Identifying Nonspecific Ligand Binding in Electrospray Ionization Mass Spectrometry Using the Reporter Molecule Method

Running title: **Reporter molecule method**

Nian Sun, Jiangxiao Sun, Elena N. Kitova and John S. Klassen*

Alberta Ingenuity Centre for Carbohydrate Science and Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2

Abstract

The application of the reporter molecule (M_{rep}) method for identifying nonspecific complexes in the ES-MS analysis of protein-ligand and DNA-ligand interactions *in vitro* is described. To test the reliability of the method, it was applied to the ES-MS analysis of protein-carbohydrate complexes originating from specific interactions in solution and from nonspecific interactions in the ES process. These control experiments confirm the basic assumptions underlying the M_{rep} method, namely that nonspecific ligand binding is a random process, and that the ES droplet histories for specific and nonspecific complexes are distinct. The application of the M_{rep} method to the ES-MS analysis of the sequential binding of the ethidium cation, a DNA intercalator, to single and double strand oligodeoxynucleotides is also described and highlights the general utility of the method.

Introduction

The association of biological molecules to form specific, non-covalent complexes is implicated in virtually all biological processes, including the immune response, inflammation and bacterial and viral infections. Detailed information regarding the structure and stability of these non-covalent complexes is essential to a complete understanding of biological processes, as well as the development of new therapeutics. Critical to achieving these goals are analytical methods capable of measuring the stoichiometry and affinity of non-covalent biological complexes *in vitro*. Recently, the direct electrospray ionization mass spectrometry (ES-MS) assay has emerged as a powerful tool for quantifying the association thermochemistry of protein-small molecule ligand interactions in solution.[1-3] With the ES-MS assay, protein-ligand binding constants (K_a) are determined from the ratio (R) of the total abundance (Ab) of bound and unbound protein ions (e.g. PL^{n+}, P^{n+}) measured in the gas phase by ES-MS for solutions of known initial concentrations of protein ($[P]_o$) and ligand ($[L]_o$), eqs 1, 2.

$$\frac{[PL]_{eq}}{[P]_{eq}} = \frac{Ab(PL^{n+})}{Ab(P^{n+})} = R \tag{1}$$

$$K_a = \frac{R}{[L]_o - \frac{R}{1+R}[P]_o} \tag{2}$$

The technique boasts a number of strengths, including its simplicity (no labeling or immobilization required), speed (measurements can usually be completed within a few seconds), and specificity (the unique ability to provide direct insight into stoichiometry and ability to study multiple binding equilibria simultaneously). Additionally, when

performed using nanoflow ES (nanoES), the ES-MS assay affords high sensitivity, normally consuming picomoles or less of analyte per analysis.

The ES-MS assay has been used to quantify a variety of protein-ligand interactions[4-10] and, in many instances, the affinities agree well with values obtained by more established analytical methods, such as isothermal titration calorimetry (ITC) and surface plasmon resonance. Despite these successful examples, there remain a number of experimental limitations to be overcome for the direct ES-MS assay to achieve its full potential. Among these is the problem of false positives, which result from the formation of nonspecific protein-ligand complexes during the ES process. These nonspecific interactions are not present in bulk solution but form in the ES droplets due to concentration effects.[11] The resulting complexes may be sufficiently stable in the gas phase that they survive until detection.[11,12] In fact, it has been shown that nonspecific interactions involving a given protein and its specific ligand can be more stable than the corresponding specific complex in the gas phase.[13] The occurrence of nonspecific ligand binding obscures the true binding stoichiometry in solution and introduces errors into the K_a values derived from ES-MS measurements. The problem of nonspecific binding is most severe in the case of weak ligand interactions ($K_a < 10^5$ M^{-1}) because high concentrations of ligand are generally required to produce detectable levels of complex.[14]

Our laboratory recently demonstrated that, for carbohydrate ligands, the distribution of ligands bound nonspecifically to proteins during the ES process is independent of protein structure and size.[15] Based on these findings, we developed a quantitative approach to correct for nonspecific protein-ligand binding in ES-MS analysis. The method involves the addition of a reference protein (P_{ref}), which does not bind

specifically to any of the solution components, to the ES solution containing the protein and ligand of interest.[15] The occurrence of nonspecific protein-ligand binding is identified by the appearance of peaks corresponding to ions of nonspecific (P_{ref} + *q*ligand) complexes in the ES mass spectrum. Additionally, the fraction of P_{ref} undergoing nonspecific ligand binding provides a quantitative measure of the contribution of nonspecific ligand binding to the measured abundance of protein and specific protein-ligand complex. As a result, errors in binding stoichiometry and K_a, introduced by nonspecific ligand binding, can be corrected.

This simple correction method for false positives has dramatically improved the reliability of the ES-MS assay for quantifying protein-small molecule ligand interactions, in particular, low affinity complexes.[14,16] However, the P_{ref} method does have certain limitations. One of the underlying assumptions in the P_{ref} method is that in-source gas phase dissociation of the nonspecific protein-ligand complexes does not occur or, if it does, it affects equally all protein complexes present.[15] Based on available data, the gas phase stability of nonspecific protein-small molecule complexes are relatively insensitive to the structure of the protein, but are sensitive to size and charge state of the protein.[13] The effects of differential gas phase dissociation can be minimized by choosing a P_{ref} that is similar in molecular weight to the protein(s) of interest. Of course, this precaution does not preclude the possibility of in-source dissociation of the specific (formed in solution) protein-ligand interactions, which may be less stable than the nonspecific interactions. Additionally, it has yet to be established whether the P_{ref} method can be extended to monitor nonspecific ligand binding to other biopolymers or macromolecules.

Recently, a qualitative method to identify the formation of nonspecific protein-protein interactions in ES-MS analysis was reported. The method involves the addition of a reporter molecule (M_{rep}), which does not bind specifically to the proteins and protein complexes of interest, to the ES solution at high concentration.[17] The high concentration promotes the formation of nonspecific interactions between M_{rep} and all of the protein species. From the measured distributions of nonspecifically bound M_{rep} it is possible to establish whether a given protein complex originates in solution or whether it forms, at least in part, from nonspecific binding during the ES process. Complexes originating from nonspecific interactions will necessarily have different droplet histories than those of the unbound protein and specific protein complexes. Specifically, the nonspecific protein complexes are formed later in the ES process from older and more concentrated ES droplets. These older droplets will be more concentrated in protein, as well as M_{rep}. As a result, the distributions of nonspecifically bound M_{rep} molecules observed for the unbound protein (if present) and specific protein complex(es) will differ from those observed for the nonspecific complexes - the nonspecific complexes will experience more extensive nonspecific binding to M_{rep}. Although the M_{rep} method does not provide a quantitative measure of the extent of nonspecific binding it can be made resistant to in-source dissociation by selecting an M_{rep} that forms strong gas phase interactions. Additionally, the method is not limited to protein-protein interactions but can be extended to protein-small molecule ligand interactions, as well as interactions involving other biopolymers or macromolecules.

Here, we describe for the first time the application of the M_{rep} method for identifying the formation of nonspecific ligand interactions with proteins and DNA in

ES-MS. The method was implemented in the ES-MS analysis of protein-carbohydrate interactions, as well the sequential binding of ethidium cation, a small DNA intercalator, to single and double strand oligodeoxynucleotides. Importantly, the basic assumptions underlying the M_{rep} method, namely that nonspecific ligand binding is a random process, and that the ES droplet histories for specific and nonspecific complexes are distinct, were shown to be generally valid.

Materials and Methods

Proteins, DNA and small molecules

The carbohydrate-binding antibody single chain fragment, scFv (MW 26 539 Da), was produced using recombinant technology.[18] The scFv was concentrated and dialyzed against deionized water using microconcentrators (Millipore Corp., Bedford, MA) with a molecular weight cut-off of 10 kDa, and lyophilized. The scFv was weighed immediately after removing it from the lyophilizer, dissolved in a known volume of aqueous 50 mM ammonium acetate and stored at –20 ^{o}C until used. The 20-mer oligodeoxynucleotides (ODN) 5′-CGCCCAACCCTCCTTCCCGC-3′ (ODN1, MW 5894.9 Da) and 5′-GCGGGAAGGAGGGTTGGGCG-3′ (ODN2, MW 6344.2 Da) were purchased from ACGT Corporation (Toronto, Canada). Stock solutions (300 μM) of each ODN were prepared by dissolving known amounts of ODN in deionized water. Portions of each stock solution were mixed in equimolar proportions and diluted with STE buffer (1 M pH 8 Tris-Cl, 3 M NaCl, 0.5 M pH 8 EDTA in water) and deionized water to obtain a final concentration of 300 μM. Duplex annealing was performed by heating the solution to 95 °C and gradually cooling to room temperature over a period of 30 minutes. The duplex was precipitated out of solution (initial volume 300 μL) by adding 30 μL of sodium

acetate (2.9 M pH 5.2) and 600 μL of 100% ethanol. The entire mixture was stored at -20 ºC overnight. The duplex pellet was then washed twice using 95% ethanol. The duplex was then dissolved in 100 μL of 150 mM ammonium acetate and dialyzed against 150 mM ammonium acetate using a microconcentrator with a molecular weight cut-off of 10 kDa. The purified synthetic trisaccharides, αTal[αAbe]αMan (**1**), 2-trimethylsilylethyl 4-O-[(4-O-α-D-galactopyranosyl)-β-D-galactopyranosyl]-β-D-glycopyranosid (**2**) were provided by D. Bundle (University of Alberta). The disaccharide trehalose (**3**) and the trisaccharide maltotriose (**4**) were purchased from Sigma-Aldrich Canada (Oakville, ON) and used without further purification. Ethidium bromide (**5**) was provided by T. Lowary (University of Alberta) and used without further purification. The structures of **1 - 5** are shown in Figure 1. The ES solutions were prepared by mixing known amounts of the scFv or DNA, ligand and M_{rep} stock solutions.

Mass spectrometry

All experiments were performed on an Apex II Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer (Bruker, Billerica, MA) equipped with an external nanoES ion source. NanoES was performed using aluminosilicate or borosilicate capillaries (1.0 mm o.d., 0.68 mm i.d.), pulled to ~5 μm o.d. at one end using a P-2000 micropipette puller (Sutter Instruments, Novato, CA). Details of the instrumental parameters employed in positive ion mode are given below. For measurements in negative ion mode, the voltage polarity was switched. The electric field required to spray the solution was established by applying a voltage of 800 V to a platinum wire inserted inside the glass nanoES tip. The solution flow rate was typically ~20 nL/min. The droplets and gaseous ions emitted from the nanoES tip were introduced into the mass

spectrometer through a stainless steel capillary (i.d. 0.43 mm) maintained at an external temperature of 66 °C. The ion/gas jet sampled by the capillary (48 – 52 V) was transmitted through a skimmer (0 – 2 V) and stored electrodynamically in an rf hexapole. A hexapole accumulation time of 1.5 - 2.0 s was used for measurements performed in positive ion mode and 2- 8 s in negative ion mode. Ions were ejected from the hexapole and accelerated to ~2700V into a 9.4 T superconducting magnet, decelerated, and introduced into the ion cell. The trapping plates of the cell were maintained at a constant potential of (1.4 - 1.8 V) throughout the experiments. The typical base pressure for the instrument was ~$5x10^{-10}$ mbar.

Data acquisition was controlled by an SGI R5000 computer running the Bruker Daltonics XMASS software, version 5.0. Mass spectra were obtained using standard experimental sequences with chirp broadband excitation. The time domain signal, consisting of the sum of 20 - 40 transients containing 128 or 256 K data points per transient, were subjected to one zero-fill prior to Fourier transformation.

Implementation of the M_{rep} method

An overview of the implementation of the M_{rep} method for distinguishing specific from nonspecific protein-ligand interactions in ES-MS measurements is given below. Considered are the following cases: *i.* ions corresponding to free protein (P) and its specific protein–ligand complex (PL) are detected by ES-MS, *ii.* ions corresponding P and both specific and nonspecific PL complexes are detected and *iii*. ions corresponding to P and only the nonspecific PL complex are detected. The same general approach can be applied to the ES-MS analysis of ligand binding to other biopolymers or macromolecules.

i. P and specific PL complex

The addition of M_{rep} at relatively high concentration to the ES solution containing P and specific PL complex will result in the nonspecific attachment of one or more M_{rep} molecules to P and PL during the ES process and the appearance of peaks in the mass spectrum corresponding to free and ligand-bound protein, and free and ligand-bound protein associated with one or more M_{rep} molecules, i.e., $P(M_{rep})_i^{n+}$ and $PL(M_{rep})_i^{n+}$, where $i = 0, 1, 2,\ldots,$. The fraction (f_i) of P and of PL bound nonspecifically to i molecules of M_{rep} (relative to all possible numbers of M_{rep}) is given by eq 3 and 4, respectively:

$$f_{P,i} = \frac{AbP(M_{rep})_i^{n+}}{\sum_i AbP(M_{rep})_i^{n+}} \tag{3}$$

$$f_{PL,i} = \frac{AbPL(M_{rep})_i^{n+}}{\sum_i AbPL(M_{rep})_i^{n+}} \tag{4}$$

where $Ab\mathrm{P}(M_{rep})_i^{n+}$ and $Ab\mathrm{PL}(M_{rep})_i^{n+}$ are the measured abundance of the P^{n+} and PL^{n+} ions, respectively, bound nonspecifically to i molecules of M_{rep}. The P^{n+} and PL^{n+} ions are expected to have identical or nearly identical ES droplet histories in the case where the PL^{n+} ions originate exclusively from specific interactions in solution.[15] As discussed above, the nonspecific association of small molecules to proteins during the ES process is expected to be independent of protein structure and size. Consequently, in the absence of in-source dissociation, the distribution of M_{rep} bound nonspecifically to the P and PL species will be identical (*i.e.*, $f_{P,i} = f_{PL,i}$), Figures 2a,b.

ii. P and both specific and nonspecific PL complex

If nonspecific protein-ligand binding contributes to signal of the PL^{n+} ions, then the P^{n+} and some of the PL^{n+} ions will have different droplet histories and the

distributions of M_{rep} bound nonspecifically to P and PL will no longer be equivalent (i.e., $f_{P,i} \neq f_{PL,i}$). Specifically, the nonspecific PL^{n+} ions form preferentially from late generation droplets that are enriched in L, as well as M_{rep}. Consequently, the PL^{n+} ions that result from nonspecific association are expected to undergo more extensive nonspecific binding to M_{rep} than the P^{n+} ions, Figures 2a,c.

iii. P and nonspecific PL complex

It is also useful to consider the situation where L does not bind specifically to P in solution. In this case, all of the detected PL_j^{n+} ions, where j = 1, 2,…, originate from nonspecific interactions during the ES process. In the presence of M_{rep}, both L and M_{rep} will bind nonspecifically to P during the ES process leading to the appearance of $PL_jM_{rep,i}^{n+}$ ions. The composition of the $PL_jM_{rep,i}^{n+}$ ions will ultimately depend on the concentrations of L and M_{rep} and their relative efficiencies of nonspecific binding to P. In the simplest case (equimolar concentrations of L and M_{rep} in bulk solution and equivalent nonspecific binding efficiencies) the overall distribution of P bound nonspecifically to L and M_{rep} (i.e., $PL_jM_{rep,i} \equiv PM_{rep,k}$, where $k = i + j$) will resemble that of a Poisson process (at least at low to moderate concentrations of these molecules) and the chemical makeup of the $PL_jM_{rep,i}$ species will be statistical (i.e., $f_{PL,1} = f_{PMrep,1}$, $2f_{PL,2} = 2f_{PMrep,2} = f_{PL,1Mrep,1}$, …), Figure 3.

In principle, any molecule can be used to play the role of M_{rep}. However, small, neutral polyfunctional molecules capable of forming strong nonspecific interactions with the target protein (or macromolecule) in the gas phase are preferred. In this case nonspecific binding of M_{rep} to the proteins and protein complexes will not alter the charge state distribution of the protein ions and will not spread the protein ion signal over

a broad range of m/z values, where differences in detection efficiency may complicate the comparison of the f_i values. For these reasons, small carbohydrates (di- and trisaccharides) are ideally suited for the role of M_{rep}.

Results and Discussion

a. ES-MS of protein-carbohydrate interactions

To test the M_{rep} method for monitoring nonspecific protein-ligand binding, the method was applied to the ES-MS analysis of several protein solutions containing either interacting or non-interacting carbohydrates. The scFv of the monoclonal antibody Se155-4 and its specific trisaccharide ligand **1** served as a model specific protein-carbohydrate complex. The scFv possesses a single binding site for **1** with a K_a of 1.2 x 10^5 M^{-1}.[19] Additionally, the M_{rep} method was tested in the situation where protein-carbohydrate interactions originated exclusively from nonspecific binding in the ES process. The disaccharide **3** and trisaccharide **4**, neither of which binds specifically to the scFv in solution (data not shown), served as model "non-interacting" carbohydrate ligands for the scFv.

It is useful to consider first the situation where the ions corresponding to protein-carbohydrate complexes originated exclusively from nonspecific interactions in the ES process. Shown in Figure 4a is an illustrative ES mass spectrum obtained for a solution of scFv (10 μM) and **3** and **4** at equimolar concentrations (150 μM). Ions corresponding to unbound scFv and scFv bound nonspecifically to one or more molecules of **3** or **4** were identified, i.e, $scFv^{n+}\mathbf{3}_i\mathbf{4}_j$ at n = 10 – 12 and i, j = 0 – 3. Plotted in Figure 4b is the normalized distribution of the $scFv\mathbf{3}_i\mathbf{4}_j$ species, as determined from the mass spectrum. Notably, the distribution determined for $scFv\mathbf{3}_i\mathbf{4}_j$ species is consistent with the

distribution expected for two molecules, at equimolar concentrations, that have identical (or nearly so) nonspecific binding efficiencies. These results support the assumption that nonspecific protein-carbohydrate binding in the ES process occurs in a random fashion, independent of protein size and structure.[11]

ES-MS analysis was also performed on solutions containing the scFv and its specific ligand, **1** at four different initial concentrations of protein and ligand (9 μM and 9 μM; 18 μM and 18 μM; 6 μM and 18 μM; 9 μM and 18 μM). The trisaccharide **2**, which does not interact with scFv in solution, served as M_{rep} for these measurements. Shown in Figure 5 are illustrative mass spectra acquired for two different solutions: one containing 9 μM of scFv, 9 μM **1** and 47 μM M_{rep} (Figure 5a); and one containing 6 μM scFv, 18 μM **1** and 47 μM M_{rep} (Figure 5b). Ions corresponding to scFv$(\mathbf{2})_i^{n+}$ and (scFv + **1**)$(\mathbf{2})_i^{n+}$, where n = 9 - 11 and i = 0 - 2, were identified in both cases. Notably, there were no (scFv + 2(**1**))$(\mathbf{2})_i^{n+}$ ions, resulting from the nonspecific association of scFv with **1**, detected at any of the concentrations investigated. Reported in Table 1 are the ratios of f_i values determined for the scFv and (scFv + **1**) species for each of the four solutions. Also listed are the corresponding K_a values for the (scFv + **1**) complex, which were calculated from the relative abundance of bound and unbound scFv ions, as determined from the mass spectra, using eq 2.

Importantly, the K_a values determined for the two equimolar solutions of scFv and **1** are in good agreement with the literature value determined at 25 °C by ITC.[19] These results suggest that nonspecific binding of **1** to scFv during the ES process did not contribute appreciably to the signal for the (scFv + **1**)$^{n+}$ ions. At the same time, the ratios of the f_i values are close to unity, which is consistent with the scFv^{n+} and (scFv + **1**)$^{n+}$

ions having similar ES droplet histories. In contrast, the K_a values obtained for solutions where the concentration of **1** was two or three times larger than that of scFv are 30 - 40% larger than the values obtained at equimolar concentrations. At the same time, the ratios of f_i values deviate noticeably from unity. The inflated K_a values and the non-equivalent f_i values are consistent with the occurrence of nonspecific protein-ligand binding during ES-MS analysis. These findings are also consistent with a previous observation, made using the P_{ref} method, of nonspecific binding between **1** and the scFv at molar ratios >2.[15]

b. ES-MS of DNA-ligand interactions

The direct ES-MS assay is also finding widespread use in the characterization of DNA-ligand interactions.[20-23] To demonstrate the general utility of the M_{rep} method for the detection of nonspecific ligand binding to macromolecules in ES-MS, it was used to monitor the nonspecific binding between the ethidium cation **5**, a small intercalator, and single (**SS**) and double strand ODNs (**DS**). Intercalators, such as **5**, interact with DNA by binding between base pairs; they exhibit low specificity and can bind to multiple sites within single and double strand DNA in solution.[24-27] The application of the ES-MS assay to probe the number of binding sites and to evaluate the equilibrium constants for the sequential binding of intercalators to DNA may be complicated by the contribution of nonspecific binding, particularly at the high ligand concentrations needed to populate extensively the multiple binding sites that exist.

To test the applicability of the M_{rep} method for identifying nonspecific DNA-drug interactions, ES-MS measurements were performed in negative ion mode on a series of aqueous solutions containing **SS** (ODN1) or **DS** (consisting of ODN1 and ODN2), at fixed concentrations (17 μM (**SS**) or 34 μM (**DS**)) and **5** at concentrations ranging from

20 to 70 μM in the case of **SS** and from 50 to 160 μM in the case of **DS**. The trisaccharide **1** (at a concentration of 83 μM) served as M_{rep} for all of the binding measurements. Shown in Figures 6a,b are illustrative ES mass spectra acquired for solutions of **SS** and **5**. At the lower ligand concentrations investigated, deprotonated ions corresponding to free **SS**, at charge states -4 to -5, and **SS** bound to a single molecule of **5** at the same charge states were detected, i.e. $\mathbf{SS}^{z-}$ and $(\mathbf{SS} + \mathbf{5})^{z-}$ ions (Figure 6a). Nonspecific attachment of up to two molecules of **1** to the $\mathbf{SS}^{z-}$ and $(\mathbf{SS} + \mathbf{5})^{z-}$ ions was also observed. Alkali metal ion adducts were also evident in the mass spectra. Increasing the concentrations of **5** resulted in an increase in the number of bound ligands; at the highest concentrations investigated, **SS** ions bound to as many as five molecules of **5** were detected (Figure 6b). Interestingly, the distributions of bound **5** at these higher concentrations are suggestive of cooperative ligand binding, with the $(\mathbf{SS} + q\mathbf{5})^{z-}$ ions at $q \geq 3$ being unusually abundant compared to $q<3$ species. Our laboratory is currently investigating this phenomenon in more detail. Shown in Figures 6c,d are illustrative ES mass spectra acquired for solutions of **DS** and **5**. Ions corresponding to free **DS**, at charge states -5 and -6, and **DS** bound to as many as five molecules of **5** were detected. The nonspecific binding of one molecule of **1** to the free and bound **DS** species was observed, as well as abundant alkali metal ion adducts.

From relative abundance of ligand-bound and unbound **SS** or **DS** ions measured by ES-MS, the apparent equilibrium constant ($K_{a,1}$) for the attachment of one molecule of **5** to **SS** and to **DS** was calculated at the different ligand concentrations investigated (Table 2). Overall, the $K_{a,1}$ values determined for **DS** are ~3 fold larger than for **SS**, which is consistent with the findings of a previous study of binding of **5** with single and

double strand DNA.[28] Notably, within the range of concentrations of **5** investigated, the magnitude of $K_{a,1}$ measured for **DS** was found to be essentially constant, ~5 x 10^4 M^{-1}, suggesting that nonspecific binding between **DS** and **5** during ES-MS analysis was insignificant. In contrast, the $K_{a,1}$ values for binding of **5** to **SS** exhibit a small but measurable dependence on ligand concentration. At the lower concentrations investigated (20 - 50 μM), the $K_{a,1}$ value is constant, ~1.2 x 10^4 M^{-1}, but increases by 15% at the highest concentration studied. Taken on their own, these results suggest that nonspecific binding of **5** to **SS** contributed to the mass spectra measured for the solutions with ligand concentrations >50 μM. Also listed in Table 2 are the ratios of f_i values determined for free and ligand-bound **SS** and **DS** species at each concentration. Notably, the f_i ratios determined for **DS** are close to unity for all of the ligand concentrations investigated, indicating that nonspecific binding of **5** to bound and unbound **DS** during the ES process was negligible. This finding is consistent with the conclusion reached based on the similarity in the $K_{a,1}$ values. In the case of **SS**, at ligand concentrations ≤50 μM, the f_i ratios are also close to unity, indicating that nonspecific binding of **5** to the **SS** species did not contribute appreciably to the mass spectra. In contrast, at ligand concentrations >50 μM, the f_i ratios deviate from unity. This observation is indicative of nonspecific ligand binding at these higher concentrations, consistent with the conclusions drawn from the concentration dependence of the $K_{a,1}$ values.

Conclusions

The present study describes the first applications of the M_{rep} method for identifying the formation of nonspecific ligand binding in the ES-MS analysis of protein-ligand and DNA-ligand interactions in solution. The reliability and sensitivity of the M_{rep}

method for identifying the occurrence of nonspecific protein-carbohydrate interactions was highlighted by its application to cases where the protein-carbohydrate interactions detected by ES-MS originated exclusively from nonspecific association during the ES process, exclusively from specific interactions in solution, and from both specific and nonspecific interactions. These control experiments confirmed that the basic assumptions underlying the M_{rep} method, namely that nonspecific ligand binding is a random process and that the ES droplet histories for specific and nonspecific complexes are distinct, are generally valid. To demonstrate that the M_{rep} method can generally be used to identify nonspecific ligand binding it was used to monitor nonspecific DNA-ligand interactions in the ES-MS analysis of the sequential binding of the ethidium cation, a DNA intercalator, to single and double strand oligodeoxynucleotides. Using a trisaccharide reporter molecule, the M_{rep} method was shown to identify correctly the absence or presence of nonspecific DNA-ligand binding.

Acknowledgement

The authors acknowledge the Natural Sciences and Engineering Research Council of Canada and the Alberta Ingenuity Centre for Carbohydrate Science for funding and Professors D. Bundle and T. Lowary (University of Alberta) for providing the ligands used in this study.

References

1. Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.
2. Fandrich, M.; Tito, M. A.; Leroux, M. R.; Rostom, A. A.;Hartl, F.U.; Dobson, C. M.; Robinson, C. V. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 14151-14155.
3. Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1-27.
4. Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605-611.
5. van den Heuvel, R. H. H.; Heck, A. J. R. *Mass. Spec. Rev*. **2004**, *23*, 368-389.
6. Berhane, B.; Kaddis, C. S.; Wooding, K. M.; Xie, Y.; Kaufman, S. L.; Chernushevich, I. V.; Loo, J. A. *J. Am. Soc. Mass Spectrom*. **2005**, *16*, 998-1008.
7. Chitta, R. K.; Rempel, D. L.; Gross, M. L. *J. Am. Soc. Mass Spectrom*. **2005**, *16*, 1031-1038.
8. Fryčák, P.; Schug, K. A.; *Anal. Chem*. **2007**, *79*, 5407-5413.
9. Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, A. D.; Bundle, D. R; Klassen, J. S. *Glycobiology* **2008**, *18*, 587-592.
10. Wortmann, A.; Jecklin, M. C.; Touboul, D.; Badertscher, M.; Zenobi, R. *J. Mass Spectrom*. **2008**, *43*, 600-608.
11. Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem*. **2005**, *77*, 3060-3071.
12. Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem*. **2003**, *75*, 4945-4955.
13. Wang, W.; Kitova, E. N.; Sun, J.; Klassen, J. S. *J. Am. Soc. Mass Spectrom*. **2005**, *16*, 1583-1594.
14. Sun, J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem*. **2007**, *79*, 416-425.

15. Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem*. **2006**, *78*, 3010-3018.

16. Shoemaker, G. K.; Soya, N.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2008**, *18*, 587-592.

17. Sun, J.; Kitova, E. N.; Sun, N.; Klassen, J. S. *Anal. Chem*. **2007**, *79*, 8301-8311.

18. (a) Zdanov, A.; Bundle, D. R.; Deng, S.-J.; MacKenzie, C. R.; Narang, S. A.; Young, M. N.; Cygler, M. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 6423-6427. (b) Bundle, D. R.; Baumann, H.; Brisson, J. R.; Gagne, S. M.; Zdanov, A.; Cygler, M. *Biochemistry* **1994**, *33*, 5183-5192.

19. Bundle, D. R. Unpublished data.

20. Mazzitelli, C. L.; Chu, Y.; Reczek, J. J.; Iverson, B. L.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 311-321.

21. Pierce, S. E.; Sherman, C. L.; Jayawickramarajah, J.; Lawrence, C. M.; Sessler, J. L.; Brodbelt, J. S. *Analytica Chimica Acta*, **2008**, *627*, 129-135.

22. Rosu, F.; Nguyen, C.; De Pauw, E.; Gabelica, V. *J. Am. Soc. Mass Spectrom*. **2007**, *78*, 1052-1062.

23. Rosu, F.; De Pauw, E.; Gabelica, V. *Biochimie*. **2008**, *90*, 1074-1087.

24. Waring, M. J. *J. Mol. Biol*. **1965**, *13*, 269-282.

25. LePecq, J. B.; Paoletti, C. *J. Mol. Biol*. **1967**, *27*, 87-106.

26. Douthart, R. J.; Burnett, J. P.; Beasley, F. W.; Frank, B. H. *Biochemistry* **1973**, *12*, 214-220.

27. Nafisi, S.; Saboury, A. A.; Keramat, N.; Neasult, J. F.; Tajmir-Riahi, H. A.; *J. Mol. Structure*. **2007**, *827*, 35-43.

28. Vardevanyan, P. O.; Antonyan, A. P.; Parsadanyan, M. A.; Davtyan, H. G.; Karapetyan, A. T. *Experimental and Molecular Medicine*, **2003**, *35*, 527-533.

Table 1. Comparison of the K_a values measured by ES-MS for the (scFv + **1**) complex and the corresponding $f_{x,i}$ terms determined for the nonspecific binding of M_{rep} (≡ **2**) to $scFv^{n+}$ (x = P) and $(scFv + \mathbf{1})^{n+}$ (x = PL) ion during ES-MS analysis. [a-d]

[scFv] (μM)	[**1**] (μM)	K_a (10^5 M^{-1})	$f_{PL,0}/f_{P,0}$	$f_{PL,1}/f_{P,1}$
18	18	1.13 ± 0.17	1.01 ± 0.04	1.06 ± 0.04
9	9	1.03 ± 0.06	1.02 ± 0.02	1.07 ± 0.04
6	18	1.41 ± 0.20	0.89 ± 0.04	1.32 ± 0.10
9	18	1.45 ± 0.14	0.95 ± 0.10	1.22 ± 0.11

a. All measurements performed at 25 °C, pH 7. b. For all experiments the concentration of **2** was 47 μM. c. Ratios calculated from average $f_{x,i}$ values taken from 4 measurements. d. Errors correspond to one standard deviation.

Table 2. Comparison of the $K_{a,1}$ values determined for the (**SS** + **5**) and the (**DS** + **5**) complexes by ES-MS and the corresponding $f_{x,i}$ terms for the nonspecific binding of M_{rep} (≡ **1**) to $\mathbf{SS}^{z-}$ or $\mathbf{DS}^{z-}$ (x = DNA), (**SS** or **DS** + **5**)$^{z-}$ (x = DNAL) and (**SS** or **DS** + 2(**5**))$^{z-}$ ions (x = DNAL2). [a-d]

DNA	[**5**] (μM)	$K_{a,1}$ (10^4 M^{-1})	$f_{DNAL,0}/f_{DNA,0}$	$f_{DNAL,1}/f_{DNA,1}$	$f_{DNAL2,0}/f_{DNA,0}$	$f_{DNAL2,1}/f_{DNA,1}$
SS	20	1.15 ± 0.02	1.01 ± 0.01	1.09 ± 0.01		
SS	30	1.21 ± 0.04	1.00 ± 0.01	1.00 ± 0.01		
SS	40	1.15 ± 0.05	0.99 ± 0.01	1.03 ± 0.01		
SS	50	1.19 ± 0.04	1.00 ± 0.01	1.00 ± 0.01		
SS	60	1.44 ± 0.14	0.98 ± 0.01	1.10 ± 0.01	0.93 ± 0.01	1.21 ± 0.01
SS	70	1.51 ± 0.12	0.94 ± 0.01	1.29 ± 0.01	0.90 ± 0.01	1.41 ± 0.01
DS	50	4.59 ± 0.01	0.99 ± 0.01	1.04 ± 0.01		
DS	70	5.97 ± 0.01	1.01 ± 0.01	0.95 ± 0.01	1.00 ± 0.04	1.02 ± 0.10
DS	160	4 ± 1	1.01 ± 0.01	0.97 ± 0.01	1.08 ± 0.04	0.9 ± 0.1

a. All measurements performed at 25 °C, pH 7. b. The concentration of **SS** was 17 μM, **DS** was 34 μM and **1** was 83 μM. c. Ratios calculated from average $f_{x,i}$ values taken from 4 - 5 measurements. d. Errors correspond to one standard deviation.

Figure captions

Figure 1. Structures of the αTal[αAbe]αMan (**1**), 2-trimethylsilylethyl 4-O-[(4-O-α-D-galactopyranosyl)-β-D-galactopyranosyl]-β-D-glycopyranoside (**2**), 1-O-(α-D-glycopyranosyl)-α-D-glycopyranose (**3**), 4-O-[4-o-(α-D-glycopyranosyl)-α-D-glycopyranosyl]-α,β-D-glycopyranose (**4**), and ethidium bromide (**5**).

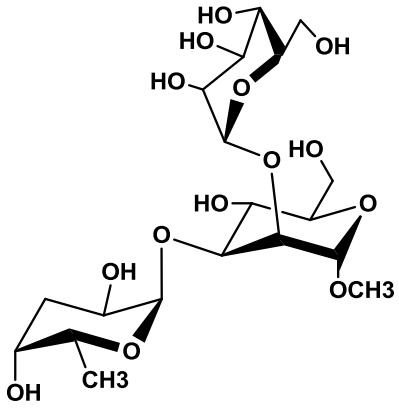
Figure 2. Simulated ES mass spectra illustrating the influence of nonspecific binding of M_{rep} molecules to a protein and its specific protein-ligand complex during the ES process. (a) Mass spectrum in case where only ions corresponding to free protein (P^{n+}) and specific protein-ligand complex (PL^{n+}) are present. (b) Mass spectrum resulting from the nonspecific binding M_{rep} to P and PL during the ES process. Notably, the distributions of M_{rep} bound to P^{n+} and PL^{n+} are identical, i.e., $f_{P,i} = f_{PL,i}$. (c) Mass spectrum resulting from the nonspecific binding M_{rep} and free L to P and PL during the ES process. In this case, the distributions of M_{rep} bound to P^{n+} and PL^{n+} are expected to be non-equivalent, i.e., $f_{P,i} \neq f_{PL,i}$.

Figure 3. Simulated ES mass spectra illustrating the influence of nonspecific binding of two non-interacting molecules, L and M_{rep}, to free protein during the ES process. (a) Mass spectrum in the case where only the free protein ion (P^{n+}) is present. (b) Mass spectrum resulting from the nonspecific binding of L and M_{rep} to P during the ES process. The distribution of PL_iM_{repj} species shown corresponds to the situation where L and M_{rep} bind nonspecifically to P with the identical efficiencies.

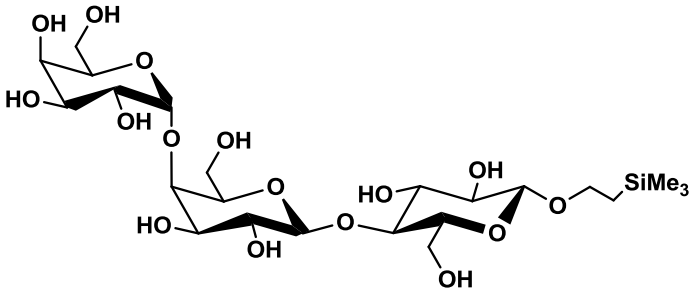
Figure 4. (a) ES mass spectrum obtained for a solution of 10 μM scFv, 0.15 mM **3** and 0.15 mM **4**. (b) Normalized distribution of scFv$\mathbf{3}_i\mathbf{4}_j$ species, as determined from the mass spectrum shown in (a).

Figure 5. ES mass spectra obtained for solutions of 47 μM **2** and (a) 9 μM scFv and 9 μM **1**, (b) 6 μM scFv and 18 μM **1**.

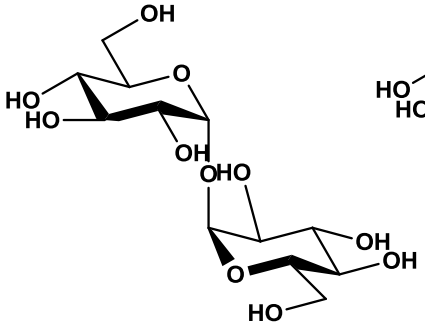
Figure 6. ES mass spectra obtained for solutions of 17 μM **SS**, 83 μM **1** and (a) 30 μM **5**, or (b) 70 μM **5**. The peaks labeled with ■ and ♦ correspond to ODN 19- and 18-mer ions produced by the loss of C and CG, respectively, from **SS** and their 1:1 complexes with **1**. The peak labeled with ● corresponds to the deprotonated cluster ion, $((2)\mathbf{1} + \mathbf{5})^-$. ES mass spectra obtained for solutions of 34 μM **DS**, 83 μM **1** with (c) 70 μM **5**, or (d) 160 μM **5**.



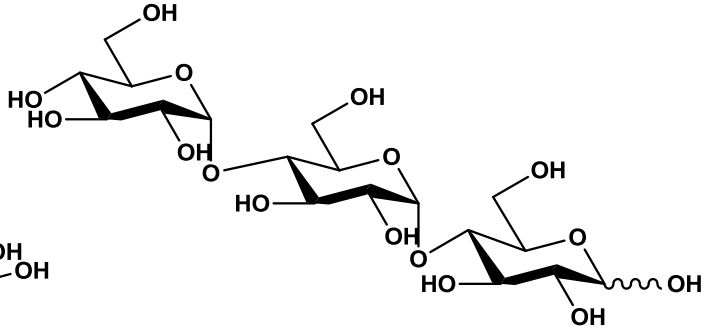
1



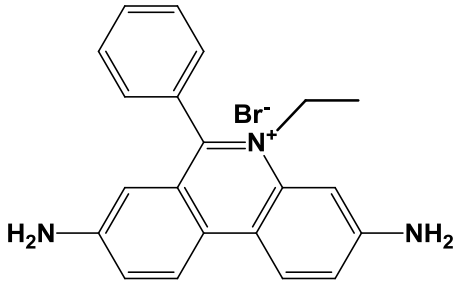
2



3



4



5

Figure 1

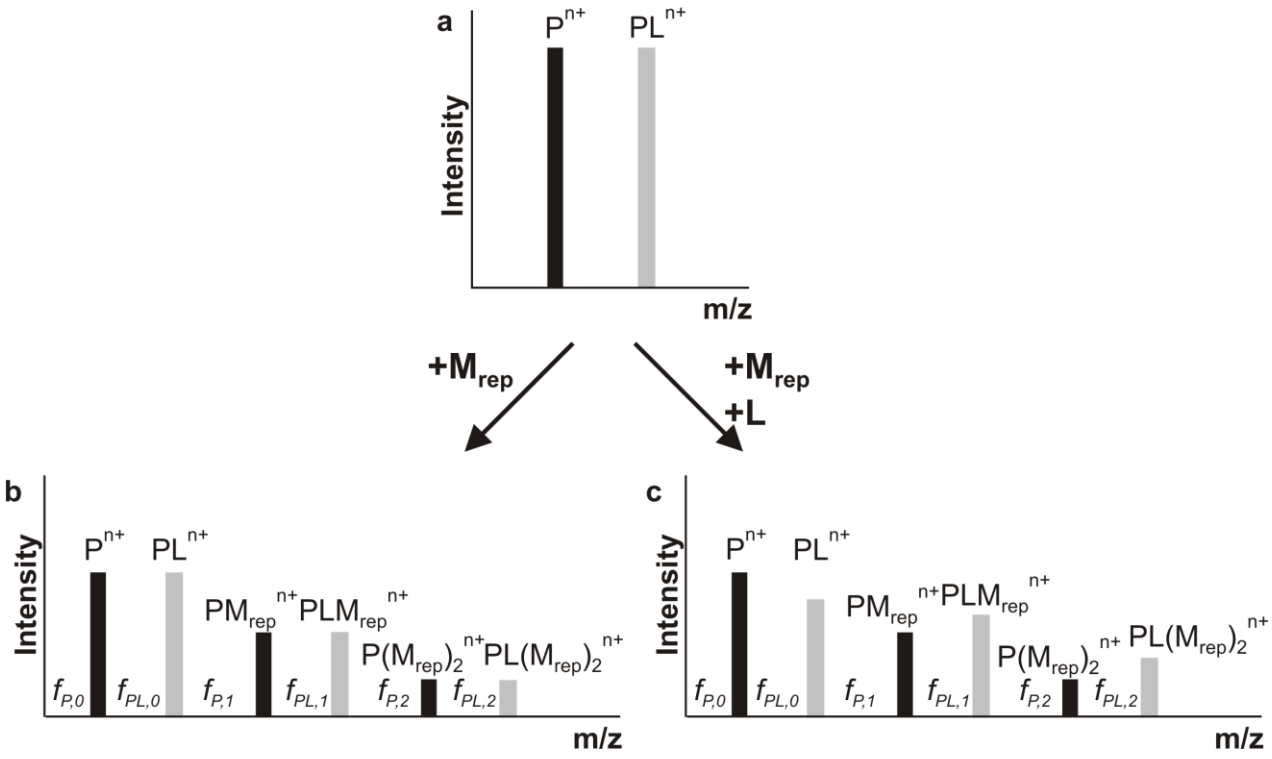

a
P^{n+}
PL^{n+}
Intensity
m/z
$+M_{rep}$
$+M_{rep}$
$+L$
b
Intensity
P^{n+}
PL^{n+}
PM_{rep}^{n+}
PLM_{rep}^{n+}
$P(M_{rep})_2^{n+}$
$PL(M_{rep})_2^{n+}$
$f_{P,0}$
$f_{PL,0}$
$f_{P,1}$
$f_{PL,1}$
$f_{P,2}$
$f_{PL,2}$
m/z
c
Intensity
P^{n+}
PL^{n+}
PM_{rep}^{n+}
PLM_{rep}^{n+}
$P(M_{rep})_2^{n+}$
$PL(M_{rep})_2^{n+}$
$f_{P,0}$
$f_{PL,0}$
$f_{P,1}$
$f_{PL,1}$
$f_{P,2}$
$f_{PL,2}$
m/z

Figure 2

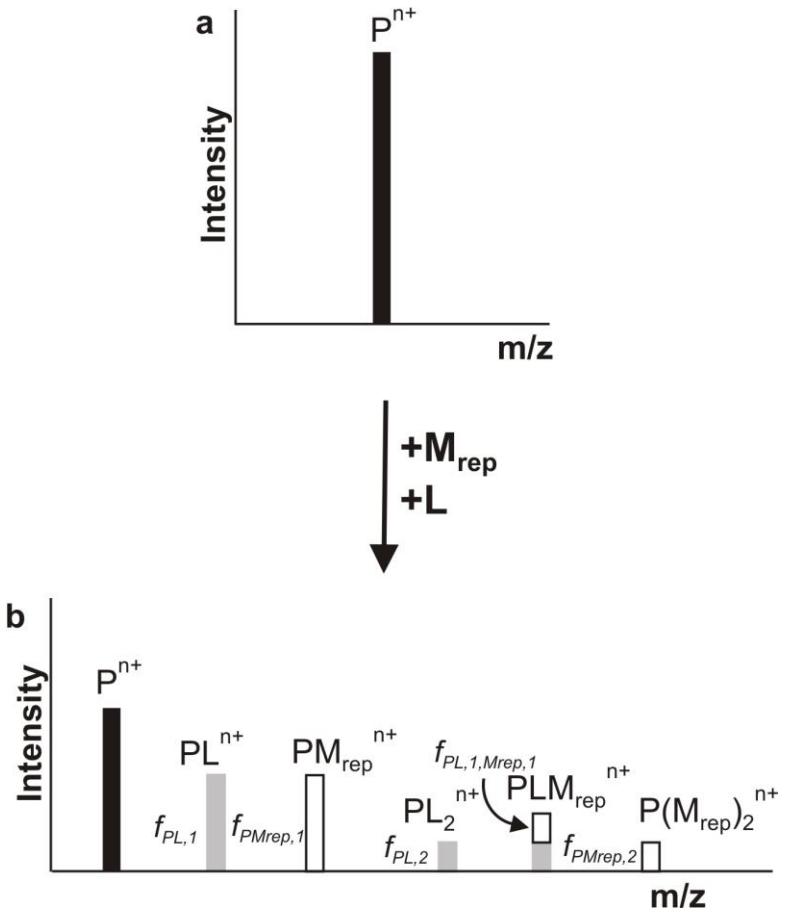

a
P^{n+}
Intensity
m/z
+M_{rep}
+L
b
Intensity
P^{n+}
PL^{n+}
PM_{rep}^{n+}
$f_{PL,1,Mrep,1}$
PLM_{rep}^{n+}
PL_2^{n+}
$P(M_{rep})_2^{n+}$
$f_{PL,1}$
$f_{PMrep,1}$
$f_{PL,2}$
$f_{PMrep,2}$
m/z

Figure 3

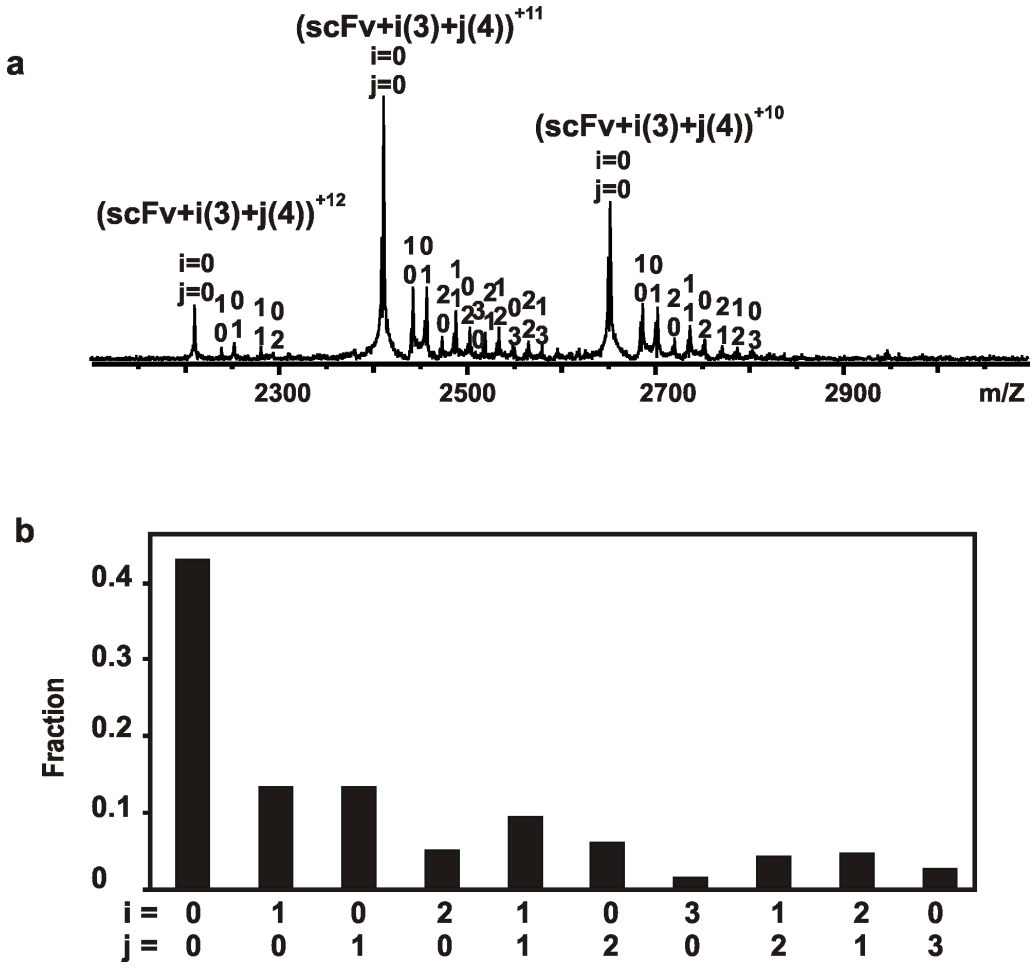

a
(scFv+i(3)+j(4))+11
i=0
j=0
(scFv+i(3)+j(4))+10
i=0
j=0
(scFv+i(3)+j(4))+12
i=0
j=0
2300
2500
2700
2900
m/Z
b
Fraction
0.4
0.3
0.2
0.1
0
i = 0 1 0 2 1 0 3 1 2 0
j = 0 0 1 0 1 2 0 2 1 3

Figure 4

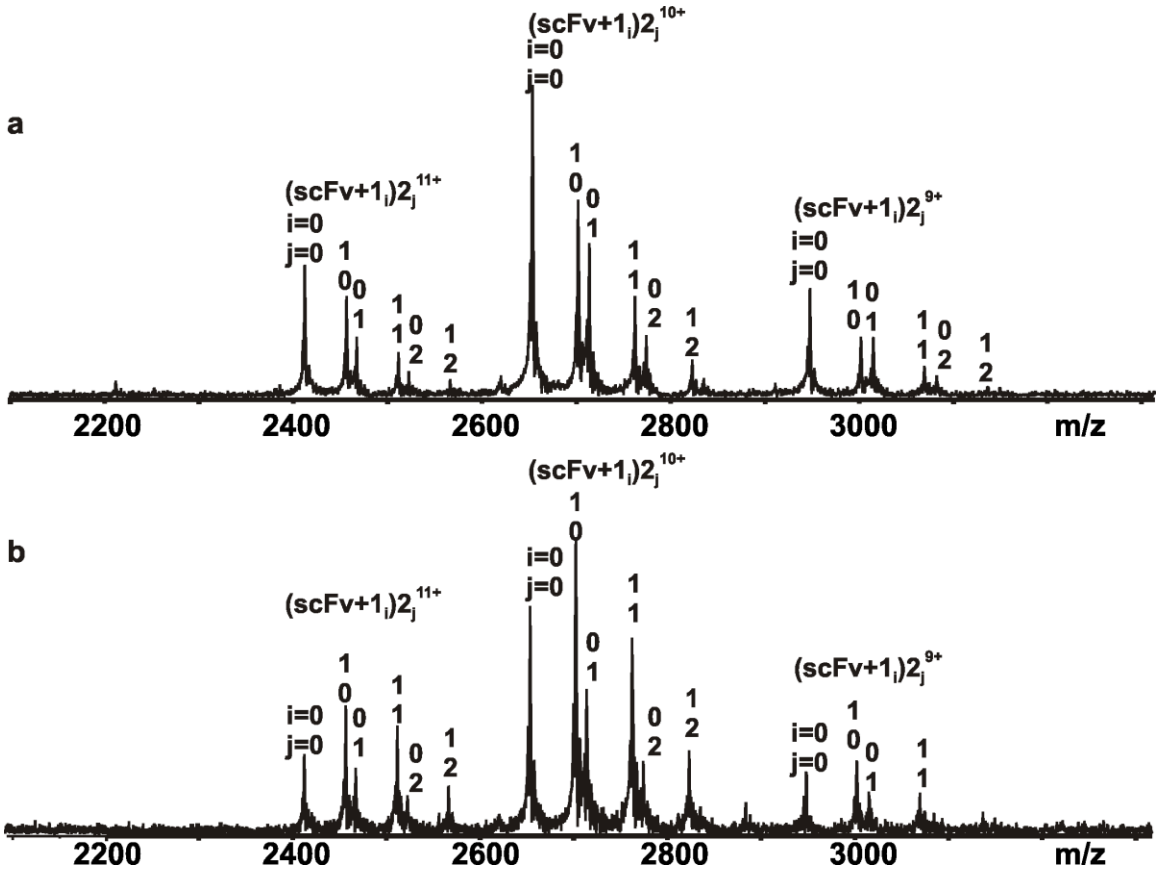

a
(scFv+1i)2j 10+
i=0
j=0
(scFv+1i)2j 11+
i=0
j=0
(scFv+1i)2j 9+
i=0
j=0
2200
2400
2600
2800
3000
m/z
b
(scFv+1i)2j 10+
i=0
j=0
(scFv+1i)2j 11+
i=0
j=0
(scFv+1i)2j 9+
i=0
j=0
2200
2400
2600
2800
3000
m/z

Figure 5

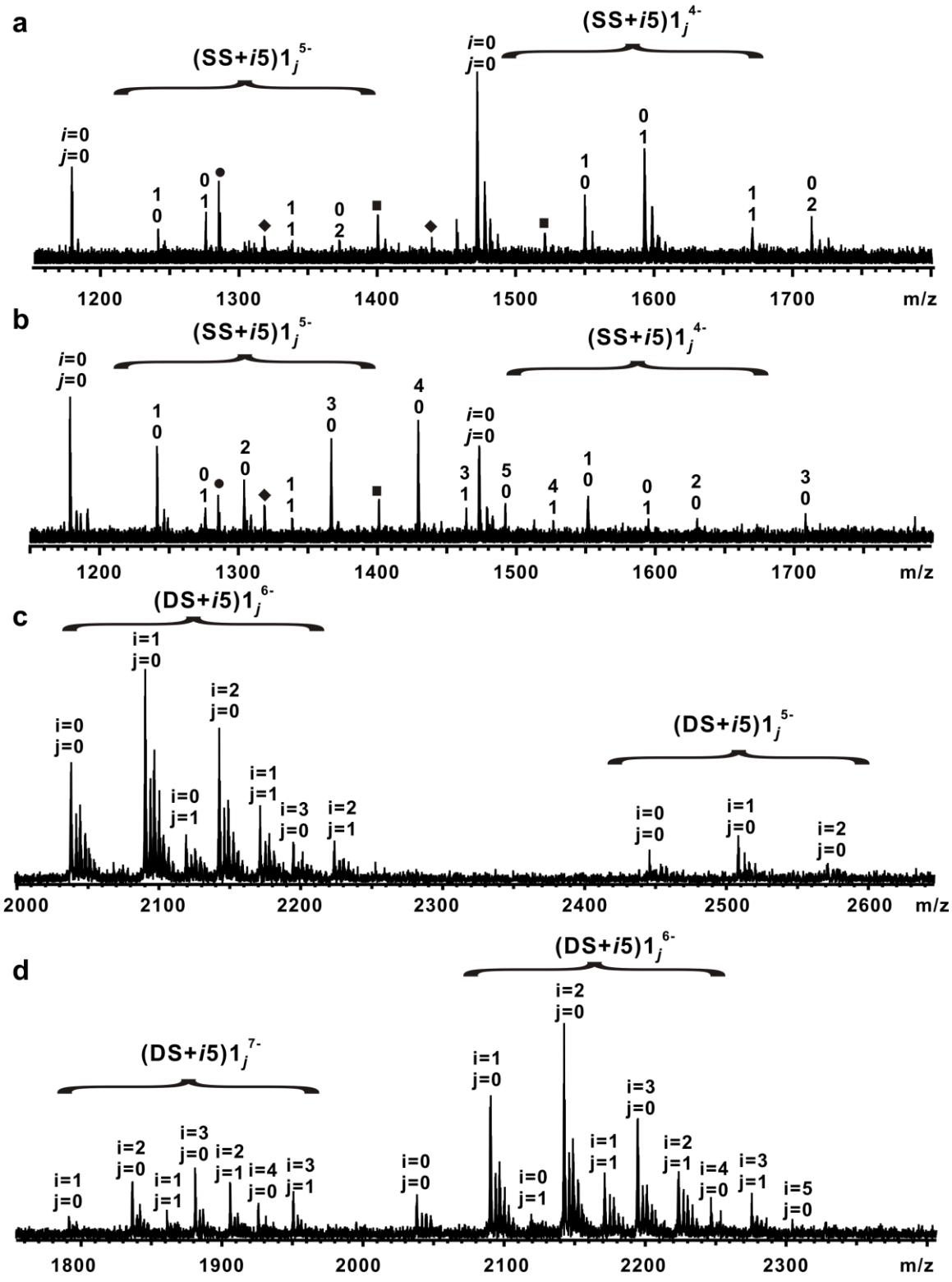

a
(SS+i5)1j 5-
(SS+i5)1j 4-
i=0 j=0
1200 1300 1400 1500 1600 1700 m/z
b
(SS+i5)1j 5-
(SS+i5)1j 4-
i=0 j=0
1200 1300 1400 1500 1600 1700 m/z
c
(DS+i5)1j 6-
(DS+i5)1j 5-
i=1 j=0
i=2 j=0
i=0 j=0
i=0 j=1
i=1 j=1
i=3 j=0
i=2 j=1
2000 2100 2200 2300 2400 2500 2600 m/z
d
(DS+i5)1j 6-
(DS+i5)1j 7-
i=2 j=0
i=1 j=0
i=3 j=0
i=1 j=1
i=2 j=1
i=4 j=0
i=3 j=1
i=0 j=0
i=0 j=1
i=3 j=1
i=5 j=0
1800 1900 2000 2100 2200 2300 m/z

Figure 6

Table S1. Summary of protein ions identified in Figure 4a.

m/z	Intensity/10^6	Ion
2212	11.2	$(scFv)^{+12}$
2241	2.79	$(scFv+(3))^{+12}$
2254	3.66	$(scFv+(4))^{+12}$
2283	3.02	$(scFv+(3)+(4))^{+12}$
2296	1.83	$(scFv+2(4))^{+12}$
2413	52.6	$(scFv)^{+11}$
2444	14.7	$(scFv+(3))^{+11}$
2459	14.8	$(scFv+(4))^{+11}$
2475	4.99	$(scFv+2(3))^{+11}$
2490	10.2	$(scFv+(3)+(4))^{+11}$
2505	6.77	$(scFv+2(4))^{+11}$
2507	2.85	$(scFv+3(3))^{+11}$
2521	5.76	$(scFv+2(3)+(4))^{+11}$
2536	6.88	$(scFv+(3)+2(4))^{+11}$
2551	2.79	$(scFv+3(4))^{+11}$
2567	3.99	$(scFv+2(3)+2(4))^{+11}$
2582	3.27	$(scFv+(3)+3(4))^{+11}$
2654	31.6	$(scFv)^{+10}$
2689	11.5	$(scFv+(3))^{+10}$
2705	10.8	$(scFv+(4))^{+10}$
2723	4.65	$(scFv+2(3))^{+10}$
2739	7.18	$(scFv+(3)+(4))^{+10}$
2755	4.47	$(scFv+2(4))^{+10}$
2773	3.13	$(scFv+2(3)+(4))^{+10}$
2789	2.81	$(scFv+(3)+2(4))^{+10}$
2806	2.31	$(scFv+3(4))^{+10}$

Table S2. Summary of protein ions identified in Figure 5a.

m/z	Intensity/10^5	Ion
2414	27.8	$(scFv)^{+11}$
2458	21.2	$(scFv+(1))^{+11}$
2469	12.8	$(scFv+(2))^{+11}$
2513	9.5	$(scFv+(1)+(2))^{+11}$
2524	5.5	$(scFv+2(2))^{+11}$
2568	3.8	$(scFv+(1)+2(2))^{+11}$
2655	65.6	$(scFv)^{+10}$
2704	41.4	$(scFv+(1))^{+10}$
2715	32.5	$(scFv+(2))^{+10}$
2764	21.2	$(scFv+(1)+(2))^{+10}$
2776	13.3	$(scFv+2(2))^{+10}$
2824	7.9	$(scFv+(1)+2(2))^{+10}$
2950	22.9	$(scFv)^{+9}$
3004	12.9	$(scFv+(1))^{+9}$
3017	12.6	$(scFv+(2))^{+9}$
3071	6.6	$(scFv+(1)+(2))^{+9}$
3084	4.6	$(scFv+2(2))^{+9}$
3138	3.2	$(scFv+(1)+2(2))^{+9}$

Table S3. Summary of protein ions identified in Figure 5b.

m/z	Intensity/10^5	Ion
2414	7.5	$(scFv)^{+11}$
2458	10.8	$(scFv+(1))^{+11}$
2469	5.5	$(scFv+(2))^{+11}$
2513	9.9	$(scFv+(1)+(2))^{+11}$
2524	3.2	$(scFv+2(2))^{+11}$
2568	4.0	$(scFv+(1)+2(2))^{+11}$
2655	22.9	$(scFv)^{+10}$
2704	28.1	$(scFv+(1))^{+10}$
2715	14.5	$(scFv+(2))^{+10}$
2764	19.7	$(scFv+(1)+(2))^{+10}$
2776	6.3	$(scFv+2(2))^{+10}$
2824	7.7	$(scFv+(1)+2(2))^{+10}$
2950	6.9	$(scFv)^{+9}$
3004	7.5	$(scFv+(1))^{+9}$
3017	4.1	$(scFv+(2))^{+9}$
3071	4.8	$(scFv+(1)+(2))^{+9}$

Table S4. Summary of DNA ions identified in Figure 6a.

m/z	Intensity/10^5	Ion
1178	7.32	$(SS)^{5-}$
1240	2.52	$(SS+(5))^{5-}$
1275	3.24	$(SS+(1))^{5-}$
1338	1.82	$(SS+(5)+(1))^{5-}$
1372	1.79	$(SS+2(1))^{5-}$
1473	11.42	$(SS)^{4-}$
1551	3.73	$(SS+(5))^{4-}$
1594	7.14	$(SS+(1))^{4-}$
1673	1.90	$(SS+(5)+(1))^{4-}$
1715	1.96	$(SS+2(1))^{4-}$

Table S5. Summary of DNA ions identified in Figure 6b.

m/z	Intensity/10^5	Ion
1178	10.61	$(SS)^{5-}$
1240	6.28	$(SS+(5))^{5-}$
1275	2.50	$(SS+(1))^{5-}$
1304	3.99	$(SS+2(5))^{5-}$
1338	1.96	$(SS+(5)+(1))^{5-}$
1367	5.62	$(SS+3(5))^{5-}$
1430	6.85	$(SS+4(5))^{5-}$
1464	1.93	$(SS+3(5)+(1))^{5-}$
1493	2.28	$(SS+5(5))^{5-}$
1527	1.44	$(SS+4(5)+(1))^{5-}$
1473	7.19	$(SS)^{4-}$
1551	3.24	$(SS+(5))^{4-}$
1594	1.62	$(SS+(1))^{4-}$
1630	1.79	$(SS+2(5))^{4-}$
1709	1.45	$(SS+3(5))^{4-}$

Table S6. Summary of DNA ions identified in Figure 6c.

m/z	Intensity/10^5	Ion
2038	11.45	$(DS)^{6-}$
2090	19.22	$(DS+(5))^{6-}$
2119	4.71	$(DS+(1))^{6-}$
2143	13.52	$(DS+2(5))^{6-}$
2172	6.91	$(DS+(5)+(1))^{6-}$
2196	4.25	$(DS+3(5))^{6-}$
2225	4.27	$(DS+2(5)+(1))^{6-}$
2447	2.99	$(DS)^{5-}$
2510	4.21	$(DS+(5))^{5-}$
2572	1.07	$(DS+2(5))^{5-}$

Table S7. Summary of DNA ions identified in Figure 6d.

m/z	Intensity/10^5	Ion
1792	1.29	$(DS+(5))^{7-}$
1837	3.59	$(DS+2(5))^{7-}$
1861	1.39	$(DS+(5)+(1))^{7-}$
1882	3.46	$(DS+3(5))^{7-}$
1907	2.60	$(DS+2(5)+(1))^{7-}$
1927	2.12	$(DS+4(5))^{7-}$
1952	2.64	$(DS+3(5)+(1))^{7-}$
2038	2.43	$(DS)^{6-}$
2090	7.68	$(DS+(5))^{6-}$
2119	1.55	$(DS+(1))^{6-}$
2143	11.41	$(DS+2(5))^{6-}$
2172	3.41	$(DS+(5)+(1))^{6-}$
2196	7.16	$(DS+3(5))^{6-}$
2225	5.10	$(DS+2(5)+(1))^{6-}$
2249	2.72	$(DS+4(5))^{6-}$
2277	3.90	$(DS+3(5)+(1))^{6-}$
2301	2.98	$(DS+5(5))^{6-}$