

On the Zwitterionic Nature of Gas-Phase Peptides and Protein Ions

Roberto Marchese¹, Rita Grandori^{2*}, Paolo Carloni^{1,3,4}, Simone Rauei^{1,5*}

1 Statistical and Biological Physics Sector, International School for Advanced Studies (SISSA-ISAS) and DEMOCRITOS, Trieste, Italy, **2** Department of Bioscience and Biotechnology, Milano-Bicocca, Milano, Italy, **3** Italian Institute of Technology (IIT), Trieste, Italy, **4** Computational Biophysics sector, German Research School for Simulation Science, FZ-Juelich and RWTH, Juelich, Germany, **5** Center for Molecular Electrocatalysis, Pacific Northwest National Laboratory, Richland, Washington, United States of America

Abstract

Determining the total number of charged residues corresponding to a given value of net charge for peptides and proteins in gas phase is crucial for the interpretation of mass-spectrometry data, yet it is far from being understood. Here we show that a novel computational protocol based on force field and massive density functional calculations is able to reproduce the experimental facets of well investigated systems, such as angiotensin II, bradykinin, and tryptophan-cage. The protocol takes into account all of the possible protomers compatible with a given charge state. Our calculations predict that the low charge states are zwitterions, because the stabilization due to intramolecular hydrogen bonding and salt-bridges can compensate for the thermodynamic penalty deriving from deprotonation of acid residues. In contrast, high charge states may or may not be zwitterions because internal solvation might not compensate for the energy cost of charge separation.

Citation: Marchese R, Grandori R, Carloni P, Rauei S (2010) On the Zwitterionic Nature of Gas-Phase Peptides and Protein Ions. *PLoS Comput Biol* 6(5): e1000775. doi:10.1371/journal.pcbi.1000775

Editor: James M. Briggs, University of Houston, United States of America

Received: November 19, 2009; **Accepted:** March 31, 2010; **Published:** May 6, 2010

Copyright: © 2010 Marchese et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by the CNR/INFN with allocation of computer time at the CINECA supercomputing center and "Fondo Ateneo per la Ricerca" (FAR) to RG. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

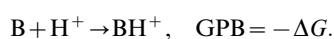
Competing Interests: The authors have declared that no competing interests exist.

* E-mail: rita.grandori@unimib.it (RG); rauei@sisa.it (SR)

Introduction

Predicting the structural properties of proteins in the gas phase is crucial to interpret mass spectrometry data, yet this is far from being understood [1–10]. So far, it has been established that (i) compact structures acquire smaller net charges than unfolded ones [11–13], (ii) secondary and tertiary structure elements play a crucial role for protein fragmentation [14–23], and (iii) hydrogen bonds (H-bonds) and salt-bridges [3,24,25] may stabilize the structures. However, how desolvation impacts on structural facets of proteins [2,3,8,26–30], peptides [3,31–35] and even single amino acids [36–45] is matter of a vivid debate.

A key point is the presence of charge separation. Whilst amino acids exist mostly in their zwitterionic form in the aqueous solution [31,36,40,46], conflicting assumptions and conclusions have been drawn for the same molecules *in vacuo* [47–50]. For peptides and proteins, the key issue of the charge state of ionizable groups, presumably different from that in solution, is even less clear [2,51–53]. One line of thought assumes neutral acidic functions for proteins analyzed in positive-ion mode (*i.e.*, generating and detecting positively charged ions) and neutral basic sites in negative-ion mode. In other words, the number of ionized groups is generally assumed to be equal to the net charge of the protein ion [54–56]. Electrostatic energy calculations based on this supposition fail to reproduce experimental values of apparent gas-phase basicity (GPB) for folded protein ions [57]. The GPB of a basic species B is defined as the negative of the free-energy change, ΔG , for the gas-phase protonation reaction



If B is the conjugate base of an acid AH, then $GPB = -GA$, where GA is the gas-phase acidity of AH. Analogously, the proton affinity is defined as the negative of the protonation enthalpy, $PA = -\Delta H$.

In contrast, an increasing number of experimental [16,24,25,58–61] and theoretical [62,63] investigations carried out on peptides and small proteins indicate that zwitterionic states may survive in the absence of solvent if the structural features allow for adequate intramolecular solvation [64–67]. Recent ultraviolet photo-dissociation [16] and fluorescence [25,61] experiments indicate the presence of stabilizing salt-bridge motifs in small biomolecules. Salt bridges exist also in protonated, gas-phase serine dimers [24] and have been predicted for arginine dimers [63,68,69]. These interactions add to other stabilizing contributions such as hydrogen bonds [3,24,25]. Molecular dynamics (MD) simulations on a minimalistic lattice model of a zwitterionic system [1] turned out to reproduce the experimental observation that compact structures acquire smaller net charges than unfolded ones [11–13]. On the basis of these simulations, it has been also proposed that steric and electrostatic shielding of charged residues in compact conformations are the major factors responsible for this structural facet. Energy calculations [2,7,52] and measurements [51,70] on several well characterized proteins in their experimentally observed, most populated charge state suggest that the presence of zwitterions depends on the specific protein structure [2,51]. Deprotonated aspartic and glutamic residues persist in the most abundant, positively charged protomer of insulin, the C-terminal domain of the ribosomal protein L7/L12 and ubiquitin, but not in tryptophan-cage and lysozyme [2].

Author Summary

In the last two decades mass spectrometry has given an impressive contribution to biochemistry, protein science, proteomics and structural biology. This technique offers powerful insights into protein structure and dynamics along with useful information on the role of solvent in protein stability as it is able to preserve non-covalent interactions and globular structures during the proteins' flight inside the mass spectrometer. Unfortunately, the key issue of the charge state of ionizable groups, presumably different from that in solution, has not been elucidated yet. So far conflicting assumptions and conclusions have been drawn by several groups. In the present work a very accurate structural and energetic analysis of the protonation state of two peptides and a small protein in the gas phase was performed. Results suggest that internal solvation can stabilize charge separation with the formation of zwitterionic states.

Prompted by the current lack of understanding of the charge state of protein ions *in vacuo*, here we have carried out an exhaustive energy analysis on three systems largely studied in the gas phase both experimentally [16,25,59,60,71–77] and theoretically [52,62,63,72]. These are the 8-residue peptide angiotensin II (AN) [74–76] and the 9-residue peptide bradykinin (BK) [16,59,62,63,71–73], as well as the tryptophan-cage (Trp-cage) [16,25,52,60,77] protein. The latter is a 20-residue mini-protein with a well defined secondary and tertiary structure in aqueous solution at ambient conditions. It consists of an α -helix and a compact hydrophobic core formed by a Trp side chain from the α -helix, surrounded by several hydrophobic residues (two prolines and one tyrosine) [78].

A complete exploration of the protomer space (*i.e.*, all of the possible charge configurations compatible with a given charge state) of these biomolecules is performed coupling force field-based molecular dynamics and density functional theory (DFT) calculations. In contrast to previous computational studies [25,31,52,62,63,79,80], all of the charge states generated by ionized and/or neutral *basic* (R, K, H, Q, N-terminus) and *acidic* groups (E, D, C-terminus), featuring more than one protomer, are taken into account.

A computational protocol apt to this task has been developed, allowing for an exhaustive exploration of the conformational space of each protomer. Based on such protocol, we suggest that low-charge states are likely zwitterions. In those cases, H-bonds and salt-bridges stabilize largely zwitterionic states, considerably reducing the differences in the apparent GPB between basic residues and the conjugated base of acidic residues. At high net charge, instead, non-zwitterion states are most likely.

Methods

Systems

The sequences of BK, AN and Trp-cage are RPPGFSPFR, DRVYIHPF, NLYIQWLKDGPPSSGRPPPS, respectively. For each system, the following protonation sites were considered: Arg₁, Arg₉, N- and C-term for BK; Asp₁, Arg₂, His₆, N- and C-term for AN (His₆ in the neutral state can be protonated either in δ or ϵ , both tautomers were considered); Gln₅, Lys₈, Asp₉, Arg₁₆, and N- and C-term for Trp-cage. In the latter, protonation of Gln₅ was considered for the $q=2+$ and $q=3+$ charge states on the basis of experimental evidences [19].

BK and AN have no determined secondary structure and all of the calculations started with an all-trans backbone and side-chain conformation. Instead, the Trp-cage initial structure was obtained by a 20-ns MD simulation in aqueous solution at ambient conditions based on the NMR structure number 1 deposited in the protein data bank (PDB code: 1L2Y) [78] (see Text S1). The most probable protonation state in water [78] was chosen.

For the chosen set of protonation sites, all of the charge states which feature more than one protomer were taken into account. For these charge states, all of the possible protomers were considered, for a total of 100 protomers (see Tables 1, 2, and 3).

Force field-based MD calculations

OPLS/AA force field-based [81,82], constant-temperature MD calculations and geometry optimizations were carried out. The cutoff of electrostatics and van der Waals interactions was fixed at 0.7nm. In the MD simulations, the equations of motion were numerically integrated with a time step of 1.5 fs. All the hydrogen-bond lengths were kept fixed using the LINCS [83] algorithm. The temperature was controlled by the Nosé-Hoover thermostat [84]. The results of force field based MD simulations depend critically on the charge state used. Therefore, we performed a simulation for each protonation state. Specifically 8-ns MD simulations at high-temperature (700K for AN and BK, 350K for Trp-cage) were performed for each protonation state. The

Table 1. Energetics and structural parameters for the lowest-energy conformers of bradykinin protomers.

N-ter	R ₁	R ₉	C-ter	ΔE	$-\Delta\Delta U$	ΔPA	ΔGB	IR	SB	sHB	iHB	HB
$q=0$												
0	+	0	–	0	380	409	368	2	1	2	5	1
0	0	+	–	10	380	409	368	2	1	2	4	3
+	0	0	–	49	542	535	540	2	1	1	5	0
0	0	0	0	57	0	0	0	0	0	0	0	4
$q=1+$												
0	+	+	–	0	380	409	368	3	2	3	2	0
0	+	0	0	46	0	0	0	1	0	0	3	4
0	0	+	0	63	0	0	0	1	0	0	4	1
+	0	+	–	67	542	535	540	3	2	2	2	3
+	+	0	–	112	542	535	540	3	1	2	4	0
+	0	0	0	152	162	126	172	1	0	0	3	3
$q=2+$												
+	+	+	–	0	542	535	540	4	3	4	3	1
0	+	+	0	77	0	0	0	2	0	0	6	1
+	+	0	0	85	162	126	172	2	0	0	5	1
+	0	+	0	93	162	126	172	2	0	0	5	2

In each row the following information is reported: protonation pattern (first column); energy difference with respect to the most stable protomer (ΔE in kJ/mol); (intrinsic) internal energy variation ($\Delta\Delta U$), proton affinity (ΔPA) and gas-phase basicity (ΔGPB) relative to the most favourable protomer (see text for a definition of these quantities; all values are in kJ/mol); ionized residues (IR); salt-bridges (SB); hydrogen bonds between salt-bridged residues (sHB); ionized hydrogen bonds where either the donor or the acceptor is ionized, $D^+ \cdots A$ or $D \cdots A^-$ H-bonds (iHB); neutral hydrogen bonds (HB). Hydrogen bonds are identified according to the donor-acceptor ($A \cdots D$) distance and the donor-acceptor H-bond angle ($D-H \cdots A$). The following geometric criterion was adopted: $d(A \cdots D) = 3.5 \text{ \AA}$ and $\angle(D-H \cdots A) = 150^\circ$. A salt-bridge is formed if the distance between any oxygen atom of the acidic residue and any protonable nitrogen atom of the basic residue is less than 4.0 Å.
doi:10.1371/journal.pcbi.1000775.t001

Table 2. Energetics and structural parameters for the lowest-energy conformers of angiotensin II protomers.

N-ter	D ₁	R ₂	H ₆	C-ter	ΔE	$-\Delta\Delta U$	ΔPA	ΔGB	IR	SB	sHB	iHB	HB
<i>q</i> = 1−													
0	−	+	ε	−	0	380	409	368	3	2	4	2	1
0	−	+	δ	−	43	380	409	368	3	2	4	2	1
0	−	0	+	−	51	447	452	423	3	2	2	2	1
0	0	0	δ	−	54	0	0	0	1	0	0	6	2
0	−	0	δ	0	73	95	105	66	1	0	0	5	2
+	−	0	δ	−	84	542	535	540	3	2	1	6	0
0	−	0	ε	0	86	95	105	66	1	0	0	4	1
0	0	0	ε	−	91	0	0	0	1	0	0	3	4
+	−	0	ε	−	100	542	535	540	3	2	1	5	0
<i>q</i> = 0													
0	−	+	+	−	0	732	756	725	4	4	5	0	0
0	0	+	δ	−	17	380	409	368	2	1	1	3	2
0	0	+	ε	−	21	380	409	368	2	1	2	1	2
0	0	0	+	−	32	447	452	423	2	0	1	4	0
0	−	+	δ	0	37	285	304	302	2	1	2	3	1
+	−	+	ε	−	39	827	839	842	4	4	5	2	4
+	−	+	δ	−	39	827	839	842	4	4	4	1	2
0	−	+	ε	0	42	285	304	302	2	1	2	3	4
+	0	0	δ	−	64	542	535	540	2	1	1	2	0
0	0	0	δ	0	69	0	0	0	0	0	0	0	7
+	0	0	ε	−	73	542	535	540	2	1	1	3	4
0	−	0	+	0	80	352	347	357	2	1	1	3	3
+	−	0	+	−	100	894	882	897	4	4	3	2	0
+	−	0	ε	0	104	447	430	474	2	1	0	4	2
0	0	0	ε	0	105	0	0	0	0	0	0	0	8
+	−	0	δ	0	106	447	430	474	2	1	1	2	2
<i>q</i> = 1+													
0	0	+	+	−	0	447	452	423	3	2	3	3	1
+	−	+	+	−	1	894	882	897	5	4	5	2	0
0	−	+	+	0	2	352	347	357	3	2	3	3	0
0	0	+	δ	0	11	0	0	0	1	0	0	4	4
0	0	+	ε	0	25	0	0	0	1	0	0	4	2
+	−	+	δ	0	41	447	430	474	3	2	2	3	0
0	0	0	+	0	42	67	43	55	1	0	0	2	3
+	0	+	ε	−	48	534	542	585	3	2	2	4	1
+	−	+	ε	0	48	542	535	540	3	2	2	2	1
+	0	+	δ	−	62	542	535	540	3	2	2	2	1
+	0	0	+	−	81	447	430	474	3	2	2	3	1
+	0	0	δ	0	85	162	126	172	1	0	0	3	1
+	−	0	+	0	117	514	473	529	3	2	2	1	2
+	0	0	ε	0	122	162	126	172	1	0	0	2	3
<i>q</i> = 2+													
0	0	+	+	0	0	0	0	0	2	0	0	4	1
+	0	+	+	−	15	542	535	540	4	3	3	4	1
+	0	+	δ	0	19	95	83	117	2	0	0	2	3
+	−	+	+	0	43	447	430	474	4	3	3	2	0
+	0	0	+	0	64	162	126	172	2	0	0	2	1
+	0	+	ε	0	95	95	83	117	2	0	0	2	2

Table 2. Cont.

In each row the following information is reported: protonation pattern (first column); energy difference with respect to the most stable protomer (ΔE in kJ/mol); (intrinsic) internal energy variation ($\Delta\Delta U$), proton affinity (ΔPA) and gas-phase basicity (ΔGPB) relative to the most favourable protomer (see text for a definition of these quantities; all values are in kJ/mol); ionized residues (IR); salt-bridges (SB); hydrogen bonds between salt-bridged residues (sHB); ionized hydrogen bonds where either the donor or the acceptor is ionized, $D^+ \cdots A$ or $D \cdots A^-$ H-bonds (iHB); neutral hydrogen bonds (HB). Hydrogen bonds are identified according to the donor-acceptor ($A \cdots D$) distance and the donor-acceptor H-bond angle ($D-H \cdots A$). The following geometric criterion was adopted: $d(A \cdots D) = 3.5 \text{ \AA}$ and $\angle(D-H \cdots A) = 150^\circ$. A salt-bridge is formed if the distance between any oxygen atom of the acidic residue and any protonable nitrogen atom of the basic residue is less than 4.0 Å. doi:10.1371/journal.pcbi.1000775.t002

chosen temperatures were selected after several careful tests. In particular, for Trp-cage, a temperature of 350K turns out to allow for an exhaustive sampling of the side chains conformations without disrupting, in the relatively short simulation time, the secondary structure. The resulting trajectories were split into 5-ps, non overlapping time windows. For each window, the geometry of the lowest-energy MD conformation was optimized by a conjugated gradient scheme up to 0.1 kJ/molÅ residual force on any atom. This simulated annealing-like procedure yielded for each protomer a large set of conformations. The geometry of structures within 100 kJ/mol (60 kJ/mol for Trp-cage) from the lowest-energy force field conformer were refined at the *ab initio* level (see Section “Identifying relevant protomers of a given charge state”). With this criterion, 60 conformers (35 for Trp-cage), were randomly selected from equally spaced energy windows, one from each window, and re-optimized at DFT/BLYP level of theory.

The GROMACS [85] software package was used for all MD calculations.

Quantum-chemical geometry optimizations

Quantum-chemical geometry optimizations were performed within the framework of DFT. The Becke exchange [86] and Lee-Yang-Parr [87] correlation functionals (BLYP) were used in a hybrid Gaussian and plane wave approach [88]. The wave function was optimized by using an orbital transformation technique [89] and analytic Goedecker-Teter-Hutter [90,91] pseudopotentials (PP). The TZV2P Gaussian basis set was used for valence electrons of all atoms, while the auxiliary electron density was expanded in plane waves up to a cutoff of 280 Ry.

The interaction between periodic images in the reciprocal space was removed according to the decoupling scheme presented in [92]. The calculations were carried out with the CP2K code [89,93,94], which has been shown to be very efficient for these systems.

The adopted DFT scheme was validated against more accurate (and more expensive) quantum chemistry methods. First, the relative energy of canonical and zwitterionic arginine conformers calculated with the present scheme agrees well with that obtained from all-electrons B3LYP, MP2, and CCSD calculations (see Table 2 in Text S1). Second, all of the 14 protomers of AN with total charge $q=1+$ underwent all-electrons, single-point energy evaluations at DFT/B3LYP level with the 6–311++G(d,p) basis set using the Gaussian03 code [95] (Angiotensin II was chosen because it is the smallest of the three molecules studied and, in particular, the charge state $q=1+$ was considered because it presents the largest set of protomers, and it is, therefore, a good benchmark case). These and the previous calculations provided the same energy ranking (see Table 3 in Text S1).

Table 3. Energetics and structural parameters for the lowest-energy conformers of Trp-cage protomers.

N-ter	Q ₅	K ₈	D ₉	R ₁₆	C-ter	ΔE	$-\Delta\Delta U$	ΔPA	ΔGB	IR	SB	sHB	iHB	HB
<i>q</i> = 0														
0	0	0	0	+	–	0	380	409	368	2	0	0	9	8
+	0	0	–	+	–	9	827	839	842	4	4	4	8	7
0	0	+	–	0	0	20	341	345	350	2	1	1	5	12
0	0	0	0	0	0	22	0	0	0	0	0	0	0	19
0	0	+	0	0	–	25	436	450	416	2	0	0	6	11
0	0	+	–	+	–	33	721	754	718	4	2	3	10	5
0	0	0	–	+	0	47	285	304	302	2	1	1	6	7
+	0	+	–	0	–	95	883	880	890	4	2	2	8	10
+	0	0	0	0	–	108	542	535	540	2	1	1	3	12
+	0	0	–	0	0	228	447	430	474	2	0	0	8	8
<i>q</i> = 1+														
+	0	+	–	+	–	wat	–	–	–	5	1	1	0	5
+	0	0	0	+	–	0	542	535	540	3	1	1	5	12
0	0	0	0	+	0	5	0	0	0	1	0	0	3	12
+	0	0	–	+	0	18	447	430	474	3	2	3	6	6
0	0	+	–	+	0	41	341	345	450	3	2	3	3	11
+	0	0	0	0	0	66	162	126	172	1	0	0	2	16
+	0	+	0	0	–	66	598	576	588	3	1	1	5	8
+	0	+	–	+	–	72	883	880	890	5	5	7	4	5
0	0	+	0	+	–	79	436	450	416	3	0	0	9	7
0	0	+	0	0	0	87	56	41	48	1	0	0	3	13
+	0	+	–	0	0	99	503	471	522	3	1	1	8	7
<i>q</i> = 2+														
+	0	+	0	+	–	0	542	535	540	4	2	1	6	8
0	0	+	0	+	0	7	0	0	0	2	0	0	6	11
+	0	0	0	+	0	57	106	85	124	2	0	0	6	14
+	+	+	0	0	–	69	674	642	647	4	1	3	7	9
+	+	+	–	+	–	74	959	946	949	6	4	6	7	4
+	0	+	–	+	0	104	447	430	474	4	1	1	11	6
+	0	+	0	0	0	105	162	126	172	2	0	0	4	9
+	+	0	0	+	–	110	618	601	599	4	2	4	5	9
0	+	+	0	+	–	116	512	516	475	4	1	2	6	8
0	+	0	0	+	0	130	76	66	59	2	0	0	7	9
0	+	+	–	+	0	144	417	411	409	4	2	3	9	8
+	+	+	–	0	0	155	579	537	581	4	1	2	8	9
0	+	+	0	0	0	132	107	107	149	2	0	0	4	11
+	+	0	0	0	0	211	238	192	231	2	0	0	6	12
+	+	0	–	+	0	264	523	496	533	4	0	0	12	7
<i>q</i> = 3+														
+	0	+	0	+	0	0	30	19	65	3	0	0	10	11
+	+	0	0	+	0	61	106	85	124	3	0	0	8	8
+	+	+	0	+	–	62	542	535	540	5	1	1	12	9
0	+	+	0	+	0	84	0	0	0	3	0	0	8	7
+	+	+	–	+	0	113	447	430	474	5	2	3	9	6
+	+	+	0	0	0	158	162	126	172	3	0	0	7	10

In each row the following information is reported: protonation pattern (first column); energy difference with respect to the most stable protomer (ΔE in kJ/mol); (intrinsic) internal energy variation ($\Delta\Delta U$), proton affinity (ΔPA) and gas-phase basicity (ΔGPB) relative to the most favourable protomer (see text for a

Table 3. Cont.

definition of these quantities; all values are in kJ/mol; ionized residues (IR); salt-bridges (SB); hydrogen bonds between salt-bridged residues (sHB); ionized hydrogen bonds where either the donor or the acceptor is ionized, $D^+ \cdots A$ or $D \cdots A^-$ H-bonds (iHB); neutral hydrogen bonds (HB). Hydrogen bonds are identified according to the donor-acceptor ($A \cdots D$) distance and the donor-acceptor H-bond angle ($D-H \cdots A$). The following geometric criterion was adopted: $d(A \cdots D) = 3.5 \text{ \AA}$ and $\angle(D-H \cdots A) = 150^\circ$. A salt-bridge is formed if the distance between any oxygen atom of the acidic residue and any protonable nitrogen atom of the basic residue is less than $4.0 \text{ \AA}$. Structural data for the most stable protonation state in aqueous solution ($q = 1+$) are also reported (data in Italics).

doi:10.1371/journal.pcbi.1000775.t003

A final concern for using DFT for non-covalent systems is the underestimation of dispersion forces [96,97]. This flaw of the current GGA functionals might influence the conformational energy, especially in the case of large molecular assemblies like those considered here. To quantify this error an estimate of the dispersion energy was performed for the DFT optimized structures using the OPLS/AA force field. The results of this calculation (see Tables 4, 5 and 6 in Text S1) indicate that the dispersion energy is not expected to change qualitatively the DFT energy ranking of protomers.

Results

Identifying relevant protomers of a given charge state

A standard procedure to identify the relevant protomers is currently lacking, even for peptides with more than a few amino-acids. On the one hand, the high complexity of the conformational space hampers an exhaustive search based on first-principle quantum chemistry (such as DFT) of the minimum-energy conformers. On the other hand, force field-based calculations [62,63,98,99], or semiempirical quantum chemical methods [52], may not be accurate enough. For instance, Merck molecular force field [100] energies have been shown to correlate poorly with those calculated at the DFT/B3LYP level [62,63]. In addition, the energies calculated by force fields do not take into account higher-order effects, which may play a role in our systems. DFT can, instead, take such effects into account.

However, if the empirically calculated conformer is much higher in energy than another (say with a ΔE_c greater than few hundreds of kJ/mol), it will be highly probable that the same ranking holds at the *ab initio* level (see Figure 1 in Text S1). Here, we seek such ΔE_c value by performing MD simulations based on the OPLS/AA, which offers the most complete set of base/conjugate acid pairs. The calculations on the three systems *in vacuo* provided several hundreds conformations, which then underwent DFT/BLYP [86,87] geometry optimizations. Such quantum chemical scheme is extremely efficient for large molecules, as those investigated here [101,102].

We found that less than 5% of the *ab initio* conformers located within 10 kJ/mol from the energy minimum fall more than $\Delta E_c = 100 \text{ kJ/mol}$ (60 kJ/mol for Trp-cage) above the OPLS/AA minimum (see Figure 2 in Text S1). Exploiting this fact, we used the ensuing protocol to identify the lowest-energy minimum for each charge state for each peptide: (i) generation of conformers for all possible protomers by OPLS/AA MD and simulated annealing-like calculations; (ii) elimination of conformers whose energy is larger than ΔE_c from the absolute minimum; (iii) DFT/BLYP geometry optimization of the conformers within ΔE_c ; (iv) ranking of the conformers based on their DFT energies.

Errors of this protocol are associated with (i) the accuracy of the DFT approach, (ii) limitations of sampling and (iii) absence of entropy contributions. This points are discussed in the following.

- (i) The accuracy of our DFT-based calculations may be assessed by comparing the energy contributions of the key quantities for the identification of the lowest free-energy minimum with high-level quantum chemical calculations. These are the gas-phase basicities (GPBs) and/or proton affinities (PAs) of the protomers. Table 4, and Tables 1, 2 and 3 in Text S1 show that our approach agrees well with more accurate calculations [62,103–105] (see Section “Quantum-chemical geometry optimizations” for further details). Similar considerations can be evinced from a comparison between our calculations of PA and those recently reported in a highly accurate quantum-chemical study on side chain interactions in proteins [99]. All of these facts establish the accuracy of our calculated energies.
- (ii) The upper-bound estimate of the uncertainty in the energy value due to the sampling can be obtained by performing different searches for the lowest-energy conformer of selected protomers. This procedure ended with either the same or very similar structures (backbone root mean square displacements less than 1.0 Å), and all were located within 10 kJ/mol from each other. Therefore, this value can be taken as an estimate of our statistical uncertainty.
- (iii) Free energies can in principle be obtained by adding zero-point-energy (ZPE) and entropy corrections (other than rotational contributions). However, these calculations are extremely expensive for large systems such as those investigated here. Fortunately, the ranking obtained with our energy-based protocol can be used to identify differences in GPBs and PAs between the protomers. Table 4 reports such data for different conformations of free amino acids. The variations in the entropic term ($\Delta\text{GPB}-\Delta\text{PA}$) among alternative structures are much smaller (less than 5%) than the differences between the corresponding energy terms (GPB and PA), similar considerations can be done based on literature data [52]. Such differences turn out to be roughly constant, with a dispersion (in terms of standard deviation from the average) smaller than 6 kJ/mol. This dispersion is smaller than the error due to the conformational sampling.

Table 4. Thermodynamic data for the protonation reaction of the side chain of some amino acids.

	Folded			Linear			Linear-Folded		
	$-\Delta U$	PA	GPB	$-\Delta U$	PA	GPB	$-\Delta\Delta U$	$\Delta(\text{PA})$	$\Delta(\text{GPB})$
Lys	1024	985	978	961	931	925	−63	−54	−53
Arg	1080	1026	1026	1046	1008	1015	−34	−18	−11
His	1013	983	971	985	953	951	−28	−30	−20
Gln	948	919	919	913	877	872	−35	−42	−47
Asp [−]	1365	1330	1328	1426	1389	1388	61	59	60
Glu [−]	1368	1358	1356	1453	1412	1420	85	54	64

Internal energy change (ΔU), proton affinity (PA), and gas-phase basicity (GPB) values (in kJ/mol) are reported for two limiting situations: (i) protonated and deprotonated species in their lowest-energy conformer (labeled as “folded”), and (ii) in an extended, all-trans conformer (labeled as “linear”). Differences between these quantities calculated for the folded and linear conformations are also given.

doi:10.1371/journal.pcbi.1000775.t004

Furthermore, the estimates of GPB and PA reported in Table 4 and Table 1 in Text S1 fall nicely into the range of experimentally determined values [105–109], and fully support the above considerations (vibrational energy corrections to enthalpy and entropy were calculated from harmonic vibrational frequencies. The effect of losing three translational degrees of freedom on going from BH^+ to B (or from AH to A) was also taken into account [103]).

We therefore conclude that the ranking obtained with our protocol provides a reliable identification of the most stable protomers.

Structural features

We discuss here the salient structural data of the low-energy protomers identified with the protocol outlined above for each system and for every charge state that features, according to our choice of ionizable residues, more than one protomer. More details and additional observations can be found in Text S1. Structural data for each protomer of the considered charge states are reported in Tables 1, 2, and 3.

Bradykinin. Four protonation sites are possible: Arg₁, Arg₉, N- and C-terminus. All of the identified lowest-energy protomers from neutral to doubly positively charged states are zwitterions (see Table 1). All of them feature an extensive H-bonding network, and a β -turn-like motif formed by the residues Ser₆-Pro₇-Phe₈-Arg₉, in agreement with Ref. [62]. However, the β -turn does not always feature an $i-i+3$ H-bond. The C-term protonation site forms salt-bridges with the others if the geometry allows to do so (see Table 1).

Neutral BK. The neutral form features two lowest-energy protomers, with (i) deprotonated C-ter and protonated Arg₉ ($\text{bk}^{(0)}[\text{n R}_1\text{R}_9^+\text{c}^-]$) and (ii) deprotonated C-ter and protonated Arg₁ ($\text{bk}^{(0)}[\text{n R}_1^+\text{R}_9\text{c}^-]$).

$[\text{BK}+\text{H}]^+$. As for $[\text{BK}+\text{H}]^+$, the lowest-energy protomer, $\text{bk}^{(1+)}[\text{n R}_1^+\text{R}_9^+\text{c}^-]$, has two Arg residues protonated and the C-ter deprotonated ($\text{bk}^{(1+)}[\text{n R}_1^+\text{R}_9^+\text{c}^-]$). The three charged sites cluster together forming two salt-bridges. The result is consistent with previous *ab initio* calculations [62,63] and experimental evidences [16,64].

$[\text{BK}+2\text{H}]^{2+}$. The lowest-energy protomer of $[\text{BK}+2\text{H}]^{2+}$ is $\text{bk}^{(2+)}[\text{n}^+\text{R}_1^+\text{R}_9^+\text{c}^-]$. The charged residues form a three-salt-bridge cluster where the C-ter is surrounded by the three positive groups (N-ter, Arg₁, and Arg₉). A similar lowest-energy zwitterionic structure for $[\text{BK}+2\text{H}]^{2+}$ is predicted also by other theoretical studies [62,110]. From an experimental point of view, no consensus is reached about the zwitterionic nature of $[\text{BK}+2\text{H}]^{2+}$. Blackbody infrared radiative dissociation experiments on bradykinin [64] could not rule out the presence of a salt-bridge between the N-terminal arginine and the carboxylate group. On the other hand, CO₂ and COOH losses by ultra-violet photodissociation suggest that the zwitterion might not be the predominant species in the gas phase [16].

Angiotensin II. Five protonation sites were chosen: N-ter, Asp₁, Arg₂, His₆ and C-ter. Although the primary sequence of AN is shorter than BK, the presence of an extra acidic residue (Asp₁) increases the number of possible protomers appreciably. In addition, the existence of two tautomers for neutral His (*i.e.*, δ and ϵ protonated forms) further enlarges the protomer space. We consider here charge states ranging from the monovalent anion to the doubly positively charged cation. All but the latter turn out to be zwitterions (see Table 2).

Lower charge states. The lowest-energy protomer of $[\text{AN}-\text{H}]^-$ is $\text{an}^{(-)}[\text{n D}_1^-\text{R}_2^+\text{H}_6^-\text{c}^-]$. It features the highest number of ionized residues with a deprotonated C-ter and Asp₁

and a protonated Arg₂. Arg₂ forms two salt-bridges, with both C-ter and Asp₁. The most stable protomer of the neutral state is $\text{an}^{(0)}[\text{n}^+\text{D}_1^-\text{R}_2^+\text{H}_6\text{c}^-]$, thus highly zwitterionic, with four salt-bridges. Arg is protonated instead of N-ter and His. The most stable structures of $[\text{AN}+\text{H}]^+$ are $\text{an}^{(1+)}[\text{n}^+\text{D}_1^-\text{R}_2^+\text{H}_6^+\text{c}^-]$, and $\text{an}^{(1+)}[\text{n}^+\text{D}_1\text{R}_2^+\text{H}_6^+\text{c}^-]$, whose relative energies are within 2 kJ/mol. Considering the overall energy ranking, also for this charge state there is a clear tendency to protonate Arg instead of N-ter and His.

The divalent cation. In the case of $[\text{AN}+2\text{H}]^{2+}$, the lowest-energy protomer, $\text{an}^{(2+)}[\text{n}^+\text{D}_1\text{R}_2^+\text{H}_6^+\text{c}^-]$, is not a zwitterion and lies 15 kJ/mol below the most stable zwitterion, $\text{an}^{(2+)}[\text{n}^+\text{D}_1\text{R}_2^+\text{H}_6^+\text{c}^-]$.

Tryptophan cage. Six protonation sites are possible, N-ter, Gln₅, Lys₈, Asp₉, Arg₁₆, and C-ter, which yield a large number of possible protomers (see Table 3).

For this system, additional 40 ns MD simulations were carried out for the lowest-energy protomers in the gas phase. Comparison is then made with the structural features of the protein in solution. All states considered feature a native-like compact structure (see Table 7 in Text S1), which does not depend appreciably on the charge state. In fact, the gyration radius (R_g) decreases slightly and in a similar way during the dynamics of all of the charge states (see Table 7 in Text S1). This contraction involves the solvent-exposed side chains, which fold onto the protein surface, in agreement with previous MD studies [2,8,14,52]. The side chain rearrangements are such to optimize at best H-bonding and salt-bridges (see Table 3). We remark that, in all of the gas-phase simulations, as a consequence of the absence of water, the Trp residue tends to move away from the hydrophobic core. This result is also consistent with previous investigations [14,52].

The lowest-energy protomer is a zwitterion for the low charge states (neutral Trp-cage, $[\text{Trp}-\text{cage}+\text{H}]^+$, and $[\text{Trp}-\text{cage}+2\text{H}]^{2+}$). However, the most probable state for $[\text{Trp}-\text{cage}+3\text{H}]^{3+}$ is not a zwitterion.

Lower charge states. The lowest-energy protomers of the neutral state are the zwitterions $\text{trp}^{(0)}[\text{n}^+\text{Q}_5\text{K}_8\text{D}_9\text{R}_{16}^+\text{c}^-]$ and $\text{trp}^{(0)}[\text{n}^+\text{Q}_5\text{K}_8\text{D}_9^-\text{R}_{16}^+\text{c}^-]$ ($\Delta E=9$ kJ/mol). $\text{trp}^{(0)}[\text{n}^+\text{Q}_5\text{K}_8\text{D}_9\text{R}_{16}^+\text{c}^-]$ does not feature any salt-bridge, but its ionized residues form a large number of strong H-bonds involving charged groups. In contrast, the highly zwitterionic $\text{trp}^{(0)}[\text{n}^+\text{Q}_5\text{K}_8\text{D}_9^-\text{R}_{16}^+\text{c}^-]$ forms four salt-bridges, arranged in a ring structure involving all the charged residues, with a distance between heavy atoms of ionized groups of about 2.85 Å.

The lowest-energy protomers of $[\text{Trp}-\text{cage}+\text{H}]^+$ are the zwitterion $\text{trp}^{(+)}[\text{n}^+\text{Q}_5\text{K}_8\text{D}_9\text{R}_{16}^+\text{c}^-]$, characterized by the N-term/C-term salt-bridge, and the non-zwitterionic species $\text{trp}^{(+)}[\text{n}^+\text{Q}_5\text{K}_8\text{D}_9\text{R}_{16}^+\text{c}^-]$ ($\Delta E=5$ kJ/mol). The presence of a zwitterionic state for $[\text{TRP}+\text{H}]^+$ in the gas-phase has been recently proposed by computational and experimental approaches [16,52]. Patriksson and coworkers [52] identify the zwitterion $\text{trp}^{(+)}[\text{n}^+\text{Q}_5\text{K}_8^+\text{D}_9^-\text{R}_{16}^+\text{c}^-]$ as the most stable protomer, which lies at 41 kJ/mol above the minimum according to our calculations. We found that the most stable protomer differs in having the N-ter charged instead of Lys₈ and the C-ter charged instead of Asp₉. In this regard, we remark that Kjeldsen *et al.* [16] identified the C-ter as the most likely carboxylate of the zwitterion identified by CO₂ photodissociation experiments on $[\text{TRP}+\text{H}]^+$.

The $[\text{TRP}+2\text{H}]^{2+}$ can form 15 protomers. The zwitterion $\text{trp}^{(2+)}[\text{n}^+\text{Q}_5\text{K}_8^+\text{D}_9\text{R}_{16}^+\text{c}^-]$ is the lowest-energy protomer. The C-ter forms two salt-bridges with N-ter and Lys₈. N-ter and C-ter are in very close contact (distance between heavy atoms of 2.6 Å) and strongly interact via H-bonds, whereas Arg₁₆ is located at about 3.0 Å from the C-ter. This species lies very close to the non-

zwitterionic protomer $\text{trp}^{(2+)}[\text{n}^+\text{Q}_5\text{K}_8^+\text{D}_9\text{R}_{16}^+\text{c}^-]$ ($\Delta E=7$ kJ/mol). Zubarev and coworkers [19] suggested a prevalently non-zwitterionic form for $[\text{TRP}+2\text{H}]^{2+}$ based on photodissociation experiments and on fragmentation patterns by Electron-Capture Dissociation (ECD) of the $[\text{TRP}+3\text{H}]^{3+}$ ion. According to these authors, $[\text{TRP}+2\text{H}]^{2+}$ has one positive charge on Arg₁₆ and a second positive charge distributed between the N-ter and Gln₅. The N-term was found to be the least basic site by ECD [17,19]. In our results, the protomer with neutral N-term is found only 7 kJ/mol above the lowest-energy protomer with protonated N-term, indicating that this group is likely only partially protonated. Gln₅ would be unfavorable. Indeed, the lowest-energy species with Gln₅ protonated, the zwitterion $\text{trp}^{(2+)}[\text{n}^+\text{Q}_5^+\text{K}_8^+\text{D}_9\text{R}_{16}^+\text{c}^-]$, is located at 69 kJ/mol. It will be shown in the next paragraph how our results can provide an alternative explanation of the ECD results.

The trivalent cation. The most probable state for $[\text{Trp}-\text{cage}+3\text{H}]^{3+}$ is the non-zwitterion $\text{trp}^{(3+)}[\text{n}^+\text{Q}_5\text{K}_8^+\text{D}_9\text{R}_{16}^+\text{c}^-]$. This result is consistent with fluorescence experiments combined with MD calculations carried out by Iavarone *et al.* [25] that led to the identification of the same protomer as the most probable for this charge state. However, our results contrast with the interpretation of recent ECD experiments [19], which points to $\text{trp}^{(3+)}[\text{n}^+\text{Q}_5^+\text{K}_8\text{D}_9\text{R}_{16}^+\text{c}^-]$ as the most probable protomer. Protonation of Gln₅ leads to high-energy protomers. The lowest-energy protomer with ionized Gln₅, $\text{trp}^{(3+)}[\text{n}^+\text{Q}_5^+\text{K}_8\text{D}_9\text{R}_{16}^+\text{c}^-]$, lies at 61 kJ/mol above the minimum. This interpretation was based on the fact that both $c-$ and $z-$ ions generated by ECD show that the z_{16} fragment (Ile₄-Gln₅ cleavage) exists in two charge states ($q=2+$ and $q=1+$, the $q=2+$ being predominant), whereas the z_{15} fragment (Gln₅-Trp₆ cleavage) is present in the $q=1+$ charge state and the z_5 fragment (Gly₁₅-Arg₁₆ cleavage) in the $q=0$ state. No change of charge in the z fragments upon Leu₇-Lys₈ cleavage is observed.

The differences between our and previous findings [19] can be here reconciled considering which interactions are lost upon fragmentation. Visual inspection of the lowest-energy protomer reveals that the charged Lys₈ side chain is H-bonded to Ile₄ and Leu₇ backbone and Asp₉ side chain (see Figure 1). In turn, Asp₉ donates a H-bond to Gln₅. Thus, Ile₄, Glu₅, Leu₇ and Asp₉ build a H-bond network that internally solvates and, therefore, stabilizes the positive charge on the lysine side chain. Dissociation of c_4 and c_5 fragments, progressively destroys this network. Such fragmentation could favor spontaneous proton transfer from Lys₈ to the departing c radical ion after the c_5 fragmentation. This explanation can also provide a rationale for the anomaly in the charge population of fragment z_{16} .

In conclusion, our results suggest that the observed effect of Gln dissociation might also be explained by its role in charge stabilization and not exclusively by its own ionization in the intact peptide.

Key stabilizing interactions for peptides and proteins in the gas phase

Our calculations suggest that most of the low-charge states are zwitterions, whilst high charge states might not. We now analyze the key factors for the stabilization of these two different states.

Zwitterionic low-charge states. Formation of charge separation between two residues is accompanied by a penalty. To a first approximation, this could be quantified in terms of the intrinsic GPB of the involved residues. The GPB of carboxylates is much higher than that of amino and guanidino groups (see Table 1 in Text S1). Therefore, charge separation between an acid, AH, and a base, B,

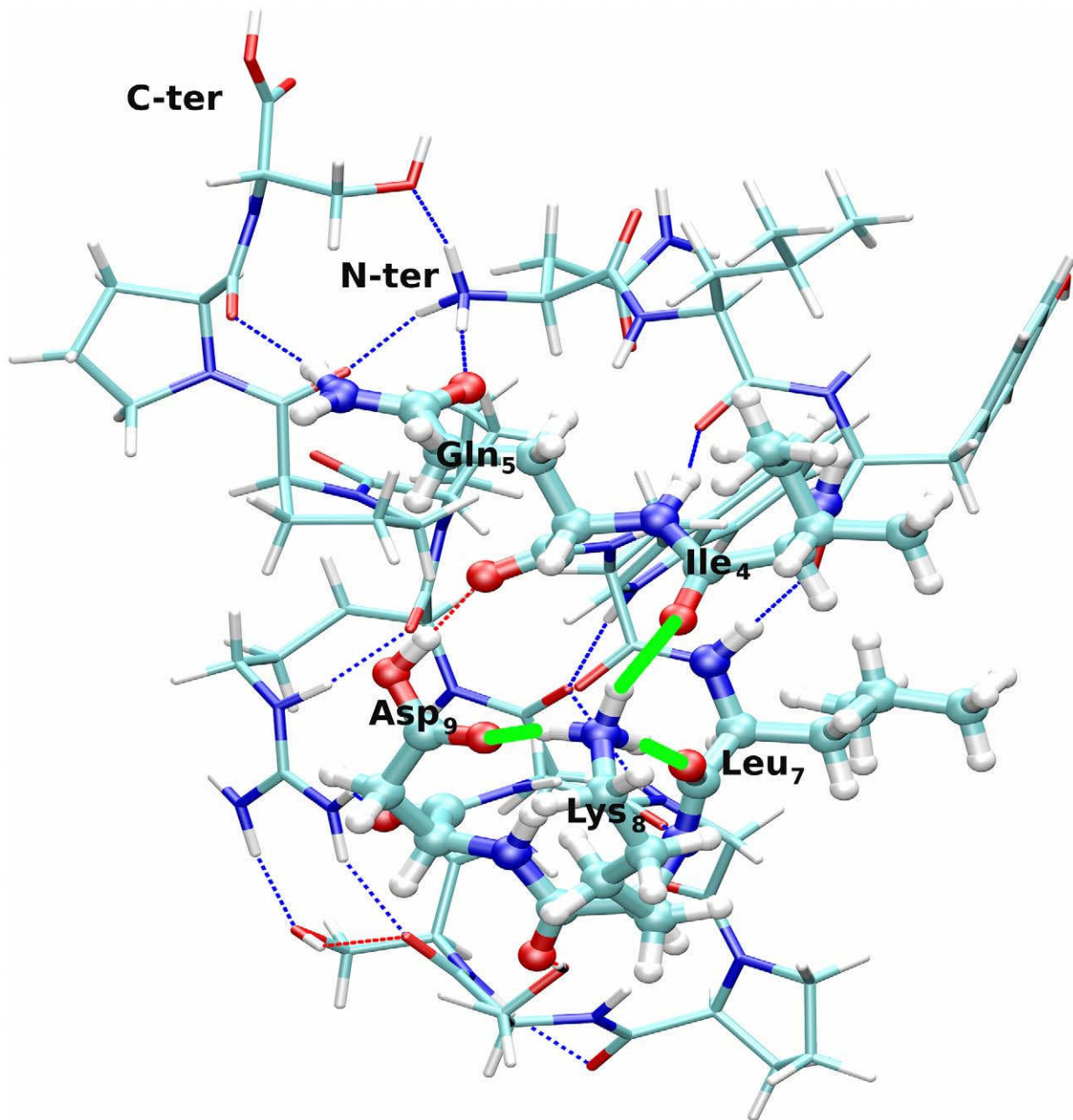


Figure 1. Ball-and-stick representation of the structure of the lowest-energy protomer $\text{trp}^{(3+)}[\text{n}^+ \text{Q}_5 \text{K}_8^+ \text{D}_9 \text{R}_{16}^+ \text{c}]$ of Trp-cage at $q=3+$. The H-bonding network of protonated Lys_8 is shown as green lines; other H-bonds are shown as dotted lines.
doi:10.1371/journal.pcbi.1000775.g001



is disfavored in the gas phase by a positive free-energy change $\Delta\text{GPB} = (\text{GPB})_{\text{A}} - (\text{GPB})_{\text{B}} (>0)$, where $(\text{GPB})_{\text{A}}$ and $(\text{GPB})_{\text{B}}$ stand for the intrinsic GPB of A^- and B, respectively. The larger the ΔGPB , the larger the expected destabilization due to charge separation. In solution, solvation of the charged moieties may counterbalance this energetic penalty [36,44]. The present results indicate that in the gas phase zwitterions can still be stabilized [3,35,38,39,51,63,111]. Indeed, the propensity of peptides to have a low energy correlates nicely with the number of ionized residues, as well as with ΔGPB (see Figure 2, first row, see also SI). For each charge state, ΔGPB values are relative to the protomer with the

lowest charge separation for which ΔGPB has been set to zero. This is caused by the fact that intramolecular interactions can counterbalance this penalty [3], stabilizing the zwitterionic forms. In fact, such interactions, including (i) salt bridges and (ii) H-bonds, may reduce differences in the apparent GPB between basic residues and the conjugated base of acidic residues.

- (i) The formation of salt-bridges correlates well with the intrinsic GPB penalty (see Figure 2, last row). A major stabilizing contribution is therefore the formation of favorable electrostatic interactions. As an example, the formation of four salt-bridges in the species $\text{an}^{(+)}[\text{n}^+ \text{D}_1^- \text{R}_2^+ \text{H}_6^+ \text{c}^-]$ and $\text{trp}^{(0)}[\text{n}^+ \text{Q}_5 \text{K}_8 \text{D}_9^- \text{R}_{16}^+ \text{c}^-]$ largely counterbalances a very

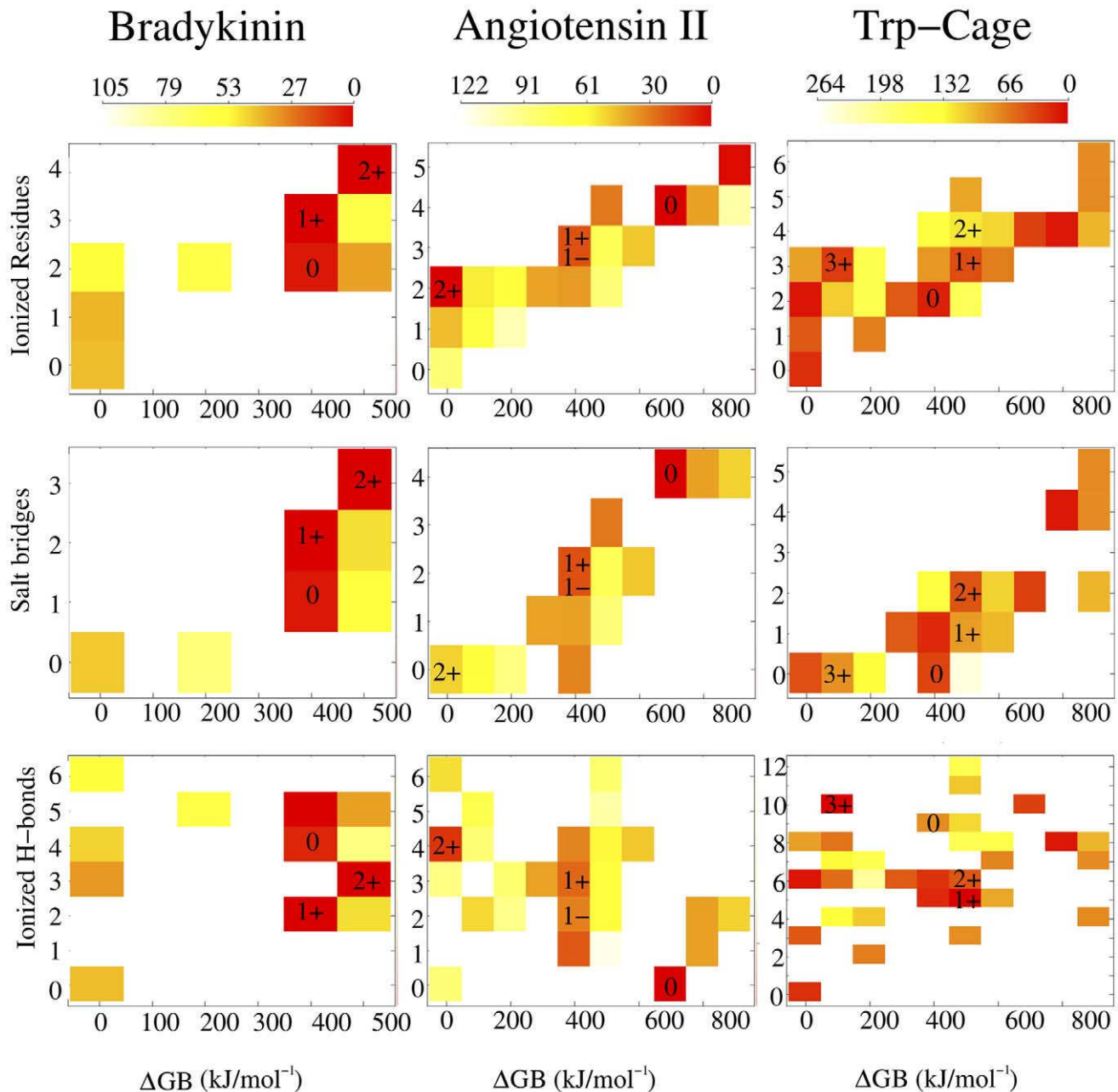


Figure 2. Energy/structure relationship for angiotensin II, bradykinin, and Trp-cage. The color scale refers to the average energy (kJ/mol) of polypeptides with the given pair of parameters. ΔGPB stands for thermodynamic penalty to create a zwitterionic state expressed in terms of residues intrinsic gas-phase basicities (see text for the definition). In each panel, numbers indicate the location of lowest-energy protomers. doi:10.1371/journal.pcbi.1000775.g002

high value of ΔGPB (897 kJ/mol and 842 kJ/mol, respectively). An examination of Tables 1, 2, and 3 suggests that also the salt-bridge geometry is relevant for stabilization. Indeed, consistently with aqueous solution analysis on proteins [112], H-bonded salt-bridges seem to have a major stabilizing effect.

- (ii) ΔGPB , instead, does not correlate either with the number of ionized H-bonds (see Figure 2, third row), or with the total number of H-bonds (see Figure 2, last row). This observation could be a consequence of the relatively small energetic contribution of H-bonds compared to salt-bridges. However, an important role of H-bonding emerges from the analysis

of some specific protomers. For instance, the lowest-energy protomer of neutral Trp-cage, $\text{trp}^{(0)}[\text{n Q}_5\text{K}_8\text{D}_9\text{R}_{16}\text{c}^-]$, features one charge separation, no salt-bridges and nine ionized H-bonds. The latter therefore must compensate a penalty in intrinsic GPB of 368 kJ/mol. Also informative, in this regard, is the comparison between the fourth and fifth lowest-energy protomers of neutral Trp-cage. Both protomers are located around 25 kJ/mol above the minimum. The former does not present any ionized residues ($\Delta\text{GPB} = 0$ kJ/mol), whereas the latter, $\text{trp}^{(0)}[\text{n Q}_5\text{K}_8^+\text{D}_9\text{R}_{16}\text{c}^-]$, has two ionized residues ($\Delta\text{GPB} = 416$ kJ/mol), which form 6 ionized H-bonds but no

salt-bridges. Other similar cases can be found in Tables 1, 2, and 3.

The role of H-bonds is further elucidated by comparing the GPB of single amino acids in their most probable (“folded”) structure and conformations in which the inner H-bonding has been removed, such as the extended, all-trans side chain conformation (We remark that the latter always corresponds to the minimum in the conformational space of alkanes in the gas phase [113]). For basic amino acids, the former is larger than the latter (see Table 4). This fact indicates a greater propensity to acquire a proton when the excess charge can be internally stabilized by H-bonding with the backbone. In acidic residues, instead the former is smaller than the latter, because intramolecular interactions stabilize the negative charge in the folded conformer. Of course, the self-solvation capability of a single isolated amino acid is quite limited. A higher number of H-bond donors and acceptors in peptides and proteins will amplify this effect. An analysis of the structures obtained in this study indicates that each ionized group tends to satisfy at best the same first solvation shell that characterizes the aqueous environment [114,115] (see Tables 8, 9, and 10 in Text S1). In particular, protonated amino moieties tend to donate three H-bonds (one for each N–H bond) and carboxylates receive four H-bonds in average. The key role of H-bonds is consistent with previous hypotheses [3,36,44]. We remark that, as discussed above, H-bonding is only one contribution. The formation of salt bridges provides further (large) energetic stabilization.

Non-Zwitterionic high charge states. As the net charge of the molecule increases, more and more basic residues are protonated. Therefore, zwitterionic states imply protonation of residues with progressively lower intrinsic GPB. Consequently, an increasing number of compensating interactions is required. Such compensation might fail because of (i) insufficient strength of interactions or (ii) topological constraints. As for (i), this is clearly the case of Trp-cage at $q=3+$, where for the creation of a zwitterionic species it is necessary to protonate residues with low basicity, *e.g.*, Gln₈, which is not counterbalanced by an adequate internal solvation. As for (ii), we remark that the location of protonation sites along the primary sequence considerably influences the possibilities of internal solvation, especially in small unstructured peptides. For instance, the two ionizable side chains in BK (Arg₁ and Arg₉) are located at the N- and C-termini, which favor the optimization of the intramolecular interactions for every charge state, because of the flexibility of the peptide backbone. In large proteins, these considerations might be less relevant since the charge-solvation possibilities increase

tremendously, and because the folded structure generally offers a favorable environment for ionized side chains.

In summary, as the net charge increases it becomes progressively more difficult to overcome the thermodynamic penalty of charge separation.

Discussion

A computational protocol aimed at identifying the most stable species of angiotensin II, bradykinin, and tryptophan-cage has been developed and may be easily extended to other systems of similar size. The protocol provides results fully consistent with the experimental data. The results suggest that most of the low-charge states are zwitterions. Intramolecular interactions can stabilize zwitterionic states considerably, by reducing the differences in apparent GPB between basic residues and the conjugated base of acidic residues. Based on a combined structural and energetic analysis, we suggest that salt-bridges provide a key energetic stabilization, in agreement with previous findings [3,38,51,63,116]. Indeed, the stabilization due to salt bridging might be such to reduce enormously the GPB of the biomolecules considered in the present study (up to 900 kJ/mol). H-bonding also has an important role in promoting charge separation. As a result, networks are formed where two (or more) salt bridges are clustered together, whenever it is possible.

Thus, we further corroborate the hypothesis that deprotonated carboxylate groups can be maintained in gas-phase peptide and protein ions produced by electrospray in positive-ion mode (and, vice-versa, protonated basic groups in negative-ion mode) [1,2,30,38,39,62,63,111]. On the other hand, the formation of zwitterionic species in high charge states requires the protonation of residues with progressively lower GPB, which is accompanied by a large thermodynamic penalty that might not be compensated by internal solvation.

Supporting Information

Text S1 Supplementary Material. Detailed structural analysis of the protomers and data to support the DFT calculations.

Found at: doi:10.1371/journal.pcbi.1000775.s001 (1.04 MB PDF)

Author Contributions

Conceived and designed the experiments: RG PC. Analyzed the data: SR. Wrote the paper: RG SR. Designed, developed code, performed, and analyzed the simulations, and drafted the article: RM. Contributed to discussion of the results and critically revised the draft: PC. Conceived and designed the simulations: RG SR.

References

- Konermann L (2007) A minimalist model for exploring conformational effects on the electrospray charge state distribution of proteins. *J Phys Chem B* 111: 6534–43.
- Patriksson A, Marklund E, van der Spoel D (2007) Protein structures under electrospray conditions. *Biochemistry* 46: 933–45.
- Jarrold MF (2000) Peptides and proteins in the vapor phase. *Annual review of physical chemistry* 51: 179–207.
- Jarrold MF (2007) Helices and sheets in vacuo. *Phys Chem Chem Phys* 9: 1659–1671.
- Kaltashov I, Mohimen A (2005) Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. *Anal Chem* 77: 5370–5379.
- Rodríguez-Cruz SE, Klassen JS, Williams ER (1999) Hydration of gas-phase ions formed by electrospray ionization. *J Am Soc Mass Spectrom* 10: 958–68.
- Steinberg MZ, Elber R, McLafferty FW, Gerber RB, Breuker K (2008) Early structural evolution of native cytochrome c after solvent removal. *Chem Bio Chem* 9: 2417–2423.
- Segev E, Wyttenbach T, Bowers MT, Gerber RB (2008) Conformational evolution of ubiquitin ions in electrospray mass spectrometry: molecular dynamics simulations at gradually increasing temperatures. *Phys Chem Chem Phys* 10: 3077–3082.
- Benesch J, Ruotolo B, Simmons D, Robinson CV (2007) Protein complexes in the gas phase: technology for structural genomics and proteomics. *Chem Rev* 107: 3544–67.
- Wyttenbach T, Bowers MT (2007) Intermolecular interactions in biomolecular systems examined by mass spectrometry. *Annual review of physical chemistry* 58: 511–533.
- Li J, Taraszka J, Counterman A, Clemmer D (1999) Influence of solvent composition and capillary temperature on the conformations of electrosprayed ions: unfolding of compact ubiquitin conformers from pseudonative and denatured solutions. *Int J Mass Spectrom* 187: 37–47.
- Chowdhury S, Katta V, Chait B (1990) Probing conformational-changes in protein by mass-spectrometry. *J Am Chem Soc* 112: 9012–9013.
- Samalikova M, Grandori R (2005) Testing the role of solvent surface tension in protein ionization by electrospray. *J Mass Spectrom* 40: 503–10.
- Patriksson A, Adams CM, Kjeldsen F, Raber J, van der Spoel D, et al. (2006) Prediction of n-c-alpha bond cleavage frequencies in electron capture

- dissociation of trp-cage dications by force-field molecular dynamics simulations. *Int J Mass Spectrom* 248: 124–135.
15. Wales TE, Engen JR (2006) Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass spectrometry reviews* 25: 158–70.
 16. Kjeldsen F, Silivra OA, Zubarev RA (2006) Zwitterionic states in gas-phase polypeptide ions revealed by 157-nm ultra-violet photodissociation. *Chem-Eur J* 12: 7920–7928.
 17. Kjeldsen F, Savitski M, Adams CM, Zubarev RA (2006) Determination of the location of positive charges in gas-phase polypeptide polycations by tandem mass spectrometry. *Int J Mass Spectrom* 252: 204–212.
 18. Robinson EW, Leib RD, Williams ER (2006) The role of conformation on electron capture dissociation of ubiquitin. *J Am Soc Mass Spectrom* 17: 1469–1479.
 19. Adams CM, Kjeldsen F, Zubarev RA, Budnik B, Haselmann K (2004) Electron capture dissociation distinguishes a single d-amino acid in a protein and probes the tertiary structure. *J Am Soc Mass Spectrom* 15: 1087–1098.
 20. Jurchen JC, Williams ER (2003) Origin of asymmetric charge partitioning in the dissociation of gas-phase protein homodimers. *J Am Chem Soc* 125: 2817–26.
 21. Jockusch R, Schnier P, Price W, Strittmatter E, Demirev P, et al. (1997) Effects of charge state on fragmentation pathways, dynamics, and activation energies of ubiquitin ions measured by blackbody infrared radiative dissociation. *Anal Chem* 69: 1119–1126.
 22. Carr S, Cassidy CJ (1997) Reactivity and gas-phase acidity determinations of small peptide ions consisting of 11 to 14 amino acid residues. *J Mass Spectrom* 32: 959–967.
 23. Boutin M, Bich C, Afonso C, Fournier F, Tabet JC (2007) Negative-charge driven fragmentations for evidencing zwitterionic forms from doubly charged coppered peptides. *J Mass Spectrom* 42: 25–35.
 24. Polreis F, Gomory I, Schlosser G, Vekey K, Solt I, et al. (2005) Mass spectrometric and quantum-chemical study on the structure, stability, and chirality of protonated serine dimers. *Chem-Eur J* 11: 5908–5916.
 25. Iavarone AT, Patriksson A, van der Spoel D, Parks JH (2007) Fluorescence probe of trp-cage protein conformation in solution and in gas phase. *J Am Chem Soc* 129: 6726–6735.
 26. Gross D, Schnier P, Rodriguez-Cruz SE, Fagerquist C, Williams ER (1996) Conformations and folding of lysozyme ions in vacuo. *Proc Natl Acad Sci USA* 93: 3143–3148.
 27. Miteva M, Demirev P, Karshikoff A (1997) Multiply-protonated protein ions in the gas phase: Calculation of the electrostatic interactions between charged sites. *J Phys Chem B* 101: 9645–9650.
 28. Artega G, Reimann C, Tapia O (2001) Proteins in vacuo: Denaturing and folding mechanisms studied with computer-simulated molecular dynamics. *Mass spectrometry reviews* 20: 402–422.
 29. Artega G, Tapia O (2001) Structural transitions in neutral and charged proteins in vacuo. *J Mol Graph Model* 19: 102–118.
 30. Grandori R (2003) Origin of the conformation dependence of protein charge-state distributions in electrospray ionization mass spectrometry. *J Mass Spectrom* 38: 11–15.
 31. Wytenbach T, Liu D, Bowers MT (2005) Hydration of small peptides. *Int J Mass Spectrom* 240: 221–232.
 32. Chung-Phillips A (2005) Polyglycine conformational analysis: Calculated vs experimental gas-phase basicities and proton affinities. *Journal of physical chemistry A* 109: 5917–5932.
 33. Liu D, Wytenbach T, Barran P, Bowers MT (2003) Sequential hydration of small protonated peptides. *J Am Chem Soc* 125: 8458–8464.
 34. Sudha R, Kohtani M, Jarrold MF (2005) Non-covalent interactions between unsolvated peptides: Helical complexes based on acid⁺/base interactions. *J Phys Chem B* 109: 6442–6447.
 35. Chakraborty T, Holm AIS, Hvelplund P, Nielsen SB, Pouilly JC, et al. (2006) On the survival of peptide cations after electron capture: Role of internal hydrogen bonding and microsolvation. *J Am Soc Mass Spectrom* 17: 1675–1680.
 36. Xu S, Nilles M, Bowen KH (2003) Zwitterion formation in hydrated amino acid, dipole bound anions: How many water molecules are required? *J Chem Phys* 119: 10696–10701.
 37. Jensen J, Gordon M (1995) On the number of water-molecules necessary to stabilize the glycine zwitterion. *J Am Chem Soc* 117: 8159–8170.
 38. Julian R, Hodyss R, Beauchamp J (2001) Salt bridge stabilization of charged zwitterionic arginine aggregates in the gas phase. *J Am Chem Soc* 123: 3577–3583.
 39. Julian R, Beauchamp J, Goddard W (2002) Cooperative salt bridge stabilization of gas-phase zwitterions in neutral arginine clusters. *J Phys Chem A* 106: 32–34.
 40. Bush MF, Prell JS, Saykally RJ, Williams ER (2007) One water molecule stabilizes the cationized arginine zwitterion. *J Am Chem Soc* 129: 13544–13553.
 41. Wytenbach T, Witt M, Bowers MT (1999) On the question of salt bridges of cationized amino acids in the gas phase: glycine and arginine. *Int J Mass Spectrom* 182: 243–252.
 42. Wytenbach T, Witt M, Bowers MT (2000) On the stability of amino acid zwitterions in the gas phase: The influence of derivatization, proton affinity, and alkali ion addition. *J Am Chem Soc* 122: 3458–3464.
 43. Bush MF, Oomens J, Williams ER (2009) Proton affinity and zwitterion stability: New results from infrared spectroscopy and theory of cationized lysine and analogues in the gas phase. *Journal of physical chemistry A* 113: 431–438.
 44. Ramackers R, Pajak J, Lambie B, Maes G (2004) Neutral and zwitterionic glycine.h₂o complexes: A theoretical and matrix-isolation fourier transform infrared study. *J Chem Phys* 120: 4182–4193.
 45. Julian R, Jarrold MF (2004) Gas-phase zwitterions in the absence of a net charge. *Journal of physical chemistry A* 108: 10861–10864.
 46. Pace CN, Treviño S, Prabhakaran E, Scholtz JM (2004) Protein structure, stability and solubility in water and other solvents. *Philos Trans R Soc Lond, B, Biol Sci* 359: 1225–34; discussion 1234–5.
 47. Heck A, van den Heuvel R (2004) Investigation of intact protein complexes by mass spectrometry. *Mass spectrometry reviews* 23: 368–389.
 48. Peschke M, Verkerk U, Kebarle P (2004) Prediction of the charge states of folded proteins in electrospray ionization. *Eur J Mass Spectrom* 10: 993–1002.
 49. Peschke M, Blades A, Kebarle P (2002) Charged states of proteins. reactions of doubly protonated alkyldiamines with nh₃: Solvation or deprotonation. extension of two proton cases to multiply protonated globular proteins observed in the gas phase. *J Am Chem Soc* 124: 11519–11530.
 50. Nguyen S, Fenn JB (2007) Free full text gas-phase ions of solute species from charged droplets of solutions. *Proc Natl Acad Sci USA* 104: 1111–1117.
 51. Touboul D, Jecklin MC, Zenobi R (2008) Investigation of deprotonation reactions on globular and denatured proteins at atmospheric pressure by esims. *J Am Soc Mass Spectrom* 19: 455–466.
 52. Patriksson A, Adams CM, Kjeldsen F, Zubarev RA, van der Spoel D (2007) A direct comparison of protein structure in the gas and solution phase: The trp-cage. *J Phys Chem B* 111: 13147–13150.
 53. Prakash H, Mazumdar S (2005) Direct correlation of the crystal structure of proteins with the maximum positive and negative charge states of gaseous protein ions produced by electrospray ionization. *J Am Soc Mass Spectrom* 16: 1409–1421.
 54. Schnier P, Gross D, Williams ER (1995) On the maximum charge-state and proton-transfer reactivity of peptide and protein ions formed by electrospray-ionization. *J Am Soc Mass Spectrom* 6: 1086–1097.
 55. Iavarone AT, Williams ER (2003) Mechanism of charging and supercharging molecules in electrospray ionization. *J Am Chem Soc* 125: 2319–27.
 56. Fenn J (1993) Ion formation from charged droplets - roles of geometry, energy, and time. *J Am Soc Mass Spectrom* 4: 524–535.
 57. Schnier P, Price WD, Williams E (1996) Modeling the maximum charge state of arginine-containing peptide ions formed by electrospray ionization. *J Am Soc Mass Spectrom* 7: 972–976.
 58. Farrugia J, O'Hair R (2003) Involvement of salt bridges in a novel gas phase rearrangement of protonated arginine-containing dipeptides which precedes fragmentation. *Int J Mass Spectrom* 222: 229–242.
 59. Freitas M, Marshall A (1999) Rate and extent of gas-phase hydrogen/deuterium exchange of bradykinins: evidence for peptide zwitterions in the gas phase. *Int J Mass Spectrom* 182: 221–231.
 60. Polfer NC, Dunbar RC, Oomens J (2007) Observation of zwitterion formation in the gas-phase h/d-exchange with ch₃od: Solution-phase structures in the gas phase. *J Am Soc Mass Spectrom* 18: 512–516.
 61. Iavarone AT, Parks JH (2005) Conformational change in unsolvated trp-cage protein probed by fluorescence. *J Am Chem Soc* 127: 8606–8607.
 62. Rodriguez C, Orlova G, Guo Y, Li X, Siu C, et al. (2006) Gaseous bradykinin and its singly, doubly, and triply protonated forms: A first-principles study. *J Phys Chem B* 110: 7528–7537.
 63. Strittmatter E, Williams ER (2000) Structures of protonated arginine dimer and bradykinin investigated by density functional theory: Further support for stable gas-phase salt bridges. *Journal of physical chemistry A* 104: 6069–6076.
 64. Schnier P, Price WD, Jockusch RA, Williams ER (1996) Blackbody infrared radiative dissociation of bradykinin and its analogues: Energetics, dynamics, and evidence for salt-bridge structures in the gas phase. *J Am Chem Soc* 118: 7178–7189.
 65. Summerfield S, Whiting A, Gaskell S (1997) Intra-ionic interactions in electrosprayed peptide ions. *Int J Mass Spectrom* 162: 149–161.
 66. Cerda BA, Wesdemiotis C (2000) Zwitterionic vs. charge-solvated structures in the binding of arginine to alkali metal ions in the gas phase. *Analyst* 125: 657–660.
 67. Rodgers M, Campbell S, Marzluff E, Beauchamp J (1995) Site-specific protonation directs low-energy dissociation pathways of dinucleotides in the gas-phase. *Int J Mass Spectrom* 148: 1–23.
 68. Lifshitz C (2004) A review of gas-phase h/d exchange experiments: The protonated arginine dimer and bradykinin nonapeptide systems. *Int J Mass Spectrom* 234: 63–70.
 69. Geller O, Lifshitz C (2003) Gas phase h/d exchange of protonated arginine monomers and dimers. *Journal of physical chemistry A* 107: 5654–5659.
 70. Touboul D, Jecklin MC, Zenobi R (2007) Rapid and precise measurements of gas-phase basicity of peptides and proteins at atmospheric pressure by electrostatic spray ionization-mass spectrometry. *J Phys Chem B* 111: 11629–11631.
 71. Price WD, Schnier P, Williams ER (1996) Tandem mass spectrometry of large biomolecule ions by blackbody infrared radiative dissociation. *Anal Chem* 68: 859–866.

72. Morgan J, Russell D (2006) Comparative studies of 193-nm photodissociation and ion-trap analysis of bradykinin analogues: The effects of charge site(s) and fragmentation timescales. *J Am Soc Mass Spectrom* 17: 721–729.
73. Mao D, Douglas D (2003) H/D exchange of gas phase bradykinin ions in a linear quadrupole ion trap. *J Am Soc Mass Spectrom* 14: 85–94.
74. Adams J, Strobel F, Reiter A, Sullards MC (1996) The importance of charge-separation reactions in tandem mass spectrometry of doubly protonated angiotensin II formed by electrospray ionization: Experimental considerations and structural implications. *J Am Soc Mass Spectrom* 7: 30–41.
75. Sullards MC, Reiter J (2000) Primary and secondary locations of charge sites in angiotensin II (m+2h)(2+) ions formed by electrospray ionization. *J Am Soc Mass Spectrom* 11: 40–53.
76. Samalikova M, Grandori R (2003) Role of opposite charges in protein electrospray ionization mass spectrometry. *J Mass Spectrom* 38: 941–947.
77. Adams CM, Kjeldsen F, Patriksson A, van der Spoel D, Graslund A, et al. (2006) Probing solution- and gas-phase structures of trp-cage cations by chiral substitution and spectroscopic techniques. *Int J Mass Spectrom* 253: 263–273.
78. Neidigh J, Fesinmeyer R, Andersen N (2002) Designing a 20-residue protein. *Nat Struct Biol* 9: 425–430.
79. Wyttenbach T, Vonghelen G, Bowers MT (1996) Gas-phase conformation of biological molecules: Bradykinin. *J Am Chem Soc* 118: 8355–8364.
80. Pallante GA, Cassidy CJ (2002) Effects of peptide chain length on the gas-phase proton transfer properties of doubly-protonated ions from bradykinin and its n-terminal fragment peptides. *Int J Mass Spectrom* 219: 115–131.
81. Jorgensen W, Maxwell D, Tirado-rives J (1996) Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118: 11225–11236.
82. Jorgensen W, Tirado-Rives J (2005) Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc Natl Acad Sci USA* 102: 6665–6670.
83. Hess B, Bekker H, Berendsen H, Fraaije J (1997) Lincs: A linear constraint solver for molecular simulations. *J Comput Chem* 18: 1463–1472.
84. Impey R, Nose S, Klein M (1983) Polymorphic phase-transitions in alkali cyanide crystals. *Mol Phys* 50: 243–246.
85. Lindahl E, Hess B, van der Spoel D (2001) Gromacs 3.0: a package for molecular simulation and trajectory analysis. *J Mol Model* 7: 306–317.
86. Becke A (1988) Density-functional exchange-energy approximation with correct asymptotic-behavior. *Phys Rev A* 38: 3098–3100.
87. Lee C, Yang W, Parr R (1988) Development of the colle-salvetti correlation-energy formula into a functional of the electron-density. *Phys Rev B* 37: 785–789.
88. Lippert R, Arias T, Edelman A (1998) Multiscale computation with interpolating wavelets. *J Comput Phys* 140: 278–310.
89. VandeVondele J, Krack M, Mohamed F, Parrinello M, Chassaing T, et al. (2005) Quickstep: Fast and accurate density functional calculations using a mixed gaussian and plane waves approach. *Comput Phys Commun* 167: 103–128.
90. Goedecker S, Teter M, Hutter J (1996) Separable dual-space gaussian pseudopotentials. *Phys Rev B* 54: 1703–1710.
91. Hartwigsen C, Goedecker S, Hutter J (1998) Relativistic separable dual-space gaussian pseudopotentials from h to rn. *Phys Rev B* 58: 3641–3662.
92. Martyna G, Tuckerman M (1999) A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *J Chem Phys* 110: 2810–2821.
93. The cp2k developers group, <http://cp2k.berlios.de/>, 2004.
94. Guidon M, Schiffmann F, Hutter J, VandeVondele J (2008) Ab initio molecular dynamics using hybrid density functionals. *J Chem Phys* 128: 214104.
95. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, et al. Gaussian 03, revision c.02. Gaussian 03, Revision C02..
96. Valdes H, Pluhackova K, Pitonak M, Rezac J, Hobza P (2008) Benchmark database on isolated small peptides containing an aromatic side chain: comparison between wave function and density functional theory methods and empirical force field. *Phys Chem Chem Phys* 10: 2747–2757.
97. Reha D, Valdes H, Vondrasek J, Hozba P, Abu-Riziq A, et al. (2005) Structure and ir spectrum of phenylalanyl-L-glycyl-L-glycine tripeptide in the gas-phase: Ir/uv experiments, ab initio quantum chemical calculations, and molecular dynamic simulations. *Chem Eur J* 11: 6803–6817.
98. Beachy M, Chasman D, Murphy R, Halgren T, Friesner R (1997) Accurate ab initio quantum chemical determination of the relative energetics of peptide conformations and assessment of empirical force fields. *J Am Chem Soc* 119: 5908–5920.
99. Berka K, Laskowski R, Riley KE, Hobza P, Vondrekasek J (2009) Representative amino acid side chain interactions in proteins. a comparison of highly accurate correlated ab initio quantum chemical and empirical potential procedures. *J Chem Theory Comput* 5: 982–992.
100. Halgren T (1996) Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff 94. *J Comput Chem* 17: 490–519.
101. Sulpizi M, Raugi S, VandeVondele J, Carloni P, Sprik M (2007) Calculation of redox properties: Understanding short- and long-range effects in rubredoxin. *J Phys Chem B* 111: 3969–3976.
102. Gervasio F, Carloni P, Parrinello M (2002) Electronic structure of wet dna. *Phys Rev Lett* 89: 108102.
103. Dinadayalan TC, Sastry GN, Leszczynski J (2006) Comprehensive theoretical study towards the accurate proton affinity values of naturally occurring amino acids. *Int J Quantum Chem* 106: 2920–2933.
104. Rak J, Skurski P, Simons J, Gutowski M (2001) Low-energy tautomers and conformers of neutral and protonated arginine. *J Am Chem Soc* 123: 11695–11707.
105. Li Z, Matus MH, Velazquez HA, Dixon DA, Cassidy CJ (2007) Gas-phase acidities of aspartic acid, glutamic acid, and their amino acid amides. *Int J Mass Spectrom* 265: 213–223.
106. Hunter E, Lias S (1998) Evaluated gas phase basicities and proton affinities of molecules: An update. *J Phys Chem Ref Data* 27: 413–656.
107. Bojesen G, Breindahl T (1994) On the proton affinity of some alpha-amino-acids and the theory of the kinetic method. *J Chem Soc Perk T 2*: 1029–1037.
108. Gorman G, Speir J, Turner C, Amster I (1992) Proton affinities of the 20 common alpha-amino-acids. *J Am Chem Soc* 114: 3986–3988.
109. Li X, Harrison A (1993) A kinetic approach to the proton affinities of amine bases. *Org Mass Spectrom* 28: 366–371.
110. Ewing NP, Pallante GA, Zhang X, Cassidy CJ (2001) Gas-phase basicities for ions from bradykinin and its des-arginine analogues. *J Mass Spectrom* 36: 875–881.
111. Lemoff A, Bush MF, Williams ER (2005) Structures of cationized proline analogues: evidence for the zwitterionic form. *J Phys Chem A* 109: 1903–1910.
112. Kumar S, Nussinov R (1999) Salt bridge stability in monomeric proteins. *J Mol Biol* 293: 1241–1255.
113. Li H, Maroncelli M (2006) Solvation and solvatochromism in co2-expanded liquids. 1. simulations of the solvent systems co2 + cyclohexane, acetonitrile, and methanol. *J Phys Chem B* 110: 21189–21197.
114. Peraro MD, Raugi S, Carloni P, Klein M (2005) Solute-solvent charge transfer in aqueous solution. *Chemphyschem* 6: 1715–1718.
115. Hugosson H, Laio A, Maurer P, Rothlisberger U (2006) A comparative theoretical study of dipeptide solvation in water. *J Comput Chem* 27: 672–684.
116. Strittmatter E, Wong R, Williams E (2000) Effects of gas-phase basicity on the proton transfer between organic bases and trifluoroacetic acid in the gas phase: Energetics of charge solvation and salt bridges. *Journal of physical chemistry A* 104: 10271–10279.