ELSEVIER

International Journal of Mass Spectrometry 219 (2002) 269–281

International Journal of Mass Spectrometry

www.elsevier.com/locate/ijms

Mass spectrometric noncovalent probing of amino acids in peptides and proteins

Sebastian D. Friess, Jürg M. Daniel, Rudolf Hartmann, Renato Zenobi*

Department of Chemistry, Swiss Federal Institute of Technology, 8093 Zürich, Switzerland

Received 22 June 2001; accepted 6 November 2001

Abstract

Mass spectrometric methods are becoming available that probe structural or topological aspects of peptides and proteins. However, most of them require extensive chemical modification or covalent crosslinking of the biomolecules. We present a novel strategy based on the formation of noncovalent complexes between basic acidic amino acid residues and selective ligands. Matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) are employed to analyze the noncovalent assemblies. Different basic sites are probed with sulfonic acid derivatives. Naphthalene-disulfonic acid selectively binds to arginine residues and free amino termini; here we provide a chemical explanation for this behavior. Acidic sites are probed with molecules carrying guanidinium groups. However, only cysteic acid residues can so far be detected in this fashion; identification of glutamic and aspartic acid still awaits further investigation. (Int J Mass Spectrom 219 (2002) 269–281)

Keywords: Noncovalent interaction; MALDI; Protein structure; Amino acid residue

1. Introduction

Mass spectrometric methods might not be an obvious choice for analyzing three-dimensional protein structures. In fact, mass spectrometry (MS) is naturally a tool for molecular weight determination, or for answering questions about stoichometry, e.g., of enzyme-inhibitor systems. Nevertheless, the analysis of peptides and proteins using soft ionization MS to derive structural information has become a popular tool: matrix-assisted laser desorption/ionization (MALDI) [1,2] and electrospray ionization (ESI) [3,4] both allow higher order assemblies to largely retain their tertiary structure. Current knowledge about peptide and protein structure in the gas phase has recently been reviewed by Jarrold [5]. The article focuses on the use of ion mobility spectroscopy in combination with soft ionization. Generally, soft ionization MS does not provide direct structural data in the same sense as nuclear magnetic resonance (NMR), circular dichroism (CD), or X-ray crystallography. However, its results are complementary to data obtained from more traditional techniques [6].

One example of how MS can address questions about protein structure is to follow hydrogen–deuterium (H/D) exchange, e.g., at amide peptide linkages; for reviews, see [7,8]. Protons buried within the hydrophobic core, or those involved in intramolecular hydrogen bonding, can have exchange rates up to

* Corresponding author. E-mail: zenobi@org.chem.ethz.ch
Dedicated to the memory of Leo A. Weissberg.

1387-3806/02/$ – see front matter
PII S1387-3806(01)00581-4

eight orders of magnitude slower than protons exposed to solvent [9]. Each exchange of a hydrogen atom with deuterium increases the mass by one, hence the method requires high mass resolution. H/D exchange is used in both MS and NMR. Mass spectrometry has several advantages, among them better sensitivity and the ability to monitor fast and slow kinetic exchange rates.

Remarkable work on protein structure determination by MS has been published by Kuntz and co-workers: they have used chemical cross-linking to identify the fold family of proteins [10]. After cross-linking, the protein is digested, and the fragments are analyzed by MS. Bifunctional cross-linkers react with specific amino acid side chains, but two residues (e.g., lysines) will be cross-linked only if they are within the distance dictated by the length of the cross-linker. This data provides very useful residue distance constraints for computational models that simulate protein folding.

Another approach of analyzing higher order protein structures involves covalent modification of specific functional groups, followed by digestion and peptide mapping of the resulting fragments. This "surface topology probing" [11,12] yields surface accessibilities of amino acid residues, i.e., it reveals which residues are exposed and which are buried within the core of the structure. Selective covalent modification reactions exist for a variety of amino acids: lysine residues can be identified after conversion to homo-arginines [13], cysteine thiols are trimethylamino-ethylated to thialaminine [14], comparable routines exist for histidine [15] and more general methods such as methylation or acetylation have been applied as well [16]. Disulfide bonds generally survive enzymatic digests, thus, disulfide-bridged fragments and their reduced free forms allow verification of the cysteine locations [17].

We present a novel strategy for obtaining topological information about biomolecules. It eliminates the need for high mass resolution (as needed in H/D experiments) and preceding chemical modification. The method is based on the formation of noncovalent complexes that survive the desorption and ionization processes in MS experiments. Many examples have been published on the use of ESI for the detection of supramolecular noncovalent assemblies; for reviews, see [18–20]. MALDI is not established equally well for this purpose, although convincing examples exist. Among them, peptide–metal complexes, peptide–protein interactions, oligomerization studies, and many others have been presented; for reviews, see [21–23]. Our group has previously shown that MALDI MS can be used to determine the number of accessible basic sites of a peptide or protein in its folded structure [24], i.e., arginine, histidine, lysine residues, and the amino terminus. The ligand for these residues is Cibacron Blue F3G-A (CCB), a dye widely used in the field of affinity chromatography [25]. It has been found that the number of CCB adducts is far greater if denaturing conditions are used compared to MALDI sample preparation with non-acidic matrices. This has been interpreted as a result of the accessibility of basic amino acid residues, which is much larger for open conformations. Thus, the number of noncovalent adducts reflects the number of exposed basic residues. For example, insulin exhibits a total of five CCB adducts ($\Delta m/z = 773$ for the free acid) in addition to the protonated molecular ion signal ($m/z = 5734$). All of the basic amino acid residues are located in the B-chain and are quite far from each other; steric hindrance is thus minor. Insulin consists of two polypeptide chains linked together via disulfide bonds and an ionic interaction between the amino terminus of the A-chain and the carboxyl terminus of the B-chain [26]. This strong ionic interaction is not disrupted during MALDI sample preparation, such that the expected number of adducts adds up to five (one Arg, one Lys, two His, one free amino terminus) in perfect agreement with the mass spectrum. A similar strategy has been used by McLuckey and co-workers: they have identified the basic residues of a wide variety of oligopeptides by counting the number of hydroiodic acid molecules that attach to the analyte ion [27,28]. The sum of the ion charge state and the maximum number of ligands (gaseous hydroiodic acid) reflects the total number of basic

sites of the analyte. The authors reported gas phase adduct formation, but their strategy is conceptually similar to ours.

We have recently studied sulfonic acid derivatives of much simpler structure and more predictable reactivity than CCB. Much to our surprise, naphthalene-sulfonic acids were found to selectively probe arginine residues plus the amino terminus [29]. A mass spectrum of insulin and naphthalene-disulfonic acid (NDS) shows only two adducts in addition to the protonated molecular ion signal (one Arg, one free amino terminus). Several control experiments have unambigously confirmed this selectivity.

This article extends our earlier work: we study these noncovalent interactions in detail to clearly understand the mode of interaction. We present control experiments employing ESI to assure that the complex formation does not depend on the ionization technique used. Additionally, we present initial experiments aimed at the detection of exposed deprotonated acidic amino acid residues (glutamic acid, aspartic acid, cysteic acid, and carboxyl terminus).

2. Experimental

2.1. MALDI MS

Most of the experiments were performed on a home-built 2 m linear time-of-flight mass spectrometer equipped with a UV nitrogen laser ($\lambda = 337$ nm, pulse width 3 ns); an acceleration potential of 24 kV was used in the source region for ion extraction. Ions were detected with a pair of microchannel plates and the signal was acquired by a 500 MHz (LeCroy 9450) digital oscilloscope. Details of this instrument are published elsewhere [30]. For higher mass resolution, one mass spectrum was acquired on a commercial TOF mass spectrometer (Voyager-DE Elite, PE Biosystems, Framingham, MA) equipped with a similar laser and delayed extraction mode (delay 350 ns). The spectrum was recorded in positive linear mode with an acceleration voltage of 25 kV. All data were transferred to a personal computer for further processing.

Spectra were acquired using 4-nitroaniline (4NA, 10^{-2} M in methanol) as a non-acidic MALDI matrix. All of the analyte and ligand solutions were prepared in water with concentrations of 10^{-4} and 10^{-3} M, respectively. To remove sodium and potassium ions, desalination of all analyte solutions was performed using the drop dialysis method [31,32]: 20 mL of the solution were dialyzed against a 100 mM solution of diammonium hydrogen citrate using a 0.025 mm pore size membrane.

MALDI samples were prepared using the two-layer technique [33]: 1 mL of matrix solution was deposited on the target and rapidly dried in vacuo to form a very thin layer of fine crystals. The analyte and the ligand solution were mixed with the matrix solution (1:1:2, v/v/v) and a total of 1.5 mL of this mixture was deposited on top of the first matrix layer, dried in vacuo and then analyzed.

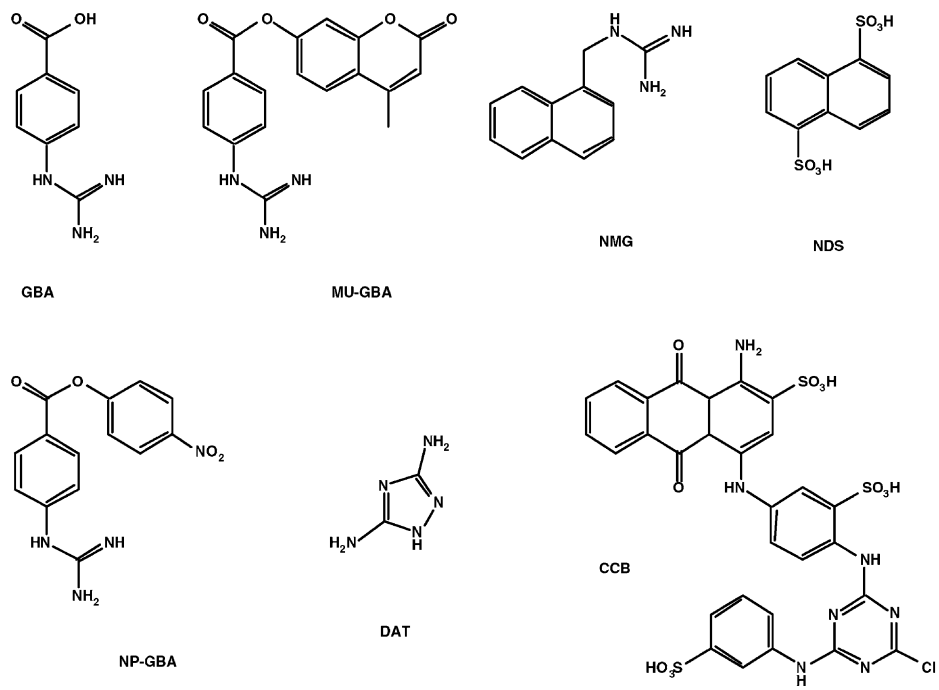
2.2. ESI-MS

Positive ion mass spectra were acquired on a home-built prototype reflectron time-of-flight mass spectrometer equipped with a commercial electrospray source (Agilent, Palo Alto, CA). The desalted protein solution was diluted with nanopure water (10^{-5} M) and mixed with ligand solution. Ions were sprayed at a potential of 4 kV and orthogonally accelerated into the flight tube (potential 7.2 kV, frequency 4 kHz). Ions were detected with an electron multiplier (Electron Multipliers AF831H). A total of 20,000 spectra were collected on a personal computer, averaged and exported for further processing.

2.3. Materials

Luteinizing hormone releasing hormone, bovine insulin (A- and B-chain), cytochrome *c*, epidermal growth factor receptor fragments 661–681, substance K, phosphate accepting polypeptide, and diammonium hydrogen citrate were obtained from Sigma (Buchs, Switzerland). 1-Naphthylmethylamine,

Ligands:



GBA	guanidino-benzoic acid	NMG	naphthylmethyl-guanidine
MU-GBA	methylumbiferryl-benzoic acid	NDS	naphthalene-disulfonic acid
NP-GBA	nitrophenyl-benzoic acid	DAT	diaminotriazole
CCB	Cibacron Blue F3G-A		

Peptides:	PAP	H_2N-Leu-Arg-Arg-Ala-Ser-Leu-Gly-COOH
	5SP	H_2N-Leu-Arg-Ala-Gly-Leu-Ala-Leu-Arg-Gly-COOH

Scheme 1. Chemical structures, names, and abbreviations of the ligands used (above). Abbreviations and sequences of the two peptides used as ligands (below).

4-guanidino-benzoic acid, methylumbelliferyl-4-guanidino-benzoic acid, 4-nitrophenyl-4-guanidino-benzoic acid, 3,5-diamino-1,2,4-triazole, thiourea, di-*tert*-butyldicarbonate, naphthalene-1,5-disulfonic acid, and Cibacron Blue F3G-A were obtained from Fluka (Buchs, Switzerland). [Pro^2]-luteinizing hormone releasing hormone and the peptide LRAGLALRG were obtained from custom synthesis (Pineda Peptide and Antibody Service, Berlin, Germany). 1-Naphtlymethylguanidine was prepared synthetically by guanylation [34] of 1-naphthylmethylamine with *N*,*N*′-bis(*tert*-butyloxy-carbonyl)thiourea, which was received after protection of thiourea with di-*tert*-butyldicarbonate [35].

All reagents were of highest purity available and solvents were of spectrophotometric grade. Structures and sequences of ligands are shown in Scheme 1.

3. Results and discussion

Ionic, hydrophobic, hydrogen bonding and/or van der Waals forces are the major factors that govern noncovalent complexes in solution. We have recently shown that noncovalent complexes between peptides–proteins and naphthalene-sulfonic acids are preformed in solution prior to desorption and ionization [29]. Nonspecific gas phase adduct formation was ruled out by suitable control experiments. Another type of control is presented here: a comparison between different soft ionization methods should eliminate the possibility that the degree of complexation depends on the ionization method used.

3.1. Comparing MALDI and ESI

The transfer of a protein–ligand complex from solution or condensed phase into the gas phase environment is usually accompanied by structural changes. However, the extent of such changes as a result of solvent removal and the interactions that govern stability of the distinct species remain poorly understood. Many studies have suggested that at least some aspects of the higher order structure of proteins can be retained after transfer into the gas phase [36,37]. Complementary or even identical results can usually be obtained from MALDI and ESI measurements: calcium-induced tetramer formation of calcium-binding proteins was monitored by first shot MALDI spectra [38] and fully confirmed by ESI experiments [39]. Earlier, the aldose reductase system and the binding to its cofactor was studied to yield comparative results: in MALDI, the cofactor lost a

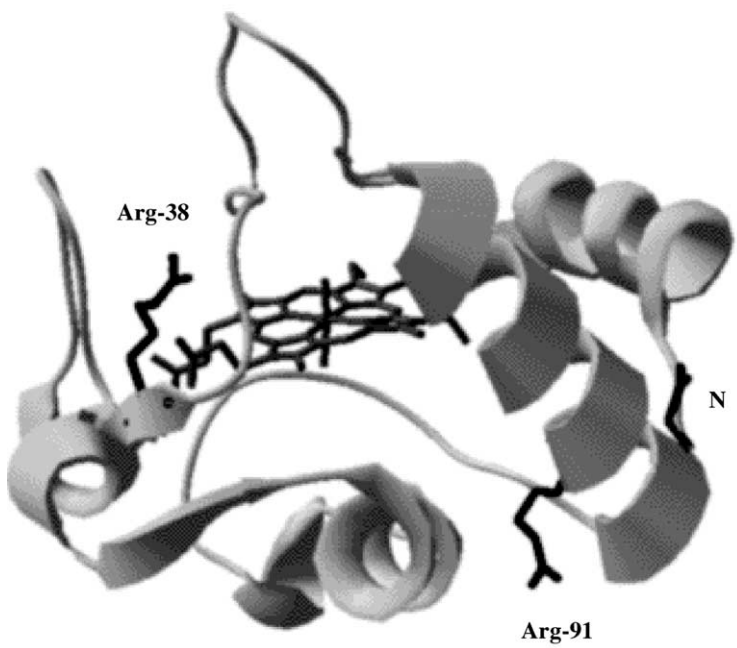


Fig. 1. Three-dimensional protein structure of cytochrome *c*. Shown in black are the two arginine residues (Arg) and the amino terminus (N); dark gray, heme cofactor. The picture was produced with the program pdb-viewer accessible through the ExPASy Molecular Biology Server of the Swiss Institute of Bioinformatics (http://www.expasy.ch/spdbv/). Coordinates were taken from the Protein Data Bank (http://www.rcsb.org/pdb, code 3CYT).

residual part [40], whereas the intact noncovalent complex was successfully detected by ESI [41].

In this article, the comparison of MALDI with ESI is exemplified by cytochrome *c*. Probing its exposed amino acid residues with NDS reveals three adducts, reflecting complex formation with both arginine residues (Arg^{38} and Arg^{91}) and the amino terminus; consider Fig. 1 for a plausible structural representation of this finding. Identical results can be obtained from both MALDI and ESI experiments: the MALDI spectrum consists of the protonated molecular ion signal of cytochrome *c* and three noncovalent complexes (see Fig. 2 lower trace). Unlike MALDI, which generates mainly singly charged ions, ESI tends to produce ions with *m*/*z* ratios of around 500–2500, indicating multiply charged species. The gas phase structure of cytochrome *c* has been extensively studied in the past [42–47], revealing a charge state distribution in ESI experiments arranged around +8 for a near-native conformation. The ESI mass spectrum of cytochrome *c* and NDS shows three distinct charge states (+6, +7 and +8) with +7 giving the most abundant signals (see Fig. 2 upper trace). This charge state distribution indicates that cytochrome

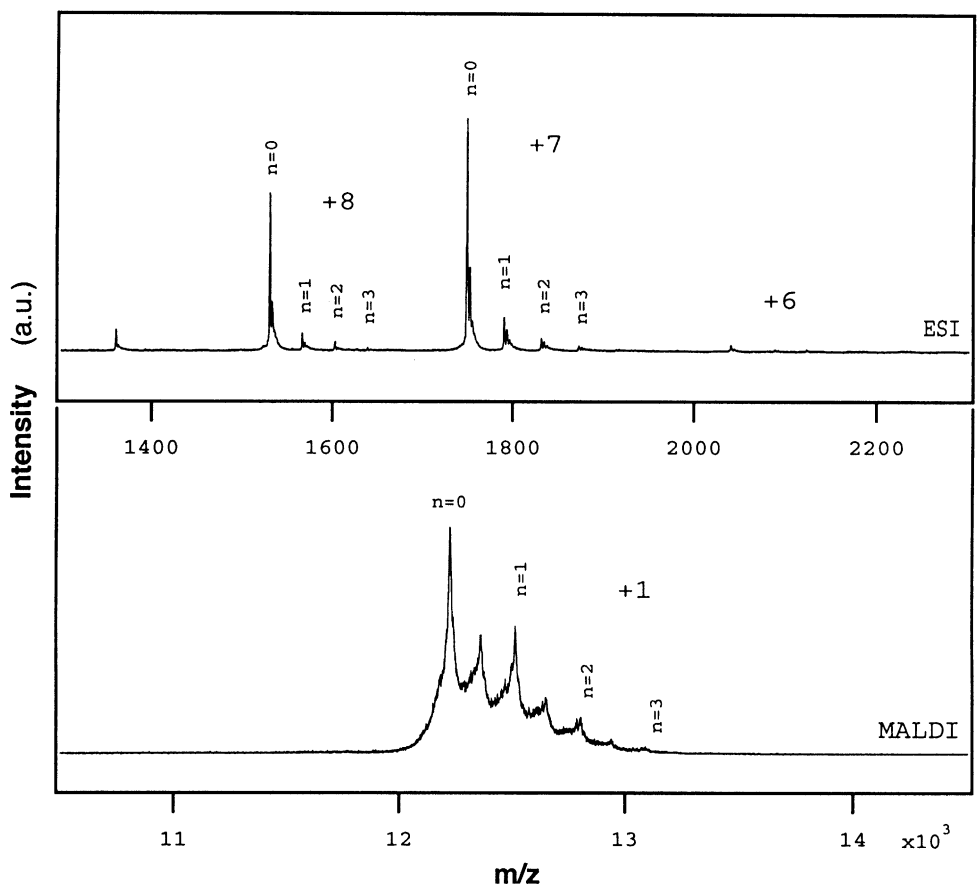
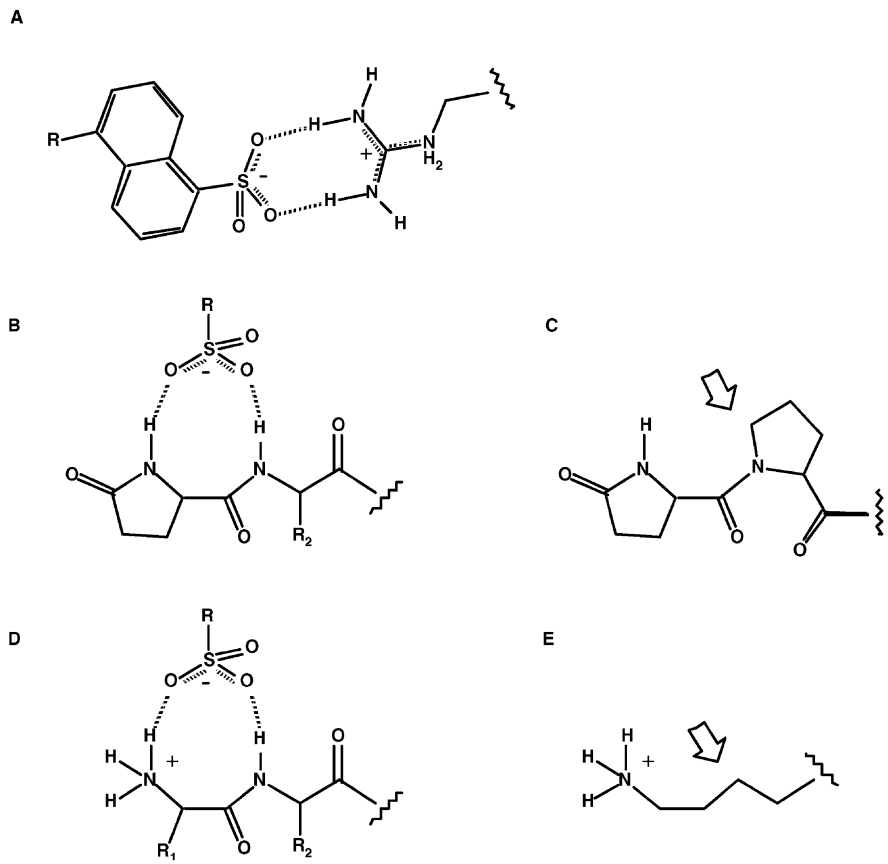


Fig. 2. Positive ion mode ESI (upper trace) and MALDI (lower trace) mass spectra of cytochrome *c* with NDS. Peak series are assigned to $[P + nL + H]^{z+}$, where P represents the protein, L represents the ligand in free acid form and *z* represents the charge state of the observed species. Weak additional signals in the MALDI spectrum result from nonspecific matrix adducts. Identical results are obtained from both ionization techniques.

c has adopted a near-native conformation and that desolvation/ionization has not dramatically altered the structure, although minor changes may have occurred. Nevertheless, triple adduct formation of NDS can be demonstrated for all three charge states. Thus, both ionization methods give identical results as expected. The binding of NDS to the arginine residue has been proposed to consist of ionic and hydrogen bonding interactions, mainly being preserved after transfer into the gas phase. Structural characterization of cytochrome *c* with CCB to probe all basic sites has previously been published [24]: up to nine ligands reacted with the same biomolecule in agreement with charge state distributions being arranged around +8 for the near-native conformation.

3.2. Elucidating the interaction between sulfonic acid ligands and peptides

The mode of interaction of sulfonic acid ligands with exposed protonated basic amino acid residues has not yet been fully explained. For example, NDS



Scheme 2. Graphical representation of postulated noncovalent interactions with sulfonic acids. (A) Naphthalene-sulfonic acid and arginine residue. (B) Naphthalene-sulfonic acid and amino terminal pyroglutamic acid. The complex is formed due to a cooperative hydrogen bond with an adjacent amide proton. (C) Absence of the amide proton (see arrow) after introduction of proline at position 2 in the sequence does not permit the formation of the noncovalent complex. (D) Naphthalene-sulfonic acid and amino terminus. The complex is formed due to a cooperative hydrogen bond in analogy to the structure presented in B. (E) Lysine residues also lack additional atoms for hydrogen bonding (see arrow), thus, they do not form the noncovalent interaction.

does detect the amino terminus but not lysine which has a similar pK_α-value [48]. Thus, our results cannot be explained by ionic interactions alone. Here we provide a chemical explanation at least for simple sulfonic acid ligands. Additional interactions, such as hydrogen bonds must thus, be taken into account. Consider Scheme 2A for the proposed noncovalent interaction between protonated arginine residues and deprotonated sulfonic acids.

We have investigated luteinizing hormone releasing hormone (LH-RH), a 10 residue oligopeptide with an amino terminal pyroglutamic acid (pGlu), which is common in many bioactive peptides such as neurotensin, bombesin, and other gastrointestinal hormones. Amino terminal pGlu prevents protonation at the amino terminus under mild conditions: pK_a-values for protonation at the pyrrolidinoyl oxygen are below zero [26] and the nitrogen lone pair is delocalized in the amide bond. Nevertheless, the noncovalent complex between the amino terminus of LH-RH and NDS can be readily observed (see Fig. 3 (lower trace)). The mass spectrum exhibits two adducts, one for the arginine residue and one for the amino terminus. Introduction of an additional lysine at position six in the sequence of LH-RH does not alter the spectrum [29], so all putative complexes have already been detected. We think that the subsequent amide proton of the backbone can be used for complex formation (see Scheme 2B). The overall complexation now comprises an ion–dipol interaction and an additional cooperative hydrogen bond, which should be sufficiently stable to survive the MALDI process. A peak at the mass of an alleged adduct is not a valid criterion to prove its specificity. We, therefore, used a synthetic peptide containing no subsequent cooperative hydrogen atom in its backbone, thus, significantly changing structural interactions. All factors for amino terminal complexation have been ruled out (see Scheme 2C). The mass spectrum of [Pro^2]-LH-RH with NDS should show the protonated molecular ion signal and only one adduct for arginine detection, but not more. The data in Fig. 3 (upper trace) confirms our model; the second NDS adduct has disappeared from the spectrum. The model also explains the discrimination in the

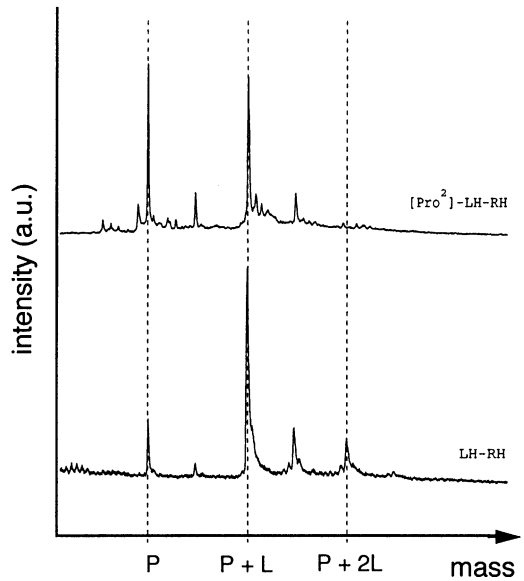
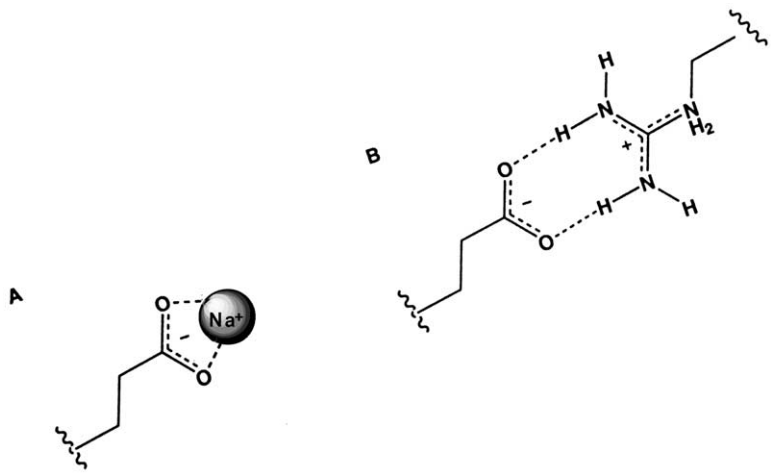


Fig. 3. Negative ion mode MALDI mass spectra of [Pro^2]-LH-RH (upper trace) and LH-RH (lower trace) with NDS. Peaks are assigned to $[P + nL - H]^-$, where P represents the peptide and L represents the ligand in free acid form. The abscissa is a mass scale, but because the two peptides differ in their molecular weight ($\Delta m = 40$ Da), the spectrum of [Pro^2]-LH-RH has been shifted up for better comparison. The expected *m*/*z* values for all three species P, P + L and P + 2L are marked by dotted lines. LH-RH exhibits one adduct more than the proline modification.

NDS complexation of lysine vs. amino terminus: the additional hydrogen bonding is only possible at the amino terminus (see Scheme 2D). Lysine residues do not have subsequent amide protons to be involved in complex formation (see Scheme 2E). All these observations lead to the conclusion that the interaction between protonated basic sites of peptides–proteins and deprotonated sulfonic acid dyes is not only a matter of ionic, but also of ion–dipole and hydrogen bonding interactions. Furthermore, shape complementarity favors the formation of these noncovalent complexes, as evidenced in the near-perfect bidentate binding to arginine residues. The mode of interaction can thus, be explained for naphthalene-sulfonic acids ligands. However, the complex formation with Cibacron Blue F3G-A (refer to Scheme 1) is still beyond a straightforward explanation, due to its more complicated structure that includes three



Scheme 3. Graphical representation of postulated noncovalent interactions between carboxyl groups and sodium (A) or a guanidinium functionality (B). The sodium complex can be detected in the mass spectra, the complex on the right has not been observed in any of the spectra.

sulfonic acid groups of completely different reactivity. Perhaps, this multifunctional dye works as a chelating ligand and thus probes all basic amino acid residues.

3.3. Probing carboxylic sites

A second set of experiments has been designed for investigating the noncovalent complex formation at deprotonated acidic residues of peptides and proteins (glutamic acid, aspartic acid and the carboxyl terminus). In addition, recombinant synthesis of compounds containing cysteine often leads to residual sulfonic acids, called cysteic acid or Cys[SO_3H], which are highly acidic sites, too [26]. Hydroxyl groups of serine and tyrosine residues and sulfhydryl groups of cysteines are usually not deprotonated under normal physiological conditions, nor are amide carboxyl termini.

The simplest imaginable ligands for deprotonated acidic residues are alkali cations: sodium or potassium ions are ubiquitous in MALDI, thus, we usually desalt all analyte solutions prior to use. In contrast, the ligand solution for this experiment consists of saturated aqueous NaCl. Following the regular MALDI preparation protocol (see Section 2), we studied the formation of complexes between acidic residues and sodium cations. Adducts can be identified by increments of $\Delta m = 22$ Da per sodium incorporated in the biomolecule (see Scheme 3A). The negative ion mode mass spectrum of insulin A-chain after complexation with sodium is shown in Fig. 4. Up to seven adducts can clearly be identified for insulin

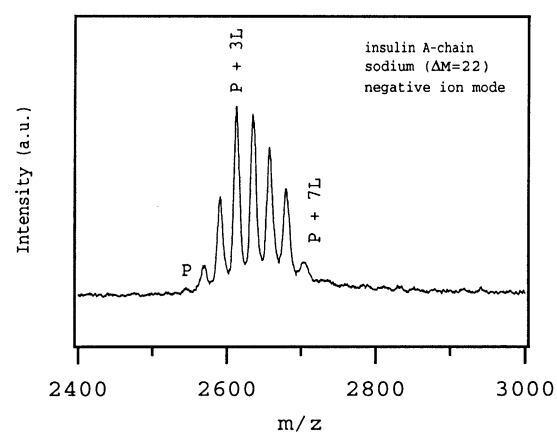


Fig. 4. Negative ion mode MALDI mass spectrum of insulin A-chain after incubation in a saturated aqueous NaCl solution. Peaks are assigned to $[P + nL - H]^-$, where P represents the peptide and L represents the ligand (sodium). Up to seven adducts can be identified.

A-chain, corresponding to the seven acidic sites found in the sequence (two Glu, four Cys[SO_3H] and one carboxyl terminus). Comparable results have been obtained for the epidermal growth factor receptor fragments 661–681 (EGFR). Three sodium adducts were detected, corresponding to three Glu residues (the carboxyl terminus of EGFR is masked; data not shown). At this point it is, however, not clear whether sodium complexes are specific adducts formed in solution or whether they originate from collisions in the MALDI plume. Cationization is a process that has often been observed as a result of (nonspecific) plume reactions [49]. Without further substantiation, we cannot unambiguously assign the sodium adducts to specific complexes formed in solution.

Ligands that allow hydrogen bond formation, e.g., organic compounds or peptides that contain guanidinium groups, may be more suitable to probe acidic residues: they are protonated under physiological conditions and thus, an inversion of our previous concept may apply. Scheme 1 lists the guanidinium ligands used in this study. Among the examples reported in the literature on peptide–peptide interactions, two are closely related to our research: ionization efficiencies of highly acidic biomolecules were shown to largely depend on the extent of noncovalent complex formation with basic polypeptides [50]. A recent study explored the interaction between basic peptides such as dynorphin and acidic peptides [51]. The authors have identified the structural elements that drive the noncovalent complex formation: the "RR-motif", a di-arginine site in the sequence of the basic polypeptide, was claimed to probe carboxylic residues. The authors describe the formation of salt bridges between dynorphin and two or more adjacent glutamic and aspartic acid sites (e.g., minigastrin, sequence LEEEEEAYGWMDF amide).

We have tried to use the RR-motif as a selective ligand for single carboxylic residues and chose the phosphate acceptor polypeptide (PAP, sequence LRRASLG) for this purpose. Fig. 5a presents the mass spectrum of EGFR (three Glu) with PAP as potential ligand: there are definitely no adducts observed. Prior experiments have proven that EGFR easily

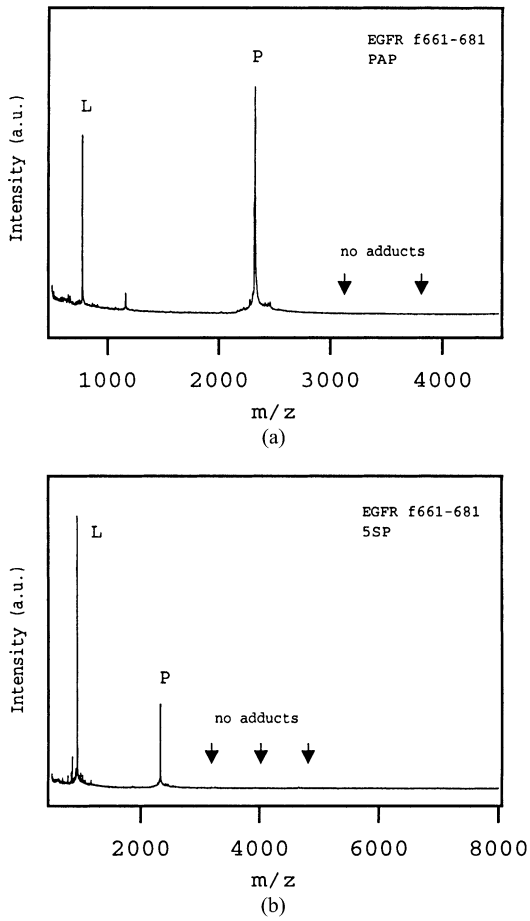


Fig. 5. Positive ion mode MALDI mass spectra of EGFR f661–681 with PAP (a) and 5SP (b). In both spectra, protonated molecular ion signals for the peptide P and the ligand L are found, but no adducts can be identified.

exchanges acidic protons against sodium, i.e., the carboxyl groups should be deprotonated under physiological conditions ($pK_a = 4.07$ [48] for γ-COOH of Glu). We conclude that the RR-motif is not a specific ligand for carboxyl groups, although it may allow to complex several adjacent acidic residues like in the sequence of minigastrin. To further elucidate the role of two adjacent arginine residues, we have synthesized a peptide containing two arginines separated by five spacing amino acids (5SP, sequence LRAGLALRG). Mass spectra of EGFR with 5SP again do not exhibit any adducts either, as evidenced in Fig. 5b. Identical

spectra have been obtained for substance K (one Glu and one free amino terminus) with both ligands and in both positive and negative ion mode (data not shown). Thus, neither two adjacent arginines nor two arginine functionalities within a sequence can probe single carboxyl groups.

We also studied simple mono-guanidyl compounds, namely guanidinium-benzoic acid (GBA) and two derivatives (MU- and NP-GBA), diaminotriazole (DAT) to exclusively test shape complementarity, and we have synthesized 1-methylnaphthyl-guanidine; see all structures in Scheme 1. None of these compounds were found to form detectable noncovalent complexes with carboxylic residues of peptides. Spectra of such systems consist of molecular ion signals of both the analyte and the ligand, but no adducts were detected, neither in positive nor in negative ion mode (data not shown). We assumed the guanidinium ligands to be closely matching in terms of shape complementarity (see Scheme 3B for the expected interaction), however, this is not supported by our experiments. Guanidinium groups, previously hypothesized to easily probe carboxylic residues [51], do not fulfill the requirements for successful detection. We are still tackling the task of detecting glutamic and aspartic acid residues by noncovalent complex formation.

3.4. Probing cysteic acid residues

In the course of our studies, we also thought of investigating complex formation at cysteic acid sites. A noncovalent complex, consisting of cysteic acid in the biomolecule and arginine in the ligand sequence, should be easily detectable: Gaskell and co-workers have studied several gas phase peptide ions and showed that their energetically preferred conformations incorporated an intramolecular ionic interaction between the arginine and the cysteic residue [52,53].

Experimentally, we found that cysteic acid sites are readily probed by arginine-containing ligands. Fig. 6 shows a MALDI mass spectrum of insulin A-chain and 5SP. Identical results were obtained with using PAP (data not shown). Up to three adducts can be identified upon complex formation with insulin A-chain

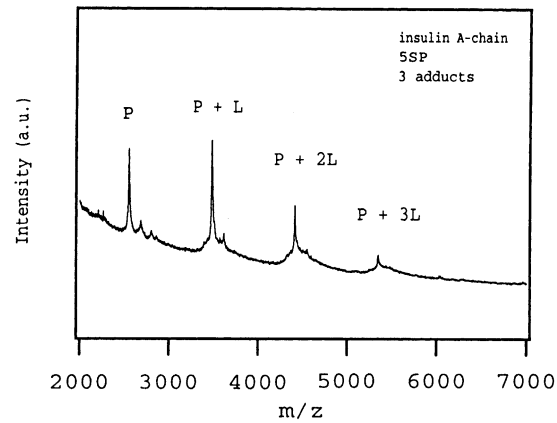


Fig. 6. Positive ion mode MALDI mass spectrum of insulin A-chain with 5SP. Peaks are assigned to $[P+nL-H]^-$, where P represents the peptide and L represents the ligand as free base. Up to three adducts can be identified.

(four Cys[SO_3H] groups). Two of these groups are directly adjacent, originally resulting from an intrachain disulfide bridge, so Coulombic repulsion may limit the number of charges on these residues. Three out of four cysteic acid groups are probably deprotonated and expectedly three noncovalent adducts are formed. Cysteic acid residues can thus be probed with arginine residues to form noncovalent peptide–peptide interactions.

4. Conclusion

Selective complexation of glutamic and aspartic residues awaits further investigation. Detecting such interactions seems to be less straightforward than expected. A simple 1:1 complex formation between guanidinium and carboxylic acid groups was not found to occur, in contrast to suggestions in the literature. On the other hand, the interaction of sulfonic acid ligands with basic amino acid residues is now much better understood and can be used to probe protein structure. As a rule, multifunctional sulfonic acid dyes such as Cibacron Blue F3G-A probe all exposed basic sites (Arg, Lys, His, and the amino terminus). More specific ligands exist for selected basic amino

acids. Naphthalene-disulfonic acid (NDS) is selective for arginine residues. It will also recognize the amino terminus if an adjacent amide proton can be involved in an additional cooperative hydrogen bond. This interaction cannot be realized for proline at position 2 in the sequence. NDS does not probe lysine nor histidine residues. Only exposed residues on the surface of the folded protein will form the noncovalent complexes. In the future, we will try to correlate this number of identified adducts to given surface accessibility parameters, known from the literature or from computational simulation algorithms. Future experiments will show whether such surface accessibilities can be directly obtained by our noncovalent strategies.

Acknowledgements

This work is supported by the Swiss National Science Foundation, grant number 2-77081-00. Wolfgang Lindner (Vienna) suggested the [Pro^2]-LH-RH experiment. The authors would like to thank Donald Hilvert, Silke Wendt, and Richard Knochenmuss for help and supporting material.

References

[1] M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, Int. J. Mass Spectrom. Ion Process 78 (1987) 53.
[2] M. Karas, F. Hillenkamp, Anal. Chem. 60 (1988) 2299.
[3] J.B. Fenn, M. Mann, C.K. Keng, S.F. Wong, C.M. Whitehouse, Mass Spectrom. Rev. 9 (1990) 37.
[4] M. Yamashita, J.B. Fenn, J. Chem. Phys. 88 (1984) 4451.
[5] M.F. Jarrold, Annu. Rev. Phys. Chem. 51 (2000) 179.
[6] R.L. Winston, M.C. Fitzgerald, Mass Spectrom. Rev. 16 (1997) 165.
[7] D.L. Smith, Y. Deng, Z. Zhang, J. Mass Spectrom. 32 (1997) 135.
[8] J.R. Engen, D.L. Smith, Anal. Chem. 73 (2001) 256A.
[9] V. Katta, B.T. Chait, Rapid. Commun. Mass Spectrom. 5 (1991) 214.
[10] M.M. Young, N. Tang, J.C. Hempel, C.M. Oshiro, E.W. Taylor, I.D. Kuntz, B.W. Gibson, C. Dollinger, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 5802.
[11] D. Suckau, M. Mak, M. Przybylski, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 5630.
[12] M.O. Glocker, C. Borchers, W. Fiedler, D. Suckau, M. Przybylski, Bioconjugate Chem. 5 (1994) 583.
[13] J.R. Kimmel, Meth. Enzymol. 11 (1967) 584.
[14] H.A. Itano, R.A. Robinson, J. Biol. Chem. 247 (1972) 4819.
[15] M. Kalkum, M. Przybylski, M.O. Glocker, Bioconjugate Chem. 9 (1998) 226.
[16] M. Galvani, M. Hamdan, P.G. Righetti, Rapid Commun. Mass Spectrom. 14 (2000) 1925.
[17] H. Belva, C. Valois, C. Lange, Rapid Commun. Mass Spectrom. 14 (2000) 224.
[18] B. Ganem, Y. Li, J. Henion, J. Am. Chem. Soc. 113 (1991) 6294.
[19] M. Przybylski, M.O. Glocker, Angew. Chem. Int. Ed. 35 (1996) 806.
[20] J.A. Loo, Mass Spectrom. Rev. 16 (1997) 1.
[21] F. Hillenkamp, Matrix-assisted laser desorption/ionization of non-covalent complexes, in: W. Ens, K. Standing, I. Chernushevich (Eds.), New Methods for the Study of Biomolecular Complexes, Kluwer Academic Publishers, Dordrecht, 1998.
[22] T.B. Farmer, R.M. Caprioli, J. Mass Spectrom. 33 (1998) 697.
[23] J. Godovac-Zimmermann, L.R. Brown, Mass Spectrom. Rev. 20 (2001) 1.
[24] B. Salih, R. Zenobi, Anal. Chem. 70 (1998) 1536.
[25] P.M. Boyer, J.T.-A. Hsu, Protein purification by dye-ligand chromatography, in: A. Fiechter (Ed.), Advances in Biochemical Engineering and Biotechnology, Springer, Heidelberg, 1993.
[26] L. Stryer, Biochemistry, 3rd Edition, New York, 1988.
[27] J.L. Stephenson Jr., S.A. McLuckey, Anal. Chem. 69 (1997) 281.
[28] S.A. McLuckey, J.M. Wells, J.L. Stephenson Jr., D.E. Goeringer, Int. J. Mass Spectrom. 200 (2000) 137.
[29] S.D. Friess, R. Zenobi, J. Am. Soc. Mass Spectrom. 12 (2001) 810.
[30] F. Dubois, R. Knochenmuss, R.J.J.M. Steenvoorden, K. Breuker, R. Zenobi, Eur. Mass Spectrom. 2 (1996) 167.
[31] R. Maruzyk, A. Sergeant, Anal. Biochem. 105 (1980) 403.
[32] H. Görisch, Anal. Biochem. 173 (1988) 393.
[33] Y.Q. Dai, R.M. Whittal, L. Li, Anal. Chem. 68 (1996) 2721.
[34] Y.F. Yong, J.A. Kowalski, M.A. Lipton, J. Org. Chem. 62 (1997) 1540.
[35] J.-G. Pons, J.-L. Fauchère, F. Lamaty, A. Molla, R. Lazaro, Eur. J. Org. Chem. 5 (1998) 853.
[36] 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) 8464.
[37] Q. Wu, J. Gao, D. Joseph-McCarthy, G.B. Sigal, J.E. Bruce, G.M. Whitesides, R.D. Smith, J. Am. Chem. Soc. 119 (1997) 1157.
[38] T. Vogt, J. Roth, C. Sorg, F. Hillenkamp, K. Strupat, J. Am. Soc. Mass Spectrom. 10 (1999) 1124.
[39] K. Strupat, H. Rogniaux, A. Van Dorsselaer, J. Roth, T. Vogl, J. Am. Soc. Mass Spectrom. 11 (2000) 780.
[40] J. Gross, A. Leisner, F. Hillenkamp, S. Hahner, M. Karas, J. Schäfer, F. Lützenkirchen, E. Nordhoff, J. Am. Soc. Mass Spectrom. 9 (1998) 866.
[41] N. Potier, P. Barth, D. Tritsch, J.F. Biellmann, A. Van Dorsselaer, Eur. J. Biochem. 243 (1997) 274.

[42] S.A. McLuckey, G.J. Van Berkel, G.J. Glish, J. Am. Chem. Soc. 112 (1990) 5668.
[43] D.S. Wagner, R.J. Anderegg, Anal. Chem. 66 (1994) 706.
[44] T.D. Wood, R.A. Chorush, F.M. Wampler, D.P. Little, P.B. O'Connor, F.W. McLafferty, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 2451.
[45] P.D. Schnier, D.S. Gross, E.R. Williams, J. Am. Chem. Soc. 117 (1995) 6747.
[46] K.B. Shelimov, D.E. Clemmer, R.R. Hudgins, M.F. Jarrold, J. Am. Chem. Soc. 119 (1997) 2240.
[47] S. Gong, E. Camara, F. He, M. Kirk Green, C.B. Lebrilla, Int. J. Mass Spectrom. 185–187 (1999) 401.
[48] R.M.C. Dawson, W.H. Elliot, K.M. Jones, Data for Biochemical Research, Oxford Science Publications, Oxford, 1986.
[49] R. Zenobi, R. Knochenmuss, Mass Spectrom. Rev. 17 (1998) 337.
[50] P. Juhasz, K. Biemann, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 4333.
[51] A.S. Woods, M.A. Huestis, J. Am. Soc. Mass Spectrom. 12 (2001) 88.
[52] K.A. Cox, S.J. Gaskell, M. Morris, A. Whiting, J. Am. Soc. Mass Spectrom. 7 (1996) 522.
[53] O. Burlet, C.-Y. Yang, S.J. Gaskell, J. Am. Soc. Mass Spectrom. 3 (1992) 337.