ELSEVIER

International Journal of Mass Spectrometry 220 (2002) 87–96

International Journal of Mass Spectrometry

www.elsevier.com/locate/ijms

Molecular recognition of arginine in small peptides by supramolecular complexation with dibenzo-30-crown-10 ether

Ryan R. Julian, Minta Akin, Jeremy A. May, Brian M. Stoltz, J.L. Beauchamp*

Beckman Institute, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA

Received 3 April 2002; accepted 16 July 2002

Abstract

The protonated, alkyl-guanidinium side chain of arginine forms a stable noncovalently bound complex with dibenzo-30-crown-10 (DB30C10) in the gas phase. This gas phase adduct is detected through the use of electrospray ionization mass spectrometry (ESI-MS). The supramolecular complex is stabilized by extensive hydrogen bonding and ion–dipole interactions between the protonated guanidinium group and the 10 oxygen atoms present in DB30C10. Competitive collision induced dissociation (CID) experiments demonstrate that DB30C10 possesses a higher affinity for alkyl-guanidinium ions than the related compound 27-crown-9 (27C9). This is attributed to the smaller size of 27C9, which does not afford extra space to accommodate the alkyl portion of the side chain of arginine. Recent experiments in our laboratory have demonstrated the molecular recognition and quantification of lysine residues in small peptides through the formation of 18-crown-6 (18C6) adducts. The molecular recognition capabilities of 18C6 and DB30C10 are mutually compatible and can be utilized to determine whether a peptide contains arginine or lysine (or both) without any prior knowledge of the peptide sequence. Competitive CID experiments demonstrate that the DB30C10/arginine interaction is stronger than the 18C6/lysine interaction when they are attached to the same peptide. In contrast to the behavior observed for 18C6 interacting with lysine containing peptides, multiple adducts of DB30C10 with a peptide containing multiple arginine residues are not observed in great abundance, limiting the utility of DB30C10 for the quantification of arginine residues. (Int J Mass Spectrom 220 (2002) 87–96)

Keywords: Arginine; Molecular recognition; Crown ether; Gas phase

1. Introduction

The field of supramolecular chemistry is devoted in part to the study of noncovalent molecular interactions (for a good general review see [1]). Among the myriad applications of supramolecular chemistry, the focus here will be on the discovery and development of specific receptors designed to recognize biologically relevant species. Recent efforts have produced substantial progress in this endeavor utilizing a variety of hosts including crown ethers [2], lariat crown ethers [3], cyclodextrins [4], calixarenes [5], cyclophanes [6], cyclofructans [7], sulfonate dyes [8], and monosaccharides [9] among others. Although the majority of these studies have been performed in solution, the number of gas phase studies of supramolecular chemistry has been increasing [10] and several of the above examples refer to gas phase work.

For the molecular recognition of a specific side chain in a peptide or protein, the charged amino acids

* Corresponding author. E-mail: jlbchamp@its.caltech.edu

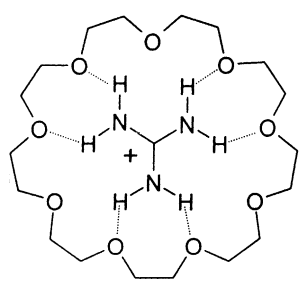
1387-3806/02/$ – see front matter
PII S1387-3806(02)00837-0

arginine, histidine, lysine, glutamic acid, and aspartic acid offer the best targets of opportunity. In the gas phase these remain distinguished as the most basic and most acidic of the twenty amino acids. The acidic side chains are difficult to distinguish due to the fact that both terminate with a carboxylic acid and differ only in the number of methylenes joining the acidic functionality to the backbone. The different chemical functionalities of the basic side chains enhances the possibility of achieving specificity. Lysine and arginine have received the most attention with regards to molecular recognition. We have recently demonstrated the excellent molecular recognition capabilities of 18-crown-6 (18C6) as a specific host for the side chain of lysine in the gas phase [2]. In these experiments, 18C6 is mixed into a solution with a lysine containing peptide which is then electrosprayed into the gas phase. The 18C6 forms a specific complex with the alkyl-ammonium group on the protonated side chain of lysine, stabilized by three hydrogen bonds. This complex forms in solution and is transferred intact to the gas phase using electrospray ionization. The quantification of lysine residues in small peptides is possible with this technique through the formation of multiple adducts [2].

In related gas phase work, Freiss and Zenobi have utilized a series of sulfonates for the molecular recognition of arginine [11]. The molecular recognition is primarily due to electrostatic attraction between the basic guanidinium group of arginine and an acidic disulfonate. The complexes are detected utilizing matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS). The N-terminus is additionally recognized by this technique, yielding one additional peak beyond the number of arginines in the peptide.

In addition, several cleverly designed molecules have been developed for the molecular recognition of arginine in solution. Dougherty and coworkers synthesized a cyclophane based host which utilizes a combination of hydrophobic, cation-π, and ion-pairing interactions to form a stable complex with arginine [6]. Schrader and coworkers have developed a series of bisphosphonate receptors that utilize a combination of hydrogen bonding, electrostatic interactions, and cation-π effects to recognize alkyl-guanidinium groups [12]. Bell et al. are responsible for the “arginine cork”, which demonstrates arginine selectivity through the formation of several specific hydrogen bonds, accompanied by electrostatic attraction [13].

In contrast, Lehn et al. first suggested that large macrocycles might serve as neutral receptors for arginine, based on the recognition of the guanidinium group by large crown ethers [14]. This idea is supported by several related experiments. In synthetic chemistry, guanidinium has been successfully used as a template for the synthesis of large macrocycles such as 27-crown-9 (27C9) as shown in structure **1** [15]. X-ray crystallography has shown that both 27C9 and 30-crown-10 (30C10) crown ethers and their derivatives are good receptors for guanidinium, forming inclusion complexes in the solid phase with a 1:1 stoichiometry [16]. In these complexes the molecules are coplanar and stabilized primarily through a network of hydrogen bonds as shown in structure **1**. For guanidinium itself, the 27-crown ethers were found to be the ideally sized hosts. These results suggest that 27C9 or 30C10 could be utilized for the molecular recognition of arginine in gas phase experiments in a manner analogous to that for 18C6 described above [2].



1

In the present work, experiments are conducted with various macrocycles to determine which is ideally suited to host the side chain of arginine. Competitive collision induced dissociation (CID) experiments are utilized to determine the relative bond dissociation energetics for the various crown/peptide interactions. The molecular recognition capabilities of 18C6

are utilized in combination with larger macrocycles, yielding additional information about the amino acid composition of peptides.

2. Methods

All data were obtained using a Finnigan LCQ ion trap quadrupole mass spectrometer without modification. Soft sampling is crucial for the detection of these non-covalent complexes. The critical instrument settings that yield adduct formation include capillary voltage 14.12 V, capillary temperature 200 °C, and tube lens offset −39 V. Higher capillary temperatures dissociate the crown ether/peptide complexes. The tube lens offset controls the acceleration of ions as they leave the capillary region. The tube lens voltage is minimized to reduce the number of energetic collisions with the He buffer gas.

Sample concentrations were typically kept in the ∼10 to 100 μM range for all species of interest, unless otherwise noted. All samples were electrosprayed in a mixture of 80:20 methanol/water. Crown ethers were added in a two- to three-fold excess to the sample and electrosprayed with the analyte in order to observe adducts. Samples were electrosprayed with a flow of 3–5 mL/min from a 500 mL Hamilton syringe for optimal signal. Silica tubing with an inner diameter of 0.005 in. was used as the electrospray tip. No acid was added to any of the samples, unless otherwise noted. All chemicals unless otherwise noted were purchased from Sigma or Aldrich and used without further purification. 27C9 was synthesized according to well established techniques [17].

All calculations were performed with the HyperChem 5.1 Professional Suite. Candidate structures were identified with molecular mechanics and were fully optimized at the PM3 semi-empirical level.

3. Results and discussion

3.1. DB30C10

The results for the complexation of DB30C10 with three peptides are shown in Fig. 1. In each case, the peptide is electrosprayed from a solution with DB30C10 present in two- to three-fold excess. The peptide KPPGFSPFR (a bradykinin analog that we will designate Kbk) forms a single DB30C10 adduct $[Kbk + DB30C10 + 2H]^{2+}$, as shown in Fig. 1a. Kbk contains only one C-terminal arginine, and the results in Fig. 1a are consistent with the molecular recognition of arginine by the DB30C10. Fig. 1b shows the results for YGGFMRGL (Enk). In this case, a single DB30C10 adduct corresponding to $[Enk + DB30C10 + 2H]^{2+}$ is observed, which is also consistent with the single internal arginine present in the peptide. These promising results suggested that the extension of this method to the recognition of multiple arginines would be possible. However, as seen in Fig. 1c, bradykinin (Bk, sequence RPPGFSPFR) does not form the expected double DB30C10 adduct. In fact, the single adduct $[Bk + DB30C10 + 2H]^{2+}$ is only observed in low abundance.

These results contrast with those obtained by the supramolecular complexation of lysine by 18C6. Fig. 2 shows the results of electrospraying a mixture of Kbk with 18C6 present in solution. As we have recently demonstrated in our labs [2], 18C6 forms stable gas phase adducts preferentially with lysine. This interaction leads to the base peak in Fig. 2, which corresponds to the $[Kbk + 18C6 + 2H]^{2+}$ adduct peak. In Fig. 2, the adduct peaks are much more intense than the uncomplexed peptide peaks. Furthermore, multiple adduct peaks are detected, despite the fact that 18C6 attaches only weakly to the N-terminus and the side chain of arginine.

One interpretation of this data would suggest that the low abundance of the DB30C10 peaks in Fig. 1 is the result of very weak gas phase binding between DB30C10 and the side chain of arginine, but further experimentation detailed afterwards shows this not to be the case. Recent related work in our labs has shown that the solution phase binding constant for crown/ion complexes is related to the abundance of the adduct peak observed in the gas phase [18]. Solution phase data on related crown ethers suggests that a low solution phase binding constant for arginine by DB30C10 may be a contributing factor to the low

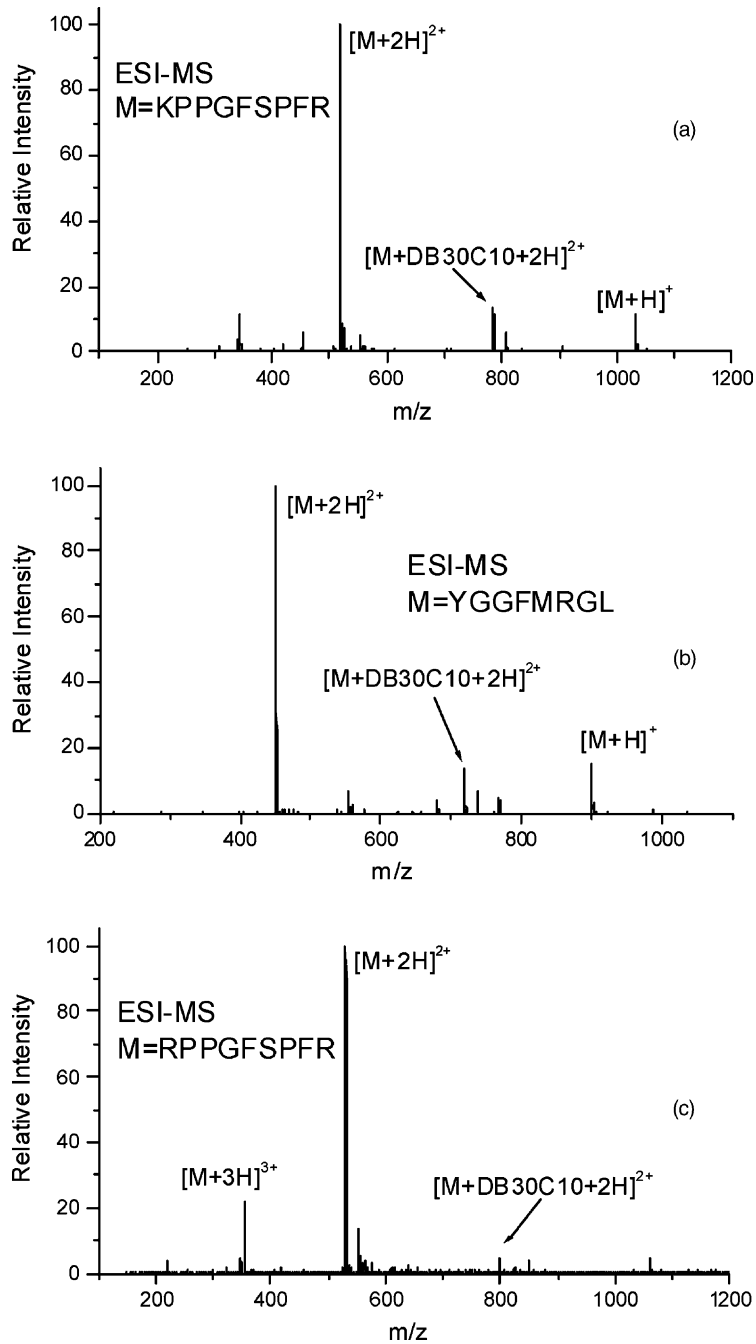


Fig. 1. (a) ESI-MS of KPPGFSPFR (Kbk) with DB30C10. This peptide forms one adduct with DB30C10, corresponding to the single arginine present. (b) ESI-MS of YGGFMRGL (Enk) with DB30C10. The single expected adduct is observed. (c) The ESI-MS of RPPGFSPFR (Bk) with DB30C10 yields only a single, low intensity adduct where two are expected.

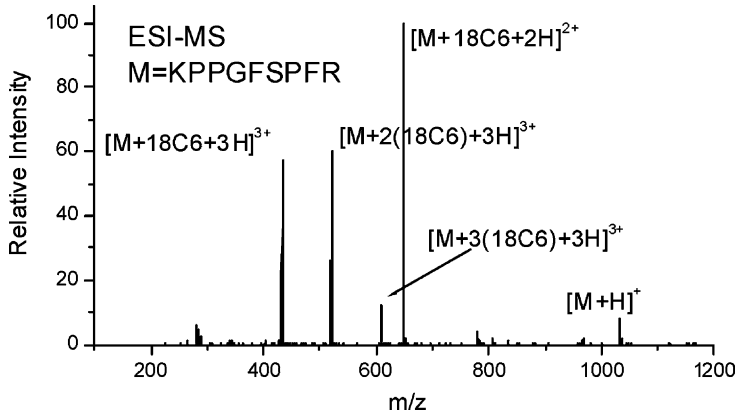
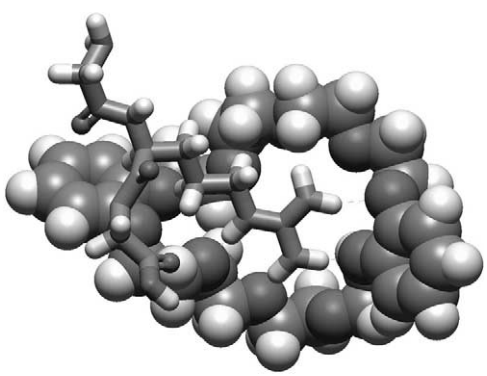


Fig. 2. ESI-MS resulting from the addition of 18C6 to a solution of KPPGFSPFR. In this case, the adducts are the most intense peaks present, and multiple 18C6 adducts are easily observed.

observed abundance of adducts in the gas phase. This may account for the low adduct abundance observed for bradykinin. Furthermore, molecular modeling demonstrates that the size of DB30C10 may lead to unfavorable steric interactions when two or more DB30C10 molecules are complexed with a peptide. The relative sizes are shown with the example of DB30C10 bound to GRG in structure **2**. Despite the low abundance of the gas phase adducts, DB30C10 forms strongly bound complexes with arginine once in the gas phase as will be detailed afterwards.



2

As discussed above, Boer et al. have suggested that the 27-crown ethers are the optimally sized hosts for guanidinium, because the crystal structure data indicates that the 30-crown ethers are slightly too large [16]. In order to ascertain the optimal host for arginine in the gas phase, we performed a competitive CID experiment in which both 27C9 and DB30C10 were attached to the dipeptide RR.[1] Fig. 3 shows the results for the CID of the $[RR + DB30C10 + 27C9 + 2H]^{2+}$ complex. Two fragments are observed which both correspond to the loss of a neutral crown ether, but the DB30C10 is retained by RR more often than 27C9 by a factor of 4. This suggests that arginine, with the larger alkyl-guanidinium side chain, prefers the larger DB30C10 host. Molecular modeling and crystal structures [16] indicate that for 27C9 to interact with arginine, the guanidinium group must pivot out of the plane of the crown ether, weakening the hydrogen bonds that stabilize the complex. Thus for arginine, the 30-crown ethers appear to be optimally sized.

3.2. DB30C10 and 18C6

The formation of DB30C10/peptide complex does not preclude the addition of 18C6 adducts if the peptide contains lysine residues in addition to arginine.

[1] Although these two arginines are not identical, it is anticipated from molecular modeling that the bond dissociation energies will be largely determined by the interaction of the crown ether with the side chain. Secondary interactions such as hydrogen bonds with the backbone should be largely equivalent for both of these crowns.

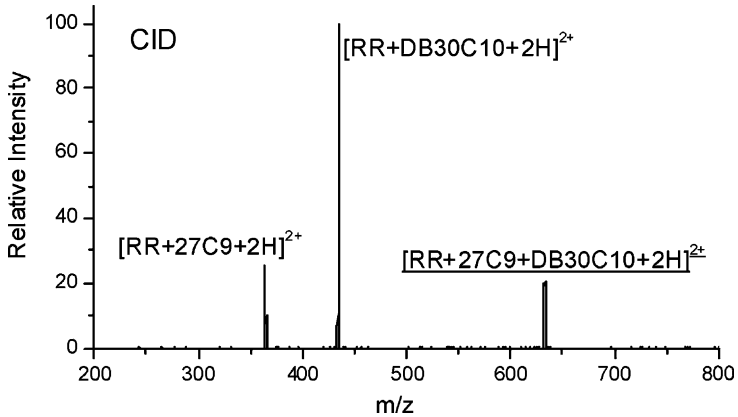


Fig. 3. CID spectrum of diarginine with 27C9 and DB30C10 both attached. The DB30C10 is retained preferentially, demonstrating a higher bond dissociation energy for the larger crown. The peak with an underlined label is being dissociated.

In other words, it is possible to form a complex between a peptide and both crown ethers simultaneously. Fig. 4 shows the results for such an experiment with the peptide TKPR. This peptide contains one lysine and one arginine, which should yield at least three different complexes, two corresponding to each crown attached individually and one complex with both crowns attached. As expected, Fig. 4 shows that $[TKPR+DB30C10+3H]^{3+}$, $[TKPR+18C6+2H]^{2+}$, and $[TKPR + DB30C10 + 18C6 + 3H]^{3+}$ are all observed. Therefore the crown ether based arginine and lysine molecular recognition techniques appear to be mutually compatible. In fact, for reasons which remain unclear, the presence of both crowns enhances the abundance of both adducts. Furthermore, a mixed peak such as $[TKPR + DB30C10 + 18C6 + 3H]^{3+}$ raises the issue of which crown ether is more strongly bound.

Fig. 5 shows the results of a series of experiments conducted to determine the relative binding strengths of 18C6 and DB30C10 to lysine and arginine, respectively. In each experiment, a peptide containing arginine and lysine is electrosprayed with both 18C6 and DB30C10. The peak corresponding to

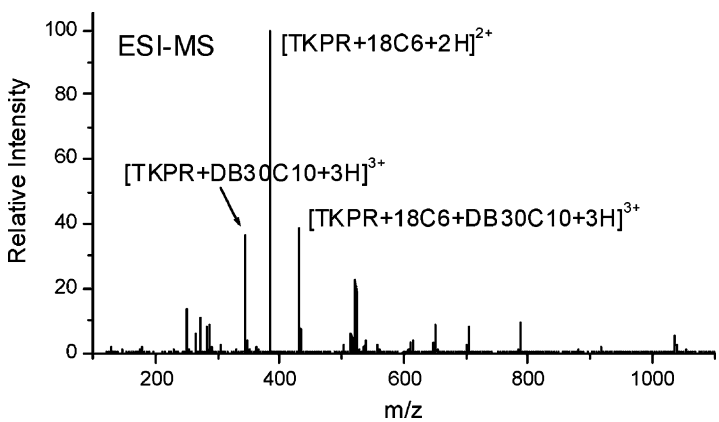


Fig. 4. ESI-MS of TKPR with both 18C6 and DB30C10. 18C6 attaches to the lysine and DB30C10 attaches to the arginine, demonstrating the compatibility of the two experiments.

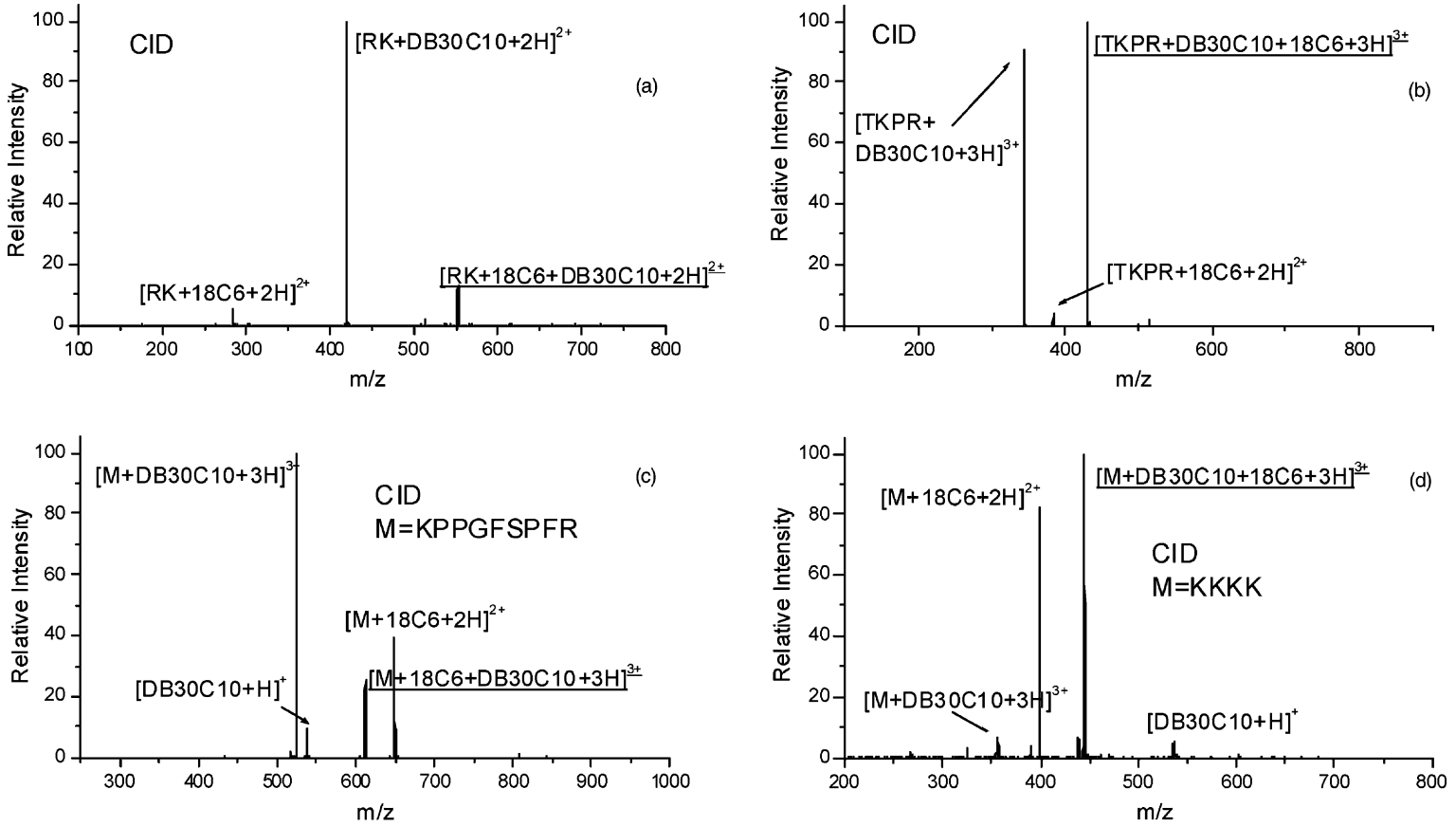


Fig. 5. CID spectra for the mixed [peptide + 18C6 + DB30C10] species for several peptides. (a) CID of the RK mixed cluster shows retention of the DB30C10, suggesting a higher binding energy for arginine recognition. (b) The same result is obtained with TKPR, with the arginine now on the C-terminus. (c) The larger peptide KPPGFSPFR also yields a higher binding affinity for arginine recognition by DB30C10. (d) Control experiment in which tetralysine (which has no arginine residues) is used. In this case, the 18C6 is retained, demonstrating a higher affinity for lysine than DB30C10. This suggests that 18C6 is specific for lysine recognition and similarly that DB30C10 specifically binds to arginine. In each case, the peak with an underlined label is being dissociated.

[peptide + 18C6 + DB30C10] is then isolated and subjected to mild CID, leading to the loss of the crown ether with the lowest gas phase binding energy. The peptides RK, TKPR, and Kbk were chosen for their range in size and relative positions of lysine and arginine. As seen in Fig. 5a–c, the DB30C10 remained attached to the peptide preferentially over 18C6 by a substantial margin in every case. Fig. 5d shows the results for the same experiment with tetralysine. In this case, the peptide contains no arginine residue. Although peptides lacking arginine will not bind DB30C10 to a significant extent, we found that it was possible to form, in low abundance, non-specific complexes with multiple, sequential lysines present. The results are reversed in Fig. 5d, with the 18C6 remaining attached to tetralysine preferentially over DB30C10.

These results further support the notion of selectivity for complexation of 18C6 and DB30C10 with lysine and arginine, respectively. If adduct formation were the result of non-specific interactions with the peptides, then the relative binding energetics would be difficult to predict. As clearly illustrated by the data in Fig. 5, this is not the case. For arginine containing peptides, DB30C10 has a higher binding energy. In the absence of arginine, lysine contain-

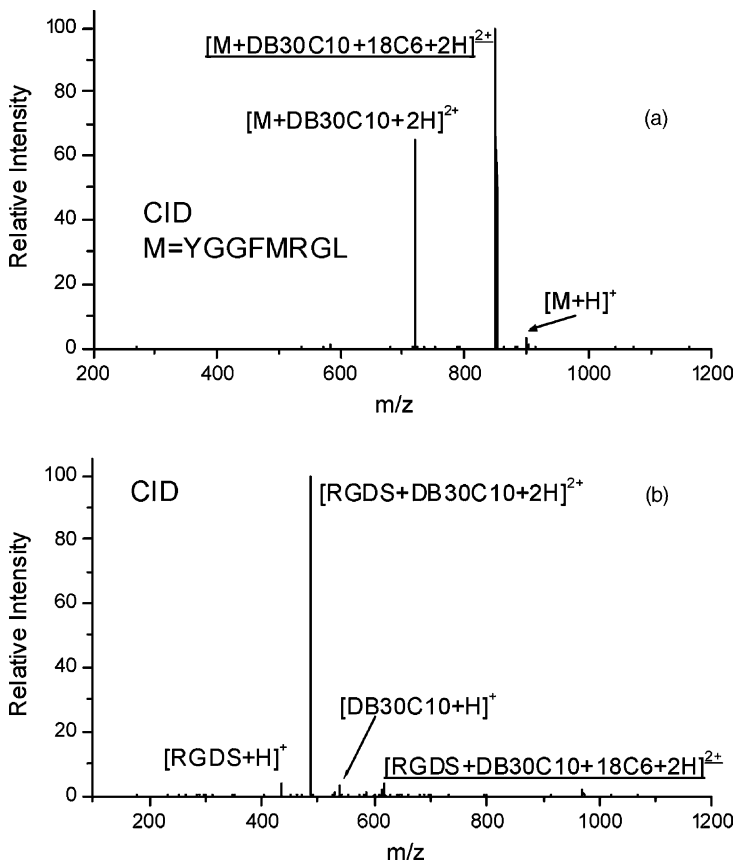


Fig. 6. In the absence of lysine, CID on the [peptide + 18C6 + DB30C10] peak yields no competitive dissociation, with all of the DB30C10 being retained over the 18C6. Similar results are obtained for both (a) YGGMRGL and (b) RGDS. This allows for the distinction of lysine vs. arginine containing peptides without prior knowledge of the sequence. In each case, the peak with an underlined label is being dissociated.

ing peptides bind 18C6 with a higher affinity. Each crown binds specifically to a particular amino acid. The bond strength of the $[18C6 + NH_4]^+$ complex is 71 ± 3 kcal/mol [19]. The experimental bond dissociation energy for 18C6 attached to a protonated primary amine in the gas phase is not known, but is estimated to be >50 kcal/mol [20]. The bond dissociation energy for DB30C10 attached to a guanidinium ion should prove to be even higher in the gas phase. This likely results from the more extensive hydrogen bonding along with favorable ion–dipole interactions (see structure **1**).

The combination of the 18C6 and DB30C10 experiments yields more information about the composition of a peptide than can be obtained from each experiment individually. If a peptide contains only lysine, then experiments such as that shown in Fig. 5d will demonstrate that there are no arginines present. If the peptide contains both arginine and lysine, then the number of lysines can be quantified and the presence of at least one arginine can be confirmed by experiments such as those shown in Fig. 5a–c. Fig. 6 shows the results for the competitive dissociation of the corresponding [peptide + 18C6 + DB30C10] peak for two peptides that contain only arginine. In the absence of lysine, 18C6 will form a weakly bound complex with the N-terminus of the peptide. Upon CID only the loss of the 18C6 is observed. The DB30C10 is retained without competition from 18C6 adducts for arginine containing peptides that do not contain lysine (Fig. 6). This allows for the discrimination between arginine and arginine and lysine containing peptides because only peptides that contain both amino acids will yield CID peaks in which both 18C6 and DB30C10 are retained. This would allow the lysine and arginine containing fragments of a tryptic digest to be identified without any prior knowledge about the sequence of a protein.

4. Conclusions

DB30C10 forms a specific, strongly bound complex with the alkyl-guanidinium group in the side chain of arginine in the gas phase. This interaction can be used to identify the presence of arginine in a peptide without any prior knowledge of its sequence. DB30C10 is a better host for arginine than 27C9 because the larger crown ether affords the extra space needed to accommodate the alkyl portion of the side chain. Comparison with lysine recognition by 18C6 demonstrates that the relative bond dissociation energy for the DB30C10 system is greater. Notwithstanding, the bond dissociation energy for 18C6 is greater than that for DB30C10 in lysine containing peptides that do not contain arginine, demonstrating the specificity of both interactions. This is the first example of the molecular recognition of arginine by a neutral host in the gas phase.

Further development in the field of molecular recognition of biological molecules will lead to advances in proteomics and our understanding of how biological systems employ molecular recognition themselves. The results presented in this paper allow for the partial sequencing of peptides in the gas phase, but the ability to use DB30C10 as a delivery agent for a wider range of site specific chemistry in the gas phase is even more promising. This can be accomplished through the development of lariat crown ethers, which have a variable side chain covalently attached. Similar work utilizing 18C6 is currently in progress in our laboratory, but the high binding energy of the DB30C10 offers an excellent opportunity to develop chemistry that can be induced on the peptide by the lariat prior to dissociation of the crown ether. Through such reagents, the selective cleavage of peptides in the gas phase, or the determination of gas phase structures could theoretically be accomplished.

Acknowledgements

The authors gratefully acknowledge funding provided by NSF (Grant CHE-9727566), the Beckman Institute Foundation, the California Institute of Technology, the Camille and Henry Dreyfus Foundation (New Faculty Award to BMS), and Abbott Labs (graduate fellowship to JAM).

References

[1] P.D. Beer, P.A. Gale, D.K. Smith, Supramolecular Chemistry, Oxford University Press, New York, 1999.
[2] R.R. Julian, J.L. Beauchamp, Int. J. Mass Spectrom. 210 (2001) 613.
[3] A. Galan, D. Andreu, A.M. Echavarren, P. Prados, J. de Mendoza, J. Am. Chem. Soc. 114 (1992) 1511.
[4] (a) S.L. Hauser, E.W. Johanson, H.P. Green, P.J. Smith, Org. Lett. 2 (23) (2000) 3575;
(b) J. Ramirez, F. He, C.B. Lebrilla, J. Am. Chem. Soc. 120 (1998) 7387.
[5] R. Ludwig, J. Fresen, Anal. Chem. 367 (2000) 103.
[6] S.M. Ngola, P.C. Kearney, S. Mecozzi, K. Russell, D.A. Dougherty, J. Am. Chem. Soc. 121 (1999) 1192.
[7] M. Sawada, M. Shizuma, Y. Takai, H. Adachi, T. Takeda, T. Uchiyama, Chem. Commun. (1998) 1453.
[8] B. Salih, R. Zenobi, Anal. Chem. 70 (1998) 1536.
[9] P. Krishna, S. Prabhakar, M. Manoharan, E.D. Jemmis, M. Vairamani, Chem. Commun. (1999) 1215.
[10] C.A. Schalley, Int. J. Mass Spectrom. 194 (2000) 11.
[11] S.D. Friess, R. Zenobi, J. Am. Soc. Mass Spectrom. 12 (7) (2001) 810.
[12] (a) S. Rensing, A. Arendt, A. Springer, T. Grawe, T. Schrader, J. Org. Chem. 66 (2001) 5814;
(b) T.H. Schrader, Tetrahedron. Lett. 39 (1998) 517.
[13] T.W. Bell, A.B. Khasanov, M.G.B. Drew, A. Filikov, T.L. James, Angew. Chem., Int. Ed. 38 (1999) 2543.
[14] J.M. Lehn, P. Vierling, R.C. Hayward, J. Chem. Soc., Chem. Commun. (1979) 296.
[15] (a) K. Madan, D.J. Cram, J. Chem. Soc., Chem. Commun. 11 (1975) 427;
(b) E.P. Kyba, R.C. Helgeson, K. Madan, G.W. Gokel, T.L. Tarnowski, S.S. Moore, D.J. Cram, J. Am. Chem. Soc. 99 (1977) 2564.
[16] J.A.A. Boer, J.W.H.M. Uiterwijk, J. Geevers, S. Harkema, D.N. Reinhoudt, J. Org. Chem. 48 (1983) 4821.
[17] R. Chenevert, L.J. D'Astous, Hetercyclic Chem. 23 (1986) 1785.
[18] R.R. Julian, J.L. Beauchamp, J. Am. Soc. Mass Spectrom. 13 (2002) 493.
[19] M. Meot-Ner, L.W. Sieck, J.F. Leibman, S. Scheiner, J. Phys. Chem. 100 (1996) 6445.
[20] A. Colorado, J. Brodbelt, J. Am. Soc. Mass Spectrom. 7 (1996) 1116.