International Journal of Mass Spectrometry 289 (2010) 84–91

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms

Effects of salts on the charge-state distribution and the structural basis of the most-intense charge-state of the gaseous protein ions produced by electrospray ionization

Halan Prakash, Bharat T. Kansara, Shyamalava Mazumdar*

Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400005, India

ARTICLE INFO

Article history:
Received 20 April 2009
Received in revised form
25 September 2009
Accepted 28 September 2009
Available online 27 October 2009

Keywords:
Electrospray ionization mass spectrometry
Charge-states of protein ions
Surface accessibility
Conjugate acid
Conjugate base

ABSTRACT

The charge-state distribution in the electrospray ionization mass spectra (ESI-MS) of different proteins in the absence and presence of various partially volatile salts such as ammonium sulphate, potassium acetate and ammonium acetate as well as in the presence of some non-volatile salts are reported. The effects of the salts on the charge-states of the proteins elucidated the protein–salt particle interactions and the charge-transfer reaction that is recognized as an important mechanism of protonation or deprotonation of proteins. These results also indicated that the surface accessibility of the polar residues is an important factor for the interaction between the protein and the salt particles. The present studies could show a nice correlation of the most-intense charge-state of the gaseous protein ions with the surface exposed free basic (SEFBR) and free acidic (SEFAR) residues obtained from the crystal structures of the proteins.

1. Introduction

Electrospray ionization (ESI) is a soft ionization method versatile for the ionization of large proteins [1–12]. Extensive ESI mass spectrometric studies on the conformational properties of various proteins have been reported, and it is now well recognized that the charge-states of the proteins in the ESI-mass spectra (ESI-MS) depend on the tertiary structure of the proteins in solution [2,4,13,14]. The ESI-MS of folded protein solutions often show a monomodal distribution of charge-states peaks with a maximum charge-state and a most-intense charge-state [13–15]. On unfolding of the protein, the charge-state distribution in the ESI-mass spectra of the protein shifts to larger positive charge-states and becomes broader leading to bimodal charge-state distribution [13–15]. Several reports have elucidated the role of solvent accessibility and coulombic repulsions to interpret the ESI-MS of proteins [14–19]. The Conformation-Dependent Charge Neutralization Theory (CDNT) recognized the role of the intramolecular interactions in folded proteins in determining the observed charge-state distributions [18]. According to the Rayleigh Limiting Charge Theory (RLCT), the maximum ESI charge-state of the protein depends on the size of the molecule [19]. On the other hand, the CDNT proposed that the native protein structures stabilize pre-existing charges of the opposite polarity to the net charge of the ion, preventing their neutralization during the electrospray process and the most-intense ESI charge-state of a folded protein was shown to correspond to the difference between the number of basic residues and acidic residues in a protein sequence [16,18]. Proteins generally show ESI-MS both in the positive and negative ion modes, and the maximum charge-state as well as the most-intense charge-state were found to be different in the positive and negative ion modes, which is yet not understood unambiguously.

We have recently shown that the maximum charge-state of the ESI-MS of the native protein solutions could be related to the number of basic and acid residues of the protein, and a direct correlation between the crystal structure and the maximum ESI-MS charge-state was observed for a large number of proteins [12]. The surface exposed free basic residues (SEFBR) were proposed to act as the charge carriers in the positive ESI ion mode and the surface exposed free acidic residues (SEFAR) act as the negative charge carriers in the negative ESI ion mode [12]. The SEFBR or SEFAR are the residues that are not involved in any interaction with the neighboring residues [12] in the protein. The maximum charge-state of many proteins in both the positive and negative ion modes could be rationalized using this model [12].

The ESI-MS of proteins obtained in the presence of salts such as ammonium acetate was earlier shown to give peaks corre-

* Corresponding author. Tel.: +91 22 2278 2363; fax: +91 22 2280 4610.
E-mail addresses: shyamal@tifr.res.in, mshyamal@gmail.com (S. Mazumdar).
URL: http://www.tifr.res.in/~shyamal (S. Mazumdar).

1387-3806/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2009.09.013

sponding to the protein–salt adduct, which indicate that there are protein–salt particle interactions and charge-transfer reactions between the proteins and dissociated ions of the salts [6,7,9,20]. We have also earlier shown that the number of SEFAR of proteins (e.g. Cytochorme *c* and myoglobin) correlates with the number of potassium adducts peaks observed for the protein in the presence of potassium salts [12]. This suggested the charge-transfer reactions between the polar residues (acidic) of the proteins and the charged particles (potassium cations) that could envelop the protein ions in the ESI and subsequent evaporation [12] processes. Felitsyn et al. [6] and Peschke et al. [9] reasoned that in presence of the volatile salt like ammonium acetate, the ammonium cations bind to the basic residues of the protein and there is proton transfer reaction from the ammonium ion to the basic residues of the protein with formation of ammonia (base) and the protonated protein ion. It was also reported that the acetate ion can bind to the protonated basic residues of the protein, which can abstract proton and readily evaporate as acetic acid (volatile). On the other hand, other non-volatile anions such as halides [6,9,16,18,20] would bind to the basic sites of the protein surface and form protein-charged particle adduct leading to formation of zwitterions containing the halide and the basic group on the protein surface. Thus, in the charge-transfer reactions between the dissociated ions of the salts and proteins, acids and bases could be formed [20]. These acids and bases could differ in their relative volatility, and could influence both the response and the charge-state distribution of the ESI-MS of proteins. Therefore, the type of the salts is important for the charging of the proteins [20]. Ammonium acetate generally is considered as a versatile salt for the ESI-MS of proteins, which is mainly attributed to the possibility of the charge-transfer by ammonium cations and acetate anions, and the high volatility of ammonia (base) and acetic acid that could be formed due to the charge-transfer reactions [20].

Apart from the relative volatility of the charged particles as mentioned above [20], the relative surface accessibility of the polar residues of the protein for the interaction with the charged particles is also an important factor for the charge-transfer reactions between the polar residues of the protein and the charged particles around the protein that could occur during the formation of the protonated/deprotonated protein ions.

In view of this, here we have carried out a systematic study of the charge-state distribution of Cytochrome *c* and myoglobin in the absence and presence of various salts of varying volatility such as ammonium acetate, ammonium sulphate, potassium acetate, along with potassium nitrate, potassium chloride, etc. We also compared the relative surface accessibility of the SEFBR and SEFAR of various proteins with the most probable charge-state, i.e., most-intense charge-state of the protein. The results of the salts effects on the charge-state distribution of the protein emphasize the protein–salt interactions and charge-transfer reactions as proposed earlier [6,9,20], and indicate that the relative volatility of the charged particles (from salt) is an important factor that could influence the formation of charge-states of the protein. Moreover, the results show that the relative surface accessibility of the SEFBR and SEFAR of various proteins and the most probable charge-state are correlated, which further supports the protein–salt interactions and charge-transfer reactions prior to the formation of the final protonated or deprotonated protein ions. A model has been proposed based on the surface accessibility of the SEFBR and SEFAR to predict the most probable charge-state of the protein.

2. Experimental

2.1. Chemicals

Horse heart Cytochrome *c*, horse myoglobin and ammonium acetate were obtained from Sigma, USA. Cytochrome *c* was purified by G-25 size exclusion chromatography with the corresponding solvent, before recording the ESI-mass spectra. Myoglobin solutions were purified by Chelating Sepharose to remove any unfolded protein impurity just before experiment. The purity of the sample was checked by SDS PAGE and absorption spectroscopy. Potassium acetate and ammonium sulphate, potassium nitrate, potassium chloride, sodium chloride and ammonium nitrate were of AR grade obtained from Qualigens, India. Milli-Q water (Millipore) was used for all the experiments.

2.2. Surface accessibility of individual atoms

The accessible molecular surface (for solvent) is that part of the van der Waals surface that can be touched by the probe (water) with a radius of 1.4 Å. This molecular accessible surface is in $Å^2$. The calculated values is obtained from by What IF server [21,22] (http://swift.cmbi.kun.nl/WIWWWI/). The crystal structures of the proteins were obtained from Research Collaboratory for Structural Bioinformatics (RCSB, http://pdbbeta.rcsb.org/pdb/). The surface exposed basic and acid residues that form ion-pair or hydrogen bonding are analyzed from the crystal structure using the What IF server software as reported earlier [12,21,22].

2.3. Electrospray ionization mass spectrometry (ESI-MS)

Mass spectra were obtained using a Thermo Finnigan LCQ Deca Electrospray quadrupole ion trap mass spectrometer. The flow rate of the protein solution was maintained at 1 μl/min. Capillary temperature was maintained at approximately 200 °C and the capillary voltage was kept at 31 V. The ion-spray voltage was maintained at 4.5 kV and the ion optics was tuned to get maximum ion count. The instrument was separately calibrated both in the normal-mass range (0–2000 *m*/*z*) and in the high-mass range (0–4000 *m*/*z*).

3. Results and discussion

The folded forms of Cytochrome *c* (∼ 3 μM) and myoglobin (∼5 μM) show monomodal distribution of multiple charge-states

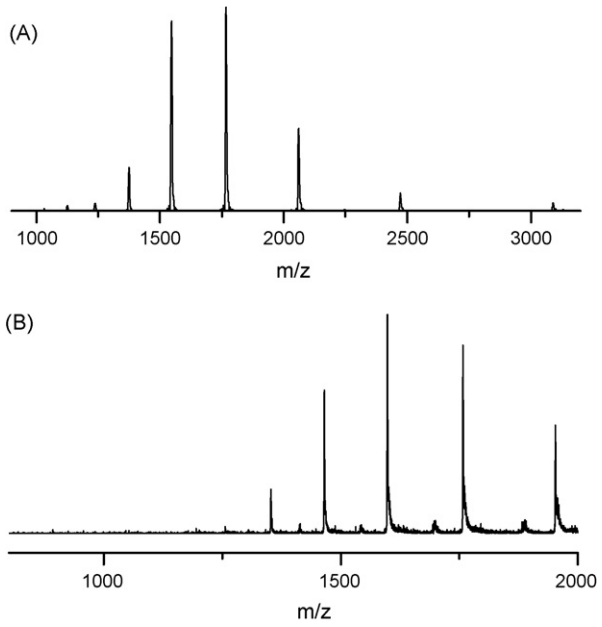


Fig. 1. Positive ion mode ESI-mass spectra of (A) Cytochrome *c* (3 μM) in water and of (B) myoglobin (∼5 μM) in water.

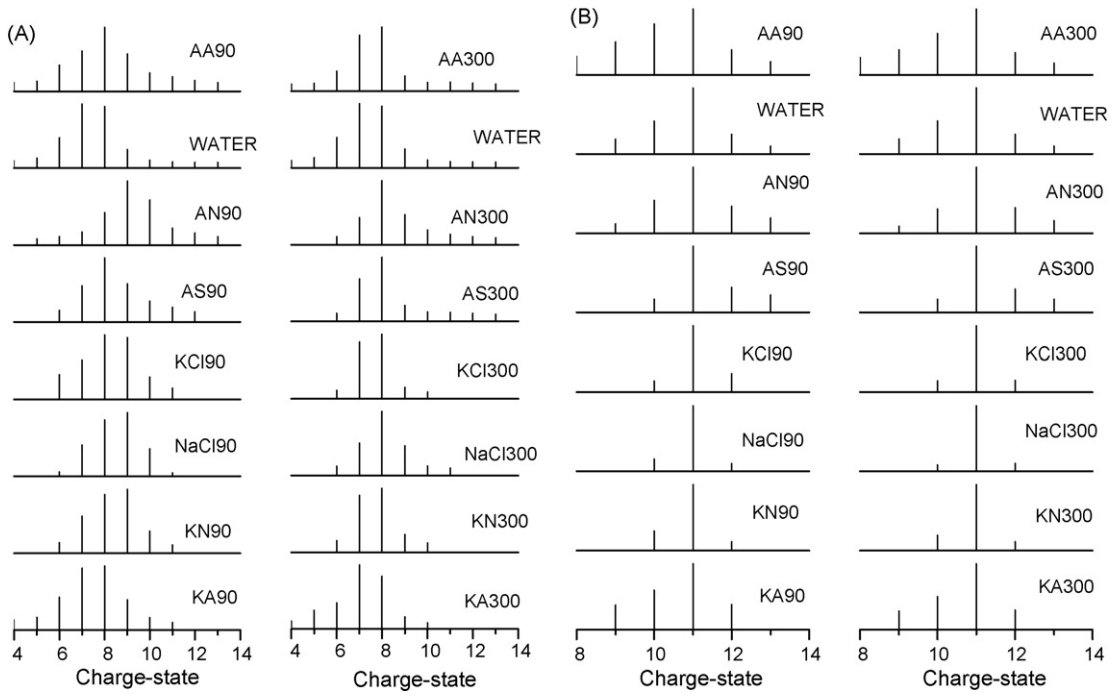


Fig. 2. (A) Schematics of positive ion mode ESI-mass spectra of Cytochrome *c* (3 μM) in ammonium acetate (AA), water, ammonium nitrate (AN), ammonium sulphate (AS), potassium chloride (KCl), sodium chloride (NaCl), potassium nitrate (KN) and potassium acetate (KA). The salt concentrations (90 and 300 μM) are indicated in after the name of the salt (e.g., AA90: 90 μM ammonium acetate). (B) Schematics of positive ion mode ESI-mass spectra of myoglobin (~5 μM) in AA, water, AN, AS, KCl, NaCl, KN and KA. The salt concentrations (90 and 300 μM) are indicated in after the name of the salt (e.g., AA90: 90 μM ammonium acetate).

as shown in Fig. 1. The ESI-MS spectrum of freshly purified Cytochrome *c* in water showed distinct charge-state peaks from +4 to +13 as reported earlier [12] with the most-intense peak at +7/+8 charge-state [14–16,18]. The intensity of the maximum charge-state (+13) is found to be less than 1% with respect to the most-intense charge-state peak (+7). Myoglobin solution often shows charge-states corresponding to the apo-protein formed by depletion of the heme from the holo-protein [12]. However, ESI-MS signals from the holo-protein were predominant in the mass spectrum of freshly purified myoglobin solution in water. Distinct charge-state peaks from +8 to +13 were detected in the ESI-MS spectrum of mytoglobin from water with the most-intense peak at +11 charge-state as reported earlier [12].

The ionization of the protein from water in the electrospray is associated with protonation of the SEFBR [12]. The solvent molecules (water) provide the requisite protons for protonation of the SEFBR's in the protein during the charging process.

$$H_2O + H^+ \rightarrow H_3O^+ \quad \text{(electro-ionizationofwater)}$$

$$\text{Protein} + H_3O^+ \rightarrow \text{Protein–}H^+ + H_2O$$

There are 12 basic SEFBR and one positively charged heme (prosthetic group) in the crystal structure of Cytochrome *c* (PDB code 1HRC), which directly correlate to the maximum charge-state of the protein as shown earlier [12]. Analogously, the crystal structure of myoglobin (PDB code: 1WLA) shows the presence of 12 SEFBR albeit there are 32 basic residues in the protein. Considering the contribution of the N-terminal amine and the heme suggests the maximum charge-state to be +14 which agrees with the present result [12]. However, the nature of the charge-state distribution of the proteins was found to depend on the property of the solvent. The effects of various types of salts on these charge-state distributions are discussed below.

3.1. Comparison of ESI-MS of Cytochrome c in water and in different types of salts

3.1.1. Effect of ammonium acetate on the charge-states

Fig. 2 schematically shows the charge-state distributions of Cytochrome *c* (Fig. 2A) and of myoglobin (Fig. 2B) in the presence of various salts at 90 and 300 μM concentrations. The presence of ammonium acetate (90–300 μM) did not have any significant effect on the charge-state distribution of the proteins. Considering the basic concept of ionization of protein in the ESI technique, the presence of ammonium in the solution would enhance the propensity of protonation of the SEFBR's while the acetate ions would decrease the protonation propensity of the SEFBR's in the protein. These two effects will compete with each other.

$$NH_4OAc \rightleftharpoons NH_4^+ + OAc^-$$

$$\text{Protein} + NH_4^+ \rightleftharpoons \text{Protein–}H^+ + NH_3 \quad \text{(volatilebase)}$$

$$\text{Protein–}H^+ + OAc^- \rightleftharpoons \text{Protein} + AcOH \quad \text{(volatileacid)}$$

Thus ammonium acetate solution would probably not have any extra effect on the charge-state distribution of the protein than that of water as mentioned above, though a possible effect of the volatile ions on the stabilization of the charge-state may not be ruled out. Thus ammonium acetate would not drastically affect the charge-state distribution of the ESI-MS of the protein as observed in Fig. 2A and B. The salt of weak acids or bases will show equilibrium dissociation and thereby may not enhance the ionic concentration significantly, which might also be responsible for the absence of any major effect of ammonium acetate on the charge-state distribution of the proteins.

3.1.2. Effect of ammonium salts with less-volatile acids

Ammonium sulphate (AS) and ammonium nitrate (AN) can be considered as salts of volatile base and non-volatile or less-volatile acids. The results shown in Fig. 2A and B suggest that the lower charge-states +4, +5 are absent in the ESI-MS of Cytochrome *c* in the presence of 90 μM ammonium sulphate (AS90) while the effect is more drastic in case of myoglobin, where even the +9 charge-state was also not detected. Increase in the salt concentration to 300 μM (AS300) does not seem to have any major effect, though the general tendency of decrease in the low-charge-states of the protein is more prominent at higher salt concentration. Similar results were also observed with ammonium nitrate (AN90 and AN300; Fig. 2A and B), though the +9 charge-state was detected in the presence of 90 μM ammonium nitrate, nevertheless the intensity of lower charge-states seem to be decreased also in the presence of ammonium nitrate.

The most-intense charge-state of Cytochrome *c* is shifted from +8 to +9 in the presence of 90 μM ammonium nitrate but such effects were not observed at higher concentrations of the salt or in case of myoglobin. It is important to note that the higher charge-states up to +13 (Fig. 2A and B) were detected in presence of the ammonium salts. The results thus show that only the lower charge-states of the gaseous protein ions are suppressed significantly by the non-volatile anions in case of their ammonium salts.

Acids and bases with different volatility could be formed in the ESI of protein solutions with salts due to the proton transfer reactions and evaporation of the salt particle bound to the protein, as proposed earlier [6,9,20]. Accordingly, both ammonium sulphate and ammonium nitrate being salts of a weak base and strong acids would be highly dissociated to form ammonium and sulphate or nitrate ions. The ammonium ion being more acidic than water, can easily transfer a proton to an SEFBR forming ammonia (volatile) and protonated protein molecule, thus enhancing the propensity of protonation of the SEFBR's in the protein leading to an increase in the relative population of the higher positive charge-states in the ESI charge-state distribution compared to that in water. This effect thus decreases the relative population of the lower charge-states of the gaseous protein ion as seen in Fig. 2A and B.

$$(NH_4)_2SO_4 \rightarrow 2NH_4^+ + SO_4^{2-}$$

$$NH_4NO_3 \rightarrow NH_4^+ + NO_3^-$$

$$\text{Protein} + NH_4^+ \rightleftharpoons \text{Protein–}H^+ + NH_3\,(\text{volatilebase})$$

The sulphate or nitrate ions being non-volatile would either form adducts with the protein by binding to basic sites or will be filtered away during the evaporation process.

Earlier, it has been shown that higher charge-states of the protein ions are formed from the unfolded protein (solution at pH less than 3.5)[14,15]. Both in the presence and absence of thee salts, the protein solution shows the highest the maximum charge-state (Fig. 2A and B) as proposed [12] from the crystal structures (see Table 1) of the proteins. This clearly indicates that the conformation of the protein in the presence of the salt (90–300 μM) is comparable to that in water and thus the changes in the charge-state distribution in the ESI-MS of the protein in the presence of the salt are not associated with any perturbation of tertiary structure of the protein in solution [12].

3.1.3. Effect of potassium acetate on the ESI charge-states

The ESI-MS spectra of Cytochrome *c* and myoglobin in the presence of potassium acetate show that the higher charge-states above +12 are drastically decreased in the presence of the salt (KA90, KA300 in Fig. 2A and B). The most-intense charge-state however remains almost unaffected (+7 or +8 for Cytochrome *c* and +11 for myoglobin) and the lower charge-states (+4 in case of 90 μM salt in Cytochrome *c* and +9 in case of myoglobin) are observed. The effect of potassium acetate on charge-states of the protein ions is thus opposite to that observed in the presence of the ammonium salts of non-volatile ions discussed above.

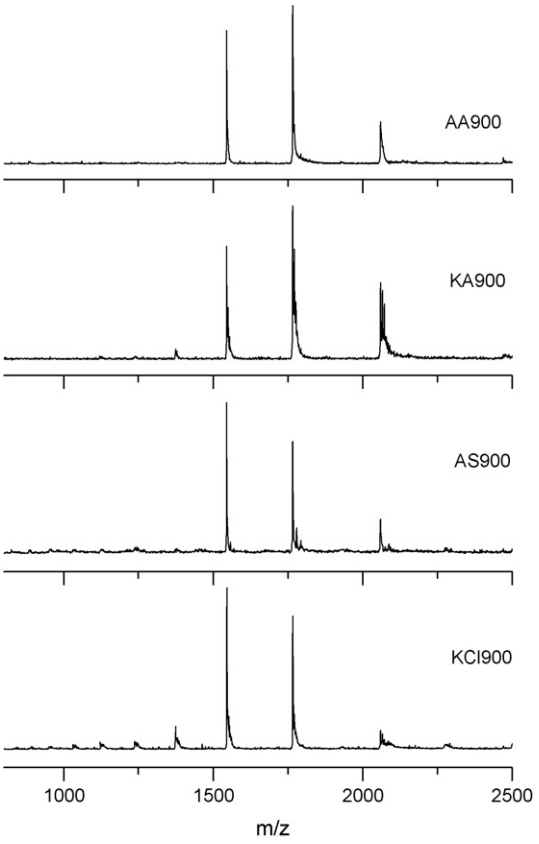


Fig. 3. Positive ion mode ESI-mass spectra of Cytochrome *c* (3 μM) in higher concentration (900 μM) of salts ammonium acetate (AA900), potassium acetate (KA900), ammonium sulphate (AA900) and potassium chloride (KCl900).

Potassium acetate being a salt of a strong base and weak acid would be highly dissociated to form potassium cations and acetate ions. The potassium ion may either bind to the acidic sites in the protein forming adducts with the protein or may filter out during the evaporation process [12,20]. The acetate ion can associate itself to an SEFBR forming acetate adduct of the protein or it can abstract proton from the SEFBR forming volatile acetic acid [12,20]. This process would effectively decrease the propensity of protonation of the SEFBR's in the protein leading to a decrease in the relative population of the higher positive charge-states in the ESI charge-state distribution compared to that in water. This effect thus increases the relative population of the lower charge-states of the gaseous protein ion as seen in Fig. 2.

$$KOAc \rightarrow K^+ + OAC^-$$

$$\text{Protein–}H^+ + OAc^- \rightleftharpoons \text{Protein} + AcOH \quad (\text{volatileacid})$$

3.1.4. Effect of salts of non-volatile acids and bases on the ESI charge-states

Chlorides and nitrates of potassium and sodium are salts of strong base and strong acid that would be highly dissociated to form the cations (K^+, Na^+) and the anions (Cl^-, NO_3^-). The ESI-MS spectra of the proteins was significantly affected by these non-volatile salts and the intensities of both the lower as well as higher charge-states of the protein ions were drastically decreased in the presence

Table 1
Prediction of the most-intense charge-state based on the mean surface accessibility of the basic and acidic residues (SEFBR/SEFAR) of various proteins.

Protein (PDB code) +ve/−ve ion spectra	MW	BR (AR)	SEFBR (SEFAR)	MIC_{Obs}	Number of SEFBR (SEFAR) with accessibility more than the mean value [mean value ($Å^2$)]	Number of SEFBR (SEFAR) with accessibility more than the mean value and the N- or C- terminal group or the prosthetic group	Reference
SrcSH2 domain (1SPR) +ve	11,935	18 (22)	8	+7	5 [13.3]	6	[27]
RNase A (1ROB) +ve	13,674	18 (10)	8	+7	5 [16.5][a]	6	[28]
RNase A (1ROB) −ve			(6)	-6	(5) [5.8]	(6)	[16]
Acyl CoA (1HB6) +ve	9,869	18 (17)	9	+8	7 [14.9]	8	[29]
Lysozyme (1DPX) +ve	14,302	18 (9)	9	+8	6 [11.4]	7	[12,25]
Lysozyme (1DPX) −ve			(5)	−6, −5	(3) [5.7]	(4)	[25]
Chorismate mutase (1DBF) +ve	14,472	19 (19)	10	+5	6 [12.4]	7	[30]
Trp repressor (2WRP) +ve	12,207	15 (16)	11	+7	6 [13.7]	7	[31]
α-Lactalbumin (1F6S) +ve	14,169	16 (20)	11	+8	7 [15.1]	8	[32]
β-Lactaglobulin (1B8E) +ve	18,265	19 (14)	12	+8	8[13.5]	9	[33]
Cytochrome *c* (1HRC) +ve	12,360	24 (12)	12	+7, +8	7 [15.7] [a]	8[A]	This work and [12]
Cytochrome *c* (1HRC) −ve			(4)	−5	(1) [8.3] [b]	(4)[B]	This work and [12]
Myoglobin (1WLA) +ve	17,567	32 (21)	12	+11, +10	6 [11.6]	8[C]	[12]
Myoglobin (1WLA) −ve			(4)	−6	(2) [6.8] [a]	(5)[B]	[12]
Inorganic pyrophosphatase (1IGP) +ve	19,544	25 (29)	14	+9	7 [11.6]	8	[34]
Human serum transferrin (1BP5) +ve	37,153	48 (39)	15	+13	9 [13.4]	10	[35]
Bovine carbonic anhydrase (1V9E) +ve	29,067	38 (30)	16	10	9 [13.7]	10	[12]

PDB: Protein Data Bank; MW: molecular weight (Dalton); BR: total no. of basic residues; AR: total no. of acidic residues, shown in parentheses; MIC: most-intense charge; SEFBR: surface exposed free basic residue; SEFAR: surface exposed free acidic residue are shown in parentheses; A: N-terminal is acetylated hence do not contribute to positive charge; ferric heme contribute to one positive charge [12]. B: C-terminal and two heme propionates contribute to the negative charge; C: N-terminal and ferric heme contribute to the positive charge. (a) One SEFBR or (SEFAR) is close to the mean value; (b) two SEFAR are close to mean value.

of these salts (Fig. 2). However, the most-intense charge-states of the proteins were not affected due to the presence of these salts.

Fig. 3 shows the ESI-MS spectra of Cytochrome *c* in the presence of 900 μM salts. Interestingly, on increasing the concentration of the salts above 0.1 mM (0.3–0.9 mM) the charge-state distribution of the protein becomes narrower (Fig. 3), compared to that in water (Fig. 1) or in the presence of lower (<0.3 mM) salts (Fig. 2). Moreover, though the efficiency of ionization of the proteins is decreased at higher concentration of the salts the most-intense charge-state remains the same as in the presence of lower concentrations of the salts for both of the two proteins studied. The ESI-MS of Cytochrome *c* in the presence of 900 μM KCl was drastically diminished compared to that in the presence of the ammonium or the acetate salts, but the most-intense charge-state still remained the same even in the presence of 900 μM KCl. High concentration of KCl was also found cause partial unfolding of the protein, and weak bimodal charge-state distribution with higher charge-states were observed (Fig. 3, KCl900). It is important to note that the ESI-MS of myoglobin is much more sensitive to the salt concentration and very weak signals were obtained in the presence of 900 μM salts. Nevertheless, the most-intense charge-state peak of myoglobin remained +11 even in the presence of the high salt concentration.

Thus, the above results further support that the salts and the protein indeed interact with each other as was proposed earlier [6,9,20]. Moreover, the variations in the effects of the salts on the charge-state distribution could be due to the differences in the relative volatility of the charged particles that could be formed on dissociation of the salts. The present investigation also clearly indicates that the relative surface accessibility of the SEFBR could be an important factor in the charge-transfer reactions involving the protonation of the protein ions. We have earlier shown that the negative ion mode ESI-MS of Cytochrome *c* gives only two negative charge-states corresponding to −4 and −5, and that of myoglobin gives −5 and −6 charge-states. These results could be rationalized based on the number of SEFAR (surface exposed free acidic residues) and the negative charges on the heme propionates in the protein [12]. Addition of salts drastically affects these two negative charge-states and the presence of non-volatile salts (such as K^+ or SO_4^{2-} containing salts) completely destroys the signals in the negative ion mode. This fact again supports the role of salts in the charging of proteins. The negative ion mass spectra seem to be more significantly affected by the salt concentrations compared to those in the positive ion mode. The negative ion mode spectra of proteins are generally very weak, indicating that the stability of the negative charge on the gaseous protein ions is lower than that of the positive ion species [9,12,17]. The ions of the electrolyte may easily quench the negative charge during the flight or at the ion trap leading to discharge of the negatively charged gaseous protein ions. This type of effect may not be predominant in case of the positive ion mass spectrometry, as the neutralization of the protein ions in such case would require transfer of protons rather than electron; hence the positive ion mass spectra of proteins are generally more intense and stable.

In view of these results, a detailed analysis of the surface accessibility of the polar residues of the protein and its relation with the most probable charge-state of the protein was carried out as discussed below.

3.2. Analysis of the surface accessibility and the correlation of the surface exposed basic or acidic residues with the most-intense charge-state

Analysis of the surface accessibility [21,22] of Cytochrome *c* from the crystal (PDB code 1HRC) structure shows that there are 12 SEFBR present in the protein [12]. The relative surface accessibility of these 12 SEFBR (lysine 5, 7, 8, 22, 25, 53, 60, 73, 86, 87, 88 and 100) are shown in Fig. 4, and the remaining 7 lysine residues in the protein are involved in close interaction with acidic residues hence are not likely to participate in the stabilization of the charge-state of the protein even though they may be exposed to the surface [12]. These 12 lysines (SEFBR) and one heme prosthetic group in Cytochrome *c* was earlier proposed to give rise to a maximum of +13 charge in the ESI-MS [12]. The average of the surface accessibility of these 12 SEFBR was found to be 15.75 Å^2 (shown by a horizontal dashed line in Fig. 4). Among the 12 SEFBR, the surface accessibility of 7 residues (lysine 5, 7,8, 22,60,73, 88) were found to be more than the mean (15.75 Å^2) surface accessibility value, and another one (lysine 86) residue has accessibility close to the mean value.

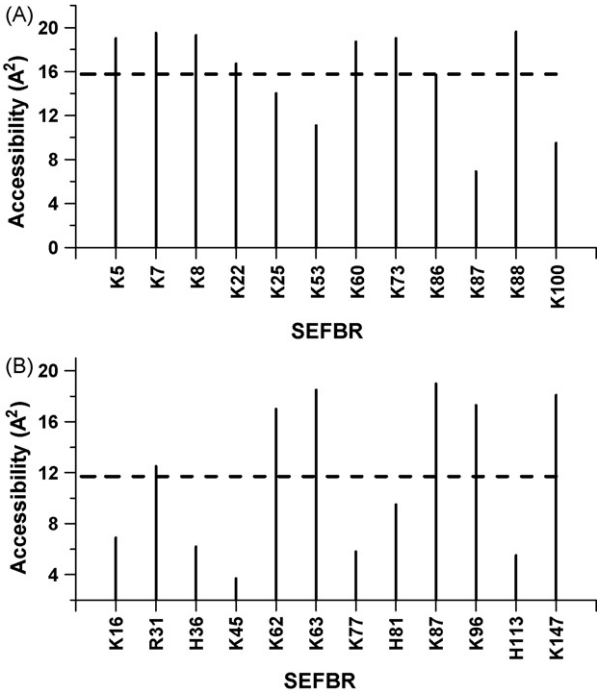


Fig. 4. Relative accessibility of the surface exposed basic groups of (A) Cytochrome *c* and (B) myoglobin obtained from the crystal structures (PDB; 1HRC and 1WLA respectively). The surface exposed free basic residues (SEFBR) which do not form ion pair or hydrogen bound are represented by vertical bars. The horizontal dotted line represents the average surface accessibility value (of the 12 SEFBR). The surface exposed basic residues that form ion-pair or hydrogen bonded are represented by vertical shaded bars.

The ESI-MS of Cytochrome *c* in various salts and in water (Figs. 2 and 3) showed that +7, +8 and +9 charge-states are the most prominent in all cases. It is thus apparent that among the 12 SEFBR's, there are 7 SEFBR's (7 lysine residues with the surface accessibility above 15.75 Å^2), which agrees with the number of most-intense charge-states (+7) in the ESI-MS of the protein. It is also important to note that the heme could contribute to the charge-state and thus considering the 7 SEFBR and the heme, the most probable charge-state of Cytochrome *c* could be 8+, which agrees with that observed experimentally (+7 or +8).

Analysis of the surface accessibility [21,22] of myoglobin from the crystal (PDB code 1WLA) structure shows that there are 12 SEFBR present in the protein [12] (K16, K45, K62, K63, K77,K87, K96, K147, R31, H36,H81, H113) and they would contribute (along with the heme and the N-terminal amine) to the positive charge-state of the protein. Analysis of the relative surface accessibility of these residues [21,22] showed that the average surface accessibility of the SEFBR's of myoglobin is 11.7 Å^2 and among the 12 SEFBR's only 6 residues (R31, K62, K63, K87, K96 and K147) have surface accessibility higher than this average value. The H81 residue

has surface accessibility close to the average value. The N-terminal amine of myoglobin is also free and accessible. Thus, the number of SEFBR which has the surface accessibility higher than the average value (11.7 Å^2), is 8 (one arginine, six lysins, and the N-terminal amino group). Moreover, the heme could also contribute to the positive charge, and the charge-state is calculated as +9. This value is lower than that observed (+11, Figs. 1 and 2B) experimentally [12]. Such discrepancy may arise due to differences in the structure of the protein in solution and in the crystal. Myoglobin being a flexible molecule with non-covalently attached heme, may have more open structure in solution that observed in the crystals, while such solvation directed structure change may not be significant in Cytochrome *c*. The most-intense charge-state based on the present analyses of the crystal structures of the proteins also seems to give slightly lower value than that reported in case of some other proteins (Table 1). The possibility of the presence of partially open structure of the protein in solution compared to that observed in the crystals may explain the differences in the experimental and the calculated values of the most-intense charge-states of the gaseous proteins in the present model.

Analogously, the most-intense negative charge-states in the negative ion ESI-MS spectrum would correspond to the numbers of the SEFAR's that are having high surface accessibility (more than the mean value of surface accessibility of all SEFAR's in the protein).

Table 1 shows the mean value of surface accessibility of the SEFBR's (or of SEFAR's) for several proteins, along with the number of SEFBR (for positive ion spectra) or SEFAR (for negative ion spectra), and the correlation of the number of SEFBR and SEFAR with the most-intense charge-state of the protein. Many positively charged proteins are known to give both positive ion mode and negative ion mode ESI-mass spectra. For example, Cytochrome *c* characterized by a net positive charge (+12) in the sequence of the protein, shows +7 (or +8, +9) as the most-intense charge-state in the positive ion mode and −5 as the most-intense charge-state in the negative ion mode [12]. Kaltashov and Mohimen [23] and Hautreux et al. [24], reported that the average charge-state of some protein was related to the total surface area of the protein. However, only the average charge-state of the protein obtained in positive ion mode was found to be related to the total surface area of the protein, and the charge-states obtained in the negative ion mode was not considered [23,24].

3.3. A proposed model for the formation of the ESI charge-states of a protein on the basis of the surface accessibility of polar residues in the protein

The soft ionization method in the electrospray ionization does not significantly affect the structure of the protein and the charged solvent droplets produced by the ESI are considered to contain the proteins in the same conformational state as that in the solution [12,14,25,26]. Thus, during the process of removal of the solvent molecules in the passage of the droplets under vacuum [1], the polar residues at the protein surface either form intramolecular hydrogen bonds/ion-pairs by interacting with nearby oppositely charged residues on the surface of the protein or remain surface exposed free residues to absorb proton (SEFBR) to form positively charged centers or release proton (SEFAR) to form negatively charged centers.

The charged groups at the core of the protein are generally neutralized by a suitable oppositely charged group and thus they would not contribute to the charge-states. The model of conformation-dependent neutralization theory (CDNT) proposed by Grandori and co-workers demonstrated that the neutralization of the opposite charges in a folded protein indeed is important for the charge-state distribution in the ESI-MS of a protein [15,16,18]. In the positive ion mode, protonation of free carboxylate groups occurs (making them neutral) and the positive charges of the protein are formed by protonation at the free basic groups (including the N-terminal amine, if free, unlike in case of Cytochrome *c*) that are not involved in interactions with any other acidic groups at the surface of the gaseous protein ion. These SEFBR would thus carry the positive charge of the protein ion and the number of these residues would then correlate to the maximum charge-state [12].

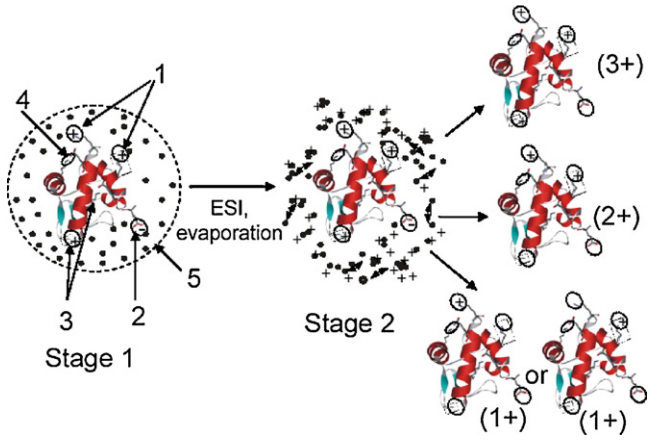
Analogously, in the negative ion mode, deprotonation of the basic residues would make them neutral and the negative charges at the surface exposed free acidic groups (SEFAR) are formed by deprotonation of the isolated carboxylic acid groups (including the C-terminal, if free) that are not involved in hydrogen bonding or salt bridge formation with any other residue. The number of such acidic residues would then correspond to the maximum negative charge of the protein in the negative ion mode [12]. This model of formation of multiple charge-states of proteins can explain the basis of observation of positive ion mass spectra of a protein with net negative charge and negative ion spectra of a protein with net positive charge and the results for several proteins were reported earlier [12].

The charge-state distributions of protein ions was earlier shown to be related to the solvent-exposed surface area of the protein in solution [23,26]. The surface area determined from the crystal structure was shown to correlate well with the average charge in ESI-mass spectra of the native folded protein. The present results highlight that the solvent-exposed surface of the protein indeed seems to be the main factor responsible for the observation of the charge-state of the protein and the average charge could be effectively related to the charged patch on the surface. As the charge-distribution of a large number of soluble proteins may be similar hence the number of SEFBR or SEFAR for a series of protein could be related to the size of the protein. Thus the number of SEFBR or SEFAR is related to a subset of the total surface area of a protein and so the observation of direct correlation of the charge-states with the size of the protein as demonstrated earlier also agrees with the present model as a special case [23,26].

The present results show that the nature of the charge-state distribution was affected by the nature of the salt and support the protein–salt interactions and charge-transfer that has been earlier proposed [12,20]. However, the most-intense charge-state of the protein seems to remain almost the same even in the presence of the salts, indicating that it is possibly an intrinsic property of the protein. We propose that the surface exposed free polar residues (e.g., SEFBR for positive ion mode) with more than the mean accessibility of all the SEFBR's, would be protonated first and that is the number of most populated charge-state of the protein. Analogously, for the negative ion mode, the SEFAR having more than the average of the accessibility of all SEFAR will give rise to the most-intense negative ion peak of the protein.

The formation of the charge-states of the protein may be depicted by a simplistic scheme as shown in Scheme 1 [1]. Before the electrospray ionization, the polar residues on the surface of the protein are solvated in the droplet (Stage 1) [1,6,12]. In the electrospray ionization and subsequent sequential evaporation and fission of the droplets, results in the formation of the protein ions surrounded by a few solvent molecules and other charge carrier ions in the Stage 2 as proposed earlier [6,7,9,10,20]. The SEFBR (or SEFAR), which are well exposed could be influenced by acids and bases (from the salts) and the excess $H_3O(+)$ ions. In Scheme 1 two basic residues (SEFBR) with surface accessibility more than the mean value and one basic residue (SEFBR) with surface accessibility less than the mean value are considered. Based on the above discussions, the formation of +2 charge-state would have higher probability than +1 or +3 charge-states of the protein.

It is important to note that the present model is based on the consideration that the surface exposed basic or acidic residues



Scheme 1. A proposed model for the formation of charge-states of protein in the ESI-MS. Stage 1 shows a representative protein molecule and its polar residues in solution before subjecting to the ESI. The labels 1 and 2 show the SEFBR (1) or SEFAR (2) that possess accessibility above the mean accessibility value, the label 3 marks the basic (or acidic) residue with accessibility below the critical value. Label 4 shows the salt bridged or ion-pair interaction between the acidic and basic residues. Label 5 represents the solvent surface. Stage 2 shows the protein and charged solvent and solute (salt) particles due to the ESI and subsequent evaporation of the droplet. The formation of various charge-states from the Stage 2 is represented by the +1, +2 and +3 charge-states. The acidic residue is neutralized in the positive ion mode. (In the negative ion mode the basic residues are deprotonated and the acidic residues contribute to the charge.)

that are engaged in electrostatic interaction are not available to contribute to the charge-state of the protein. The electrostatic interaction between the oppositely charged residues should neutralize each other as demonstrated earlier [16,18]. Nevertheless, intramolecular electrostatic interactions may also stabilize configurations with a net charge different than zero and charged hydrogen bonds may stabilize the charged form of ionisable residues. The present model assumes that such processes are absent or not predominant in the systems studied so that the experimental results agree well with that proposed by this simple model.

4. Conclusions

The effects of various types of salts on the ESI-MS charge-state distribution of Cytochrome *c* and myoglobin indicate that salts and proteins interact and there are charge-transfer reactions between the charged particles and the proteins prior to the formation of the final protein ion. These results clearly indicate that the factors such as the relative volatility of the charged-particles and the surface accessibility of the polar residues of the protein towards the charged particles are important in the formation of the charge-states of the protein. The surface accessibility of SEFBR and SEFAR were determined from the crystal structure and compared with the most probable charge-state (most-intense obtained in the positive ion and the negative ion mode ESI-MS spectra of the proteins. The results indicated that the number of the SEFBR that have surface accessibility above the mean surface accessibility of all the SEFBR's in the protein, corresponds to the most-intense charge-state of the protein in the positive ion mode. Analogously, the number of the SEFAR that have surface accessibility above the mean surface accessibility of all SEFAR's in the protein would give the most-intense negative ion charge-state of the protein. Hence, the most-intense charge-states both in the positive ion as well as in the negative ion modes for a large number of proteins could be predicted based on this simple model from surface accessibility of the protein.

Acknowledgement

This work was supported by the Tata Institute of Fundamental Research, Mumbai.

References

[1] J.B. Fenn, Angew. Chem. Int. Ed. Engl. 42 (2003) 3871–3894.
[2] C.S. Hoaglund-Hyzer, A.E. Counterman, D.E. Clemmer, Chem. Rev. 99 (1999) 3037–3080.
[3] A.T. Iavarone, O.A. Udekwu, E.R. Williams, Anal. Chem. 76 (2004) 3944–3950.
[4] M. Peschke, U.H. Verkerk, P. Kebarle, Eur. J. Mass Spectrom. (Chichester, Engl.) 10 (2004) 993–1002.
[5] G. Wang, R.B. Cole, Anal. Chem. 66 (21) (1994) 3702–3708.
[6] N. Felitsyn, M. Peschke, P. Kebarle, Int. J. Mass Spectrom. 219 (2002) 39–62.
[7] U.H. Verkerk, M. Peschke, P. Kebarle, J. Mass Spectrom. 38 (2003) 618–631.
[8] M.H. Amad, N.B. Cech, G.S. Jackson, C.G. Enke, J. Mass Spectrom. 35 (2000) 784–789.
[9] M. Peschke, A. Blades, P. Kebarle, J. Am. Chem. Soc. 124 (2002) 11519–11530.
[10] A. Blades, M. Peschke, U. Verkerk, P. Kebarle, J. Phys. Chem. A 106 (42) (2002) 10037–10042.
[11] T.L. Constantopoulos, G.S. Jackson, C.G. Enke, J. Am. Soc. Mass Spectrom. 10 (1999) 625–634.
[12] H. Prakash, S. Mazumdar, J. Am. Soc. Mass Spectrom. 16 (2005) 1409–1421.
[13] L. Konermann, D.J. Douglas, Biochemistry 36 (1997) 12296–12302.
[14] R. Grandori, Protein Sci. 11 (2002) 453–458.
[15] M. Samalikova, I. Matecko, N. Muller, R. Grandori, Anal. Bioanal. Chem. 378 (2004) 1112–1123.
[16] M. Samalikova, R. Grandori, J. Mass Spectrom. 38 (2003) 941–947.
[17] V.J. Nesatyy, M.J. Suter, J. Mass Spectrom. 39 (2004) 93–97.
[18] R. Grandori, J. Mass Spectrom. 38 (2003) 11–15.
[19] J. Fernandez de la Mora, Anal. Chim. Acta 406 (2000) 93–104.
[20] U.H. Verkerk, P. Kebarle, J. Am. Soc. Mass Spectrom. 16 (2005) 1325–1341.
[21] R. Rodriguez, G. Chinea, N. Lopez, T. Pons, G. Vriend, CABIOS 14 (1998) 523–528.
[22] G. Vriend, J. Mol. Graph 8 (1990) 52–56.
[23] I.A. Kaltashov, A. Mohimen, Anal. Chem. 77 (2005) 5370–5379.
[24] M. Hautreux, N. Hue, A.D.F.d. Kerdaniel, A. Zahir, V. Malec, O. Laprévote, Int. J. Mass Spectrom. 231 (2004) 131–137.
[25] L. Konermann, D.J. Douglas, J. Am. Soc. Mass Spectrom. 9 (1998) 1248–1254.
[26] A.J. Heck, R.H. Van Den Heuvel, Mass Spectrom. Rev. 23 (2004) 368–389.
[27] J.A. Loo, H. Peifeng, P. McConnell, W.T. Mueller, T.K. Sawyer, V. Thanabal, J. Am. Soc. Mass Spectrom. 8 (1997) 234–243.
[28] S. Zhang, C.K. Van Pelt, D.B. Wilson, Anal. Chem. 75 (2003) 3010–3018.
[29] C.V. Robinson, E.W. Chung, B.B. Kragelund, J. Knudsen, R.T. Aplin, F.M. Poulsen, C.M. Dobson, J. Am. Chem. Soc. 118 (1996) 8646–8653.
[30] S. Wendt, G. McCombie, J. Daniel, A. Kienhofer, D. Hilvert, R. Zenobi, J. Am. Soc. Mass Spectrom. 14 (2003) 1470–1476.
[31] N. Potier, L.J. Donald, I. Chernushevich, A. Ayed, W. Ens, C.H. Arrowsmith, K.G. Standing, H.W. Duckworth, Protein Sci. 7 (1998) 1388–1395.
[32] S.S. Ray, S.K. Singh, P. Balaram, J. Am. Soc. Mass Spectrom. 12 (2001) 428–438.
[33] K.R. Babu, A. Moradian, D.J. Douglas, J. Am. Soc. Mass Spectrom. 12 (2001) 317–328.
[34] J.A. Loo, Int. J. Mass Spectrom. 200 (2000) 175–186.
[35] D.R. Gumerov, I.A. Kaltashov, Anal. Chem. 73 (2001) 2565–2570.