gives parameters that were derived from a series of 43 $[Xxx_nLys + H]^+$ peptides, where $n = 8$–14 (the 9-15-Lys parameterization set). Using these parameters, calculated cross sections for 33 of 43 peptides (~77%) are within 2% of experiment. Another set of parameters was derived from a series of 81 $[Xxx_nLys + H]^+$ peptides, where $n = 2$–4. These are referred to as the 3-5-Lys set. Fifty-eight of 81 peptides (~72%) are retrodicted to within 2% for these parameters. Finally, parameters were derived from a series of 41 $[Xxx_nArg + H]^+$ peptides, where $n = 4$–9 (the 5-10-Arg parameters); 38 of the 41 peptides (~93%) are within 2% of experiment. The 5-10-Arg set is the only parameterization set obtained from arginine-terminated peptides because of the limited number of these peptides. No parameter sets are presented for peptides from incomplete digestion (where multiple basic residues are present) because of small numbers of related sequences.

One means of assessing the usefulness of intrinsic size parameters for prediction of cross sections can be obtained by comparing the retrodicted cross sections with values determined by a polynomial fit of the cross sections as a function of molecular weights. Figure 5a shows the ratios of calculated to experimental cross sections for singly-charged peptides over an extended molecular weight range. In this figure, the cross sections for peptides containing 2–4, 5–10, and 11–15 amino acid residues were calculated using the 3-5-Lys, 5-10-Lys,

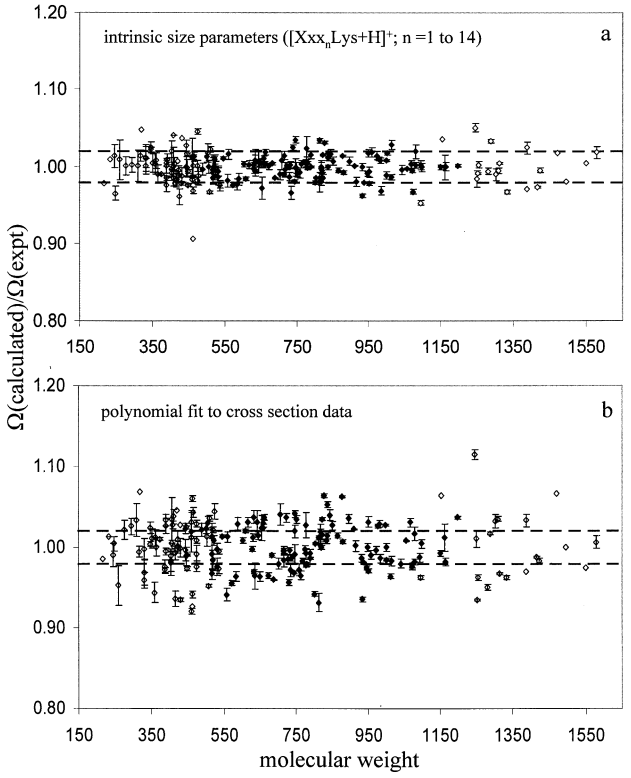

Figure 5. (**a**) Plot of calculated cross sections divided by experimental values as a function of molecular weight for $[(Xxx)_nLys + H]^+$ database peptides. Here Xxx refers to any naturally occurring amino acid except Cys, His, Lys, or Arg and n = 1–3 (open diamonds), 4–9 (solid diamonds), and 10–14 (open diamonds) amino acid residues. Cross sections for peptides containing 2–4, 5–10, and 11–15 amino acid residues were calculated using intrinsic size parameters obtained from the 3-5-Lys, 5-10-Lys, and 9-15-Lys sets (Table 3), respectively. (**b**) Plot of calculated cross sections divided by experimental cross sections as a function of the molecular weight for the same peptides shown in (**a**). The calculated cross sections were obtained from polynomial fits to the cross section data over the different molecular weight ranges. In both plots, dotted lines indicate the region where predicted values agree to within 2% of experimental values.

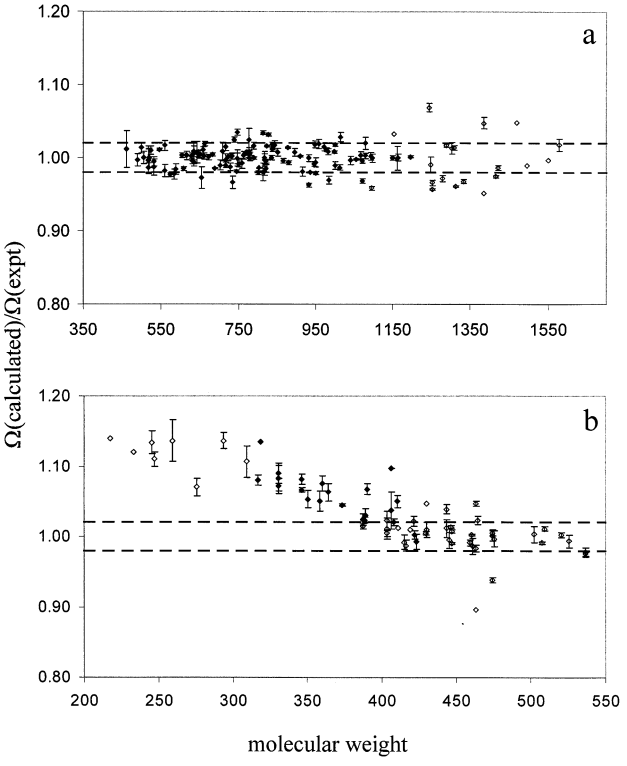

Figure 6. (**a**) Plot of calculated cross sections divided by experimental cross sections as a function of molecular weight for $[(Xxx)_nLys + H]^+$ peptide ions, where n = 4–9 (solid diamonds) and n = 10–14 (open diamonds). (**b**) Plot of predicted cross sections divided by experimental values as a function of molecular weight for $[(Xxx)_nLys + H]^+$ database peptides, where n = 1 (open diamonds), n = 2 (solid diamonds), and n = 3 (open diamonds). Size parameters from the 5-10-Lys parameterization set (Table 3) were used for calculations. In both plots, dotted lines indicate the region where predicted cross sections are found to be within 2% of experimental values.

and 9-15-Lys parameterization sets, respectively (Table 3). Figure 5b shows a plot of the cross sections determined from a polynomial fit to the experimental values divided by the individual cross sections; peptides of different lengths (2–4, 5–10, and 11–15 residues) were fit separately. The percentages of predicted cross sections within 2% of the experiment are ~70%, 90%, and 60% (for the parameterization models) and ~40%, 50%, and 30% (for the polynomial fits) for peptides containing 2–4, 5–10, and 11–15 amino acid residues, respectively. These results show that the parameterization method significantly improves the accuracy of prediction for a large fraction of the peptides.

It is instructive to assess the limits over which predictions remain accurate. Figure 6a shows a plot of cross sections that were calculated using the 5-10-Lys parameters divided by experimental cross sections for peptides containing 5–15 amino acid residues. The percentage of successful predictions for 11–15 residue peptides (which are not length homologues of the peptides used to determine parameters) is much smaller; ~40% are within 2% of experimental values. Figure 6b shows a ratio of calculated (5-10-Lys parameters) to experimental cross sections for peptides containing 2–4 amino acid residues; ~70% of the four residue peptides are predicted to within 2% of the corresponding experimental values. Only ~15% of the three residue peptides and none of the two residue peptides have calculated cross sections within 2% of the respective experimental values. The 5-10-Lys parameterization set overestimates the sizes of smaller peptides. Similar studies tested the efficacy of the 9-15-Lys parameters in predictions of $[Xxx_nLys + H]^+$ (n = 1–7) peptides and found that these parameters underestimate the experimental cross sections; only ~25% of calculated values were within 2% of the experimental values.

Bona Fide Cross Section Predictions

Extensive work involving sequence-to-structure prediction methods in condensed phase have been carried out

Table 4. Experimental and predicted cross sections for expected peptides from tryptic digests of sperm whale myoglobin and yeast enolase

Number of residues	Peptide sequence[a]	peptide source	MW[b]	Experimental cross section (Å^2)[c]	Predicted cross section (Å^2)[d]
5	DIAAK	myo_sw	516.60	155.37 (2.98)	156.63
5	TGQIK	enolase_yst	545.64	157.92 (0.59)	158.60
6	AAAAEK	enolase_yst	559.63	159.57 (1.96)	164.58
6	ANIDVK	enolase_yst	658.76	178.62 (0.20)	177.84
6	ASEDLK	myo_sw	661.72	175.16 (0.58)	177.37
6	NPNSDK	enolase_yst	673.69	Not observed	171.03
6	TEAEMK	myo_sw	707.80	181.27 (1.72)	186.57
6	NVPLYK	enolase_yst	732.89	195.63 (0.22)	191.80
6	YDLDFK	enolase_yst	799.89	198.72 (1.11)	201.58
7	IATAIEK	enolase_yst	744.89	203.73 (0.67)	200.66
8	AADALLLK	enolase_yst	814.00	221.35 (0.24)	218.01

[a]Peptide sequences correspond to fragments expected from tryptic digests as obtained from peptide-mass.pl (http://expasy.hcuge.ch/sprot/peptide-mass.html). For cross section calculations, we have only included peptides of the general form $[Xxx_nLys + H]^+$ (Xxx corresponds to any naturally occurring amino acid and $n = 4$ to 14).
[b]Molecular weights are reported as isotopic averages.
[c]Cross sections correspond to an average of three data sets. Uncertainties, given parenthetically, correspond to one standard deviation. One sequence was not observed experimentally.
[d]Reduced cross sections for each peptide were obtained using the 5-10-Lys parameterization set (Table 3) and eq 1. Calculated collision cross sections were obtained by multiplying the reduced cross sections with a polynomial fit to the polyalanine cross sections as a function of molecular weight ($-2.724 \times 10^{-5}\, x^2 + 2.141 \times 10^{-1}\, x + 4.080 \times 10^{1}$, where x = peptide molecular weight).

[14, 26, 27]; it is often noted that the true test of methods for predicting structure from sequence is to make the prediction prior to the experimental measurement [14]. We have done this for 11 lysine-terminated peptides containing 5–10 residues that are expected from tryptic digestion of sperm whale myoglobin and yeast enolase and have listed the predictions in Table 4. Only 10 of the 11 expected peptides were observed. Experimental cross sections are also provided in Table 4 [37]. Considering the simplicity of this model, the agreement between these predictions and experiment is remarkable. Eight predictions of the 10 peptides (80%) observed experimentally are within 2% of experiment. All calculated cross sections are within 3.2% of the experimental values. This is substantially better than predictions based on a polynomial fit to cross section data, where five are within 2% of experimental values, and five deviate by more than 3.5%.

Variations in Size Parameters

It is clear that the chemical and physical properties of amino acid residues influence ion cross section [11]. Peptide length, composition, and sequence should also affect the types of gas-phase structures formed. The parameterization sets described above were chosen to represent peptides differing in composition and length. The parameter sets show large changes in the size parameter for lysine: 1.23 ± 0.04 (5-10-Lys), 0.76 ± 0.27 (9-15-Lys), 0.87 ± 0.04 (3-5-Lys). When parameters for other residues are examined, it appears that other residues (within a parameter set) show smaller changes that often appear to oppose the Lys contribution. For example, comparison of the 5-10-Lys and 3-5-Lys parameter sets shows that in the latter set, the ~30% decrease in the Lys parameter is accompanied by small increases (3%–15%) in the other residues. Because the Lys residue occurs in every sequence, it can mediate changes that occur in all other parameters. In the present system, the variation in the lysine size parameter has a pronounced influence on the ability of the different parameter sets to predict accurate cross sections, especially for smaller peptides. This is expected because the fraction of the prediction based on lysine increases with decreasing number of peptide residues. Additionally, one sees that if there are gross changes in structure that occur over a small size range, the analysis that we have presented will find only the best average parameters; the interesting structural properties of outliers will be missed with this analysis. Clearly a larger database where groups of deviating peptides are observed will be useful for discerning structural properties.

Finally, we note that, overall, the trends in intrinsic size parameters for different regions of the database are similar to results that we obtained in our initial report [11]. Table 3 shows that polar residues are smaller than nonpolar residues for all of the parameterization sets that were derived [11]. This is because the more attractive long-range interactions between polar groups lead to more compact conformations. Additionally, for several parameterization sets, the contributions to size increase with increasing side chain length. This result can be understood in terms of the larger number of conformational degrees of freedom that are associated with the longer sidechains.

Summary and Conclusions

The recent development of an ESI/ion trap/ion-mobility/time-of-flight mass spectrometry instrument [5a, 6] has enabled measurements of cross sections for large

numbers of peptide ions contained in mixtures. Data from mixtures of peptides obtained upon digestion of 34 common proteins have been compiled into a structural database containing 660 entries (420 singly-charged species and 240 doubly charged species). These data have been used to determine intrinsic amino acid size parameters which can be used to calculate cross sections based solely on peptide composition for peptides that have not been measured. The usefulness of intrinsic size parameters for prediction of cross section has been assessed by comparing predictions and experimental values for different types of peptides using different parameterization sets. Parameters for some residues vary substantially for different peptide lengths. Bona fide cross section predictions for a related series of peptides are found to be reasonably accurate; eight of 10 cross section predictions are within 2% of experimental values and all 10 are within 3.2%. In the absence of any experimental measurements, the tables of amino acid size parameters for peptides containing 3–5, 5–10, and 9–15 residues should be valuable for prediction of cross section. At this early stage, predicted values should be used cautiously. These parameters will be most useful for sequences that are similar to those found in the proteins used in this study. Clearly, sequences with specifically designed structures such as helices or sheets may not be represented by this composition-based model; in these systems, sequence is also an important structural factor. Initial molecular modeling studies of 103 of the $[Xxx_nLys + H]^+$ ($n = 4-8$) peptides show that most of these peptides have globular conformations where the charge is solvated by electronegative groups [13]. A useful complement of parameter-based predictions of cross sections would be calculations of cross sections [8, 38, 39] for trial conformations generated by molecular modeling [8a, 10, 40]; ultimately, parameter-based cross section predictions may be a useful constraint in molecular modeling strategies for calculating gas-phase ion structure. Studies of the general properties of the doubly charged ions are currently underway in our laboratory.

Acknowledgments

We are grateful for the contributions of Cherokee S. Hoaglund-Hyzer and Penelope A. Lewis. We gratefully acknowledge financial support from the National Institute of Health (grant no. 1R01GM55647-01).

References

1. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989,** *246*, 64.
2. For discussions of ion traps see, for example: Paul, W.; Steinwedel, H. U.S. Patent 2939952, 1960. March, R.; Huges, R. *Quadrupole Storage Mass Spectrometry*; Wiley: New York, 1989. Cooks, R. G.; Kaiser, R. E., Jr. *Acc. Chem. Res.* **1990,** *23*, 213. Glish, G. L.; McLuckey, S. A. *Int. J. Mass Spectrom. Ion Processes* **1991,** *106*, 1. March, R. E. *J. Mass Spectrom.* **1997,** *32*, 351.
3. Hoaglund, C. S.; Valentine, S. J.; Clemmer, D. E. *Anal. Chem.* **1997,** *69*, 4156.
4. For reviews of ion mobility studies see: (a) Hagen, D. F. *Anal. Chem.* **1979,** *51*, 870. (b) St. Louis, R. H.; Hill, H. H. *Crit. Rev. Anal. Chem.* **1990,** *21*, 321. (c) von Helden, G.; Hsu, M. T.; Bowers, M. T. *J. Chem. Phys.* **1991,** *93*, 3835. (d) Jarrold, M. F. *J. Phys. Chem.* **1995**. (e) Clemmer, D. E.; Jarrold, M. F. *J. Mass Spectrom.* **1997,** *32*, 577. (f) Liu, Y.; Valentine, S. J.; Counterman, A. E.; Hoaglund, C. S.; Clemmer, D. E. *Anal. Chem.* **1997,** *69*, 728A.
5. (a) Hoaglund, C. S.; Valentine, S. J.; Sporleder, C. R.; Reilly, J. P.; Clemmer, D. E. *Anal. Chem.* **1998,** *70*, 2236. (b) Valentine, S. J.; Counterman, A. E.; Hoaglund, C. S.; Reilly, J. P.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **1998,** *9*, 1213.
6. Henderson, S. C.; Valentine, S. J.; Counterman, A. E.; Clemmer, D. E. *Anal. Chem.* **1999,** *71*, 291.
7. The term "collision cross section," as determined by ion mobility techniques, is strictly a measure of a collision integral which depends on the dynamics of momentum transfer for all ion-buffer gas collisions.
8. (a) Wyttenbach, T.; von Helden, G.; Bowers, M. T. *J. Am. Chem. Soc.* **1996,** *118*, 8355. (b) Wyttenbach, T.; Bushnell, J. E.; Bowers, M. T. *J. Am. Chem. Soc.* **1998,** *120*, 5098.
9. Counterman, A. E.; Valentine, S. J.; Srebalus, C. A.; Henderson, C. S.; Hoaglund, C. S.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **1998,** *9*, 743.
10. (a) Hudgins, R. R.; Ratner, M. A.; Jarrold, M. F. *J. Am. Chem. Soc.* **1998,** *120*, 12974. (b) Hudgins, R. R.; Mao, Y.; Ratner, M. A.; Jarrold, M. F. *Biophys. J.* **1999,** *76*, 1591.
11. Valentine, S. J.; Counterman, A. E.; Hoaglund-Hyzer, C. S.; Clemmer, D. E. *J. Phys. Chem. B* **1999,** *103*, 1203.
12. Leon, S. J. *Linear Algebra with Applications, 3rd ed.*; Macmillan: New York, 1990, pp 208–211.
13. Counterman, A. E.; Clemmer, D. E. *J. Am. Chem. Soc.* **1999,** *121*, 4031.
14. Benner, S. A.; Cannarozzi, G.; Gerloff, D.; Turcotte, M.; Chelvanayagam, G. *Chem. Rev.* **1997,** *97*, 2725.
15. Winter, B. E.; Light-Wahl, K. J.; Rockwood, A. L.; Smith, R. D. *J. Am. Chem. Soc.* **1992,** *114*, 5897. Suckau, D.; Shi, Y.; Beu, S. C.; Senko, M. W.; Quinn, J. P.; Wampler, F. M.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **1993,** *90*, 790. Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1994,** *116*, 9765. Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1995,** *117*, 12840. Wood, T. D.; Chorush, R. A.; Wampler, F. M.; Little, D. P.; O'Connor, P. B.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **1995,** *92*, 2451. Cassady, C. J.; Carr, S. R. *J. Mass Spectrom.* **1996,** *93*, 3143. McLafferty, F. W.; Guan, Z.; Haupts, U.; Wood, T. D.; Kelleher, N. L. *J. Am. Chem. Soc.* **1998,** *120*, 4732. Freitas, M. A.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1998,** *9*, 1012. Freitas, M. A.; Marshall, A. G. *Int. J. Mass Spectrom. Ion Processes* **1999,** *185/186/187*, 565.
16. Williams, E. R. *J. Mass Spectrom.* **1996,** *31*, 831. Green, M. K.; Lebrilla, C. B. *Mass Spectrom. Rev.* **1997,** *16*, 53. Harrison, A. G. *Mass Spectrom. Rev.* **1997,** *16*, 201.
17. Klaasen, J. S.; Blades, A. T.; Kebarle, P. *J. Phys. Chem.* **1995,** *99*, 15509. Stephenson, J. L., Jr.; McLuckey, S. A. *J. Am. Chem. Soc.* **1996,** *118*, 7390. Stephenson, J. L., Jr.; McLuckey, S. A. *Anal. Chem.* **1996,** *68*, 4026. Rodriguez-Cruz, S. E.; Klassen, J. S.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **1997,** *8*, 565. Woenckhaus, J.; Mao, Y.; Jarrold, M. F. *J. Phys. Chem. B* **1997,** *101*, 847. Woenckhaus, J.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997,** *119*, 9586. Fye, J. L.; Woenckhaus, J.; Jarrold, M. F. *J. Am. Chem. Soc.* **1998,** *120*, 1327. Lee, S.; Freivogel, P.; Schindler, T.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1998,** *120*, 11758.
18. Kaltashov, I. A.; Fenselau, C. C. *J. Am. Chem. Soc.* **1995,** *117*,

9906. Adams, J.; Strobel, F.; Reiter, A. *J. Am. Soc. Mass Spectrom.* **1996,** *7*, 30. Kaltashov, I. A.; Fenselau, C. C. *Proteins* **1997,** *27*, 165.
19. Sullivan, P. A.; Axelsson, J.; Altmann, S.; Quist, A. P.; Sunqvist, B. U. R.; Reimann, C. T. *J. Am. Soc. Mass Spectrom.* **1996,** *7*, 329. Reimann, C. T.; Sullivan, P. A.; Axelsson, J.; Quist, A. P.; Altmann, S.; Roepstorff, P.; Velazquez, I.; Tapia, O. *J. Am. Chem. Soc.* **1998,** *120*, 7608.
20. Covey, T. R.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **1993,** *4*, 616. Collings, B. A.; Douglas, D. J. *J. Am. Chem. Soc.* **1996,** *118*, 4488.
21. Cox, K. A.; Julian, R. K.; Cooks, R. G.; Kaiser, R. E. *J. Am. Soc. Mass Spectrom.* **1994,** *5*, 127.
22. Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1995,** *117*, 10141. Shelimov, K. B.; Jarrold, M. F. *J. Am. Chem. Soc.* **1996,** *118*, 10313. Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997,** *119*, 2240. Shelimov, K. B.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997,** *119*, 2987. Shelimov, K. B.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997,** *119*, 9586.
23. (a) Valentine, S. J.; Clemmer, D. E. *J. Am. Chem. Soc.* **1997,** *119*, 3558. (b) Valentine, S. J.; Anderson, J.; Ellington, A. D.; Clemmer, D. E. *J. Phys. Chem. B* **1997,** *101*, 3891. (c) Valentine, S. J.; Counterman, A. E.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **1997,** *8*, 954.
24. Schnier, P. D.; Price, W. D.; Jockusch, R. A.; Williams, E. R. *J. Am. Chem. Soc.* **1996,** *118*, 7178.
25. See, for example: Martyna, G. J. *J. Chem. Phys.* **1996,** *104*, 2018. Samuelson, S. O.; Tobias, D. J.; Martyna, G. J. *J. Phys. Chem. B* **1997,** *101*, 7592. Reimann, C. T.; Velazquez, I.; Tapia, O. *J. Phys. Chem. B* **1998,** *102*, 2277.
26. Fasman, G. D. *Prediction of Protein Structure and the Principles of Protein Conformation*; Plenum: New York, 1989.
27. *Protein Structure Prediction: A Practical Approach*; Sternberg, M. J. E., Ed.; Oxford University Press: New York, 1996.
28. Li, J.; Taraszka, J. A.; Counterman, A. E.; Clemmer, D. E. *Int. J. Mass Spectrom. Ion Processes* **1999,** *185/186/187*, 37.
29. Mao, Y.; Woenckhaus, J.; Kolafa, J.; Ratner, M. A.; Jarrold, M. F. *J. Am. Chem. Soc.* **1999,** *121*, 2712.
30. McLafferty, F. W.; Guan, Z.; Haupts, U.; Wood, T.; Kelleher, N. L. *J. Am. Chem. Soc.* **1998,** *120*, 4732.
31. Gross, D. S.; Schnier, P. D.; Rodriguez-Cruz, S. E.; Fagerquist, C. K.; Williams, E. R. *Proc. Natl. Acad. Sci. USA* **1996,** *93*, 3143.
32. Colby, S. M.; Reilly, J. P. *Anal. Chem.* **1996,** *68*, 1419.
33. MATLAB Version 4; The Math Works, Inc.: Natick, MA, 1997.
34. Mason, E. A.; McDaniel, E. W. *Transport Properties of Ions in Gases*; Wiley: New York, 1988.
35. Samuelson, S. O.; Martyna, G. J. *J. Chem. Phys.* **1998,** *109*, 11061. Samuelson, S. O.; Martyna, G. J. *J. Phys. Chem. B* **1999,** *103*, 1752; Clemmer, D. E., unpublished results.
36. *Protein Sequencing: A Practical Approach*; Findlay, J. B. C.; Gelsow, M. J., Eds.; IRL Press: Oxford, 1989, p 43.
37. It is important to note that parameters used for cross section calculations of the 11 peptides were slightly different than the 5-10-Lys parameterization set (Table 3) because the data for sequences obtained from tryptic digests of sperm whale myoglobin and yeast enolase were not originally available.
38. Wyttenbach, T.; von Helden, G.; Batka, J. J., Jr.; Carlat, D.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1997,** *8*, 275.
39. Shvartsburg, A. A.; Jarrold, M. F. *Chem. Phys. Lett.* **1996,** *261*, 86. Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. *J. Phys. Chem.* **1996,** *100*, 16082. Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. *J. Chem. Phys.* **1998,** *108*, 2416.
40. Hoaglund, C. S.; Liu, Y.; Ellington, A. D.; Pagel, M.; Clemmer, D. E. *J. Am. Chem. Soc.* **1997,** *119*, 9051.